Ramsey numbers of sparse digraphs

Jacob Fox* Xiaoyu He† Yuval Wigderson‡

Abstract

Burr and Erdős in 1975 conjectured, and Chvátal, Rödl, Szemerédi and Trotter later proved, that the Ramsey number of any bounded degree graph is linear in the number of vertices. In this paper, we disprove the natural directed analogue of the Burr–Erdős conjecture, answering a question of Bucić, Letzter, and Sudakov. If H is an acyclic digraph, the *oriented Ramsey number* of H, denoted $\overrightarrow{r_1}(H)$, is the least N such that every tournament on N vertices contains a copy of H. We show that for any $\Delta \geq 2$ and any sufficiently large n, there exists an acyclic digraph H with n vertices and maximum degree Δ such that

$$\overrightarrow{r_1}(H) \geq n^{\Omega(\Delta^{2/3}/\log^{5/3}\Delta)}.$$

This proves that $\overrightarrow{r_1}(H)$ is not always linear in the number of vertices for bounded-degree H. On the other hand, we show that $\overrightarrow{r_1}(H)$ is nearly linear in the number of vertices for typical bounded-degree acyclic digraphs H, and obtain linear or nearly linear bounds for several natural families of bounded-degree acyclic digraphs.

For multiple colors, we prove a quasi-polynomial upper bound $\overrightarrow{r_k}(H) = 2^{(\log n)^{O_k(1)}}$ for all bounded-degree acyclic digraphs H on n vertices, where $\overrightarrow{r_k}(H)$ is the least N such that every k-edge-colored tournament on N vertices contains a monochromatic copy of H. For $k \geq 2$ and $n \geq 4$, we exhibit an acyclic digraph H with n vertices and maximum degree 3 such that $\overrightarrow{r_k}(H) \geq n^{\Omega(\log n/\log\log n)}$, showing that these Ramsey numbers can grow faster than any polynomial in the number of vertices.

1 Introduction

The k-color Ramsey number of a (simple undirected) graph H, denoted $r_k(H)$, is the minimum N such that every k-edge-coloring of the complete graph K_N contains a monochromatic copy of H. Broadly speaking, the main question in graph Ramsey theory is to understand how $r_k(H)$ depends on H and k. The most well-studied case is that of two colors, $k = 2$. For H on n vertices, it is known [10, 36] that $r_2(H)$ grows exponentially in n if and only if H has $\Omega(n^2)$ edges.

However, it has long been observed that the Ramsey number of a sparse graph H is much smaller than exponential in $|V(H)|$. In their foundational paper on the topic, Burr and Erdős [6] conjectured that this phenomenon is quite general and that any sparse graph has linear Ramsey number. Here, the appropriate notion of sparsity is *degeneracy:* H is said to be d-degenerate if every subgraph of H has a vertex of degree at most d, and the degeneracy of H is the minimum d such that H is d-degenerate. Burr and Erdős conjectured that $r_k(H) = O_{k,d}(n)$ for any n-vertex graph H with degeneracy d. Here and throughout we use the standard asymptotic notation where the implicit constant is allowed to depend only on the subscripts of $O(\cdot)$. Major progress towards this conjecture was made by Chvátal, Rödl, Szemerédi, and Trotter [8], who proved the Burr–Erdős conjecture under the stronger assumption that H has bounded degree (rather than bounded degeneracy), that is, that $r_k(H) = O_{k,\Delta}(n)$ for any n-vertex graph H with maximum degree Δ. Finally,

*Department of Mathematics, Stanford University, Stanford, CA 94305, USA. Email: jacobfox@stanford.edu. Research supported by a Packard Fellowship and by NSF award DMS-185563.

†Department of Mathematics, Stanford University, Stanford, CA 94305, USA. Email: alkjash@stanford.edu. Research supported by NSF GRFP Grant DGE-1656518.

‡Department of Mathematics, Stanford University, Stanford, CA 94305, USA. Email: yuvalwig@stanford.edu. Research supported by NSF GRFP Grant DGE-1656518.

building on many prior developments (e.g. [19, 28]), the full Burr–Erdős conjecture was proved by Lee [30] in 2017.

There are many analogous questions and results for directed graphs (henceforth *digraphs*). We assume all digraphs are simple and oriented, so they do not contain self-loops, parallel edges or anti-parallel edges. For a digraph H, define the k-color *oriented Ramsey number* $\overrightarrow{r_k}(H)$ to be the minimum N such that any k-edge-colored tournament on N vertices contains a monochromatic copy of H. Note that if H contains a directed cycle, then H does not appear in any transitive tournament, so $\overrightarrow{r_k}(H)$ only exists for acyclic H. Henceforth, we work exclusively with acyclic digraphs H.

Unlike undirected Ramsey numbers, oriented Ramsey numbers are interesting even in the case of one color, $k = 1$, where $\overrightarrow{r_1}(H)$ is simply the minimum N such that any tournament on N vertices contains a copy of H. Let TT_n denote the transitive tournament on n vertices. The study of oriented Ramsey numbers was initiated by Stearns [35] in 1959 and Erdős and Moser [18] in 1964, who showed the upper and lower bounds, respectively, in

$$2^{n/2-1} \le \overrightarrow{r_1}(\mathrm{TT}_n) \le 2^{n-1}. \tag{1.1}$$

The exponential constants in these bounds have not been improved, similar to the classical case of the diagonal undirected Ramsey number $r_2(K_n)$ [10, 33, 34]. This may not be surprising, given that $\overrightarrow{r_1}(\mathrm{TT}_n) \le r_2(K_n)$. Thus, improving the lower bound in (1.1) is at least as difficult as improving the lower bound on $r_2(K_n)$.

Somewhat more is known about the oriented Ramsey number $\overrightarrow{r_k}(H)$ when H is sparse. When $H = P_n$ is the directed path on n vertices, Chvátal [7] and Gyárfás and Lehel [22] determined that $\overrightarrow{r_k}(P_n) = (n-1)^k+1$ using the Gallai–Hasse–Roy–Vitaver theorem [20, 24, 32, 38]. In the case of one color, it was more generally conjectured by Sumner in 1971 that for any oriented tree T on $n \ge 2$ vertices, $\overrightarrow{r_1}(T) \le 2n-2$. Sumner's conjecture has received a considerable amount of attention over the years (see e.g. [15, 16, 23, 25, 26, 29, 37]); it was proven for n sufficiently large by Kühn, Mycroft, and Osthus [29], and Dross and Havet [15] showed that $\overrightarrow{r_1}(T) \le \frac{21}{8}n - \frac{47}{16}$ for all $n \ge 2$. In more colors, it was shown by Bucić, Letzter, and Sudakov [4] that $\overrightarrow{r_k}(T) = O_k(|V(T)|^k)$ for any oriented tree T and all $k \ge 1$. In the same paper, they asked a natural directed analogue of the classical Burr–Erdős problem.

Problem 1.1 ([4])**.** Is it true that $\overrightarrow{r_1}(H) = O_\Delta(n)$ for every acyclic digraph H with n vertices and maximum degree Δ?

Here we write $N^+(v)$ for the out-neighborhood and $N^-(v)$ for the in-neighborhood of a vertex $v \in V(H)$, and say that a digraph H has maximum degree Δ if $\max_{v\in V(H)}(|N^+(v)| + |N^-(v)|) = \Delta$.

Yuster [40] recently initiated the study of the special case of Problem 1.1 when $H = P_n^\ell$ is the ℓ-th power of a directed path P_n, which is the digraph on vertex set $[n]$ whose edges are the ordered pairs (i, j) satisfying $1 \le j - i \le \ell$. This case was recently settled by Draganić et al. [14], who showed that $\overrightarrow{r_1}(P_n^\ell) = O_\ell(n)$. Letting the *bandwidth* of an acyclic digraph H on n vertices be the minimum ℓ such that P_n^ℓ contains a copy of H, this aforementioned result implies that $\overrightarrow{r_1}(H) = O_\ell(n)$ if H has n vertices and bandwidth ℓ.

In this paper, we answer Problem 1.1 in the negative, and show that in fact $\overrightarrow{r_1}(H)$ can grow faster than any fixed power of n, as long as the maximum degree is a sufficiently large constant.

Theorem 1.2. *For any $\Delta \ge 2$ and n sufficiently large in terms of Δ, there exists an acyclic digraph H on n vertices and maximum degree at most Δ for which*

$$\overrightarrow{r_1}(H) \ge n^{\Omega(\Delta^{2/3}/\log^{5/3}\Delta)}.$$

The power $\Delta^{2/3}$ seems to be best possible with our method, but the power of $\log\Delta$ can be improved. Although the answer to Problem 1.1 is negative, we prove an almost linear upper bound on $\overrightarrow{r_1}(H)$ for almost all H, in the following sense. Define $\overrightarrow{G}(n, d)$ to be the orientation of the random regular graph $G(n, d)$ on vertex set $[n]$ with all edges pointing to the right.

Theorem 1.3. *If $d \geq 2$ is fixed and $H = \overrightarrow{G}(n,d)$, then w.h.p.*[1] *(as $n \to \infty$)*

$$\overrightarrow{r_1}(H) \leq n(\log n)^{4\log d}.$$

It is not difficult to extend Theorem 1.3 to show $\overrightarrow{r_1}(H) = n(\log n)^{O_\Delta(1)}$ w.h.p. if H is the forward acyclic orientation of a uniformly random graph with any fixed degree sequence $\Delta = d_1 \geq d_2 \geq \cdots \geq d_n$, and therefore also for the forward acyclic orientation of a uniformly random bounded-degree graph. We also prove a similar bound $\overrightarrow{r_1}(H) \leq n(\log n)^{O_d(1)}$ when $H = \overrightarrow{G}(n,p)$ is the forward acyclic orientation of an Erdős–Rényi random graph of constant average degree $d = pn$.

Although we are able to show that $\overrightarrow{r_1}(H)$ is w.h.p. almost linear for a random bounded-degree acyclic digraph, we have not determined the worst-case behavior of this Ramsey number. For general acyclic digraphs H on n vertices with maximum degree Δ, the best upper bound we are able to prove is $\overrightarrow{r_1}(H) \leq n^{O_\Delta(\log n)}$ (see Theorem 1.7 below). Nonetheless, we are able to prove stronger (and in some cases linear) upper bounds on $\overrightarrow{r_1}(H)$ in case H lies in certain natural families. We now give two examples of such families.

Let the *height* of H be the number of vertices on the longest directed path in H. Equivalently, the height can be seen as a directed analogue of the chromatic number: H has height at most h if and only if $V(H)$ can be partitioned into independent sets $S_1, \ldots, S_h$ such that every edge between S_i and S_j is oriented from S_i to S_j, for every $i < j$. For acyclic digraphs of bounded height and bounded degree, we are able to prove the following linear bound on $\overrightarrow{r_1}(H)$.

Theorem 1.4. *If H is an acyclic digraph on n vertices with maximum degree Δ and height h, then*

$$\overrightarrow{r_1}(H) \leq (\Delta h)^{10\Delta \log h} n.$$

In particular, $\overrightarrow{r_1}(H) = O_{h,\Delta}(n)$.

Note that this theorem also implies the aforementioned $n^{O_\Delta(\log n)}$ upper bound on $\overrightarrow{r_1}(H)$ for any bounded-degree acyclic digraph H, since the height of an acyclic digraph is at most its vertex count.

Next, we say that an acyclic digraph H of height h is *graded* if its vertex set can be partitioned into h independent sets $S_1, \ldots, S_h$ such that every edge in H is directed from some S_i to S_{i+1}. Equivalently, H is graded if for every pair of vertices (u,v), all directed paths from u to v have the same length (the equivalence of the definitions follows e.g. from [31, Proposition 4.4]). A natural example of a graded digraph is a grid (in any dimension) with all edges oriented towards the first orthant. In general, a graded digraph can be obtained from any graded lattice (in the sense of partially ordered sets) L by orienting every edge $x < y$ of the Hasse diagram of L from x to y. For graded digraphs of bounded degree, we are able to prove a polynomial bound on $\overrightarrow{r_1}(H)$.

Theorem 1.5. *If H is a graded digraph on n vertices with maximum degree Δ and height h, then*

$$\overrightarrow{r_1}(H) \leq h^{10\Delta \log \Delta} n.$$

In particular, since $h \leq n$, we have that $\overrightarrow{r_1}(H) \leq n^{O(\Delta \log \Delta)}$.

Our methods are motivated by those used by Conlox, Fox, Lee, and Sudakov [11] to prove bounds on ordered Ramsey numbers, and the two problems are especially closely related when the number of colors is at least 2. Using this connection, we are able to give a super-polynomial lower bound for $\overrightarrow{r_k}(H)$ when $k \geq 2$.

Theorem 1.6. *For any $n \geq 4$, there exists an acyclic digraph H on n vertices with maximum degree 3 for which*

$$\overrightarrow{r_k}(H) \geq n^{\log n / 20 \log \log n}$$

for all $k \geq 2$.

Thus, for acyclic digraphs of bounded degree, $\overrightarrow{r_k}(H)$ can grow super-polynomially if $k \geq 2$. In the other direction, for any number of colors we have the following quasi-polynomial upper bound.

[1] As usual, we say that an event $\mathcal{E}$ happens *with high probability (w.h.p.)* if $\Pr(\mathcal{E}) \to 1$ as $n \to \infty$.

Theorem 1.7. *If $k \geq 1$ and H is any acyclic digraph with n vertices and maximum degree Δ, then*

$$\overrightarrow{r_k}(H) \leq 2^{O_{k,\Delta}\left((\log n)^{2^{2k-1}}\right)}.$$

For one color, there is still a gap between the polynomial lower bound and the super-polynomial upper bound.

We remark that there is another well-studied analogue of ordinary Ramsey numbers in the directed setting, namely the *directed Ramsey number* $\overleftrightarrow{r_k}(H)$, introduced by Bermond [2]. This is defined as the minimum N such that a monochromatic copy of H exists in every k-coloring of the edges of $\overleftrightarrow{K_N}$, the digraph with edges in both directions between all pairs of vertices. There are similarities and differences between the two theories (see e.g. [4]), and several of our results on oriented Ramsey numbers can be extended to the setting of directed Ramsey numbers. We expand on these connections in our concluding remarks, Section 6.

To conclude this introduction, we remark on an interesting phenomenon brought to light by our results. An important "meta-question" driving many advances in Ramsey theory asks which graph parameters roughly determine a Ramsey number. In the Ramsey theory of undirected graphs, this question has been more or less resolved, at least in a qualitative sense. Namely, for an undirected graph H, the degeneracy and the number of vertices are the main parameters that determine the growth of $r_2(H)$ (we focus on the two-color case to be concrete). This is easiest to see in the lower bounds: if H has n vertices, then certainly $r_2(H) \geq n$. Additionally, if H has degeneracy d, then it is a simple exercise to show that a random 2-edge-coloring on $2^{d/2}$ vertices does not contain a monochromatic copy of H w.h.p., implying that $r_2(H) \geq 2^{d/2}$. Putting these two facts together, we find that

$$\log r_2(H) = \Omega(d + \log n).$$

Conlon, Fox, and Sudakov [12, Conjecture 2.16] conjectured that this bound is tight up to the implicit constant, namely that

$$\log r_2(H) = \Theta(d + \log n)$$

for any graph H with n vertices and degeneracy d. Moreover, this conjecture is known to be true up to a multiplicative factor of $\log^2 d$ [19, Theorem 3.1]. Because of these results, one can say that the degeneracy and the vertex count of H roughly determine its Ramsey number.

For acyclic digraphs, we do not know what parameters determine the growth order of $\overrightarrow{r_1}(H)$ (indeed, we don't even know if $\overrightarrow{r_1}(H)$ is polynomial or super-polynomial in n when H has bounded degree). Nonetheless, our results indicate that one parameter, which we call "multiscale complexity," is relevant. Namely, suppose we order the vertices of H as $v_1, \ldots, v_n$ so that every edge is oriented to the right, that is $v_i \to v_j$ is an edge only if $i < j$ (such an ordering is often called a *topological sorting* of H). Under this ordering we may assign every edge $v_i \to v_j$ a *length* $j - i$. Here, if $u, v \in V(H)$, we write $u \to v$ to signify that there is a directed edge from u to v, and similarly $u \leftarrow v$ for an edge in the other direction.

In graphs of bounded bandwidth, every edge is short and has length $O(1)$. At the other extreme, if H has bounded height, then most edges of H are long and have length $\Omega(n)$. The same is true in the random model $\overrightarrow{G}(n, d)$, where most edges have length $\Omega(n)$. Some other acyclic digraphs have intermediate edge-length statistics, such as the *directed grid* whose vertex set is $[\sqrt{n}]^2$ and all edges are oriented towards the lexicographically larger ordered pair. For every acyclic ordering of such a grid, there are many edges in most dyadic length scales between 1 and $\Omega(\sqrt{n})$.

Loosely, let us say that H has *high multiscale complexity* if, for most dyadic intervals $I_t = [2^t, 2^{t+1})$ with $0 \leq t \leq \log n$, there are many edges in H whose length is in I_t. Conversely, if most edge lengths of H are concentrated in a small number of dyadic intervals, then we loosely say that H has *low multiscale complexity*. At a high level, all of our upper bound results prove upper bounds on $\overrightarrow{r_1}(H)$ in terms of n, Δ, and the multiscale complexity of H; if the multiscale complexity is low, then these bounds are stronger, and one can prove linear, nearly linear, or polynomial bounds on $\overrightarrow{r_1}(H)$ (depending on the precise context). Conversely, our lower bound construction for Theorem 1.2 is a family of digraphs which we call *interval meshes*. We delay the precise definition to Section 2, but interval meshes are in some sense designed to maximize multiscale complexity among all graphs of maximum degree at most Δ.

We stress that we have not fully solved the problem of which natural parameters determine the growth order of $\overrightarrow{r_1}(H)$, and we think this problem deserves further research. Nonetheless, our results do make it clear that some notion of multiscale complexity is one of these parameters, and we believe this notion is fundamental. As such, we state and prove many of our technical upper bounds in greater generality than is needed to deduce our main theorems, in order to demonstrate how notions of multiscale complexity arise naturally from our techniques.

The rest of the paper is laid out as follows. Section 2 carries out the construction of interval meshes to prove the lower bound Theorem 1.2. Section 3 uses the greedy embedding technique to prove the main technical lemmas needed for the upper bounds Theorems 1.3, 1.4, and 1.5. Section 4 completes the proofs of these results, as well as a more general upper bound in terms of "multiscale complexity." Using the connection to ordered Ramsey numbers, Section 5 proves Theorems 1.6 and 1.7. Finally, in our concluding remarks, Section 6, we collect a few of the tantalizing open problems that remain in this area.

Notation and terminology. By an *embedding* $H \hookrightarrow T$, we mean an injective function $V(H) \to V(T)$ such that edges of H are mapped to edges of T, with edge orientations preserved. We say that a digraph T is H-free if there is no embedding $H \hookrightarrow T$. All logarithms are to base 2. For the sake of clarity of presentation, we sometimes omit floor and ceiling signs when they are not crucial.

2 Lower bounds

In this section, we prove the lower bound Theorem 1.2, which states that for any $\Delta \geq 1$ there exists a family of acyclic digraphs $\{H_n\}_{n\geq 1}$ with maximum degree Δ for which $|V(H_n)| = n$ and $\overrightarrow{r_1}(H_n) \geq n^{\Omega(\Delta^{2/3}/\log^{5/3}\Delta)}$. The lower bound consists of three ingredients: constructing a bounded degree acyclic digraph H that is hard to embed, constructing a Ramsey tournament T that is hard to embed H into, and proving that there is no embedding $H \hookrightarrow T$.

The next three subsections separately provide these three ingredients. Section 2.1 defines a class of digraphs H with edges "at all scales," which we call *interval meshes,* and proves the existence of bounded-degree H with this property. Section 2.2 defines the Ramsey tournament T as a lexicographic power of a tournament R with no large transitive subtournament, and shows that embeddings $H \hookrightarrow T$ correspond to certain highly constrained walks on R which we call (R, f, s)-*walks.* Section 2.3 completes the proof by showing that long (R, f, s)-walks do not exist, and therefore large interval meshes H cannot be embedded into small powers $T = R^m$.

2.1 Interval meshes

Our proof of Theorem 1.2 is motivated by the lower bound construction for ordered Ramsey numbers proved by Conlon, Fox, Lee, and Sudakov [11]. They prove a lower bound on the ordered Ramsey number of a random matching on $[n]$; see Theorem 5.1 and the surrounding discussion for details. The main property they need of the random matching is that most pairs of long intervals have an edge between them. We need the following stronger property of the same form for our acyclic digraph H.

Definition 2.1. If $f : \mathbb{N} \to \mathbb{R}_{>0}$ is a nondecreasing function, we define an f-*interval mesh* to be an acyclic digraph H whose vertex set is an interval $I \subseteq \mathbb{N}$ such that any edge $(i, j) \in E(H)$ has $i < j$ and for all pairs of intervals $(a_1, b_1], (a_2, b_2] \subseteq I$ with $b_1 \leq a_2$ and

$$a_2 - b_1 \leq f(\min(b_1 - a_1, b_2 - a_2)), \tag{2.1}$$

there exists an edge in H between $(a_1, b_1]$ and $(a_2, b_2]$. When the function f is clear from context, we simply call H an *interval mesh.*

The notion of an interval mesh is one way of formalizing the notion from the introduction of "high multiscale complexity", since interval meshes have many edges in every length scale. We construct interval meshes of bounded degree using a greedy algorithm.

Lemma 2.2. *For any nondecreasing function $f : \mathbb{N} \to \mathbb{R}_{>0}$ with $S = \sum_{m \geq 0} f(2^{m+2}) \cdot 2^{-2m} < \infty$, there exists an f-interval mesh H on vertex set $\mathbb{N}$ with maximum degree at most $2S + 17$.*

Proof. Starting from an empty digraph on $\mathbb{N}$, we construct H by using a greedy algorithm to add edges. All edges introduced point to the right, so the resulting digraph must be acyclic. Define $I_{m,j}$ to be the dyadic interval $(j \cdot 2^m, (j+1) \cdot 2^m]$ with length 2^m.

Let $m \geq 0$ range through the nonnegative integers. On subroutine m, we iterate through all pairs $(i, j) \in \mathbb{N}^2_{\geq 0}$ satisfying

$$1 \leq j - i \leq f(2^{m+2}) \cdot 2^{-m} + 4 \tag{2.2}$$

and add an edge between the (currently) lowest degree vertex of $I_{m,i}$ and the (currently) lowest degree vertex of $I_{m,j}$, if an edge does not exist between $I_{m,i}$ and $I_{m,j}$ already. Writing $d(U)$ for the sum of the degrees of the vertices in a set U, we see that for any given i, subroutine m increases $d(I_{m,i})$ by a total of at most

$$2[f(2^{m+2}) \cdot 2^{-m} + 4] = 2^{-m+1} \cdot f(2^{m+2}) + 8.$$

Let H be the digraph produced from this infinite process. Explicitly, H is the edge union of all the graphs $H^{(m)}$, where $H^{(m)}$ is the graph produced after subroutine m. It has the property that every pair of dyadic intervals $I_{m,i}$, $I_{m,j}$ satisfying (2.2) has an edge between them.

We first check that H has maximum degree at most $2S + 17$. Since $I_{m,i}$ is a union of 2^{m-k} dyadic intervals of length 2^k, we have that after subroutine m,

$$d(I_{m,i}) \leq \sum_{k=0}^{m} 2^{m-k}(2^{-k+1} \cdot f(2^{k+2}) + 8) \leq 2^{m+1} \cdot S + 2^{m+1} \cdot 8 = 2^m(2S + 16).$$

In particular, the average degree in $I_{m,i}$ is less than $2S + 17$ at the end of subroutine m. However, subroutine m only adds edges incident to vertices of $I_{m,i}$ which have at most average degree, so no new vertex of degree at least $2S + 18$ can appear on subroutine m. Thus, no vertex of degree at least $2S + 18$ ever appears, as desired.

Next, we check that H is an f-interval mesh. Suppose two intervals $(a_1, b_1]$, $(a_2, b_2]$ with $b_1 \leq a_2$ satisfy (2.1) and let m be the largest positive integer such that $(a_1, b_1]$ and $(a_2, b_2]$ both contain dyadic intervals of length 2^m. Note that in any interval $(a, b]$ of integers, a longest dyadic subinterval $I_{\ell,i}$ has length $2^\ell \in [\frac{b-a}{4}, b - a]$, so in particular $2^{m+2} \geq \min(b_1 - a_1, b_2 - a_2)$. Next, pick i maximal and j minimal such that $I_{m,i} \subseteq (a_1, b_1]$ and $I_{m,j} \subseteq (a_2, b_2]$.

Using (2.1), we find

$$j - i \leq 4 + \frac{a_2 - b_1}{2^m} \leq 4 + \frac{f(\min(b_1 - a_1, b_2 - a_2))}{2^m} \leq 4 + \frac{f(2^{m+2})}{2^m},$$

and so there is an edge in H between $I_{m,i}$ and $I_{m,j}$, and therefore between $(a_1, b_1]$ and $(a_2, b_2]$ as well. □

The acyclic digraphs H_n we use are induced subgraphs on intervals of the infinite interval mesh constructed in Lemma 2.2, for an appropriate choice of f.

2.2 Walks in tournaments

Next, we construct the large tournament T which is difficult to embed H into. Again, the construction is motivated by the lower bound argument of Conlon, Fox, Lee, and Sudakov [11] for ordered Ramsey numbers, although its analysis requires new techniques.

Recall that the *lexicographic product* $G \cdot H$ of two digraphs G and H is the digraph on vertex set $V(G) \times V(H)$ with an edge $(g_1, h_1) \to (g_2, h_2)$ iff $g_1 \to g_2$ in G or $g_1 = g_2$ and $h_1 \to h_2$ in H. Henceforth, write G^m for the m-fold lexicographic product of G with itself. Note that lexicographic powers of tournaments are tournaments. By (1.1), there exists for any $r \geq 3$ a tournament R on r vertices with no transitive

subtournament of size $2\log r + 2$. We take $T = R^m$ and show that an interval mesh H is difficult to embed into T.

To this end, let H be an interval mesh. We relate embeddings $H \hookrightarrow R^m$ to certain constrained walks on the tournament R.

Definition 2.3. For a tournament R, a nondecreasing function $f : \mathbb{N} \to \mathbb{R}_{>0}$, and $s \geq 1$, define an (R, f, s)-*walk* to be a sequence of ordered pairs $\{(v_i, a_i)\}_{i=1}^{\ell}$ (ℓ possibly infinite) where for each i, $v_i \in V(R)$, $1 \leq a_i \leq s$, $v_i \neq v_{i+1}$, and for any pair $i < j$ for which $v_i \leftarrow v_j$ in R, we have

$$a_{(i,j)} > f(\min(a_i, a_j)),$$

where $a_I \coloneqq \sum_{k \in I} a_k$ if I is an interval of integers and the empty sum equals 0. We define the *length* of an (R, f, s)-walk to be $a_{[1,\ell]}$.

Let $L_{R,f}(s)$ be the length of the longest (R, f, s)-walk if such a maximum exists, and $+\infty$ otherwise. The next lemma reduces Theorem 1.2 to showing asymptotic upper bounds on $L_{R,f}(s)$.

Lemma 2.4. *Suppose there exists $r \geq 1$, a nondecreasing $f : \mathbb{N} \to \mathbb{R}_{>0}$, and a tournament R on r vertices for which $\limsup_{s\to\infty} L_{R,f}(s)s^{-1} = \alpha$. If H is an f-interval mesh on n vertices, then*

$$\overrightarrow{r_1}(H) \geq n^{\frac{\log r}{\log \alpha} - o(1)}.$$

Proof. For each $n \geq 1$, let $m = m(n)$ be the minimum positive integer for which there exists an f-interval mesh H with vertex set $[n]$ and an embedding $\phi : H \hookrightarrow R^m$.

Let $\pi : R^m \to R$ be the projection onto the first coordinate. Consider the sequence $\{\pi \circ \phi(j)\}_{j=1}^{n}$. Consecutive terms of this sequence may repeat, so we block the sequence into consecutive runs of identical vertices. Suppose there are ℓ total runs $I_1 \sqcup \cdots \sqcup I_\ell = [n]$ and run I_i consists of a_i repetitions of vertex $v_i \in V(R)$. We claim that $\{(v_i, a_i)\}_{i=1}^{\ell}$ is an (R, f, s)-walk, where $s \coloneqq \max(a_i)$. It is easy to see that $1 \leq a_i \leq s$ for all i, and that $v_i \neq v_{i+1}$ since we already blocked consecutive identical vertices together. It remains to check the key condition, that if $i < j$ and $v_i \leftarrow v_j$ in R, we must have

$$a_{(i,j)} > f(\min(a_i, a_j)).$$

Suppose this is not true, so there exists some $i < j$ with $v_i \leftarrow v_j$ and $a_{(i,j)} \leq f(\min(a_i, a_j))$. By the definition of v_i, v_j, we have that $\pi(\phi(I_i)) = \{v_i\}$, $\pi(\phi(I_j)) = \{v_j\}$. By the definition of the lexicographic power, if $v_i \leftarrow v_j$ then $w_i \leftarrow w_j$ for all $w_i \in \pi^{-1}(v_i), w_j \in \pi^{-1}(v_j)$. Thus, all edges between $\phi(I_i)$ and $\phi(I_j)$ are oriented from $\phi(I_j)$ to $\phi(I_i)$. For ϕ to be a homomorphism, this means that H has no edge oriented from I_i to I_j. On the other hand, $|I_i| = a_i$, $|I_j| = a_j$, and the distance between the two intervals is $a_{(i,j)} \leq f(\min(a_i, a_j))$, so since H is an f-interval mesh there is such an edge. This is a contradiction, and we see that $\{(v_i, a_i)\}_{i=1}^{\ell}$ is an (R, f, s)-walk of length n.

Fix any $\varepsilon > 0$. We are given that $L_{R,f}(s) \leq (\alpha + \varepsilon)s$ for sufficiently large s, so we get $s \geq (\alpha + \varepsilon)^{-1}n$ using the fact that $n \leq L_{R,f}(s)$ since we have found an (R, f, s)-walk of length n. Since $s = \max(a_i)$, this means that there is some subinterval $I_i \subseteq [n]$ of length at least $(\alpha + \varepsilon)^{-1}n$ for which $\pi \circ \phi$ is constant on I_i, i.e. the image $\phi(I_i) \subseteq \{v_i\} \times R^{m-1}$ lies in a copy of R^{m-1}. Putting this together, we have shown that for large enough n, the existence of an embedding $\phi : H \hookrightarrow R^m$ implies the existence of an embedding $\phi' : H' \hookrightarrow R^{m-1}$ for some f-interval mesh H' on $(\alpha + \varepsilon)^{-1}n$ vertices. In other words,

$$m(n) \geq m((\alpha + \varepsilon)^{-1}n) + 1$$

for all n sufficiently large. Applying this recursively, we obtain that $m(n) \geq \frac{\log n}{\log(\alpha+\varepsilon)} - O_\varepsilon(1)$. By the definition of $m(n)$, any f-interval mesh H on n vertices has no embedding into $R^{m(n)-1}$, and so

$$\overrightarrow{r_1}(H) > |V(R^{m(n)-1})| \geq r^{\frac{\log n}{\log(\alpha+\varepsilon)} - O_\varepsilon(1)} = \Omega_\varepsilon\left(n^{\frac{\log r}{\log(\alpha+\varepsilon)}}\right),$$

which proves the lemma. ∎

To finish the proof of Theorem 1.2, it remains to bound the asymptotic growth rate of $L_{R,f}(s)$, which we do in the next section.

2.3 Completing the proof

The next lemma is a simple observation that is helpful for analyzing the structure of (R, f, s)-walks.

Lemma 2.5. *If R is a tournament without a transitive t-subtournament, then any sequence $v_1, \dots, v_t$ of t vertices either contains a back-edge $v_i \leftarrow v_j$ with $i < j$ or two consecutive elements $v_i = v_{i+1}$.*

Proof. If the vertices are all distinct, then since R has no transitive t-subtournament there must exist a back-edge. Suppose $v_i = v_j$ and $i < j$ but $j \neq i + 1$. Then either $v_i \leftarrow v_{i+1}$ or $v_{i+1} \leftarrow v_j$ is a back-edge. □

We now prove a recursive upper bound on $L_{R,f}(s)$. Given an implicit parameter $t \geq 3$, tournament R, and nondecreasing function $f : \mathbb{N} \to \mathbb{R}_{>0}$, we say that a positive integer s is *short* if $L_{R,f}(s) \leq 2st$ and $L_{R,f}(s'') \leq f(s'')$ for all $s'' \leq s$.

Lemma 2.6. *Suppose $s \geq 1$, $t \geq 3$, R is a tournament without a transitive t-subtournament, and $f : \mathbb{N} \to \mathbb{R}_{>0}$ is a nondecreasing function satisfying $f(s) > 6st^2$. If s is short, then every $s' \in [2st, 4st]$ is short.*

Proof. Suppose $s < s' \leq 4st$, $\{(v_i, a_i)\}_{i=1}^{\ell}$ is an (R, f, s')-walk, and let $i_1 < \cdots < i_u$ be the sequence of all indices where $a_{i_j} \geq s$.

Our first goal is to show that $u < t$. If not, by Lemma 2.5 there is either a back-edge $v_{i_x} \leftarrow v_{i_y}$ with $x < y \leq t$ or some $x \leq t - 1$ for which $v_{i_x} = v_{i_{x+1}}$. We show that neither of these situations is possible.

In the first case, there is an edge $v_{i_x} \leftarrow v_{i_y}$ with $x < y \leq t$. By the definition of an (R, f, s')-walk,

$$a_{(i_x, i_y)} > f(\min(a_{i_x}, a_{i_y})) \geq f(s). \tag{2.3}$$

On the other hand,

$$a_{(i_x, i_y)} = \sum_{j \in (x,y)} a_{i_j} + \sum_{j \in [x,y)} a_{(i_j, i_{j+1})} \leq s't + t \cdot L_{R,f}(s) \leq 6st^2,$$

since for each $j = x, \dots, y - 1$, the subsequence $\{(v_i, a_i)\}_{i=i_j+1}^{i_{j+1}-1}$ is an (R, f, s)-walk with length at most $L_{R,f}(s) \leq 2st$. But $f(s) > 6st^2$, so this contradicts (2.3) and the back-edge $v_{i_x} \leftarrow v_{i_y}$ cannot exist.

Next, suppose for some $x \leq t - 1$ that $v_{i_x} = v_{i_{x+1}}$. The subsequence $\{(v_i, a_i)\}_{i=i_x+1}^{i_{x+1}-1}$ is an (R, f, s'')-walk where s'' is the maximum value of a_i in this subsequence. Pick some z for which $a_z = s''$. Either $v_{i_x} \leftarrow v_z$ or $v_z \leftarrow v_{i_{x+1}}$ is a back-edge, and without loss of generality assume it is the former. Applying the definition of the (R, f, s'')-walk on the two indices i_x and z,

$$a_{(i_x, z)} > f(\min(a_{i_x}, a_z)) = f(s'').$$

On the other hand, this sum is bounded above by $a_{(i_x, i_{x+1})}$. We know $a_{(i_x, i_{x+1})} \leq L_{R,f}(s'') \leq f(s'')$ because $\{(v_i, a_i)\}_{i=i_x+1}^{i_{x+1}-1}$ is an (R, f, s'')-walk and $s \geq s''$ is short, so we have another contradiction. Thus $u < t$.

We obtain

$$a_{[1,\ell]} \leq \sum_{j=1}^{u} a_{i_j} + \sum_{j=0}^{u} a_{(i_j, i_{j+1})} \leq s't + t \cdot L_{R,f}(s) \leq s't + 2st^2, \tag{2.4}$$

for all $s < s' \leq 4st$ and any (R, f, s')-walk $\{(v_i, a_i)\}_{i=1}^{\ell}$. Here we let $i_0 = 0$ and $i_{u+1} = \ell + 1$ for convenience.

Inequality (2.4) implies that $L_{R,f}(s') \leq 2s't$ for all $s' \in [2st, 4st]$. It also implies that $L_{R,f}(s'') \leq 6st^2 < f(s) \leq f(s'')$ for all $s'' \in (s, 4st]$. We have verified both conditions for s' to be short for every $s' \in [2st, 4st]$, as desired. □

It remains to pick a function f for which Lemma 2.6 bootstraps successfully. Define

$$f(s) \coloneqq \begin{cases} \frac{10s^2 t^{3/2} \log t}{\log^2 s} & s \geq 4 \\ 40t^{3/2} \log t & s < 4, \end{cases} \tag{2.5}$$

where the values of $f(1), f(2), f(3)$ are chosen just to make f nondecreasing. Recall that all logarithms are to base 2.

Lemma 2.7. *If $t \geq 10^6$, $f : \mathbb{N} \to \mathbb{R}_{>0}$ is defined by (2.5), and R is a tournament without a transitive t-subtournament, then $L_{R,f}(s) \leq 2st$ for all s sufficiently large.*

Proof. We apply Lemma 2.6 inductively to show that s is short for all s sufficiently large. This implies the desired result.

For the base case, we claim that s is short for all $s \leq s_0 \coloneqq 40t^{1/2}\log t$. Indeed, suppose $s \leq s_0$ and $\{(v_i, a_i)\}_{i=1}^{\ell}$ is an (R, f, s)-walk. By Lemma 2.5, if $\ell \geq t$ then there is either a back-edge $v_i \leftarrow v_j$ with $i < j \leq t$ or two consecutive elements $v_i = v_{i+1}$. The latter contradicts the definition of an (R, f, s)-walk, so assume the former holds. But $a_{(i,j)} < st \leq f(1) \leq f(\min(a_i, a_j))$, so this contradicts the definition of an (R, f, s)-walk again. We have shown that $\ell < t$, so $L_{R,f}(s) < st \leq f(s)$ for all $s \leq s_0$. This proves s is short for every $s \leq s_0$.

Next, we check that $f(s) > 6st^2$ for all $s \geq s_0$. Indeed, $f(s)/s$ is increasing for $s \geq s_0$, and $\log s_0 \leq \log t$ since $t \geq 10^6$. We get

$$f(s_0) = \frac{10s_0^2 t^{3/2}\log t}{\log^2 s_0} > 1000t^{5/2}\log t > 6s_0t^2,$$

proving that the conditions of Lemma 2.6 are satisfied. By induction, Lemma 2.6 then implies that s is short for $s \in [(2t)^k s_0, (4t)^k s_0]$ for all $k \geq 0$. All sufficiently large integers lie in some such interval, so $L_{R,f}(s) \leq 2st$ for all sufficiently large s, as desired. □

Putting everything together, we have a proof of the general lower bound.

Proof of Theorem 1.2. We may assume that Δ is sufficiently large as we always have $\overrightarrow{r_1}(H_n) \geq n$, which proves the theorem for small Δ by picking the implicit constant factor in the exponent appropriately. Let $t = \Delta^{2/3}/(200\log^{2/3}\Delta)$; we may assume $t \geq 10^6$. Define $f : \mathbb{N} \to \mathbb{R}_{>0}$ by (2.5). We have

$$S = \sum_{m \geq 0} f(2^{m+2}) \cdot 2^{-2m} \leq f(1) \cdot \sum_{m \geq 1} m^{-2} \leq 80t^{3/2}\log t.$$

Lemma 2.2 implies that there exists an f-interval mesh H on $\mathbb{N}$ with maximum degree at most $2S + 17 \leq 200t^{3/2}\log t \leq \Delta$. For any $n \geq 1$, let H_n be the induced subgraph of H on the interval $[n]$, so that H_n has n vertices and is also an f-interval mesh of maximum degree at most Δ.

There exists a tournament R on $r = 2^{\Omega(t)}$ vertices with no transitive t-subtournament. By Lemmas 2.4 and 2.7 applied with these choices of R, f and H_n, we find that

$$\overrightarrow{r_1}(H_n) > n^{\frac{\log r}{\log(2t)} - o(1)} \geq n^{\Omega(t/\log t)},$$

which proves the theorem, by our choice of t. □

We remark that the polylogarithmic factor in $\Delta = \Omega(t^{2/3}/\log^{2/3}\Delta)$ can be easily improved. Indeed, the growth rate of $f(s) = \Theta(s^2/\log^2 s)$ is chosen so that

$$\sum_{m \geq 0} f(2^{m+2}) \cdot 2^{-2m} < \infty,$$

and we may take $f(s) = \Theta(s^2/\log^{1+\varepsilon} s)$ for any fixed $\varepsilon > 0$ instead, leading to a slightly smaller Δ.

3 Greedy embedding

In this section we prove the main lemmas needed for all of the upper bounds in this paper. We use the greedy embedding technique, motivated by similar arguments for ordered graphs from [11].

Framework. Let H be an acyclic digraph on n vertices $v_1, \ldots, v_n$. We would like to find an embedding $\phi : H \hookrightarrow T$ into an ambient tournament T. In addition we specify n sets $U_1, \ldots, U_n \subseteq V(T)$ and aim to satisfy $\phi(v_i) \in U_i$ for all i. Embedding then proceeds in n rounds, where round t determines the image $\phi(v_t)$.

After round t, we keep track of the shrinking sets $U_1^{(t)}, \dots, U_n^{(t)}$ of "valid candidates" for each vertex, where initially $U_i^{(0)} = U_i$ for all i. On round t, $\phi(v_t)$ is a carefully chosen vertex in $U_t^{(t-1)}$, and $U_t^{(t)} \coloneqq \{\phi(v_t)\}$. The other candidate sets are updated as follows:

$$U_j^{(t)} \coloneqq \begin{cases} [U_j^{(t-1)} \cap N^+(v_t)] \setminus \{\phi(v_t)\} & \text{if } v_t \to v_j, \\ [U_j^{(t-1)} \cap N^-(v_t)] \setminus \{\phi(v_t)\} & \text{if } v_j \to v_t, \\ U_j^{(t-1)} \setminus \{\phi(v_t)\} & \text{else.} \end{cases}$$

The process fails if there is an empty $U_t^{(t-1)}$, as there would be no valid choice for $\phi(v_t)$. Otherwise, it succeeds if $\phi(v_t)$ is chosen successfully for each $t \le n$. Note that after round t, $\{\phi(v_1), \dots, \phi(v_t)\}$ is an embedded copy of $H[\{v_1, \dots, v_t\}]$ in T, and these vertices have been removed from the other candidate sets, so the update rule guarantees that $U_j^{(t)} = \{\phi(v_j)\}$ remains a singleton when $t \ge j$. If the greedy embedding process succeeds, it exhibits the existence of a copy of H in T. See Figure 3.1 for a schematic illustration of the greedy embedding process.

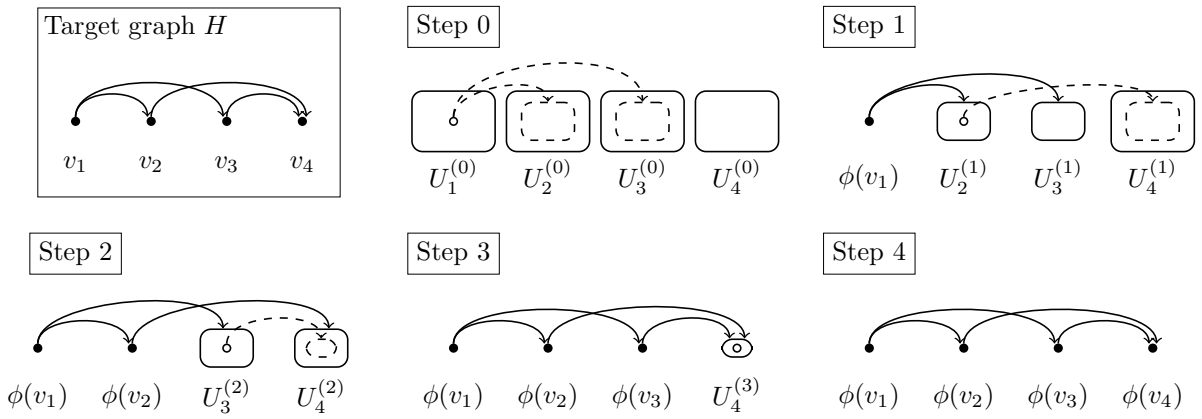

Figure 3.1: Illustration of the greedy embedding process for an acyclic orientation of the four-cycle. A directed arrow from a vertex to a set indicates that the vertex is complete to the set. All edges point forward in this example, but we do not always make this assumption.

For the upper bounds described in the introduction (Theorems 1.3, 1.4, and 1.5), we apply greedy embedding in two separate stages, which we call the outer stage and the inner stage. Roughly speaking, in the outer stage, we run the greedy embedding procedure many times to show that if T is H-free, then T contains a large "approximate blowup of H." In the inner stage, we use greedy embedding one final time within this "approximate blowup" to guarantee the existence of a copy of H in T. In either case, we can conclude that T contains a copy of H—either it is found directly by the greedy embedding strategy, or else the failure of the greedy embedding yields the "approximate blowup" of H, in which a copy of H can be found directly.

This section is split into three subsections. The first covers the basic results that follow from the greedy embedding framework described above, namely how a failure to greedily embed H in a tournament T implies that T contains a certain nice structure, namely a pair of large vertex sets such that most edges between them have the same orientation. The tools built in this first subsection are then used as basic building blocks and iterated in the subsequent subsections. In the second subsection, we use them to build the outer stage of the embedding. In the third subsection, we explain how to use this outer stage construction of an "approximate blowup of H" to finally embed H itself.

In some of these greedy embedding arguments, we are concerned with partitioning an acyclic digraph H into a number of parts and embedding the parts one at a time, so the following definition will be useful.

Definition 3.1. If H is an acyclic digraph, we say that a collection $\{P_i\}_{i=1}^r$ of vertex subsets of H is a *directed partition* of H if $P_1 \sqcup P_2 \sqcup \cdots \sqcup P_r = V(H)$, and any edge of H between two distinct parts P_i, P_j with $i < j$ is oriented from P_i to P_j.

In particular, the height of H is exactly the least h for which there exists a directed partition of H into h independent sets.

3.1 The basic greedy embedding building blocks

Recall that an undirected graph G is said to be *d-degenerate* if there exists an ordering $v_1, \ldots, v_n$ of the vertices of G such that each $|N(v_i) \cap \{v_1, \ldots, v_{i-1}\}| \leq d$, and such an ordering is called a *d-degenerate ordering* of G. A d-degenerate ordering is a natural order for greedily embedding an undirected graph G, since each candidate set $U_j^{(t)}$ in the greedy embedding strategy only shrinks in size by more than one at most d times. We say that a digraph H is d-degenerate if its underlying undirected graph is d-degenerate. Note that if H has maximum degree Δ then it is Δ-degenerate, but d-degenerate digraphs can have arbitrarily large maximum degree.

Define a *δ-dense pair* (W_1, W_2) in a tournament T to be a pair of vertex subsets such that at least $\delta|W_1||W_2|$ of the edges between W_1 and W_2 point from W_1 to W_2. The *size* of the pair is defined to be $\min(|W_1|, |W_2|)$. We do not require W_1 and W_2 to be disjoint, although the assumption of δ-density implies that $W_1 \cap W_2$ cannot be too large if δ is close to 1.

The first lemma in this subsection uses greedy embedding to show that if T doesn't contain a copy of a given d-degenerate H, then T contains some large dense pair. The undirected analogue of this lemma is well-known, and goes back at least to work of Erdős and Hajnal [17, Lemma 1.5].

Lemma 3.2. *Let H be a d-degenerate digraph with n vertices and maximum degree Δ, and let $0 < c \leq \frac{1}{2}$. If T is an H-free tournament on $N \geq 2\Delta c^{-d} n$ vertices, then T contains a $(1-c)$-dense pair (W_1, W_2) with size at least $c^d N/(2\Delta)$.*

Proof. We use the greedy embedding framework described above. Let us label the vertices of H according to the d-degenerate ordering as $v_1, v_2, \ldots, v_n$. We initialize $U_i^{(0)} = V(T)$ for all $1 \leq i \leq n$. We now attempt to embed the vertices of H one at a time in T, in the d-degenerate ordering. For $j > t$, let $N_t(v_j)$ denote the set of vertices v_i with $i \leq t$ such that v_i and v_j are connected by an edge (in some direction). We inductively pick the values of $\phi(v_t) \in V(T)$ and maintain vertex sets $U_i^{(t)}$ with the following properties.

1. For every $i \leq t$, we have $U_i^{(t)} = \{\phi(v_i)\}$.
2. For every $j > t$, we have $|U_j^{(t)}| \geq c^{|N_t(v_j)|} N - t$.
3. For $i \leq t$, if $v_i \to v_j$ then $\phi(v_i) \to x$ for every $x \in U_j^{(t)}$, and if $v_i \leftarrow v_j$ then $\phi(v_i) \leftarrow x$ for every $x \in U_j^{(t)}$.

From these properties, we see that if the process continues through step $t = n$, then we will have embedded a copy of H in T, contradicting our assumption that T is H-free. Moreover, all three properties are vacuously true for $t = 0$. Now, suppose we have maintained this process up through step $t-1$. Suppose there exists $w_t \in U_t^{(t-1)}$ such that for every $j > t$ with $v_t \to v_j$ (resp. $v_t \leftarrow v_j$), we have $|N^+(w_t) \cap U_j^{(t-1)}| \geq c|U_j^{(t-1)}|$ (resp. $|N^-(w_t) \cap U_j^{(t-1)}| \geq c|U_j^{(t-1)}|$). Then we may define $\phi(v_t) = w_t$, $U_t^{(t)} = \{\phi(v_t)\}$, and update the remaining sets as

$$U_j^{(t)} \coloneqq \begin{cases} [U_j^{(t-1)} \cap N^+(v_t)] \setminus \{\phi(v_t)\} & \text{if } v_t \to v_j, \\ [U_j^{(t-1)} \cap N^-(v_t)] \setminus \{\phi(v_t)\} & \text{if } v_j \to v_t, \\ U_j^{(t-1)} \setminus \{\phi(v_t)\} & \text{else,} \end{cases}$$

for all $j > t$. Properties 1 and 3 above continue to hold automatically after round t, and all that remains to check is Property 2. For those $j > t$ for which v_j is not adjacent to v_t, $N_t(v_j) = N_{t-1}(v_j)$ and at most one vertex is removed from $U_j^{(t-1)}$ to obtain $U_j^{(t)}$. Therefore,

$$|U_j^{(t)}| \geq |U_j^{(t-1)}| - 1 \geq c^{|N_{t-1}(v_j)|}N - (t-1) - 1 = c^{|N_t(v_j)|} - t.$$

On the other hand, if v_j is adjacent to v_t, then $|N_t(v_j)| = |N_{t-1}(v_j)| + 1$, and therefore

$$|U_j^{(t)}| \geq c|U_j^{(t-1)}| - 1 \geq c \cdot (c^{|N_{t-1}(v_j)|}N - (t-1)) - 1 \geq c^{|N_t(v_j)|}N - t.$$

Thus, all three properties are maintained if such a w_t exists.

Since we assumed that T was H-free, this process cannot continue until step $t = n$, and therefore must fail at some step $1 \leq t \leq n-1$. Let $W_0 = U_t^{(t-1)}$. Since the process fails at this step, for every $w \in W_0$, we can assign some $j > t$ such that either $v_t \to v_j$ and $|N^+(w) \cap U_j^{(t-1)}| < c|U_j^{(t-1)}|$, or $v_t \leftarrow v_j$ and $|N^-(w) \cap U_j^{(t-1)}| < c|U_j^{(t-1)}|$. Since v_t has at most Δ neighbors in total, at least $|W_0|/\Delta$ choices of w are assigned the same $j > t$ by the pigeonhole principle. Fix such a "popular" j.

Suppose first that $v_t \to v_j$. Let $W_2 = U_j^{(t-1)}$, and let W_1 be the set of all $w \in W_0$ which have fewer than $c|W_2|$ out-neighbors in W_2. Then (W_1, W_2) is a $(1-c)$-dense pair. Similarly, if $v_j \in N^-(v_t)$, then we would similarly find that (W_2, W_1) is a $(1-c)$-dense pair.

It remains to verify the lower bound on the sizes of W_1 and W_2. Recall that the greedy embedding process succeeded up through step $t-1$, meaning that

$$|U_t^{(t-1)}| \geq c^{|N_{t-1}(v_t)|}N - (t-1) \geq c^d N - n \geq \frac{c^d N}{2},$$

and similarly for $U_j^{(t-1)}$, where we use the d-degeneracy assumption to conclude that $|N_{t-1}(v_t)| \leq d$, and our assumption that $t < n \leq c^d N/2$. Since $|W_2| = |U_j^{(t-1)}|$ and $|W_1| \geq |U_t^{(t-1)}|/\Delta$, this completes the proof. □

The second lemma proves a much stronger bound when H is 1-degenerate and weakly connected, i.e. some orientation of a tree (recall that a digraph is called *weakly connected* if its underlying undirected graph is connected). Note that a 1-dense pair in a tournament is just a pair of sets W_1, W_2 with all edges directed from W_1 to W_2.

Lemma 3.3. *Let H be a weakly connected 1-degenerate digraph with maximum degree Δ on m vertices $v_1, \ldots, v_m$, and let T be an arbitrary tournament. If there exist sets $U_1, \ldots, U_m \subseteq V(T)$, each of size $M \geq 2m\Delta$ such that if there is no embedding $\phi : H \hookrightarrow T$ satisfying $\phi(v_i) \in U_i$, then T contains a 1-dense pair with size at least $M/(m(\Delta+1))$.*

Proof. We begin by picking subsets $V_1 \subseteq U_1, \ldots, V_m \subseteq U_m$, each of size M/m, such that $V_1, \ldots, V_m$ are pairwise disjoint. We can do this greedily, by first picking an arbitrary subset $V_1 \subseteq U_1$ of size M/m, then picking an arbitrary subset $V_2 \subseteq U_2 \setminus V_1$ of size M/m, then picking $V_3 \subseteq U_3 \setminus (V_1 \cup V_2)$, and so on. At the ith step, we have deleted at most $(i-1)M/m \leq (m-1)M/m$ elements from U_i, so at least M/m elements remain, from which we pick V_i arbitrarily.

The remainder of the proof is very similar to that of the previous lemma, though it will be more convenient to work with a slightly different setup than before. Let the vertices of H be ordered so that each v_i has at most one neighbor v_j with $j > i$ (this is the reverse of the usual degenerate ordering). For each t, let S_t denote the set of vertices v_i with $i \leq t$ that are connected to v_i by a path of vertices whose indices are monotonically increasing, including v_t itself. In other words, we set $S_1 = \{v_1\}$, let $N_t(v_t)$ denote the set of vertices v_i adjacent to v_t (in either orientation) with $i < t$, and set

$$S_t = \{v_t\} \cup \bigcup_{i \in N_t(v_t)} S_i.$$

We define $W_t \subseteq V_t$ to be the set of all $w \in V_t$ for which there exists an embedding $\phi : H[S_t] \hookrightarrow T$ mapping each v_i into V_i for $v_i \in S_t$, and mapping v_t to w. We know $W_n = \varnothing$, since otherwise there is an embedding of H into the sets $V_1, \dots, V_m$, contradicting our assumption. Let $t \leq n$ be the minimum index such that $|W_t| < M/(m(\Delta+1))$; note that $t > 1$ since $W_1 = V_1$ has size M/m. Let $N_t(v_t) = \{v_{i_1}, \dots, v_{i_s}\}$. Our maximum degree assumption implies $s \leq \Delta$, while if $s = 0$, then $W_t = V_t$, contradicting our assumption that $|W_t| < M/(m(\Delta+1))$. Thus, $1 \leq s \leq \Delta$.

By definition, W_t is precisely the set of $w \in V_t$ adjacent in the appropriate orientation to at least one vertex in each of $W_{i_1}, \dots, W_{i_s}$. Let $X_1, \dots, X_s \subseteq V_t$ denote the choices of w which do not have any edge in the appropriate orientation to $W_{i_1}, \dots, W_{i_s}$, respectively. We get $V_t = W_t \cup X_1 \cup \cdots \cup X_s$. Since $s \leq \Delta$ and $|W_t| \leq M/(m(\Delta+1))$, we see that there exists some j for which $|X_j| \geq M/(m(\Delta+1))$. Moreover, since t was taken to be minimal, we have that $|W_{i_j}| \geq M/(m(\Delta+1))$ as well. This yields a pair of sets (X_j, W_{i_j}) where all edges between them are oriented the same way, and both sets have size at least $M/(m(\Delta+1))$. This is the desired 1-dense pair. □

We remark that we expect the m dependence in Lemma 3.3 to be unnecessary, and proving this would improve our results for random digraphs; for details, see Conjecture 6.5 and the surrounding discussion.

We do not use Lemma 3.3 directly, but only the following simple corollary. It allows us to find a large 1-dense pair in T whenever T does not contain a copy of some fixed oriented forest with small maximum degree and small weakly connected components.

Lemma 3.4. *Let H be a 1-degenerate digraph with maximum degree Δ and vertices $v_1, \dots, v_n$, and suppose that every weakly connected component of H has at most m vertices. Let T be an arbitrary tournament. For any collection of sets $V_1, \dots, V_n \subseteq V(T)$, each of size $M \geq 3n\Delta$, either there is an embedding $\phi : H \hookrightarrow T$ satisfying $\phi(v_i) \in V_i$, or T contains a 1-dense pair with size at least $M/(4m\Delta)$.*

Proof. Let the weakly connected components of H be $C_1, \dots, C_r$. We prove the result by induction on r. The base case, $r = 1$, follows immediately from Lemma 3.3, since $n \geq m$ and $\Delta + 1 \leq 2\Delta$. For the inductive step, let H' be the induced subgraph of H consisting of the weakly connected components $C_1, \dots, C_{r-1}$. By the inductive hypothesis, either T contains a 1-dense pair of size at least $M/(4m\Delta)$, in which case we are done, or else there is an embedding ϕ of H' into T satisfying that $\phi(v_i) \in V_i$ for all $v_i \in V(H')$. Let T' be the tournament obtained by deleting the image of ϕ from T, and similarly let $V_1', \dots, V_n'$ be obtained by deleting the image of ϕ from $V_1, \dots, V_n$. Since we have deleted at most $n \leq n\Delta$ vertices, each V_i' has size at least $M' \geq 2n\Delta \geq 2m\Delta$. Therefore, by Lemma 3.3, either T' contains a 1-dense pair of size at least $M'/(2m\Delta) \geq M/(4m\Delta)$, in which case so does T, or else there is an embedding of $H[C_r]$ into T' mapping each $v_i \in C_r$ into V_i'. Since we deleted the image of ϕ from T to define T', such an embedding yields an embedding of H into T, completing the induction. □

Our final basic greedy embedding lemma is the following, which shows that if we assume appropriate density conditions on the sets in which we are trying to greedily embed H, then we are guaranteed not to fail in the embedding. In Subsection 3.3, we use this lemma as a basic building block. Again, the undirected analogue of this lemma is well-known, e.g. [21, Lemma 2].

Lemma 3.5. *Let H be an acyclic digraph with maximum degree Δ on n vertices $v_1, \dots, v_n$, and let T be a tournament containing subsets $V_1, \dots, V_n \subseteq V(T)$, each of size at least $4n$. If for every edge $v_i \to v_j$ in H, (V_i, V_j) is a $(1 - \frac{1}{8\Delta^2})$-dense pair, then T contains a copy of H.*

Note that the sets V_i are allowed to overlap[2] or even be identical. This is crucial; for instance, if H has height h then our choice of V_i's will take only h distinct values.

Proof. Every acyclic digraph has a vertex ordering where all edges point forward, so we may reorder the vertices to assume all edges $v_i \to v_j$ satisfy $i < j$. We now run the greedy embedding process for H into T with the given U_i, using essentially the same framework as we used in Lemma 3.2. The key difference

[2]However, if $v_i \to v_j$, then the assumption that (V_i, V_j) is $(1 - \frac{1}{8\Delta^2})$-dense implies that $V_i \cap V_j$ cannot be too large.

is that before, the greedy embedding process could fail and terminate prematurely, whereas the additional assumptions here guarantee that greedy embedding runs to completion.

We begin by refining the sets $V_1, \dots, V_n$. Namely, for every edge $v_i \to v_j$ in H, let $V_{i,j} \subseteq V_i$ be the set of vertices $w \in V_i$ with $|N^-(w) \cap V_j| \geq |V_j|/(4\Delta)$. Since the pair (V_i, V_j) is $(1 - \frac{1}{8\Delta^2})$-dense, there are at most $|V_i||V_j|/(8\Delta^2)$ edges directed from V_j to V_i in total, so

$$|V_{i,j}| \cdot \frac{|V_j|}{4\Delta} \leq \frac{|V_i||V_j|}{8\Delta^2},$$

and thus $|V_{i,j}| \leq |V_i|/(2\Delta)$. Each v_i has at most Δ out-neighbors v_j, which implies that $|\bigcup_j V_{i,j}| \leq |V_i|/2$. Therefore, if $U_i \coloneqq V_i \setminus (\bigcup_j V_{i,j})$, then $|U_i| \geq |V_i|/2$, and every vertex in U_i has at most $|V_j|/(4\Delta) \leq |U_j|/(2\Delta)$ in-neighbors in U_j for every j such that $v_i \to v_j$.

Now, we exhibit an embedding $\phi : H \hookrightarrow T$ by picking $\phi(v_t) \in U_t$ inductively for each $t \in [n]$, as follows. Having picked $\phi(v_1), \dots, \phi(v_{t-1})$, we claim that there is at least one valid candidate for $\phi(v_t) \in U_t$ which is consistent with the previous choices. Indeed, there are at most Δ choices of $i < t$ such that $v_i \to v_t$. For every such i, we have that $\phi(v_i) \in U_i$, which means that $\phi(v_i)$ has at most $|U_t|/(2\Delta)$ in-neighbors in U_t. Hence, at least $|U_t|/2$ vertices in U_t are out-neighbors of $\phi(v_i)$ for all i such that $v_i \to v_j$. Moreover, at most $t-1$ vertices of U_j have been picked as outputs of ϕ, and thus the number of possible candidates for $\phi(v_t)$ is at least

$$\frac{|U_t|}{2} - (t-1) \geq \frac{|V_t|}{4} - n + 1 \geq 1,$$

by our assumption that $|V_j| \geq 4n$. This shows that at every step $t \leq n$ we can pick vertex $\phi(v_t)$ such that $\{\phi(v_1), \dots, \phi(v_t)\}$ is a copy of $H[\{v_1, \dots, v_t\}]$ in T, and so T contains a copy of the entirety of H, as desired. □

We remark that one can prove a strengthening of Lemma 2.4, where the assumption is weakened to each pair (V_i, V_j) being merely $(1 - \Omega(\frac{1}{\Delta}))$-dense. This can be done by replacing the greedy embedding argument by a random embedding technique, using the Lovász local lemma. For details, we refer the reader to [13, Lemma 4.5], where the analogous undirected result is proved using this technique.

3.2 The outer stage

In this subsection, we show how two of our basic building blocks, Lemmas 3.2 and 3.4, can be iterated to construct, in any H-free tournament T, an "approximate blowup" of H. This will be a large collection of vertex sets which correspond to the vertices of H, such that the edges between sets are mostly oriented in the correct direction. Before stating the result precisely, we need some definitions.

Let $\{0,1\}^*$ denote the set of all finite binary strings. Recall that a *prefix code* is a set $C \subset \{0,1\}^*$ with the property that no element of C is a prefix of another element of C. The *depth* of C is defined as the maximum length of an element of C. Let $\prec$ denote the lexicographic ordering on $\{0,1\}^*$, namely the ordering in which $x \prec y$ if x is a proper prefix of y or if $x_i < y_i$ where i is the first index for which $x_i \neq y_i$.

Definition 3.6. Given an acyclic digraph H, a *prefix labeling* of H is a surjective map $\rho : V(H) \to C$ for some prefix code $C \subset \{0,1\}^*$, with the property that if $v_i \to v_j$ is an edge of H, then either $\rho(v_i) = \rho(v_j)$ or $\rho(v_i) \prec \rho(v_j)$. The map ρ naturally defines a graph structure on C, where we say that two codewords x, y are *adjacent under* ρ if there exists some edge between the sets $\rho^{-1}(x)$ and $\rho^{-1}(y)$. By the *maximum degree of* ρ, we mean the maximum degree of this graph on C. If $\rho^{-1}(x)$ is an independent set for every $x \in C$, then we call ρ a *prefix coloring.* Less stringently, we call ρ a *forest prefix labeling* if $\rho^{-1}(x)$ is a directed forest (or equivalently, a 1-degenerate digraph) for every $x \in C$. By the *maximum component size* of ρ, we mean the maximum number of vertices of any weakly connected component in $\rho^{-1}(x)$, over all $x \in C$. Thus, ρ is a prefix coloring if and only if its maximum component size is 1.

Thus, we see that prefix colorings of H correspond to colorings of the underlying undirected graph of H, with the property that the palette of colors C is a prefix code, and that the lexicographic order on C is

consistent with the edge directions in H. Similarly, a forest prefix labeling is in particular a partition of the underlying undirected graph into sets which induce forests, which corresponds to the undirected problem of *vertex arboricity*. However, for both concepts, we will crucially use the additional structure given both by the edge directions of H and by the structure of the prefix code C.

For a binary string x, let us denote by $x0$ and $x1$ the strings obtained by appending a 0 or 1, respectively, to the end of x. For a prefix code C, we denote by $C_0(x)$ the set of all elements $y \in C$ which have $x0$ as a prefix, and by $C_1(x)$ the set of elements that have $x1$ as a prefix. Suppose that $\rho : V(H) \to C$ is a prefix labeling of an acyclic digraph H. For binary string x, let us denote by $a_0(x)$ the number of codewords $y \in C_0(x)$ which are adjacent under ρ to some codeword $z \in C_1(x)$. Similarly, $a_1(x)$ is the number of codewords $z \in C_1(x)$ that are adjacent under ρ to some codeword $y \in C_0(x)$. Finally, we let

$$a(x) = \begin{cases} 1 & \text{if } a_0(x) = 0 \text{ or } a_1(x) = 0, \\ \min(a_0(x), a_1(x)) & \text{otherwise.} \end{cases}$$

In particular, $a(x) = 1$ if x is not a proper prefix of any element of C. With this notation, we can now define the key parameters of ρ that we later use to bound Ramsey numbers.

Definition 3.7. Let H be an acyclic digraph and $\rho : V(H) \to C$ some prefix labeling of H. We define the *dyadic complexity* of ρ to be the quantity

$$\operatorname{comp}(\rho) := \max_{y \in C} \prod_{x \text{ a prefix of } y} a(x).$$

Additionally, the *depth* of ρ, denoted $\operatorname{depth}(\rho)$, is defined as the depth of C, i.e. the maximum length of an element of C.

To understand these definitions, it is helpful to think of $\{0,1\}^*$ as the vertices of the infinite binary tree. In this setup, the elements of a prefix code C correspond to the leaves of some subtree. A prefix labeling $\rho : V(H) \to C$ is then a partition of the vertices of H into sets labeled by the leaves of this subtree. Two codewords (leaves) are adjacent under ρ if there is an edge between the corresponding vertex subsets of H. For a binary string x, which should be thought of as a non-leaf vertex of the subtree, $a(x)$ roughly records the "cost" of separating the descendants of x: it measures how many pairs of codewords adjacent under ρ there are between its descendants on the left and on the right. Because of the structure of the proof, this cost function is somewhat unnatural: it is the minimum of two quantities, each of which is the number of descendants on one side which are adjacent under ρ to any number of descendants on the other side. Moreover, this cost function should be thought of as multiplicative, so that the ultimate cost of the whole labeling—namely the dyadic complexity $\operatorname{comp}(\rho)$—is the product of the costs of all ancestors of y, maximized over all $y \in C$.

We remark that the dyadic complexity of a prefix coloring of H is one possible formalization of the notion of multiscale complexity discussed in the introduction. Indeed, if every prefix coloring of H has high dyadic complexity, then H has many edges at "many different dyadic scales". All of our upper bounds on oriented Ramsey numbers depend on the dyadic complexity of some prefix labeling of H, making formal the intuition that the strength of our upper bound results depends on whether H has high or low multiscale complexity.

In order to embed some acyclic digraph H in a tournament T, we first build a certain structure of vertex subsets of T with high forward edge density between many of the pairs. Then, in the inner stage of the embedding process, we use such a structure to find a copy of H. The structure we build depends on a parameter $\delta \in [0,1]$, as well as on a prefix labeling ρ of H. We now define this structure, and next prove a lemma showing how to find such a structure.

Definition 3.8. Let $\delta \in [0,1]$ be some parameter, let H be an acyclic digraph, and let $\rho : V(H) \to C$ be some prefix labeling of H, for some prefix code C. Let T be any tournament. A (ρ, δ)-*skeleton* is a collection $\{V_x\}_{x \in C}$ of (not necessarily disjoint) vertex subsets of T, indexed by the codewords in C, with the property that if $x \prec y$ are elements of C that are adjacent under ρ, then (V_x, V_y) is a δ-dense pair. We define the *size* of a (ρ, δ)-skeleton to be $\min_{x \in C} |V_x|$.

Our next lemma shows how to iterate Lemma 3.2 to construct a (ρ, δ)-skeleton in any sufficiently large H-free tournament T. Roughly speaking, since the structure of a (ρ, δ)-skeleton is based on the binary tree structure of C, we may construct such a skeleton by performing a depth-first search, and applying Lemma 3.2 every time we need to split an existing node into two daughter nodes in this binary tree.

Lemma 3.9. *Let $c \in (0,1)$ be some parameter, let H be a d-degenerate acyclic digraph with maximum degree Δ, and let $\rho : V(H) \to C$ be some prefix labeling of H, for some prefix code C. Suppose that T is an H-free tournament on N vertices, with*

$$N \geq \left(\frac{4^{d+1}\Delta}{c^d}\right)^{\mathrm{depth}(\rho)} \mathrm{comp}(\rho)^d n.$$

Then T contains a $(\rho, 1-c)$-skeleton of size at least $\left(\frac{c^d}{4^{d+1}\Delta}\right)^{\mathrm{depth}(\rho)} \mathrm{comp}(\rho)^{-d} N$.

Proof. For every binary string $x \in \{0,1\}^*$ which is a prefix of some codeword in C, let H_x denote the subgraph of H induced by the vertices v for which x is a prefix of $\rho(v)$. We will construct, for every such $x \in \{0,1\}^*$, a vertex subset $V_x \subseteq V(T)$, with

$$|V_x| \geq \left(\prod_{z \text{ a proper prefix of } x} \frac{c^d}{4^{d+1}\Delta a(z)^d}\right) N.$$

We initialize $V_\varnothing = V(T)$, where $\varnothing \in \{0,1\}^*$ denotes the empty string, and observe that this property holds vacuously for $V_\varnothing$ since $\varnothing$ has no proper prefixes. In order to construct these vertex sets for other x, we proceed via a depth-first search along the binary tree, as follows. Recall that for any binary string x, the numbers $a_0(x)$ and $a_1(x)$ are the number of codewords in C beginning with $x0$ and $x1$, respectively, such that some vertex labeled by that codeword has an edge to H_{x1} and H_{x0}, respectively. If $\min(a_0(x), a_1(x)) = 0$, then we define $V_{x0} = V_{x1} = V_x$, and observe that our desired inequality holds, since $a(x) = 1$ in this case. In particular, if $x \in C$ is a codeword, then we stop the inductive process, since we only wished to define such vertex subsets for strings x that are the prefix of some codeword. Now, suppose that $\min(a_0(x), a_1(x)) \geq 1$, and assume without loss of generality[3] that $a_0(x) \leq a_1(x)$. We will first show how to define $V_{x0} \subseteq V_x$ satisfying the desired inequality on its cardinality. We will then proceed to recursively define vertex subsets $V_y \subseteq V_{x0}$ for every binary string y prefixed by $x0$. Note that we have not yet defined the set V_{x1}: we are proceeding in a depth-first fashion, so we will not define V_{x1} until we have defined V_y for every y prefixed by $x0$. This will eventually happen, since we already described above how to define V_y if y is a codeword of C; therefore, we eventually reach the bottom of our depth-first search (namely a codeword $y \in C$), at which point we stop going down the tree, and begin to retrace our steps and traverse back up the tree.

Recall that we assumed that $a_0(x) \leq a_1(x)$, and let $c_x = c/(4a_0(x))$. Since

$$\begin{aligned}
|V_x| &\geq \left(\prod_{z \text{ a proper prefix of } x} \frac{c^d}{4^{d+1}\Delta a(z)^d}\right) N \\
&= \frac{4\Delta(4a(x))^d}{c^d} \left(\prod_{z \text{ a proper prefix of } x0} \frac{c^d}{4^{d+1}\Delta a(z)^d}\right) N \\
&\geq \frac{\left(c^d/(4^{d+1}\Delta)\right)^{\mathrm{depth}(\rho)} \mathrm{comp}(\rho)^{-d}}{c^d/(4\Delta(4a(x))^d)} N \\
&\geq 4\Delta c_x^{-d} n, \qquad (3.1)
\end{aligned}$$

and since T contains no copy of H, we may apply Lemma 3.2 with parameter c_x. Then this lemma says that V_x contains a $(1-c_x)$-dense pair (W_0, W_1), where $\min(|W_0|, |W_1|) \geq c_x^d |V_x|/(2\Delta)$. Let $V_{x0} \subseteq W_0$ denote the

[3] If $a_1(x) < a_0(x)$, then we swap the roles of 0 and 1 and the roles of in- and out-neighbors in this construction.

set of vertices in W_0 whose in-degree to W_1 is at most $2c_x|W_1|$; since there are at most $(1-c_x)|W_0||W_1|$ edges directed from W_1 to W_0, we see that $|V_{x0}| \geq |W_0|/2$. Therefore,

$$|V_{x0}| \geq \frac{c_x^d}{4\Delta}|V_x| = \frac{c^d}{4^{d+1}\Delta a(x)^d}|V_x| \geq \left(\prod_{z \text{ a proper prefix of } x0} \frac{c^d}{4^{d+1}\Delta a(z)^d}\right) N, \tag{3.2}$$

since the proper prefixes of $x0$ are just x, in addition to all the proper prefixes of x itself.

As discussed above, we can now recursively define V_y for all y prefixed by $x0$. It thus only remains to define $V_{x1} \subseteq V_x$, under the assumption that we have defined $V_y \subseteq V_{x0} \subseteq W_0$ for all y prefixed by $x0$.

To do so, let $y_1, \ldots, y_{a_0(x)}$ denote the set of codewords prefixed by $x0$ with the property that some vertex in $\rho^{-1}(y_i)$ is adjacent to some vertex in H_{x1}, noting that there are exactly $a_0(x)$ such codewords by the definition of $a_0(x)$. Let $W_1^{(i)}$ denote the subset of W_1 consisting of vertices in W_1 whose out-degree to V_{y_i} is at least $c|V_{y_i}|$. Since every vertex in $V_{y_i} \subseteq V_{x0}$ has at most $2c_x|W_1|$ in-neighbors in W_1, we see that the total number of edges directed from W_1 to V_{y_i} is at most $2c_x|W_1||V_{y_i}|$, and therefore $|W_1^{(i)}| \leq (2c_x/c)|W_1|$. Because of this, we have that

$$\left|\bigcup_{i=1}^{a_0(x)} W_1^{(i)}\right| \leq \sum_{i=1}^{a_0(x)} |W_1^{(i)}| \leq a_0(x)\frac{2c_x}{c}|W_1| = \frac{|W_1|}{2}.$$

Therefore, we can define $V_{x1} = W_1 \setminus \left(\bigcup_{i=1}^{a_0(x)} W_1^{(i)}\right)$, and we have that $|V_{x1}| \geq |W_1|/2$. By the same computation as in equation (3.2), we see that this definition of V_{x1} satisfies our desired lower bound on the size of V_{x1}.

We claim that in this construction, if $y \prec z$ are codewords that are adjacent under ρ, then the pair (V_y, V_z) is $(1-c)$-dense. To see this, let x be the longest common prefix of y and z. In the construction at level x, we first proceeded to either V_{x0} or V_{x1} in the depth-first search, depending on the relative sizes of $a_0(x)$ and $a_1(x)$. In the former case, we ensured that every vertex in V_{x1} had out-degree at most $c|V_y|$ to V_y, while in the latter case, we ensured that every vertex in V_{x0} had in-degree at most $c|V_z|$ to V_z. In either case, we see that (V_y, V_z) is $(1-c)$-dense, since $V_y \subseteq V_{x0}$ and $V_z \subseteq V_{x1}$. Additionally, by the same computation as in equation (3.1), we see that $|V_y| \geq \left(\frac{c^d}{4^{d+1}\Delta}\right)^{\operatorname{depth}(\rho)} \operatorname{comp}(\rho)^{-d} N$ for every $y \in C$. This verifies all the properties of a $(\rho, 1-c)$-skeleton, and concludes the proof. □

Our next two lemmas are very similar to Lemmas 3.2 and 3.9. Namely, the first shows us how to find a 1-dense pair in an H-free tournament T, and the second then iterates the first to form a skeleton of 1-dense pairs. The main difference between these and the previous results are that for these lemmas, we need strengthened assumptions on H (namely that it has a directed partition into forests). Moreover, the first step in the proof is an application of Lemma 3.9, and we find the 1-dense pair by failing to greedily embed H in the skeleton given by Lemma 3.9.

Lemma 3.10. *Let $d \geq 2$, and suppose that H is a d-degenerate acyclic digraph on n vertices with maximum degree Δ, and suppose that $\rho : V(H) \to C$ is some forest prefix labeling with maximum degree A and maximum component size m. Let T be an H-free tournament on $N \geq (2^{10}A\Delta^2)^{d\operatorname{depth}(\rho)} \operatorname{comp}(\rho)^d n$ vertices. Then T contains a 1-dense pair (U_1, U_2) of size at least $m^{-1}(2^{10}A\Delta^2)^{-d\operatorname{depth}(\rho)} \operatorname{comp}(\rho)^{-d} N$.*

Proof. We first apply Lemma 3.9 to the prefix labeling ρ and with parameter $c = 1/(16A\Delta)$. This lemma outputs a $(\rho, 1 - \frac{1}{16A\Delta})$-skeleton $\{V_x\}_{x \in C}$ in T of size at least

$$\left(\frac{c^d}{4^{d+1}\Delta}\right)^{\operatorname{depth}(\rho)} \operatorname{comp}(\rho)^{-d} N \geq 12\Delta n.$$

For every pair of codewords $x \prec y$ that are adjacent under ρ, let $V_{x,y}$ denote the set of vertices in V_x whose in-degree to V_y is at least $\frac{1}{8\Delta}|V_y|$. Since at most $\frac{1}{16A\Delta}|V_x||V_y|$ of the edges between V_x and V_y are directed

from V_y to V_x, we see that $|V_{x,y}| \le \frac{1}{2A}|V_x|$. Therefore, if we define $V'_x = V_x \setminus \left(\bigcup_{y \succ x} V_{x,y}\right)$, then we see that $|V'_x| \ge \frac{1}{2}|V_x|$, since there are at most A choices for $y \succ x$ with x, y adjacent under ρ. Additionally, every vertex in V'_x has in-degree at most $\frac{1}{8\Delta}|V_y| \le \frac{1}{4\Delta}|V'_y|$ from any V'_y with $y \succ x$ such that x, y are adjacent under ρ. We now attempt to greedily embed H in these sets $\{V'_x\}_{x \in C}$.

Let the codewords of C be $x_1, \dots, x_r$, sorted so that $i < j$ if and only if $x_i \prec x_j$. Let $P_i = \rho^{-1}(x_i)$, so that P_i is an oriented forest, each of whose weakly connected components has at most m vertices. Additionally, $P_1 \sqcup \cdots \sqcup P_r$ forms a directed partition of $V(H)$, which we recall means that every edge of H is oriented from P_i to P_j where $i \le j$. For every vertex $v \in P_i$, we initialize a set of candidates $U^{(0)}_{i,v} = V'_{x_i}$. Inductively, having defined $U^{(t)}_{i,v}$ for each $i > t$ and each $v \in P_i$, we attempt to pick an embedding $\phi_t : H[P_t] \hookrightarrow V'_{x_t}$, such that for every $v \in P_t$, $\phi_t(v) \in U^{(t-1)}_{t,v}$. If such a ϕ_t exists, then for each $i > t$ and each $v \in P_i$, we let

$$U^{(t)}_{i,v} \coloneqq \{u \in U^{(t-1)}_{i,v} \setminus \phi_t(P_t) : \forall w \in N^-(v) \cap P_t, \phi_t(w) \to u\}.$$

Note that by the structure of the sets $V'_{x_1}, \dots, V'_{x_r}$, we only change $U^{(t)}_{i,v}$ as follows. First, in at most Δ steps, we embed an in-neighbor of v, and we decrease $|U^{(t)}_{i,v}|$ by at most $\frac{1}{4\Delta}|V'_{x_i}|$. Additionally, we remove at most n additional vertices from $U^{(t)}_{i,v}$, corresponding to vertices that were picked as images of ϕ_t. In total, we remove at most $\Delta \cdot \frac{1}{4\Delta}|V'_{x_i}| + n \le \frac{1}{2}|V'_{x_i}|$ vertices. We thus see that $|U^{(t)}_{i,v}| \ge \frac{1}{2}|V'_{x_i}|$ for all t.

If we are able to run this process for all $1 \le t \le r$, then we have embedded a copy of H in T, so we may assume that the process fails at some step t. This means that there is no embedding $\phi_t : H[P_t] \hookrightarrow T$ such that every vertex $v \in P_t$ is mapped to $U^{(t-1)}_{t,v}$. Therefore, by applying Lemma 3.4 to $H[P_t]$, we conclude that T contains a 1-dense pair (U_1, U_2) of size at least

$$\frac{|U^{(t-1)}_{t,v}|}{4m\Delta} \ge \frac{|V_{x_t}|}{16m\Delta} \ge \frac{1}{m}(2^{10}A\Delta^2)^{-d\,\mathrm{depth}(\rho)}\,\mathrm{comp}(\rho)^{-d}N. \qquad \square$$

Our next result shows how we can iterate the previous lemma to construct many 1-dense pairs. The iteration is nearly identical to the one in Lemma 3.9, where we iterated the construction of one dense pair to the construction of a $(\rho, 1-c)$-skeleton. This lemma takes as input two (not necessarily distinct) prefix labelings on H, one of which is a forest labeling. The reason to have two separate labelings is that it may be useful to use the failure of embedding of H according to one labeling to construct a good embedding structure for the other labeling.

Lemma 3.11. *Let $d \ge 2$, and suppose that H is a d-degenerate acyclic digraph on n vertices with maximum degree Δ. Let $\rho_1 : V(H) \to C_1$ and $\rho_2 : V(H) \to C_2$ be two prefix labelings, such that ρ_1 is a forest prefix labeling of maximum degree A_1 and maximum component size m_1. Let*

$$\gamma = \left(m_1(2^{10}A_1\Delta^2)^{d\,\mathrm{depth}(\rho_1)}\,\mathrm{comp}(\rho_1)^d\right)^{-1}.$$

If T is an H-free tournament on $N \ge \gamma^{-\mathrm{depth}(\rho_2)}n$ vertices, then T contains a $(\rho_2, 1)$-skeleton of size at least $\gamma^{\mathrm{depth}(\rho_2)}N$.

Proof. As in the proof of Lemma 3.9, we construct our skeleton by assigning a set $V_x \subseteq V(T)$ for every $x \in \{0,1\}^*$ that is a prefix of some codeword in C_2, with the property that V_{x0}, V_{x1} are both subsets of V_x. We guarantee inductively that

$$|V_x| \ge \gamma^{|x|}N, \tag{3.3}$$

where $|x|$ is the length of x. To begin the induction, we set $V_\varnothing = V(T)$, which satisfies our size hypothesis since $|\varnothing| = 0$. Inductively, suppose we've defined V_x. If $x \in C_2$, we stop. If not, we apply Lemma 3.10 to the induced subtournament on V_x, which we may do since

$$|V_x| \ge \gamma^{|x|}N \ge \gamma^{\mathrm{depth}(\rho_2)-1}N \ge (2^{10}A_1\Delta^2)^{d\,\mathrm{depth}(\rho_1)}\,\mathrm{comp}(\rho_1)^d n.$$

This allows us to find a 1-dense pair (U_1, U_2) of size at least $\gamma|V_x|$. We then set $V_{x0} = U_1$ and $V_{x1} = U_2$, which we see satisfy (3.3) inductively. We continue in this way until we define V_x for every $x \in C_2$. To conclude, suppose that $y, z \in C_2$ are adjacent under ρ_2, and let x be their longest common prefix. Then $V_y \subseteq V_{x0}$ and $V_z \subseteq V_{x1}$, and we know that every edge between V_{x0} and V_{x1} is oriented from V_{x0} to V_{x1}, which implies that (V_y, V_z) is 1-dense, as claimed. □

3.3 The inner stage

In this subsection, we will see how to use the various structures built in the previous subsection to successfully embed a copy of H in any sufficiently large tournament T. The basic idea is that the (ρ, δ)-skeletons constructed in Lemmas 3.9 and 3.11 are precisely the structures we need in order to apply Lemma 3.5 and find a copy of H.

Recall that a prefix coloring is a prefix labeling where the preimage of every codeword is an independent set. Our first result here shows a general upper bound on $\overrightarrow{r_1}(H)$ in terms of the depth and complexity of a prefix coloring of H.

Theorem 3.12. *Let H be an acyclic digraph on n vertices with maximum degree $\Delta \geq 1$. Then for any prefix coloring $\rho : V(H) \to C$, we have $\overrightarrow{r_1}(H) \leq N$, where*

$$N = \left(2^{5\Delta+4}\Delta^{2\Delta+1}\right)^{\operatorname{depth}(\rho)} \operatorname{comp}(\rho)^{\Delta} n.$$

Proof. Let T be a tournament on N vertices, and suppose for contradiction that T is H-free. By Lemma 3.9 applied with $d = \Delta$ and $c = \frac{1}{8\Delta^2}$, we can find in T a $(\rho, 1 - \frac{1}{8\Delta^2})$-skeleton $\{V_x\}_{x \in C}$ of size at least

$$\left(\frac{(8\Delta^2)^{-\Delta}}{4^{\Delta+1}\Delta}\right)^{\operatorname{depth}(\rho)} \operatorname{comp}(\rho)^{-\Delta} N = \left(2^{5\Delta+2}\Delta^{2\Delta+1}\right)^{-\operatorname{depth}(\rho)} \operatorname{comp}(\rho)^{-\Delta} N \geq 4n.$$

Now, for any vertex $v_i \in V(H)$, let $V_i = V_{\rho(v_i)}$. Then since $\rho^{-1}(x)$ is an independent set for any $x \in C$, we see that if $v_i \to v_j$ is an edge of H, then (V_i, V_j) is a $(1 - \frac{1}{8\Delta^2})$-dense pair. Therefore, by Lemma 3.5, we conclude that T contains a copy of H. □

The second result of this subsection uses the $(\rho, 1)$-skeletons we constructed in case we are given a forest prefix labeling. Using this skeleton, we are able to prove the following result, which takes as input a forest prefix labeling and any arbitrary prefix labeling (which may be the same as the forest prefix labeling). The output is a bound on the Ramsey number, in terms of the depth and complexity of the first labeling, and in terms of the maximum Ramsey number of the parts in the partition induced by the second prefix labeling.

Theorem 3.13. *Let $d \geq 2$, and let H be a d-degenerate acyclic digraph on n vertices with maximum degree Δ. Suppose that $\rho_1 : V(H) \to C_1$ is a forest prefix labeling of H and that $\rho_2 : V(H) \to C_2$ is any prefix labeling. Let A_1 and m_1 be the maximum degree and maximum component size of ρ_1, respectively. Then $\overrightarrow{r_1}(H) \leq N$, where*

$$N = \left(m_1(2^{10}A_1\Delta^2)^{d\operatorname{depth}(\rho_1)} \operatorname{comp}(\rho_1)^d\right)^{\operatorname{depth}(\rho_2)} \max\left(n, \max_{x \in C_2} \overrightarrow{r_1}(H[\rho_2^{-1}(x)])\right).$$

Proof. Let T be a tournament on N vertices, and suppose that T is H-free. We apply Lemma 3.11 to T, which allows us to find a $(\rho_2, 1)$-skeleton $\{V_x\}_{x \in C_2}$ where

$$|V_x| \geq \max_{y \in C_2} \overrightarrow{r_1}(H[\rho_2^{-1}(y)]) \geq \overrightarrow{r_1}(H[\rho_2^{-1}(x)])$$

for all $x \in C_2$. In other words, this is a collection of disjoint sets $\{V_x\}_{x \in C_2}$ such that if x precedes y in the lexicographic ordering, then every edge is oriented from V_x to V_y. By the definition of the Ramsey number, we see that the induced subtournament $T[V_x]$ must contain a copy of $H[\rho_2^{-1}(x)]$ for all $x \in C_2$. We pick such a copy arbitrarily for each $x \in C_2$, and observe that their union forms a copy of H in T. □

4 Upper bounds on oriented Ramsey numbers

4.1 Upper bounds in terms of height

Theorem 3.12, which bounds $\overrightarrow{r_1}(H)$ in terms of the dyadic complexity and depth of a prefix coloring, allows us to prove bounds on $\overrightarrow{r_1}(H)$ in terms of other, more natural, parameters. For instance, the next lemma relates the height of an acyclic digraph to its dyadic complexity and depth.

Lemma 4.1. *If H is an acyclic digraph of height h, then there exists a prefix coloring $\rho : V(H) \to C$ with $\mathrm{depth}(\rho) \le \lceil \log h \rceil$ and $\mathrm{comp}(\rho) \le h^{\lceil \log h \rceil}$.*

Proof. Recall that if H has height h, then H has a directed partition $P_0, \dots, P_{h-1}$ into h independent sets. This partition naturally yields a prefix coloring using the prefix code C consisting of all binary strings of length $\lceil \log h \rceil$. Namely, we can label each vertex in P_i by the base-2 representation of i, and this yields a prefix coloring with depth $\lceil \log h \rceil$. To estimate the dyadic complexity of this prefix coloring, note that for any binary string x of length $\ell \le \lceil \log h \rceil$, we have that $a(x) \le 2^{\lceil \log h \rceil - \ell}$, since there are at most $2^{\lceil \log h \rceil - \ell}$ codewords in C prefixed by x. Therefore,

$$\mathrm{comp}(\rho) \le \prod_{\ell=0}^{\lceil \log h \rceil} 2^{\lceil \log h \rceil - \ell} = 2^{\lceil \log h \rceil^2 - \sum_{\ell=0}^{\lceil \log h \rceil} \ell} \le h^{\lceil \log h \rceil}. \qquad \square$$

Then Theorem 1.4 follows as a simple corollary.

Proof of Theorem 1.4. The result is immediate if $\Delta = 1$, so assume $\Delta \ge 2$. Let ρ be the prefix coloring from Lemma 4.1, which has $\mathrm{depth}(\rho) \le \lceil \log h \rceil$ and $\mathrm{comp}(\rho) \le h^{\lceil \log h \rceil}$. By Theorem 3.12,

$$\overrightarrow{r_1}(H) \le \left(2^{5\Delta+4}\Delta^{2\Delta+1}\right)^{\mathrm{depth}(\rho)} \mathrm{comp}(\rho)^{\Delta} n \le 2^{7\Delta \log h} \Delta^{3\Delta \log h} h^{2\Delta \log h} n \le (\Delta h)^{10\Delta \log h} n. \qquad \square$$

Similarly, given a graded digraph, one can find a prefix coloring with small depth and dyadic complexity.

Lemma 4.2. *If H is a graded acyclic digraph of height h, then there exists a prefix coloring $\rho : V(H) \to C$ with $\mathrm{depth}(\rho) \le \lceil \log h \rceil$ and $\mathrm{comp}(\rho) = 1$.*

Proof. The proof is identical to that of Lemma 4.1, except that $a(x) \le 1$ for every binary string x, since the only codeword prefixed by $x0$ that can have edges to a codeword prefixed by $x1$ is the codeword $(x, 0, 1, 1, \dots, 1)$. $\square$

As before, Theorem 1.5 follows as a corollary.

Proof of Theorem 1.5. We may again assume that $\Delta \ge 2$. Let ρ be the prefix coloring from Lemma 4.2, with $\mathrm{depth}(\rho) \le \lceil \log h \rceil$ and $\mathrm{comp}(\rho) = 1$. Then from Theorem 3.12,

$$\overrightarrow{r_1}(H) \le \left(2^{5\Delta+4}\Delta^{2\Delta+1}\right)^{\mathrm{depth}(\rho)} \mathrm{comp}(\rho)^{\Delta} n \le 2^{7\Delta \log h} \Delta^{3\Delta \log h} n \le h^{10\Delta \log \Delta} n. \qquad \square$$

Recall from the introduction that the bandwidth of an n-vertex acyclic digraph H is the least ℓ so that P_n^ℓ contains H. Using the same argument, one can obtain a bound of $\overrightarrow{r_1}(H) \le n^{O_\ell(1)}$ for any n-vertex acyclic digraph of bandwidth at most ℓ, using the fact that the same binary-representation prefix coloring have dyadic complexity at most $n^{O(\log \ell)}$. However, since the Ramsey number of bounded-bandwidth acyclic digraphs is known to be linear [14], we omit the proof of this weaker result.

4.2 Upper bounds for random digraphs

In this section, we prove Theorem 1.3, showing that if $d \geq 1$ is bounded, then w.h.p. $\overrightarrow{r_1}(H)$ is nearly linear when $H = \overrightarrow{G}(n,d)$. Additionally, we prove a nearly-linear upper bound when $p = d/n$ and $H = \overrightarrow{G}(n,p)$.

We will need the following result of Dross and Havet [15] mentioned in the introduction. They only state this result for orientations of trees, but one can easily extend it to orientations of forests by adding edges to join distinct connected components.

Theorem 4.3. *Let H be a 1-degenerate digraph (i.e. an orientation of an undirected forest) on $n \geq 2$ vertices. If T is any tournament on at least $\frac{21}{8}n - \frac{47}{16}$ vertices, then T contains a copy of H.*

We recall that the *vertex arboricity* of an undirected graph is the minimum number of subsets needed to cover the vertices of the graph, such that each subset induces a forest (see e.g. [5] for more on this undirected graph parameter). We first define a natural directed analogue of this quantity, though for technical reasons we also keep track of the maximum component size in such a partition.

Definition 4.4. The *directed vertex arboricity* $\overrightarrow{va}(H)$ of an acyclic digraph H is the minimum size r of a directed partition of H into sets $\{P_i\}_{i=1}^r$ where $H[P_i]$ is 1-degenerate for all i.

Additionally, if H has a directed partition into sets $\{P_i\}_{i=1}^r$ such that $H[P_i]$ is 1-degenerate for all i, and every weakly connected component in $H[P_i]$ has at most m vertices, then we say that H has an (r,m)-*forest partition*.

Both of our upper bound results for random digraphs follow from the following theorem, which yields an upper bound on $\overrightarrow{r_1}(H)$ in terms of the degeneracy, maximum degree, and forest partition size of H.

Theorem 4.5. *Let H be an acyclic digraph on n vertices with maximum degree Δ, degeneracy $d \geq 2$, and with an (r,m)-forest partition. Then $\overrightarrow{r_1}(H) \leq (r\Delta)^{6d(\log r)^2} m^{2\log r} n$.*

Proof. Let $s = \lceil \log r \rceil$, and let C denote the prefix code consisting of all strings of length s. Fix a partition $P_0 \sqcup \cdots \sqcup P_{r-1}$ of $V(H)$ into directed forests such that every edge between P_i and P_j is oriented from P_i to P_j for all $i < j$, and such that $H[P_i]$ has weakly connected components of size at most m. Let $\rho : V(H) \to C$ be the forest prefix labeling mapping P_i to the binary representation of i. Then the maximum degree of ρ is at most $|C| = 2^s \leq 2r$, the maximum component size of ρ is at most m, and $\operatorname{depth}(\rho) = s$. Moreover, we can bound the dyadic complexity of ρ as $\operatorname{comp}(\rho) \leq r^s$, since every binary string is the prefix of at most r codewords in C. We now now apply Theorem 3.13 with $\rho_1 = \rho_2 = \rho$. We recall that since $\rho^{-1}(x)$ is a forest for every x, we have that $\overrightarrow{r_1}(H[\rho^{-1}(x)]) \leq 3n$ by Theorem 4.3. Therefore,

$$\begin{aligned}
\overrightarrow{r_1}(H) &\leq \left(m_1(2^{10}A_1\Delta^2)^{d\operatorname{depth}(\rho_1)}\operatorname{comp}(\rho_1)^d\right)^{\operatorname{depth}(\rho_2)} \max\left(n, \max_{x\in C_2} \overrightarrow{r_1}(H[\rho_2^{-1}(x)])\right) \\
&\leq \left(m(2^{11}r\Delta^2)^{ds}r^{ds}\right)^s \cdot (3n) \\
&\leq (r\Delta)^{6d(\log r)^2} m^{2\log r} n.
\end{aligned}$$

□

It remains to check that both $\overrightarrow{G}(n,d)$ and $\overrightarrow{G}(n,p)$ satisfy the conditions of Theorem 4.5 for appropriate Δ, d, m, and r. Maximum degree and degeneracy are both easy to control for these graphs, so the nontrivial part is bounding their directed vertex arboricity, or more precisely the parameters of a forest partition.

Consider the random digraph $\overrightarrow{G}(n,p)$, where $np = d \geq 1$. The idea for bounding $\overrightarrow{va}(H)$ is to equitably divide $[n]$ into $5d$ intervals I_i, and note that $H[I_i] \sim \overrightarrow{G}(n/5d, p)$ is in the subcritical regime of the Erdős–Rényi random graph process, so w.h.p. is a union of trees and unicyclic components, each comprising $O(\log n)$ vertices. Each interval I_i can be further divided in two to break the unicyclic components, which shows that w.h.p. we have a $(10d, O(\log n))$-forest partition. The analysis is broadly similar for $H = \overrightarrow{G}(n,d)$ but somewhat more technical.

Recall that if $d \geq 1$ and nd is even, a uniformly random undirected d-regular graph $G(n,d)$ can be generated using the "pairings model" (also known as the "configurations model") of Bollobás [3], see the

survey of Wormald [39]. The pairings model of $G(n,d)$ generates a random d-regular multi-graph (with self-loops allowed) by taking a uniformly random perfect matching (a "pairing") on nd points divided into n d-sets, and then contracting each d-set into a single vertex. There are a total of

$$P(nd) = \frac{(nd)!}{(nd/2)! \cdot 2^{nd/2}}$$

such pairings, and any simple d-regular graph is equally likely to be generated. If $d \geq 1$ is fixed and $n \to \infty$, it is known that the probability a pairing generates a simple graph is asymptotic to $e^{(1-d^2)/4}$. To generate an honest $G(n,d)$, repeatedly sample from the above model (expected constant number of samples) until a simple graph is found.

Lemma 4.6. *If $d \geq 2$ is fixed, nd is even, and G is the induced subgraph of $G(n,d)$ on a fixed set of $\frac{n}{5d}$ vertices, then w.h.p. every connected component of G has order at most $2\log n$ and contains at most one cycle.*

Proof. We first show that w.h.p. each component contains at most one cycle.

It is not difficult to show that any *minimal* graph H on k vertices with at least two cycles is formed from a path of length k by adding edges from each of its ends, see e.g. [27, Theorem 5.5]. Thus, there are at most $k^2 \cdot k!$ labelled graphs on k vertices with this property. For a fixed such H, we bound the probability it appears among k fixed vertices $v_1, \ldots, v_k$ in the pairings model for $G(n,d)$. The total number of pairings giving such an H (without giving rise to multi-edges) can be bounded above by

$$d^{2(k+1)} \cdot P(nd - 2(k+1)),$$

since there are at most $(d^2)^{k+1}$ ways to choose the edges between d-sets that correspond to the $k+1$ edges of H, and then at most $P(nd-2(k+1))$ ways to pair the remaining points. Let $\mathcal{E}$ be the event that the pairings model generating a simple graph, which occurs with probability $(1+o(1))e^{(1-d^2)/4}$. We get that

$$\begin{aligned}\Pr[\{v_1,\ldots,v_k\} \text{ is a copy of } H \mid \mathcal{E}] &\leq (1+o(1))e^{-(1-d^2)/4} \cdot \frac{d^{2(k+1)}P(nd-2(k+1))}{P(nd)} \\ &\leq (1+o(1))e^{-(1-d^2)/4} \cdot \frac{d^{k+1}}{n^{k+1}}.\end{aligned}$$

Thus,

$$\Pr[\text{some such } H \text{ appears in } G] \leq (1+o(1))e^{-(1-d^2)/4} \sum_{k \geq 4} \binom{n/(5d)}{k} k^2 \cdot k! \cdot \frac{d^{k+1}}{n^{k+1}} = o(1),$$

which completes the proof that every component of G contains at most one cycle w.h.p.

Next, we show that w.h.p. every connected component has order at most $2\log n$ with a similar computation. For a given set size k, the number of labelled trees H on k vertices is k^{k-2} by Cayley's theorem. We bound the probability that a fixed such H appears among k fixed vertices $v_1, \ldots, v_k$. Similarly to before, we obtain that the total number of pairings giving such an H without multi-edges is bounded above by

$$d^{2(k-1)} \cdot P(nd - 2(k-1)),$$

since there are at most $(d^2)^{k-1}$ ways to choose the edges between d-sets that correspond to the $k-1$ edges of the tree H, and at most $P(nd-2(k-1))$ ways to pair the remaining points. Conditioning on the pairings model generating a simple graph, which occurs with probability $(1+o(1))e^{(1-d^2)/4}$, we get

$$\begin{aligned}\Pr[\{v_1,\ldots,v_k\} \text{ is a copy of } H \mid \mathcal{E}] &\leq (1+o(1))e^{-(1-d^2)/4} \cdot \frac{d^{2(k-1)}P(nd-2(k-1))}{P(nd)} \\ &\leq (1+o(1))e^{-(1-d^2)/4} \cdot \frac{d^{k-1}}{n^{k-1}}.\end{aligned}$$

Taking a union bound over all choices of $v_1, \dots, v_k$,

$$\begin{aligned}\Pr[\text{a tree on } k \text{ vertices appears in } G] &\le (1+o(1))e^{-(1-d^2)/4}\binom{n/(5d)}{k}k^{k-2}\cdot\frac{d^{k-1}}{n^{k-1}}\\ &\le (1+o(1))e^{-(1-d^2)/4}\frac{n^k\cdot k^{k-2}\cdot d^{k-1}}{(5d)^k\cdot(k/e)^k\cdot n^{k-1}}\\ &\le n\cdot(e/5+o(1))^k,\end{aligned}$$

which is $o(1)$ for $k = 2\log n$. This implies that w.h.p. every component of G has order at most $2\log n$, and completes the proof. □

We can now prove Theorem 1.3.

Corollary 4.7. *For any $d \ge 2$, we have that w.h.p., $\overrightarrow{r_1}(\overrightarrow{G}(n,d)) \le n(\log n)^{4\log d}$ as $n \to \infty$.*

Proof. By Lemma 4.6, we know that w.h.p., any fixed $n/(5d)$ vertices of $\overrightarrow{G}(n,d)$ span a disjoint union of trees and unicyclic components, each of which has order at most $2\log n$. Therefore, by applying the union bound to $5d$ events, we conclude that w.h.p., $\overrightarrow{G}(n,d)$ has a directed partition into $5d$ parts all with this property. Dividing each part in two to split every cycle, we conclude that w.h.p. $\overrightarrow{G}(n,d)$ has a $(10d, 2\log n)$-forest partition. Additionally, since the underlying undirected graph of $\overrightarrow{G}(n,d)$ is d-regular, we see that $\overrightarrow{G}(n,d)$ is d-degenerate and has maximum degree d. So by Theorem 4.5, we conclude that w.h.p. as $n \to \infty$,

$$\overrightarrow{r_1}(\overrightarrow{G}(n,d)) \le (10d^2)^{6d(\log(10d))^2}(2\log n)^{2\log(10d)}n,$$

which is upper-bounded by $n(\log n)^{4\log d}$ for any fixed d and sufficiently large n. □

Let $\overrightarrow{G}(n,p)$ denote the orientation of the Erdős–Rényi random graph $G(n,p)$ on vertex set $[n]$ where all edges are oriented to the right. Similarly to the above, we can prove a nearly-linear upper bound on $\overrightarrow{r_1}(\overrightarrow{G}(n,p))$.

Corollary 4.8. *For any $d \ge 2$, we have that w.h.p., $\overrightarrow{r_1}(\overrightarrow{G}(n,p)) \le n\cdot(\log n)^{O(d(\log d)^2)}$, where $p = d/n$.*

Proof. It is easy to show that if $p = d/n$, then $G(n,p)$ is $O(d)$-degenerate (see [19, Theorem 4.8] for a proof of a stronger result). Additionally, it is well-known that the maximum degree of $G(n,p)$ is $O(\log n/\log\log n)$ for any fixed $d \ge 2$. Finally, by the easier version of Lemma 4.6 (e.g. [27, Theorem 5.5]), we see that $\overrightarrow{G}(n,p)$ has a $(10d, 2\log n)$-forest partition. Hence, Theorem 4.5 implies that $\overrightarrow{r_1}(\overrightarrow{G}(n,p)) \le n\cdot(\log n)^{O(d(\log d)^2)}$. □

5 Multiple colors and ordered Ramsey numbers

In this section we study oriented Ramsey numbers of more than one color, proving Theorems 1.6 and 1.7.

We define an *ordered graph* G to be an undirected graph whose vertex set comes with a total order. If the vertex set is a subset of $\mathbb{N}$, then the total order is assumed to be that of $\mathbb{N}$. If $G_1, \dots, G_k$ are ordered graphs, then the *ordered Ramsey number* $r_<(G_1, \dots, G_k)$ is the minimum N such that any edge-coloring of the complete graph on $[N]$ in colors $1, \dots, k$ contains a monochromatic copy of G_i in color i for some i. Here an *ordered copy* of G_i is a subgraph isomorphic to G_i with vertices appearing in the same order. We write $r_<(G;k) := r_<(G, \dots, G)$ when all the graphs G_i are the same.

5.1 The lower bound

To prove Theorem 1.6, we need the following theorem from Conlon, Fox, Lee, and Sudakov [11, Theorem 2.3].

Theorem 5.1. *If M is a random matching on vertex set $[n]$, then w.h.p.,*

$$r_<(M;2) > n^{\log n/20 \log\log n}.$$

We remark that the existence of ordered matchings M with $r_<(M;2) > n^{\log n/20\log\log n}$ was proven independently by Balko, Cibulka, Král, and Kynčl [1]. We will only need this weaker result.

If H is an acyclic digraph with a Hamiltonian (directed) path, we say H is *Hamiltonian*. It has a unique vertex ordering $v_1, \ldots, v_n$ where consecutive vertices are adjacent and all edges point forwards. We assign to H a natural ordered graph H^+ on $[n]$ where $i \sim j$ if and only if $v_i \to v_j$ in H.

Lemma 5.2. *If $k \geq 1$ and H is an acyclic Hamiltonian digraph, then*

$$\overrightarrow{r_k}(H) \geq r_<(H^+;k).$$

Proof. If $N = r_<(H^+;k) - 1$, there exists a k-edge-coloring χ of the complete graph on $[N]$ in which there is no monochromatic copy of H^+. Let T be the transitive tournament on $[N]$ with all edges oriented forwards (i.e. $i \to j$ if and only if $i < j$), and the edge between $i \to j$ also colored $\chi(i,j)$. Since H has a Hamiltonian path and all edges in T point forwards in $[N]$, any copy of H in T corresponds to an ordered copy of H^+ in χ. By construction, there is no monochromatic copy of H in T, as desired. □

The two results above together can be used to prove Theorem 1.6.

Proof of Theorem 1.6. Since $\overrightarrow{r_k}(H)$ is nondecreasing in k, it suffices to prove the result for $k = 2$. By Theorem 5.1, there exists a matching M on $[n]$ which satisfies $r_<(M;2) > n^{\log n/20\log\log n}$. Define H to be the acyclic digraph on $[n]$ where $i \to j$ if $i < j$ and either $j = i+1$ or (i,j) is an edge of M. By Lemma 5.2, $\overrightarrow{r_2}(H) \geq r_<(H^+;2)$. On the other hand, M is a ordered subgraph of H^+, so $r_<(H^+;2) \geq r_<(M;2)$. It follows that

$$\overrightarrow{r_2}(H) \geq r_<(M;2) > n^{\log n/20\log\log n},$$

as desired. Note that H is the edge union of a path and a matching, so it has maximum degree 3. □

5.2 The upper bound

To prove Theorem 1.7, we upper bound k-color oriented Ramsey numbers by $2k$-color ordered Ramsey numbers. Let H^- be the ordered graph obtained from H^+ by reversing the vertex order.

Lemma 5.3. *If $k \geq 1$ and H is an acyclic digraph, then*

$$\overrightarrow{r_k}(H) \leq r_<(\underbrace{H^+, \ldots, H^+}_{k}, \underbrace{H^-, \ldots, H^-}_{k}).$$

Proof. Let T be a tournament on

$$N = r_<(\underbrace{H^+, \ldots, H^+}_{k}, \underbrace{H^-, \ldots, H^-}_{k})$$

vertices, with an edge-coloring χ using colors $1, \ldots, k$. Arbitrarily identify $V(T)$ with $[N]$ and define a $(2k)$-edge-coloring χ' of K_N by

$$\chi'(i,j) = \begin{cases} \chi(i,j) & \text{if } i < j \text{ and } i \to j \\ \chi(i,j) + k & \text{else.} \end{cases}$$

By the definition of N, there is either some color $c \leq k$ where χ' has an ordered copy of H^+ in color c, or some color $c > k$ where χ' has an ordered copy of H^- in c. In the former case, T contains a monochromatic copy of H in color c with all vertices pointed forwards in the arbitrary ordering, and in the latter case T contains a monochromatic copy of H in color $c-k$ with all edges pointed backwards. In either case, T contains a monochromatic copy of H, as desired. □

To prove an upper bound on $\overrightarrow{r_k}(H)$, it remains to generalize the following upper bound of [11] on 2-color ordered Ramsey numbers to more colors.

Theorem 5.4 ([11, Theorem 3.6])**.** *If H is an ordered graph on at most n vertices with degeneracy $d \geq 2$, then*

$$r_<(H, K_n) \leq 2^{O(d \log^2(2n/d))}.$$

The multicolor bound can now be obtained by iterating the above theorem.

Theorem 5.5. *If $k, d \geq 2$ and $H_1, \ldots, H_{k-1}$ are d-degenerate ordered graphs on at most n vertices, then*

$$r_<(H_1, \ldots, H_{k-1}, K_n) \leq 2^{O_{k,d}(\log^{2^{k-1}} n)}.$$

Proof. We prove the theorem by induction on k. The base case $k = 2$ is just Theorem 5.4. For the inductive step, note that if $M = r_<(H_2, \ldots, H_{k-1}, K_n) \leq 2^{O(\log^{2^{k-2}} n)}$ then one obtains

$$r_<(H_1, H_2, \ldots, H_{k-1}, K_n) \leq r_<(H_1, K_M),$$

by combining the last $k-1$ colors into one "super-color." It follows by applying the base case that

$$r_<(H_1, H_2, \ldots, H_{k-1}, K_n) \leq r_<(H_1, K_M) \leq 2^{O(\log^2 M)} \leq 2^{O(\log^{2^{k-1}} n)},$$

as desired. □

Theorem 1.7 follows by combining Lemma 5.3 with Theorem 5.5.

6 Concluding Remarks

In this section we collect a few appealing open problems on the Ramsey numbers of digraphs. For $k, \Delta \geq 1$ and $n > \Delta$, let $H_{k,n,\Delta}$ be an acyclic digraph H with n vertices and maximum degree Δ maximizing the value of $\overrightarrow{r_k}(H)$. Much of this paper was devoted to understanding the growth rate of $\overrightarrow{r_k}(H_{k,n,\Delta})$ for fixed k and Δ.

We first consider the one-color case. Theorem 1.2, Lemma 5.3, and Theorem 5.4 together show

$$n^{\Omega(\Delta^{2/3}/\log^{5/3}\Delta)} \leq \overrightarrow{r_1}(H_{1,n,\Delta}) \leq 2^{O(\Delta \log^2(2n/\Delta))} \tag{6.1}$$

While we do not know whether the above Ramsey number grows polynomially or super-polynomially in n for $k = 1$, we also showed that for $k \geq 2$ and $\Delta \geq 3$, $\overrightarrow{r_k}(H_{k,n,\Delta})$ is at least $n^{\Omega(\log n/\log\log n)}$. We conjecture that a super-polynomial growth rate is also possible for $k = 1$.

Conjecture 6.1. *There exists an absolute constant Δ such that*

$$\overrightarrow{r_1}(H_{1,n,\Delta}) \geq n^{\omega(1)}.$$

In the case of $k \geq 2$ colors, Theorems 1.6 and 1.7 together imply

$$n^{\Omega(\log n/\log\log n)} \leq \overrightarrow{r_k}(H_{k,n,\Delta}) \leq 2^{O(\log^{2^{2k-1}} n)}. \tag{6.2}$$

It would be interesting to determine even the logarithmic order of $\overrightarrow{r_k}(H_{k,n,\Delta})$.

Problem 6.2. For any fixed $k \geq 1$ and $\Delta \geq 2$, determine the order of growth of $\log \overrightarrow{r_k}(H_{k,n,\Delta})$.

To our knowledge, the only solved case of Problem 6.2 is $k = 1$, $\Delta = 2$. Here, $H_{1,n,2}$ must be a disjoint union of arbitrarily oriented paths and non-strongly oriented cycles (an acyclic H cannot contain a strongly oriented cycle). Thomason [37] proved that if n is large enough, $\overrightarrow{r_1}(C_n) = n$ for any non-strongly oriented cycle C_n on n vertices, which can be used to show $\overrightarrow{r_1}(H_{1,n,2}) = n + O(1)$.

It would be interesting to improve either side of (6.2) substantially. We tentatively conjecture that neither side is close to the truth, in the following quantitative form.

Conjecture 6.3. *There exist $\Delta \geq 3$ and constants $c_k = \omega_k(1)$ and $C_k = 2^{o(k)}$ such that*

$$\Omega(\log^{c_k} n) \leq \log \overrightarrow{r_k}(H_{k,n,\Delta}) \leq O(\log^{C_k} n).$$

In the one-color case, another open problem we find interesting is to determine more precisely the Ramsey number of a sparse random digraph. Theorem 1.3 shows that for fixed d, the Ramsey number of $\overrightarrow{G}(n,d)$ is w.h.p. bounded above by $n(\log n)^{O_d(1)}$. We expect that the answer is in fact linear.

Conjecture 6.4. *If $d \geq 2$ is fixed and $H = \overrightarrow{G}(n,d)$, then w.h.p. $\overrightarrow{r_1}(H) = O_d(n)$.*

This would follow from our techniques if one could prove the following strengthening of Lemma 3.3, in which the size of the 1-dense pair depends only on the maximum degree Δ (and not on the number m of vertices).

Conjecture 6.5. *For every $\Delta \geq 1$, there exists $C_\Delta > 0$ such that the following holds. Let H be a 1-degenerate digraph with maximum degree Δ on m vertices $v_1, \ldots, v_m$, and let T be an arbitrary tournament. If there exist sets $U_1, \ldots, U_m \subseteq V(T)$, each of size $M \geq C_\Delta m$, such that there is no embedding $\phi : H \hookrightarrow T$ satisfying $\phi(v_i) \in U_i$, then T contains a 1-dense pair with size at least M/C_Δ.*

Indeed, if one could prove Conjecture 6.5, then it would imply that Lemma 3.10, Lemma 3.11, Theorem 3.13, and Theorem 4.5 would all no longer depend on m, the maximum size of a tree component in a directed partition of H into oriented forests. In particular, this would imply Conjecture 6.4. Moreover, it is entirely possible that Conjecture 6.5 could be proven using similar greedy embedding arguments, since the conjecture is not hard to prove if the injectivity constraint on ϕ is removed.

We would also like to highlight one powerful digraph embedding technique that we have not used in this paper, but are hopeful can be incorporated into our arguments to prove more general results. The *median ordering* of a tournament T is the vertex ordering $v_1, \ldots, v_n$ maximizing the number of forward edges. To see the power of the median ordering, note that $v_i \to v_{i+1}$ for every $1 \leq i \leq n-1$ in this ordering, so this immediately exhibits a Hamiltonian path in T. Previous work showing linear upper bounds on $\overrightarrow{r_1}(H)$ when H is an oriented tree (e.g. [16]) or an acyclic digraph of bounded bandwidth [14] all depend on embedding H into a tournament T in some iterative manner according to its median ordering. We were not able to reproduce these upper bounds using greedy embedding arguments, which seem primarily suited for embedding digraphs H without long paths.

Finally, we recall from the introduction the *directed Ramsey number* $\overleftrightarrow{r_k}(H)$, which is defined as the least N such that every k-coloring of the edges of $\overleftrightarrow{K_N}$ contains a monochromatic copy of H. It is easy to see that $\overrightarrow{r_k}(H) \geq \overleftrightarrow{r_k}(H)$ for any k and any acyclic H. Indeed, if $N = \overrightarrow{r_k}(H)$, then given a k-edge-coloring of $\overleftrightarrow{K_N}$, we may ignore one edge from each anti-parallel pair to obtain a k-edge-colored N-vertex tournament, which contains a monochromatic H. There is also an inequality in the other direction, whose proof is identical to that given in Lemma 5.3, and which states that $\overrightarrow{r_k}(H) \leq \overleftrightarrow{r_{2k}}(H, \ldots, H, H', \ldots, H')$, where there are k copies of H and k of H', and H' is obtained from H by reversing all the edges. Thanks to these connections, one can convert many of our results on oriented Ramsey numbers to results on directed Ramsey numbers. For example, Theorem 1.7 immediately implies a quasi-polynomial upper bound on $\overleftrightarrow{r_k}(H)$ for any bounded-degree acyclic digraph H. In the other direction, since reversing the edges of any interval mesh yields another interval mesh, Theorem 1.2 shows the existence of a bounded-degree acyclic digraph H with $\overleftrightarrow{r_2}(H)$ which grows faster than any fixed polynomial in the number of vertices of H. For $k \geq 4$ colors, we can similarly use Theorem 1.6 to produce a bounded-degree acyclic digraph H such that $\overleftrightarrow{r_k}(H)$ grows super-polynomially. However, there is an interesting intermediate case at $k = 3$ colors, and we end with the following conjecture, which may be easier than Conjecture 6.1.

Conjecture 6.6. *There is an absolute constant Δ and an infinite sequence $\{H_n\}$ of n-vertex acyclic digraphs with maximum degree at most Δ and $\overleftrightarrow{r_3}(H_n) \geq n^{\omega(1)}$.*

Acknowledgments. We would like to thank Jasmine Yan for producing Figure 3.1, and the anonymous referee for many helpful comments.

References

[1] M. Balko, J. Cibulka, K. Král, and J. Kynčl, Ramsey numbers of ordered graphs, *Electron. J. Comb.* **27** (2020), P1.16.

[2] J.-C. Bermond, Some Ramsey numbers for directed graphs, *Discrete Math.* **9** (1974), 313–321.

[3] B. Bollobás, A probabilistic proof of an asymptotic formula for the number of labelled regular graphs, *European J. Combin.* **1** (1980), 311–316.

[4] M. Bucić, S. Letzter, and B. Sudakov, Directed Ramsey number for trees, *J. Combin. Theory Ser. B* **137** (2019), 145–177.

[5] S. A. Burr, An inequality involving the vertex arboricity and edge arboricity of a graph, *J. Graph Theory* **10** (1986), 403–404.

[6] S. A. Burr and P. Erdős, On the magnitude of generalized Ramsey numbers for graphs, in: *Infinite and Finite Sets I*, Colloq. Math. Soc Janos Bolyai 10, North-Holland, Amsterdam (1975), 214–240.

[7] V. Chvátal, Monochromatic paths in edge-colored graphs, *J. Combin. Theory Ser. B* **13** (1972), 69–70.

[8] V. Chvátal, 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. Conlon, A new upper bound for diagonal Ramsey numbers, *Ann. of Math.* **170** (2009), 941–960.

[10] D. Conlon, The Ramsey number of dense graphs, *Bull. Lond. Math. Soc.* **45** (2013), 483–496.

[11] D. Conlon, J. Fox, C. Lee, and B. Sudakov, Ordered Ramsey numbers, *J. Combin. Theory Ser. B* **122** (2017), 353–383.

[12] D. Conlon, J. Fox, and B. Sudakov, Recent developments in graph Ramsey theory, *London Math. Soc. Lecture Note Ser.* **424** (2015), 49–118.

[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] 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.

[15] F. Dross and F. Havet, On the unavoidability of oriented trees, *Electron. Notes Theor. Comput. Sci.* **346** (2019), 425–436.

[16] A. El Sahili, Trees in tournaments, *J. Combin. Theory Ser. B* **92** (2004), 183–187.

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

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

[19] J. Fox and B. Sudakov, Two remarks on the Burr–Erdős conjecture, *European J. Combin.* **30** (2009), 1630–1645.

[20] T. Gallai, On directed paths and circuits, *Theory of graphs (Proc. Colloq., Tihany, 1966)* (1968), 115–118.

[21] R. L. Graham, V. Rödl, and A. Ruciński, On graphs with linear Ramsey numbers, *J. Graph Theory* **35** (2000), 176–192.

[22] A. Gyárfás and J. Lehel, A Ramsey-type problem in directed and bipartite graphs, *Period. Math. Hungar.* **3** (1973), 299–304.

[23] R. Häggkvist and A. Thomason, Trees in tournaments, *Combinatorica* **11** (1991), 123–130.

[24] M. Hasse, Zur algebraischen Begründung der Graphentheorie. I, *Mathematische Nachrichten* **28** (1965), 275–290.

[25] F. Havet, Trees in tournaments, *Discr. Math.* **243** (2002), 121–134.

[26] 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.

[27] S. Janson, T. Łuczak, and A. Ruciński, *Random graphs*, Wiley, New York, 2000.

[28] A. Kostochka and B. Sudakov, On Ramsey numbers of sparse graphs, *Combin. Probab. Comput.* **12** (2003), 627–641.

[29] 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.

[30] C. Lee, Ramsey numbers of degenerate graphs, *Ann. of Math.* **185** (2017), 791–829.

[31] L. Lovász, *Graphs and Geometry*, American Mathematical Society Colloquium Publications, American Mathematical Society, Providence, RI, 2019.

[32] B. Roy, Nombre chromatique et plus longs chemins d'un graphe, *Rev. Française Informat. Recherche Opérationnelle* **1** (1967), 129–132.

[33] A. Sah, Diagonal Ramsey via effective quasirandomness, 2020. Preprint available at arXiv:2005.09251.

[34] J. Spencer, Ramsey's theorem—a new lower bound, *J. Combin. Theory Ser. A* **18** (1975), 108–115.

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

[36] B. Sudakov, A conjecture of Erdős on graph Ramsey numbers, *Adv. Math.* **227** (2011), 601–609.

[37] A. Thomason, Paths and cycles in tournaments, *Trans. Amer. Math. Soc.* **296** (1986), 167–180.

[38] L. M. Vitaver, Determination of minimal coloring of vertices of a graph by means of Boolean powers of the incidence matrix, *Dokl. Akad. Nauk SSSR* **147** (1962), 728.

[39] N. C. Wormald, Models of random regular graphs. In *Surveys in combinatorics,* 1999 (Canterbury), volume 267 of *London Math. Soc. Lecture Note Ser.*, pages 239–298. Cambridge Univ. Press, Cambridge, 1999.

[40] R. Yuster, Paths with many shortcuts in tournaments, *Discrete Math.* **334** (2021), 112–168.