Combinatorics, Probability and Computing (2020), **29**, pp. 318–345
doi:10.1017/S0963548319000373

CAMBRIDGE UNIVERSITY PRESS

ARTICLE

Monochromatic trees in random tournaments

Matija Bucić[1], Sven Heberle[1], Shoham Letzter[2,†] and Benny Sudakov[1,*,‡]

[1]Department of Mathematics, ETH Zürich, Raemistrasse 101, 8092 Zürich, Switzerland and [2]ETH Institute for Theoretical Studies, ETH Zürich, Clausiusstrasse 47, 8092 Zürich, Switzerland
*Corresponding author. Email: benjamin.sudakov@math.ethz.ch

(Received 18 September 2018; revised 6 August 2019; first published online 7 November 2019)

Abstract
We prove that, with high probability, in every 2-edge-colouring of the random tournament on n vertices there is a monochromatic copy of every oriented tree of order $O(n/\sqrt{\log n})$. This generalizes a result of the first, third and fourth authors, who proved the same statement for paths, and is tight up to a constant factor.

2010 MSC Codes: 05C05, 05C55, 05C80

1. Introduction

Ramsey theory consists of a considerable amount of mathematical results, which, roughly speaking, say that there is no completely chaotic structure, that is, any sufficiently large structure is guaranteed to have a large well-organized substructure. For instance, the famous theorem of Ramsey [15] states that for any fixed graph H, every 2-edge-colouring of a sufficiently large complete graph contains a monochromatic copy of H. The smallest order of a complete graph with this property is called the *Ramsey number of H*.

In this paper we study an analogous phenomenon for oriented graphs. An *oriented graph* is a directed graph G obtained by orienting the edges of a simple undirected graph, which is called the *underlying graph* of G.

A *tournament* is an oriented graph whose underlying graph is complete. Given oriented graphs G, H, K, we write $G \to (H, K)$ whenever in every 2-colouring of the edges of G there is a blue copy of H or a red copy of K. In the special case that $H = K$, we write $G \to H$. The *oriented Ramsey number* of H is defined to be the smallest N for which every tournament G on N vertices satisfies $G \to H$.

Note that unlike the standard Ramsey numbers, which are always finite, in the oriented setting, if H contains a directed cycle then its oriented Ramsey number may be infinite. To see this, consider the following colouring: fix an ordering of the vertices and colour all forward edges blue and all backward edges red. This 2-coloured tournament does not contain any monochromatic directed cycles. In particular, it does not have a monochromatic copy of H if H contains a directed cycle. Moreover, it is easy to see that every 2-edge-coloured tournament on N vertices contains a monochromatic transitive tournament on roughly $\log_4 N$ vertices, from which it follows that

†Research supported by Dr Max Rössler, the Walter Haefner Foundation and the ETH Zurich Foundation.
‡Research supported in part by SNSF grant 200021-175573.

it contains a monochromatic copy of every acyclic graph on at most $\log_4 N$ vertices. Hence, the oriented Ramsey number is finite if and only if H is acyclic.

Let us start by investigating Ramsey numbers of directed paths. Denote the directed path on n vertices by $\overrightarrow{P_n}$, where by a directed path we mean an oriented graph obtained from a path by orienting all its edges in the same direction. The celebrated Gallai–Hasse–Roy–Vitaver theorem [10, 14, 17, 19] says that any directed graph whose underlying graph has chromatic number at least n contains a $\overrightarrow{P_n}$ as a subgraph. It follows that $G \to \overrightarrow{P_n}$ for every tournament G of order at least $(n-1)^2+1$; indeed, given a red and blue colouring of G, either the graph of red edges or the graph of blue edges has chromatic number at least n, implying the existence of a monochromatic path on at least n vertices. This statement is sharp. To see this, consider a *transitive tournament* on $(n-1)^2$ vertices. We partition the vertices into sets A_i, each of size $n-1$, while preserving the order. We colour all edges inside some set A_i blue, and all other edges red. It is easy to see that there is no monochromatic path on n vertices in this colouring. This shows that the oriented Ramsey number of $\overrightarrow{P_n}$ is $(n-1)^2+1$.

It is interesting to consider oriented Ramsey numbers of further acyclic graphs, and the natural next example is that of trees. It turns out that oriented trees behave similarly to paths in terms of their oriented Ramsey numbers: Bucić, Letzter and Sudakov [4] proved that, given any (oriented) tree T on n vertices and any tournament G on cn^2 vertices (where c is a positive constant), we have $G \to T$, that is, the oriented Ramsey number of any tree of order n is at most cn^2.

This result resolves (up to a constant factor) the question of, given n, finding the smallest N such that every 2-colouring of every tournament of order N is guaranteed to have a monochromatic copy of T for any tree T of order at most n. However, intuitively it seems that examples of tournaments for which the bound is tight are close to being transitive. Therefore, it is natural to ask whether in tournaments that are 'far from being transitive' larger monochromatic trees are guaranteed; this question was asked implicitly, for paths, by Ben-Eliezer, Krivelevich and Sudakov [2]. A natural candidate for such a tournament is the *random tournament*, in which the orientation of each edge is chosen independently and uniformly at random. They showed that, with high probability, every 2-colouring of a random tournament on N vertices contains a monochromatic directed path of length at least $cN/(\log N)$. They also showed that every tournament of order N can be 2-coloured without creating monochromatic paths of length $3N/\sqrt{\log N}$, using the following 2-colouring of a given tournament G of order N. It is well known and easy to see that any tournament of order N has a transitive subtournament of order $\log N$. Using this, we can partition the vertices of G into transitive subtournaments A_i of order $(\log N)/2$ and a remainder A_0 of at most $\sqrt{N}$ vertices. We now 2-colour each of A_i, as described above, to ensure that the longest monochromatic path within A_i is of length $\sqrt{|A_i|}$, and we colour the edges in A_0 arbitrarily. We then colour all edges from A_i to A_j blue if $i < j$ and red if $i > j$. In this colouring, the longest monochromatic path has length at most

$$\frac{2N}{\log N}\sqrt{\frac{\log N}{2}} + \sqrt{N} \leqslant \frac{3N}{\sqrt{\log N}}.$$

In a later paper Bucić, Letzter and Sudakov [5] showed that, with high probability, any 2-colouring of a random tournament on N vertices contains a monochromatic directed path of order at most $cN/\sqrt{\log N}$, which is tight up to a constant factor, due to the above upper bound from [2]. They also showed that the same result holds for *oriented paths*, which are paths in which edges are not required to follow the same direction. Following up in this direction, they asked whether the same holds for general oriented trees. The main result of this paper answers this question in the affirmative.

Theorem 1.1. *There is a constant $c > 0$ such that, with high probability, a random tournament G on N vertices satisfies $G \to T$, where T is any oriented tree on at most $cN/\sqrt{\log N}$ vertices.*

Note that unlike for the standard Ramsey numbers, where the ground graph is complete on N vertices, the oriented Ramsey numbers allow any tournament on N vertices as a ground graph. This suggests that taking the ground graph to be the complete directed graph is perhaps a more natural directed analogue of the standard Ramsey theory, where the *complete directed graph* on n vertices, denoted by $\overleftrightarrow{K}_N$, is the graph in which, between any two vertices $i \neq j$, both directed edges, ij and ji, are present. Harary and Hell [13] and Bermond [3] introduced the notion of the *directed Ramsey number* of an oriented graph H, which is defined to be the least N such that every 2-edge-colouring of $\overleftrightarrow{K}_N$ contains a monochromatic copy of H. The directed Ramsey numbers of directed paths were determined by Gyárfás and Lehel [12], based on a result of Raynaud [16], and, independently, by Williamson [20]. Bucić, Letzter and Sudakov [4] generalized these results to oriented trees, and also to the r-coloured variant. This result plays a role in our argument of the proof for Theorem 1.1.

Note that while the problem for random tournaments is seemingly more similar to the oriented Ramsey numbers, as the base graphs in both cases are tournaments, it turns out that the directed Ramsey numbers are more relevant for our arguments. The main reason is that for random tournaments and complete directed graphs between any two not-too-small sets of vertices A and B there are many edges from A to B. However, because this does not hold for smaller sets (*i.e.* of order at most about $\log N$), the bound for random tournaments is somewhat worse than for complete graphs. Our proof of Theorem 1.1 relies only on a property of this kind, so the conclusion of Theorem 1.1 actually holds for any sufficiently pseudorandom tournament; we refer the reader to Section 3 for more details.

1.1 Organization of the paper

In the next section we give an overview of the proof of Theorem 1.1. In Section 3 we introduce some results that we will need throughout the rest of the paper. We then turn to the proof of the asymmetric generalization of Theorem 1.1, which we split into two parts. The first part is presented in Section 4 and deals with the special case when one of the trees is assumed to be a directed path. The second part of the argument, presented in Section 5, shows how to use this special case to obtain the general result.

We do not make any effort to optimize the constants presented in this paper. We also neglect rounding whenever it is not relevant for the argument. Given a 2-colouring of a graph, we call the colours red and blue. When we consider paths and trees we always assume they are oriented, that is, between two vertices there is at most one directed edge. Logarithms are always taken in base 2, unless stated otherwise.

2. Overview

In this section we give an overview of our arguments. Our aim is to prove that, given n and m, a random tournament G on N vertices satisfies $G \to (T, S)$ for every oriented tree T and S of order n and m, respectively, where $N \geqslant c(n + m + \sqrt{nm}\log(n+m))$ and c is an absolute positive constant. Our proof is divided into two main parts: in the first, we prove it under the assumption that one of T and S is a directed path, and in the second we deduce the general result. In the remainder of this section, we outline the arguments we use in each of these cases.

Tree versus path. This is the longest part of the proof, and more difficult; here T is assumed to be a *directed tree* (*i.e.* its edges are directed from a root or *vice versa*) on m vertices, and S is a directed path $\overrightarrow{P_n}$. We first prove the desired result under the assumption that T has not-too-many leaves (namely, at most $m^{1/6}$ leaves). Our aim is to find a red copy of T or a blue copy of $\overrightarrow{P_n}$.

We distinguish three types of cycle: long cycles (length at least $bm^{1/3}$), short cycles (length at most $am^{1/3}$) and medium cycles (all remaining cycles). We now consider two cases: when

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

there exist many pairwise vertex-disjoint medium or long blue cycles, or when there is a large set spanning no medium or long blue cycles. It is easy to see that one of these cases holds.

Case 1: many disjoint medium or long blue cycles. In this case we aim to find a specific structure, which we call *red–blue pairs*. These consist of many pairwise disjoint sets, $A_1, B_1, \ldots, A_t, B_t$, of suitable size, such that all $A_i - B_i$ edges are red and each set A_i is contained in a blue path P_i, where the P_i are pairwise vertex-disjoint (see Figure 4 below).

We show how to use this structure to find the red tree or the blue path of desired length. To this end, we construct a 2-edge-coloured auxiliary complete directed graph, where the edge ij is coloured blue if there are many blue edges going from A_i to A_j in G and red otherwise. Applying the directed Ramsey result for trees (see Theorem 4.5) to this auxiliary graph, we find a long blue path or a certain carefully chosen red tree (this is obtained from a suitable split of the tree into smaller subtrees, which we call a *tree-split*: see Section 4.1 and Figure 1 below).

If we find a blue path, we lift it to a blue $\overrightarrow{P_n}$ in G, making use of the blue paths P_i from our structure. If, instead, we find the red tree, we make use of the red bipartite graphs $G[A_i, B_i]$ to embed a subtree of T within it and connect these embeddings in an appropriate fashion to obtain the full T.

Finally, we explain how to find red–blue pairs by exploiting assumptions on the blue cycle structure in each of the following two subcases.

1(a) *Many disjoint medium blue cycles.* We define an auxiliary 2-coloured complete directed graph H whose vertices are medium blue cycles, and for cycles C_1 and C_2, edge C_1C_2 is blue if a constant fraction of the vertices in C_1 have a blue out-neighbour in C_2, and otherwise the edge is red.

It is easy to see that there is either a large red–red matching in H, which translates into the desired red–blue pairs structure, or there is a long blue path, which translates into a blue $\overrightarrow{P_n}$ in the original tournament.

1(b) *Many disjoint long blue cycles that span no medium blue cycles.* In this case we observe that we can find many disjoint blue cycles with no long blue chords. This allows us to obtain a red–blue pairs structure, by letting the sets A_i and B_i be intervals of the long blue cycles.

Case 2: a large set of vertices spanning no medium or long blue cycle. We first show that, in this case, there exist many pairwise disjoint sets $U_1, \ldots, U_\ell$ of suitable size such that very few of the edges from U_i to U_j, with $i < j$, are blue. Using the version of Theorem 1.1 for paths, which was proved in [5], each set U_i contains many pairs of vertices joined by a long blue path in U_i, or many pairs joined by a long red path; in the former case we say that the set U_i is blue, and otherwise we say that it is red. We now consider two subcases.

2(a) *Most of the sets are red.* In this case we consider a split of the tree T into subpaths (in Section 4.1 we show how to obtain such a *path-split*). We embed each subpath within a specific U_i, where we exploit the fact that we have many options for both start- and end-vertex of the subpath, and the fact that most of the forward edges between the U_i are red, to embed and connect the paths and obtain a red T.

2(b) *Most of the sets are blue.* We define an auxiliary 2-coloured complete directed graph K whose vertices are the blue U_i, and an edge U_iU_j is coloured blue if $i > j$ and if there is a blue edge from every large subset of U_i to every large subset of U_j, and red otherwise.

As before, we note that K contains either a large red–red matching, or a long blue directed path. In the latter case we lift the path to a blue $\overrightarrow{P_n}$ in G. If the former holds, we find many large bipartite graphs, corresponding to edges of the matching, such that almost

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

all of their edges are red. We use these graphs and the fact that almost all forward edges between sets U_i are red to embed a red T, similarly to the first case.

Removing the restriction on the number of leaves. Throughout Section 4.2 we were assuming that T has at most $m^{1/6}$ leaves, which was necessary in order to control the number of subtrees we obtain in various splits of T. In Section 4.3 we show how to remove this assumption. For this we introduce another kind of split of T, which we call the *core-split* (see Section 4.1 for details) which splits T into not too many subtrees, each of which has at most $m^{1/6}$ leaves. Assuming there is no blue $\overrightarrow{P_n}$, we find a short sequence of large sets such that each has a large number of red out-neighbours in the next set of the sequence. This we can do because otherwise we show there is a set which has a lot of blue edges which allow us to find the blue $\overrightarrow{P_n}$. Finally, we iteratively find parts of the core-split (or find a blue $\overrightarrow{P_n}$) within these sets using the result from the previous subsection, where we use the large red out-degree towards the next set to ensure we can join all the pieces into a red copy of T.

Tree versus tree. The rest of the argument consists of three intermediate steps, which generalize the result obtained in the previous section, with the final goal being a version of Theorem 1.1 for general trees T and S.

Step 1: directed tree versus directed tree with $O(1)$ leaves. Let T be out-directed with $O(1)$ leaves, and let S be a directed tree. We observe that if we remove paths from a directed tree T, that start at any leaf and stop right before a branching vertex or the root, then the resulting tree T' has at most half the number of leaves of T. We iterate a procedure which reduces the search for a red T or a blue S to a search of red T' or blue S, using the previous case of path versus tree.

Step 2: directed tree versus directed tree. Let T and S be out-directed trees. Our aim is to iterate a procedure that reduces the search of a red T or a blue S to a search for a red T_1 or a blue S_1, where the order of T_1 and S_1 is smaller than the order of T and S by at least a constant factor. To that end, we consider the k-core of a tree T, which is the subtree T' consisting of vertices whose number of descendants is at least $|T|/k$. One can show that T' has at most k leaves and that the trees in the forest $T \setminus V(T')$ have order at most $|T|/k$ (see Definition 3.4). We make use of the previous step, which tells us that we can find a red T' or a blue S if T' is the k-core of T, where $k = O(1)$. Subsequently, we try to embed the trees in $T \setminus V(T')$ within the correct out-neighbourhoods. If we succeed, we have found a red T; otherwise the tree at which we fail is our T_1. We repeat in blue to obtain S_1 and iterate until one of the trees drops to constant size, when we once again appeal to the previous result.

Step 3: tree versus tree. Here we rely on the following idea: if A and B are sets such that every vertex in A has large out-degree in B and the vertices in B have large in-degree in A, then given a general tree in T, we can aim to embed in-directed subtrees of T in A and out-directed subtrees of T in B, using the large degrees between the two sets to connect such subtrees. This idea allows us to go from the previous step, where we search for monochromatic directed trees, to a search for a red directed tree or a blue general tree. We then use this idea again to obtain the desired result for two general trees.

3. Prerequisites

In this section we mention some useful facts which we shall use throughout the proof. First, we introduce the notion of pseudorandomness. Let G be an oriented graph. For two disjoint subsets A, B of the vertices, we let $e_G(A, B)$ denote the number of edges directed from A towards B; when the graph G is clear from the context, we omit the subscript G. For a vertex v we denote the out- and in-degree of v by $d^+(v)$ and $d^-(v)$.

Definition 3.1. Let $0 < \varepsilon < 1/2$ and let k be an integer. An oriented graph G is (ε, k)-*pseudorandom* if, for any disjoint sets $A, B \subseteq V(G)$ of size at least k, we have $e(A, B) \geqslant \varepsilon|A||B|$.

It is easy to see, for example by Chernoff's inequality, that a random tournament is pseudorandom with high probability, as stated in the following lemma (see Lemma 6 in [5]). In fact, this is the only property of a random tournament that we shall use in our argument.

Lemma 3.2. *Let $0 < \varepsilon < 1/2$. There exists a constant σ such that a random tournament T is $(\varepsilon, \sigma \log |T|)$-pseudorandom, with high probability.*

We shall investigate 2-colourings of graphs, where the colours are called red and blue. Therefore, we extend the notation related to edges by an index b for blue and r for red edges. For example, $e_r(A, B)$ denotes the number of red edges going from A to B and similarly $d_b^+(v)$ is the blue out-degree of vertex v.

The following lemma gives a lower bound on the number of blue edges in a subset of the vertices which contains no red copy of a particular tree. This will allow us to find large sets where every vertex has many red neighbours.

Lemma 3.3. *Let G be an $(\varepsilon, \sigma \log N)$-pseudorandom 2-coloured tournament on N vertices. Suppose that $U \subseteq V(G)$ has the following properties:*

(i) *the induced graph $G[U]$ has at most $(\varepsilon^2/32)|U|^2$ blue edges, and*
(ii) $(\varepsilon/4)|U| \geqslant \sigma \log N$.

Then the graph G contains a red copy of any tree of size $(\varepsilon/4)|U|$.

Proof. Let us consider two sets

$$X^+ = \left\{ v \in U \,\middle|\, d_r^+(v) < \frac{3\varepsilon}{4}|U| \right\} \quad \text{and} \quad X^- = \left\{ v \in U \,\middle|\, d_r^-(v) < \frac{3\varepsilon}{4}|U| \right\},$$

where the degrees are with respect to the induced subgraph $G[U]$. We are going to show that both sets have size at most $(\varepsilon/4)|U|$. If this is the case then the induced graph $G[U \setminus (X^+ \cup X^-)]$ has minimum red in- and out-degree at least

$$\frac{3\varepsilon}{4}|U| - \frac{2\varepsilon}{4}|U| = \frac{\varepsilon}{4}|U|,$$

and then we can greedily find any red tree of size at most $(\varepsilon/4)|U|$.

So let us assume that $|X^+| \geqslant (\varepsilon/4)|U|$; the argument for X^- is analogous. Let us pick any $(\varepsilon/4)|U|$ vertices from X^+ and denote this set by Y. Then

$$e_r(Y, U \setminus Y) < |Y| \cdot \frac{3\varepsilon}{4}|U| = \frac{3\varepsilon^2}{16}|U|^2.$$

By the first assumption on U, we have

$$e(Y, U \setminus Y) \leqslant e_r(Y, U \setminus Y) + e_b(Y, U \setminus Y) < \left(\frac{3}{16} + \frac{1}{32}\right)\varepsilon^2|U|^2 = \frac{7\varepsilon^2}{32}|U|^2. \tag{3.1}$$

However, by pseudorandomness and the lower bound on $|U|$, we have

$$e(Y, U \setminus Y) \geqslant \varepsilon|Y|(|U| - |Y|) \geqslant \frac{\varepsilon^2}{4}\left(1 - \frac{\varepsilon}{4}\right)|U|^2 \geqslant \frac{7\varepsilon^2}{32}|U|^2,$$

where the last inequality follows as $\varepsilon < 1/2$. This is a contradiction to (3.1), which implies that

$$|X^+| < \frac{\varepsilon}{4}|U|,$$

as required. □

A *rooted tree* is a tree with a special vertex which we call the *root*. By removing a vertex v in a rooted tree T we obtain a forest F. The *descendants* of v are the vertices of T that are not in the tree in F which contains the root; note that each vertex is a descendant of itself.

Definition 3.4. Let T be a rooted tree on n vertices and let $k > 1$. The *k-core* of T is the subtree of T consisting of vertices that have more than n/k descendants in T.

Observation 3.5. Let T be a tree on n vertices and let T' be its k-core for $k > 1$. Then T' has at most k leaves and every tree of the forest $T \setminus V(T')$ has order at most n/k.

Proof. Suppose that T' has k non-root leaves (the root of T' is the root of T). The sets of descendants in T of each leaf of T' are disjoint and have size greater than n/k. This implies that T has order greater than n, a contradiction. Therefore, T' has at most $k - 1$ non-root leaves, so in total it has at most k leaves.

Let S be a tree in the forest $T \setminus V(T')$. Suppose $|S| > n/k$, then the root v of S has more than n/k descendants, but then v should be in T', a contradiction. □

The next result makes it possible to bound the number of vertices with degree at least 3 in the underlying graph; we call such vertices *branching*. Note that a tree is a path if and only if it has no branching vertices. Let $\mathrm{lf}(T)$ be the number of leaves in a tree T.

Lemma 3.6. *The number of branching vertices is at most* $\mathrm{lf}(T) - 1$.

Proof. We argue by induction on the number of leaves $k := \mathrm{lf}(T)$. If $k \leqslant 2$, the tree is a path, and paths do not have any branching vertices.

For the induction step we assume that the statement holds for all trees with $k - 1$ leaves. Let v be any leaf of T and P a path from v to the first vertex adjacent to a branching vertex w, which exists as $k \geqslant 3$. Then $T \setminus V(P)$ is a tree with $k - 1$ leaves and by induction has at most $k - 2$ branching vertices. It follows that the number of branching vertices in T is at most $k - 1$ (as w is the only vertex that is branching in T but need not be branching in $T \setminus V(P)$). □

We call an oriented tree T *out-directed* if there is a vertex v, which we call the *root* of T, such that all the edges in T are directed away from v. Similarly we define an *in-directed* tree to have all edges directed towards v. A *directed* tree is an out-directed tree or an in-directed tree.

Observation 3.7. Let T be a directed tree on n vertices. Then it is a subgraph of any transitive tournament G on at least n vertices.

Proof. We assume, without loss of generality, that T is out-directed. Let $N = |G|$. Since G is transitive, there exists an ordering of the vertices $u_1, u_2, \ldots, u_N$, such that all edges are directed towards the higher index. Let $v_1, v_2, \ldots, v_n$ be an ordering of the vertices of T obtained by a depth-first search algorithm starting at the root v of T. Since T is out-directed, the ordering has the property that all edges of T are directed towards a higher index and we can embed v_i in u_i for every $i \in [n]$. □

A leaf of an oriented tree is an *out-leaf* if its out-degree is 0 and an *in-leaf* if its in-degree is 0. Note that for an out-directed tree the only in-leaf is the root and all other leaves are out-leaves. In fact every out-leaf is a leaf itself.

4. Tree versus path

In this section we prove a special case of Theorem 1.1 for a directed tree versus a directed path; here the random tournament is replaced by a pseudorandom tournament.

Theorem 4.1. *Given $0 < \varepsilon < 1/2$ and $\sigma > 0$, there exists a constant $c > 0$ such that the following holds. Let G be a tournament on N vertices which is $(\varepsilon, \sigma \log N)$-pseudorandom. Then $G \to (\overrightarrow{P_n}, T)$, where T is any directed tree on m vertices, as long as $n, m \leqslant N/c$ and $nm \leqslant N^2/(c^2 \log N)$.*

As an intermediate result we prove it first for trees with relatively few leaves (see Section 4.2); we then prove Theorem 4.1 in Section 4.3. Before turning to the proofs, we discuss three types of tree-splits, which we shall use in the proofs.

4.1 Tree-splits

Our proofs in this section make use of several tree-splits; we present them here.

The $(\mathbf{c}, \boldsymbol{\alpha})$-tree-split. Let T be an out-directed tree and let T' be a subtree of T. An *extending-leaf* of T' with respect to T is an out-leaf (*i.e.* a non-root leaf) of T', which is not a leaf of T. Whenever T is clear from the context we do not mention it.

Lemma 4.2. *Let $c \geqslant 2$ and $0 < \alpha \leqslant (2c)^{-1}$. Let T be an out-directed tree on m vertices with at most m^α leaves. Then there is a partition of the vertices into subtrees $T_1, \ldots, T_\ell$ such that the following properties hold.*

- (i) *For every $i \in [\ell]$ there is at most one in-edge towards a vertex of T_i in T, and if present it is towards the root of T_i.*
- (ii) *The only vertices with out-edges leaving T_i are extending-leaves of T_i.*
- (iii) *Each extending-leaf of T_i lies in an even level (i.e. its distance to the root of T_i is even) and it has out-degree exactly one in T.*
- (iv) *$|T_i| \leqslant 6m^{c\alpha}$, for all $i \in [\ell]$.*
- (v) *$\ell \leqslant 2m^{1-c\alpha}$.*

Given a partition of T into subtrees $T_1, \ldots, T_\ell$ as in Lemma 4.2, if we contract each subtree T_i to a single vertex, the resulting graph T' is again an out-directed tree with no multiple edges. We call this graph a *(c,α)-tree-split* of T (see Figure 1); note that this split need not be unique. A subtree T_i in such a split that does not have extending-leaves, that is, the vertex corresponding to T_i in T' is a leaf, is called a *leaf-tree*.

Proof of Lemma 4.2. For each $i \leqslant 2m^{1-c\alpha}$ we construct a subtree T_i in two stages. At step i we assume we have already found $T_1, \ldots, T_{i-1}$ for which the conditions (i)–(iv) hold, and $V(T_1) \cup \cdots \cup V(T_{i-1})$ induces a subtree T' of T with the same root. In the first stage we choose the root of T_i and ensure that T_i is big enough (or a leaf-tree), so that we are later able to deduce (v), and in

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

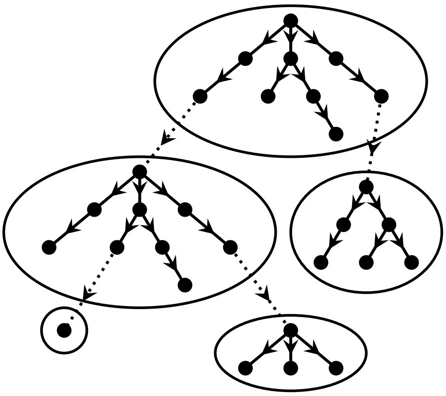

Figure 1. A tree-split of a tree.

the second stage we extend T_i further to ensure that it satisfies (i)–(iv). We stop the process when all the vertices of T are covered by the subtrees $T_1, \ldots, T_i$, and for such i we denote $\ell := i$.

Stage 1. First we choose the root v of T_i. For $i = 1$ we take the root of T and for $i > 1$ we pick the only out-neighbour of an extending-leaf (there is only one out-neighbour by (iii)) of T' (this is the subtree of T induced by $V(T_1) \cup \cdots \cup V(T_{i-1})$).

Assume first that v has at most $m^{c\alpha}$ descendants in T. Then we let T_i be the subtree consisting of all descendants of v in T. In this case T_i is a leaf-tree of order at most $m^{c\alpha} + 1$ and there is no second stage.

Otherwise, we start with a subtree T_i' consisting only of the vertex v. As long as $|T_i'| < m^{c\alpha}$, we pick an extending-leaf of T_i' and add all its children to T_i'. Note that such an extending-leaf always exists because there are more than $m^{c\alpha}$ descendants of v in T and each non-leaf vertex of T_i' has all its children from T in T_i'. Since the maximum out-degree of T is bounded from above by the number of out-leaves of T (every out-neighbour eventually leads to a different out-leaf by following out-edges), in each step of the construction of T_i' we add at most m^{α} vertices. This implies that when we stop (*i.e.* right after $|T_i'| \geqslant m^{c\alpha}$ holds), we have the following:

$$m^{c\alpha} \leqslant |T_i'| \leqslant m^{c\alpha} + m^{\alpha} \leqslant 2m^{c\alpha}.$$

Stage 2. We start with $T_i = T_i'$, produced by the first stage. Call an extending-leaf contained in T_i *bad* if it lies in an odd level or has out-degree not equal to one in T. As long as there is a bad extending-leaf in T_i, we add all its children to T_i. Eventually there are no bad extending-leaves left, since by going deep enough we reach a leaf of T, which by definition is not an extending-leaf.

Note that during the procedure of both stages (i) is always satisfied by the choice of the root. Moreover, at the end of the procedure, (ii) and (iii) hold as well, since there are no bad extending-leaves in T_i.

Now let us prove that condition (iv) holds, *i.e.* that $|T_i| \leqslant 6m^{c\alpha}$. From the first stage we know that $|T_i'| \leqslant 2m^{c\alpha}$. Furthermore, every vertex in $T_i \setminus T_i'$ is either a leaf in T or it was a bad vertex for some T_i'. In the latter case such a vertex is either branching or has out-degree 1 in T and is on an odd level; in the second case its child is either a branching vertex or a leaf of T itself. Recall that there are at most m^{α} leaves in T, so by Lemma 3.6 there are at most m^{α} branching vertices. Finally, as each vertex has a unique parent, the number of vertices of $T_i \setminus T_i'$ of the last type (*i.e.* vertices on odd levels of T whose out-degree is 1) is bounded by the number of leaves of T plus the number of branching vertices of T. This implies that

$$|T_i| = |T_i'| + |T_i \setminus T_i'| \leqslant 2m^{c\alpha} + m^{\alpha} + m^{\alpha} + 2m^{\alpha} \leqslant 6m^{c\alpha}.$$

To see that the last condition (v) holds, we note that each leaf-tree contains at least one out-leaf of T. Thus, the number of leaf-trees is bounded by m^{α}. In addition to that we can bound the number of non-leaf-trees by $m/m^{c\alpha}$, since each one has order at least $m^{c\alpha}$. This implies that

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

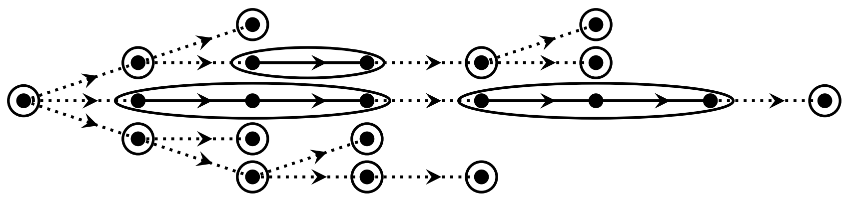

Figure 2. The path-split of a tree.

$$\ell \leqslant m^{\alpha} + m^{1-c\alpha} \leqslant 2m^{1-c\alpha},$$

where the last inequality follows from $c \geqslant 2$ and $0 < \alpha \leqslant (2c)^{-1}$. □

The α-path-split. In the following lemma we are interested in a similar split, but this time we want the subtrees in the split to be paths. The graph obtained by contracting the paths in the following lemma will be called an *α-path-split* (see Figure 2). We call a vertex of a tree a *junction* if it is a leaf or a branching vertex.

Lemma 4.3. *Let $0 < \alpha \leqslant 1/4$. Let T be an out-directed tree on m vertices with at most m^{α} leaves. Then there is a partition of the vertices into subpaths $P_1, \ldots, P_{\ell}$ such that the following properties hold.*

- (i) *If P_i contains a junction then $|P_i| = 1$.*
- (ii) *For every $i \in [\ell]$ there is at most one in-edge towards P_i, which is directed towards the start-vertex of P_i. Furthermore, unless P_i is a junction, there is at most one out-edge away from P_i which is directed from the end-vertex of P_i.*
- (iii) $|P_i| \leqslant m^{3\alpha}$.
- (iv) $\ell \leqslant 5m^{1-3\alpha}$.

Proof. We define the paths in the split as follows. First we let each junction be a separate trivial path of size 1. We then remove all junctions from the tree; thus we are left with a collection of disjoint subpaths, which we call *long subpaths*. Finally, we split each such path into smaller subpaths, called *short subpaths*, such that each has order at most $m^{3\alpha}$. We let these shorter paths be the remaining subpaths of our split. We now show that this split satisfies the desired conditions.

First note that the number of junctions is at most $2m^{\alpha}$, since by assumption there are at most m^{α} leaves and therefore at most m^{α} branching vertices, by Lemma 3.6. If we consider a graph whose vertex set is the set of junctions and put an edge between a pair of junctions whenever they are joined by a long subpath, we obtain a forest. Therefore, the number of long subpaths, denoted by d, satisfies the following:

$$d \leqslant \text{ \# of junctions } - 1 \leqslant 2m^{\alpha}.$$

For $i \in [d]$, let m_i denote the order of the ith long subpath. Then we split the ith long subpath into

$$r_i := \left\lceil \frac{m_i}{m^{3\alpha}} \right\rceil \leqslant \frac{m_i}{m^{3\alpha}} + 1$$

shorter subpaths of order at most $m^{3\alpha}$ each. Hence, the total number of paths used is bounded from above by

$$\sum_{i \in [d]} r_i + 2m^{\alpha} \leqslant \sum_{i \in d} \frac{m_i}{m^{3\alpha}} + d + 2m^{\alpha} \leqslant m^{1-3\alpha} + 4m^{\alpha} \leqslant 5m^{1-3\alpha},$$

as required. □

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

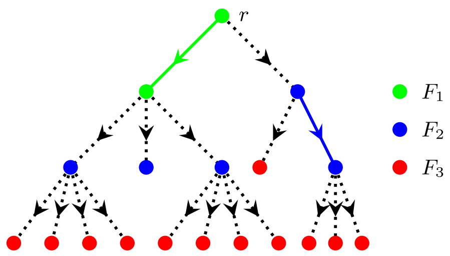

Figure 3. The 3-core-split of a tree on 19 vertices.

The *k*-core-split. We now define the *k-core-split* $F_1, \ldots, F_\ell$ of a tree T (see Figure 3); this will be used in Section 4.3 to remove the requirement on the number of leaves of T in order to prove Theorem 4.1. Let T' be the k-core of T. We set $F_1 = T'$. For $i > 1$, let $S_1, S_2, \ldots, S_{\ell_i}$ be the trees of the forest $T \setminus \bigcup_{j\in[i-1]} F_j$. For every $j \in [\ell_i]$ we define T_j to be the k-core of S_j and set the forest $F_i := \bigcup_{j\in[\ell_i]} T_j$.

Proposition 4.4. *Let $F_1, \ldots, F_\ell$ be the k-core-split of a tree T. If T' is a tree in the forest F_i then*

(i) *T' has order less than $|T|/k^{i-1}$ and*
(ii) $\mathrm{lf}(T') \leqslant k$.

Proof. Let us prove (i) by induction over i. The statement is clearly true for $i = 1$. Let us assume that it holds for some $i < l$; then every tree S from F_i has size less than $|T|/k^{i-1}$ and is a k-core of some subtree of T. Therefore, every tree in F_{i+1} has order at most $|T|/k^{i-1} \cdot k^{-1} = |T|/k^i$, using Observation 3.5, as desired. Property (ii) follows, because every tree in F_i is a k-core of some subtree of T. □

4.2 Tree with few leaves

We will make use of the following theorem (Theorem 3.17 in [4]). In fact, we will only use a special case of this theorem, when one of the trees is a path.

Theorem 4.5. *There exists a constant c such that, for any oriented trees T_1 and T_2, in any 2-colouring of the complete directed graph on $c(|T_1| + |T_2|)$ vertices there exists a red T_1 or a blue T_2.*

Furthermore, our proof makes use of the special case of Theorem 4.1 for path versus path, proved in [5] as Theorem 12.

Theorem 4.6. *Given $0 < \varepsilon < 1/2$ and $\sigma > 0$, there is a constant $c > 0$ such that the following holds. Let G be an $(\varepsilon, \sigma \log N)$-pseudorandom tournament on N vertices. Then $G \to (\overrightarrow{P_n}, \overrightarrow{P_m})$ provided $n, m \leqslant N/c$ and $nm \leqslant N^2/(c^2 \log N)$.*

This theorem is tight up to the constant factor. This result is the main bottleneck for our proof, as it will rely on this result as a black box.

The following theorem is an intermediate result, leading to the proof of Theorem 4.1. Its proof introduces some interesting new ideas about how to deal with trees in place of paths.

Theorem 4.7. *Given $0 < \varepsilon < 1/2$ and $\sigma > 0$, there exists a constant $c > 0$ such that the following holds. Let G be a tournament on N vertices, which is $(\varepsilon, \sigma \log N)$-pseudorandom. Then $G \to (\overrightarrow{P_n}, T)$,*

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

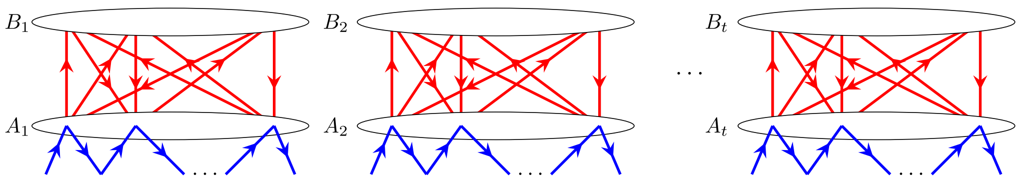

Figure 4. Red–blue pairs.

where T is any directed tree on m vertices, with at most $m^{1/6}$ leaves, as long as $n, m \leqslant N/c$ and $nm \leqslant N^2/(c^2 \log N)$.

Proof. Let $\alpha = 1/6$. Without loss of generality we may assume that T is out-directed, because once we prove this, since directed paths are both in-directed and out-directed, we can apply it to the tournament with opposite orientation of every edge to conclude the other case.

Consider a fixed 2-colouring of G. Write $a = 128\varepsilon^{-2}$, $b = 8a$ and make the following definition:

$$\text{a cycle } C \text{ is } \begin{cases} \text{short} & \text{if } |C| < am^{2\alpha}, \\ \text{medium} & \text{if } am^{2\alpha} \leqslant |C| \leqslant bm^{2\alpha}, \\ \text{long} & \text{if } bm^{2\alpha} < |C|. \end{cases}$$

We will prove the theorem under assumptions that $N \geqslant c^2$ and $n, m \geqslant N/(c \log N)$, for some value of c. We start by arguing how to conclude the theorem in general, assuming we know it under these conditions. To see this with only the first assumption, let $n' = \max(n, N/(c \log N))$ and $m' = \max(m, N/(c \log N))$. It is easy to see that $n', m' \leqslant N/c$ and $n'm' \leqslant N^2/(c^2 \log N)$ still hold. The required result for m and n now follows from the result for n' and m', which satisfy both assumptions. So there is a c such that the general result holds, given $N \geqslant c^2$. Now if $N < c^2$, then $c^2 > N \geqslant cn, cm$ implying $n, m < c$, so the result follows, with a larger constant, by appealing to the fact that oriented Ramsey numbers of acyclic graphs are finite.

So, from now on we assume $N \geqslant c^2$ and $n, m \geqslant N/(c \log N)$. In particular, this implies $m^\alpha \geqslant \sigma \log N$.

Case 1: many disjoint medium or long blue cycles. We shall consider two subcases.

1(a) There is a collection of vertex-disjoint medium blue cycles covering at least $N/4$ vertices.

1(b) There is a collection of vertex-disjoint long blue cycles covering at least $N/4$ vertices, such that they span no medium blue cycle and all the large blue cycles are as short as possible.

We resolve both subcases by finding the following structure. Given k and t, a (k, t)-*red–blue pair* is a collection of pairwise disjoint subsets of vertices $A_1, B_1, \ldots, A_t, B_t$, each of size k, such that for every $i \in [t]$ the following holds (see Figure 4):

(i) the bipartite graph $G[A_i, B_i]$ contains only red edges,

(ii) for every $i \in [t]$ there exists a blue path P_i that contains all vertices of A_i, such that P_i and P_j are vertex-disjoint for all $j \neq i$.

We start by showing how to conclude the argument once we find this structure, and then we show how to find it in each of the two cases.

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

Proposition 4.8. *Let $\frac{3}{4}am^{2\alpha} \leqslant k \leqslant am^{2\alpha}$ and $t \geqslant N/(256k)$. If G contains (k,t)-red–blue pairs then $G \to (\overrightarrow{P_n}, T)$.*

Proof. Let us define an auxiliary complete directed graph K on vertex set $[t]$. We colour the edge ij blue if at least $(1-\varepsilon/4)k$ vertices in A_i have at least $(\varepsilon/2)k$ blue out-neighbours in A_j and red otherwise.

Theorem 4.5 implies that in the auxiliary graph K there is a directed blue path of order $t/(2c_1)$ or any oriented red tree of order $t/(2c_1)$, for some positive constant c_1. We claim that the constant c can be chosen such that the following two inequalities hold:

(1) $t \geqslant 8c_1/(\varepsilon k)n$,
(2) $t \geqslant 4c_1 m^{1-2\alpha}$.

Inequality (1) follows, since, by the assumptions of Proposition 4.8 and Theorem 4.7, $t \geqslant N/(256k)$ and $N \geqslant cn \geqslant 256 \cdot 8\varepsilon^{-1}c_1 n$, where the last inequality holds under the assumption that $c \geqslant 2^{11}c_1\varepsilon^{-1}$. We obtain inequality (2) by using $N \geqslant cm$, under the assumption that $c \geqslant 2^{10}ac_1 = 2^{17}\varepsilon^{-2}c_1$ as follows:

$$t \geqslant \frac{N}{256k} \geqslant \frac{N}{256am^{2\alpha}} \geqslant \frac{cm^{1-2\alpha}}{256a} \geqslant 4c_1 m^{1-2\alpha}.$$

Suppose that there exists a blue path $P = i_1 i_2 \cdots i_\ell$ of order $\ell := t/(2c_1)$ in K. We will explain how to lift this path to a blue path in G by using subpaths of order at least $s := (\varepsilon/4)k$ from each of the blue paths P_{i_j} associated to A_{i_j} for $j \in [\ell]$.

For every $j \in [\ell]$, let $E_j \subseteq A_{i_j}$ denote the set of vertices with at least $2s = (\varepsilon/2)k$ blue out-neighbours in $A_{i_{j+1}}$. Note that $|E_j| \geqslant (1-\varepsilon/4)k$, since the edge $i_j i_{j+1}$ is blue in K.

We start with an initial subpath P' of the path P_{i_1} associated to A_{i_1} which ends in the last vertex from P_{i_1} contained in E_1. Note that $|P'| \geqslant |E_1| \geqslant s$, since P_{i_1} covers A_{i_1}, so also E_1.

Suppose that we already have a path P' of order at least $s(j-1)$ in $\bigcup_{r\in[j-1]} A_{i_r}$ whose last vertex v is contained in E_{j-1}.

Let S_j be the blue out-neighbourhood of v in A_{i_j} and let $u_1, u_2, \ldots, u_k$ denote the vertices of the blue path P_{i_j} in A_{i_j} ordered according to their order in P_{i_j}. We extend P' by the path Q consisting of $u_p, u_{p+1}, \ldots, u_q$, where p is the smallest index among the vertices in S_j and q the largest index among the vertices in E_j. Note that Q has order at least s, since

$$|E_j \cap S_j| = |E_j| + |S_j| - |E_j \cup S_j| \geqslant |E_j| + |S_j| - |A_{i_j}| \geqslant \left(1 - \frac{\varepsilon}{4}\right)k + \frac{\varepsilon}{2}k - k = \frac{\varepsilon}{4}k = s.$$

The blue path in K has order $t/(2c_1)$, so this process produces a blue path in G of order at least

$$\frac{ts}{2c_1} \geqslant \frac{8c_1 n/(\varepsilon k) \cdot \varepsilon k/4}{2c_1} = n,$$

where we used inequality (1); this completes the proof of Proposition 4.8 in the case where K contains a long blue path.

Otherwise, *i.e.* if K does not have a blue path of order $t/(2c_1)$, then K contains a red copy of a $(2,\alpha)$-tree-split of T (obtained from Lemma 4.2), since $t/(2c_1) \geqslant 2m^{1-2\alpha}$, by inequality (2). We now explain how to lift the tree-split from K to a red copy of T in G. We relabel the vertices of K in such a way that the vertices of the tree-split we found in K are $[\ell]$, for ℓ being the order of the tree-split, and vertex i representing a subtree T_i of T. Our aim is to embed the subtrees T_i in $A_i \cup B_i$. To that end we pick 'candidate sets' $D_i \subseteq A_i$ which satisfy some useful properties.

Claim 4.9. *Suppose that for each vertex i of the tree-split we have a non-empty set of candidates $D_i \subseteq A_i$, such that for any $v \in D_i$ there is a tree $T(v)$ with the following properties:*

(i) *$T(v)$ is a red copy of T_i embedded in $A_i \cup B_i$ and rooted at v,*
(ii) *each vertex u in $T(v)$ that corresponds to an extending-leaf w in T_i (as a subtree of T) is in A_i and has a red out-edge towards D_j if j is such that there is an edge from w to T_j in T.*

Then we can find a red copy of tree T inside the tournament G.

Proof. Let T' denote the subtree of the tree-split containing the root v of T. From the set of candidates for the root of T' we can pick any vertex we want and set $T' := T(v)$. By property (ii) we can choose the roots of the adjacent subtrees in the corresponding candidate sets and by (i) we can embed the subtrees themselves as well. Note that as all A_i and B_i are disjoint, we do not use any of the vertices twice. Repeating this argument eventually produces a red copy of T in the tournament G. □

We now show how to construct appropriate candidate sets consisting of at least $k/2$ vertices. For this we begin with the leaves of the tree-split and then make our way up, in the sense that we deal with the candidate set of a particular vertex from the tree-split only if we have already defined the candidate sets for all its out-neighbours.

Let us define the candidate set D_i, where the out-neighbours of vertex i in the tree-split are j_s for $s \in [h]$ (if T_i is a leaf-tree then $h = 0$). Note that by the assumption on the number of leaves of the tree T and the definition of a tree-split, $h \leqslant m^{\alpha}$. Furthermore, by construction, the candidate sets D_{j_s} have already been defined (this condition also holds, in particular, for leaf-trees).

For each $s \in [h]$, let $X_s \subseteq A_i$ be the set of vertices with at least one red out-neighbour in D_{j_s}. These sets will host the extending-leaves and guarantee property (ii) of Claim 4.9. Let Y be the set of vertices in B_i that send at least $|T_i| + \sigma \log N$ red edges into each set X_s; if there are no such sets X_s (*i.e.* T_i is a leaf-tree), we let Y be the set of vertices in B_i with at least $|T_i| + \sigma \log N$ red out-neighbours in A_i. Finally, let D_i be the set of vertices in A_i that send at least $|T_i|$ red edges into Y.

Claim 4.10. $|D_i| \geqslant |A_i| - \sigma \log N \geqslant k/2$.

Proof. Firstly, we show that $|X_s| \geqslant (\varepsilon/8)k$ for every $s \in [h]$. Indeed, as $|D_{j_s}| \geqslant k/2 \geqslant \sigma \log N$, by pseudorandomness, all but at most $\sigma \log N$ vertices of A_i send at least $\varepsilon|D_{j_s}| \geqslant (\varepsilon/2)k$ edges into D_{j_s}. Since ij_s is a red edge in the auxiliary graph K, at most $(1 - \varepsilon/4)k$ of these vertices send at least $(\varepsilon/2)k$ blue edges into D_{j_s}. It follows that there are at least $|A_i| - (1 - \varepsilon/4)k - \sigma \log N \geqslant (\varepsilon/8)k$ vertices in A_i with at least one red out-neighbour in D_{j_s}, *i.e.* $|X_s| \geqslant (\varepsilon/8)k$, as claimed.

We now claim that $|Y| \geqslant k/2$. Indeed, since all edges between A_i and B_i are red, by pseudorandomness, all but at most $\sigma \log N$ vertices of B_i send at least $\varepsilon|X_s| \geqslant (\varepsilon^2/8)k \geqslant |T_i| + \sigma \log N$ red edges into X_s, for every $s \in [h]$. Hence,

$$|Y| \geqslant |B_i| - h \cdot \sigma \log N \geqslant k - m^{\alpha} \cdot \sigma \log N \geqslant k/2.$$

A similar argument shows that $|Y| \geqslant k/2$ if T_i is a leaf-tree.

Finally, by pseudorandomness and since all edges from A_i to B_i are red, we find that $|D_i| \geqslant |A_i| - \sigma \log N \geqslant k/2$, as required. □

Now let us explain why each vertex in D_i is a candidate for the root of T_i. Note that the bipartite graph $G[D_i, Y]$ has minimum red out-degree at least $|T_i|$, allowing us to greedily embed a copy T' of the subgraph of T_i obtained by removing its extending-leaves. By the property that all

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

extending-leaves lie at even distance from the root of their corresponding subtree, the parent of such a leaf u is embedded in Y and has at least $|T_i|$ red out-neighbours in the corresponding set X_s. Since we have not embedded all the vertices of T_i yet, we can embed u in X_s.

These candidate sets D_i satisfy conditions of Claim 4.9, so we can find a red copy of the desired tree in G. This completes the proof of Proposition 4.8. □

In order to complete the proof of Theorem 4.7 in Case 1, it now remains to show how to find (k, t)-red–blue pairs in Cases 1(a) and 1(b).

Case 1(a): many disjoint medium blue cycles. We assume that there is a collection of vertex-disjoint medium blue cycles $C_1, C_2, \ldots, C_{t'}$ which cover at least $N/4$ vertices. Since the medium cycles have length at most $bm^{2\alpha}$, we get that $t' \geqslant N/(4bm^{2\alpha})$.

Let H be an auxiliary 2-coloured complete directed graph on vertex set $[t']$. We colour the edge ij blue if at least $(a/4)m^{2\alpha}$ vertices in C_i have a blue out-neighbour in C_j and red otherwise. Now we consider a maximal red–red matching M, namely a matching that consists of edges that are red in both directions.

First we suppose that the matching M covers at most $t'/2$ vertices. Since this matching is maximal, for every two vertices i and j not covered by M, at least one of the directed edges ij and ji is blue. In particular, there is a blue subtournament on at least $t'/2$ vertices. Since every tournament contains a directed Hamiltonian path, we thus find a blue directed path of order $t'/2$ in the auxiliary graph H. The following claim explains how to lift this path to the tournament G.

Claim 4.11. *Let G be an oriented graph with pairwise vertex-disjoint cycles $C_1, C_2, \ldots, C_k$ such that, for each $i < k$, there are at least r vertices in C_i that have an out-neighbour in C_{i+1} and $|C_k| \geqslant r$. Then G contains a directed path of order $k \cdot r$.*

Proof. If $k = 1$ the claim is immediate since we assume $|C_1| \geqslant r$. We show that for every $1 \leqslant i \leqslant k$ there is a path P_i of order at least $r(i-1)$ whose vertices are in $\bigcup_{j\in[i-1]} V(C_j)$ and whose last vertex has an out-neighbour w_i in C_i. Indeed, for $i = 1$ we can take P_1 to be the empty path. For $1 \leqslant i < k-1$, suppose that P_i and w_i satisfy the required properties, let Q_{i+1} be the path that starts at w_i and follows C_i until the last vertex of C_i that has an out-neighbour in C_{i+1}, let w_{i+1} be one such out-neighbour of the last vertex in Q_{i+1}, and form P_{i+1} by concatenating P_i and Q_{i+1}. It is easy to check that P_{i+1} and w_{i+1} satisfy the required properties.

Now, given a path P_k and a vertex w_k as above, extend P_k by a path that starts at w_k and follows C_k until the last vertex before w_k. This extends P_k by at least r vertices, as $|C_k| \geqslant r$, to give a path in G of order at least $k \cdot r$, as required for the claim. □

By applying Claim 4.11 to the induced blue subgraph of G we find a blue path of order at least

$$\frac{a}{4}m^{2\alpha} \cdot \frac{t'}{2} \geqslant \frac{a}{32b}N \geqslant \frac{N}{256} \geqslant n$$

(where the last inequality holds by assuming that $c \geqslant 256$), as desired.

Therefore, we can assume that the matching M covers at least $t'/2$ vertices of H. This corresponds to $t := t'/4$ disjoint pairs of cycles (C_{i_s}, C_{j_s}), for $s \in [t]$, where at least $k := (3/4)am^{2\alpha}$ vertices in C_{i_s} do not have a blue out-neighbour in C_{j_s} and *vice versa*. Hence we can find subsets A_s and B_s of C_{i_s} and C_{j_s}, respectively, of size k each, with only red edges between them. Each A_s lies in a different cycle C_{i_s} from the original collection of disjoint medium blue cycles. Thus, we can define P_s to be the cycle C_{i_s} minus one edge. This way P_s contains all vertices of A_s and the paths P_s are pairwise disjoint.

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

Note that

$$t \geqslant \frac{3N}{512k},$$

since

$$t' \geqslant \frac{N}{4bm^{2\alpha}} \geqslant \frac{N}{32am^{2\alpha}} = \frac{3N}{128k}.$$

So Proposition 4.8 applies and concludes the proof of Theorem 4.7 in this case.

Case 1(b): a large set with many disjoint long blue cycles but no blue medium cycle. Suppose Case 1(a) does not hold; thus there exists a set U of at least $3N/4$ vertices which does not contain any medium blue cycle.

Let us consider the following process, which starts with $U' := U$. As long as there exists a long blue cycle in U' we pick a shortest one, say C, and define $U' = U' \setminus C$. This process eventually terminates and produces a sequence of disjoint long blue cycles $C_1, C_2, \ldots, C_{t'}$. In this case we are going to assume that these cycles cover at least $N/4$ vertices. Note that $t' \leqslant N/(4bm^{2\alpha})$, since each long cycle contains at least $bm^{2\alpha}$ vertices.

Note that, for every $i \in [t']$, all chords in C_i of length at least $am^{2\alpha}$ are red, since otherwise we would find a blue cycle inside C_i which is either a medium cycle, or a shorter long cycle, contradicting our choice of C_i as the shortest remaining long cycle.

Write $C_i = (v_1 v_2 \cdots v_r)$ and $k := am^{2\alpha}$. Define sets

$$A_i = \{v_1, v_2, \ldots, v_{r/2-2k}\} \quad \text{and} \quad B_i = \{v_{r/2-k}, v_{r/2-k+1}, \ldots, v_{r-k}\}.$$

By the argument above we have that $G[A_i, B_i]$ spans red edges only. Let $A_{i,1}, \ldots, A_{i,r(i)}$, where $r(i) = \lfloor |A_i|/k \rfloor$, be pairwise disjoint sets of k consecutive vertices (with respect to C_i) in A_i, and let $B_{i,1}, \ldots, B_{i,r(i)} \subseteq B$ be defined similarly. The sets $A_{i,j}, B_{i,j}$ cover all but at most $4k$ vertices of each cycle C_i, hence they cover at least $N/4 - t' \cdot 4k$ vertices in total. Since the number of sets $A_{i,j}$ and sets $B_{i,j}$ is the same, and each set has size k, it follows that the number t of sets $A_{i,j}$ satisfies

$$t \geqslant \frac{N}{8k} - 2t' \geqslant \frac{N}{8k} - \frac{N}{2bm^{2\alpha}} = \frac{N}{8k} - \frac{N}{16k} = \frac{N}{16k},$$

where the first equality follows from the choice $b = 8a$ which implies that $bm^{2\alpha} = 8k$. Note that the collection of pairs $(A_{i,j}, B_{i,j})$ forms a (k, t)-red–blue pairs structure, as each set $A_{i,j}$ contains a spanning blue path (which is a part of the cycle C_i), which are mutually disjoint. Proposition 4.8 can now be used to complete the proof of Theorem 4.7 in this case as well.

Case 2: a large set of vertices spanning no blue medium or long cycle. In the remaining case, the process of picking the sequence of disjoint long blue cycles in Case 1(b) terminated before it covered at least $N/4$ vertices. Hence, we are left with a set U that covers at least $N/2$ vertices and spans neither medium nor long blue cycles.

Let us start this case with an elementary observation.

Observation 4.12. Every directed graph G with minimum out-degree d contains a cycle of length at least $d + 1$.

Proof. Let $v_1 \cdots v_\ell$ be a longest directed path in G. By the maximality of this path we get that v_ℓ has no out-neighbour outside this path. Since the out-degree of v_ℓ is at least d it has at least d out-neighbours among $v_1, \ldots, v_{\ell-1}$. Let s be the smallest index among these out-neighbours of v_ℓ. Then $(v_s v_{s+1} \cdots v_\ell)$ is a cycle of length at least $d + 1$. □

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

This observation allows us to obtain an ordering of the vertices in U with 'few' blue edges going forward.

Claim 4.13. *There exists an ordering $u_1, u_2, \ldots, u_{|U|}$ of the vertices in U such that, for every i, there are at most $am^{2\alpha}$ indices $j > i$ such that there is a blue edge from u_i to u_j.*

Proof. Suppose that there exists a subgraph of $G[U]$ which has minimum blue out-degree at least $am^{2\alpha}$. Then by Observation 4.12 we find a blue cycle of order at least $am^{2\alpha}$, a contradiction. Therefore, any subgraph of $G[U]$ has a vertex of blue out-degree at most $am^{2\alpha}$.

In particular, there exists a vertex $u_1 \in U$ with blue out-degree at most $am^{2\alpha}$. Now suppose that $u_1, u_2, \ldots, u_{i-1}$ are defined. We define u_i to be a vertex with blue out-degree at most $am^{2\alpha}$ in $G[U']$, where $U' = U \setminus \{u_1, u_2, \ldots, u_{i-1}\}$. We repeat this as long as $i \leqslant |U|$. The resulting ordering $u_1, u_2, \ldots, u_{|U|}$ satisfies the requirement of the claim. □

Let $k = N/(32m^{1-3\alpha})$. We set $t := |U|/k$ and denote $U_i = \{u_{(i-1)k+1}, \ldots, u_{ik}\}$ for $i \in [t]$. We claim that we can choose the constant c such that the following two inequalities hold:

(1) $t \geqslant 16m^{1-3\alpha}$,
(2) $k \geqslant 128\varepsilon^{-2}am^{3\alpha}$.

Indeed, inequality (1) follows independently from c, as

$$t = \frac{|U|}{k} \geqslant \frac{N}{2k} = 16m^{1-3\alpha}.$$

We obtain inequality (2) from $N \geqslant cm$, given $c \geqslant 2^{12}\varepsilon^{-2}a$, as

$$k = \frac{N}{32m^{1-3\alpha}} \geqslant \frac{cm}{32m^{1-3\alpha}} = \frac{c}{2^5}m^{3\alpha} \geqslant 2^7\varepsilon^{-2}am^{3\alpha}.$$

Let c_2 be the constant from Theorem 4.6 with parameters ε and $\sigma/(3\alpha)$. By choosing $c \geqslant 128c_2$, we obtain from $N \geqslant cn, cm$ that

$$k/4 = \frac{N}{128m^{1-3\alpha}} \geqslant c_2\frac{n}{m^{1-3\alpha}}, \quad c_2m^{3\alpha}.$$

Similarly, we get

$$k/4 \geqslant c_2\sqrt{\frac{n}{m^{1-3\alpha}}m^{3\alpha}\log(k/4)},$$

from $N \geqslant c\sqrt{nm\log N}$.

Also note that

$$k/4 = \frac{N}{128m^{1-3\alpha}} \geqslant N^{3\alpha}$$

so $\sigma\log(k/4) \geqslant 3\alpha\sigma\log N$. Thus any subtournament of G of order $k/4$ is $(\varepsilon, (\sigma/(3\alpha))\log(k/4))$-pseudorandom. Therefore, Theorem 4.6 applies for paths of order $n/(m^{1-3\alpha})$ and $m^{3\alpha}$, within any subset of vertices of size at least $k/4$.

Claim 4.14. *One of the following holds, for each set U_i.*

(i) *There are at least $k/8$ pairwise disjoint pairs of vertices in U_i that are joined by a blue path, contained in U_i, of order $n/(m^{1-3\alpha})$.*
(ii) *For each $2 \leqslant \ell \leqslant m^{3\alpha}$, there are at least $k/4$ pairwise disjoint pairs of vertices in U_i that are joined by a red path, contained in U_i, of order ℓ.*

Proof. Consider the following process. As long as there is a blue path of order $n/(m^{1-3\alpha}) \geqslant 2$ in U_i (where the inequality follows since $n \geqslant N/(c \log N)$ and $m \leqslant N/c$), we remove its first and last vertex. If this process runs for at least $k/8$ rounds then (i) holds.

Otherwise, there is a subset $W \subseteq U_i$ of size at least $\frac{3}{4}k$ with no blue path of order $n/(m^{1-3\alpha})$. Consider the following process. As long as there are $k/4$ vertices left in W we can apply Theorem 4.6 to find a red path of order ℓ (since $\ell \leqslant m^{3\alpha}$) and remove its first and last vertex. Since we remove only two vertices in each round, this process runs for at least $k/4$ rounds. Thus (ii) holds. □

If (i) holds, we say that U_i is blue; otherwise, we say that U_i is red. We now distinguish two cases depending on the majority colour of the sets U_i.

Case 2(a): most of the sets U_i are red. In this case there are at least $t/2$ red sets U_i so, while preserving the ordering, we rename $t/2$ red sets U_i as $V_1, V_2, \ldots, V_{t/2}$. Note that when $i < j$ we have by Claim 4.13 that every vertex in V_i has at most $am^{2\alpha}$ blue out-neighbours in V_j. Let us view $V_1, V_2, \ldots, V_{t/2}$ as vertices of a transitive tournament with edges pointing always towards the bigger index. Let T' be an α-path-split of T (see Lemma 4.3). By Observation 3.7, we can find a copy of T' inside this transitive tournament, since inequality (1) implies that $t/2 \geqslant 5m^{1-3\alpha}$.

We now show that if we define appropriate candidate sets for each start-vertex of a path in the path-split, then we can greedily find a red copy of T in G, in a similar manner to Case 1. Let P_i denote the path corresponding to the vertex i of the embedded path-split.

Claim 4.15. *Suppose that for each vertex i of the path-split we have a non-empty set of candidates $D_i \subseteq V_i$, such that, for any $v \in D_i$, there is a subpath $P(v)$ of T which satisfies the following:*

(i) *$P(v)$ is a red copy of P_i embedded within V_i with start-vertex v,*
(ii) *the end-vertex u of $P(v)$ has a red out-edge towards D_j for each j which is a child of i in the path-split.*

Then we can find a red copy of tree T inside the tournament G.

Proof. Use a greedy embedding, analogous to that used in the proof of Claim 4.9. □

We now define such candidate sets, each of size at least $k/8$. We start with the leaves of the path-split and then move upwards, in such a way that we are always defining the candidate set for the vertex all of whose out-neighbours have already had their candidate sets defined.

If i is a leaf of the path-split, then P_i is a leaf of T, and we can set $D_i := V_i$.

In the case of i being a non-leaf we apply Claim 4.14 with $\ell = |P_i|$, and define $E_i, S_i \subseteq V_i$ to be the sets of end- and start-vertices of a red path of length $|P_i|$, such that $|E_i| = |S_i| \geqslant k/4$ (note that if P_i is a singleton, then we can take $E_i = S_i = V_i$). We distinguish two cases for each non-leaf i in the path-split, depending on whether or not i is a branching vertex of T'.

Suppose that P_i corresponds to a non-branching vertex of T'. Then its end-vertex has out-degree exactly 1 in T; denote this out-neighbour by j. Let X be the subset of E_i, consisting of vertices that have at least one red out-neighbour in the candidate set D_j. We define the candidate set D_i to be the set of vertices in S_i that correspond to the vertices in X. In this case it remains to show that $|X| \geqslant k/8$. Indeed, by pseudorandomness, all but at most $\sigma \log N$ vertices in E_i send at least $\varepsilon|D_j| \geqslant \varepsilon k/8 > am^{2\alpha}$ edges to D_j. Recall that by the choice of the ordering of the vertices, every vertex in E_i has at most $am^{2\alpha}$ blue out-neighbours in D_j, hence all but at most $\sigma \log N$ vertices in E_i have a red out-neighbour in D_j, *i.e.* $|X| \geqslant |E_i| - \sigma \log N \geqslant k/8$.

Now suppose that P_i is a branching vertex in the path-split, *i.e.* it corresponds to a branching vertex v in T. The maximum out-degree of T is bounded by the number of leaves, so i has at most

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

m^α out-neighbours in T'; denote them by $j_1, \ldots, j_h$ (so $h \leqslant m^\alpha$). Let D_i be the set of vertices in V_i which have a red out-neighbour in each of the sets D_{j_s} for $s \in [h]$. As before, all but at most $\sigma \log N$ vertices in V_i have at least one red out-neighbour in D_{j_s} for each s. Hence $|D_i| \geqslant |V_i| - h \cdot \sigma \log N \geqslant k/8$.

We defined candidate sets required by Claim 4.15, so in the case when most of the sets U_i are red, we find a red copy of T.

Case 2(b): most of the sets U_i are blue. In this case we assume that at least $t/2$ of the U_i are blue; let us now rename $t/2$ blue U_i as $V_1, V_2, \ldots, V_{t/2}$, while preserving the ordering, and let $E_i, S_i \subseteq V_i$ be the sets of end- and start-vertices of the (blue) paths given by Claim 4.14; then $|E_i| = |S_i| \geqslant k/8$ for every i.

Define an auxiliary complete directed graph K on vertex set $[t/2]$, where vertex i corresponds to V_i. We define the following 2-colouring of its edges. Every edge ij with $i < j$ is coloured red. We colour an edge ij with $i > j$ blue if, for every choice of subsets $W_i \subseteq V_i$ and $W_j \subseteq V_j$ of size at least $k/16$, there is a blue edge from W_i to W_j; otherwise, we colour the edge red.

Let M be a maximal red–red matching in K. We now distinguish two cases: M covers at least $t/4$ of the vertices of K, or there is a blue directed path of order at least $t/4$ (we have seen in Case 1(a) that one of these possibilities occurs).

There is a long blue path in K. In this case, we assume that there is a blue path $i_1 i_2 \cdots i_{t/4}$ in K. Let $X_j \subset E_{i_j}$ denote the set of vertices which are end-vertices of blue paths of order at least $j \cdot \ell$ in $\bigcup_{r\in[j]} V_{i_r}$, where $\ell := n/(m^{1-3\alpha})$.

Claim 4.16. *For every $j \in [t/4]$, we have $|X_j| \geqslant k/16$.*

Proof. We prove this by induction. In the case $j = 1$ every vertex in E_{i_1} is an end-vertex of a path of order ℓ in V_{i_1}. So, let us assume that the statement is true for some $j \geqslant 1$. Let $Y_{j+1} \subseteq S_{i_{j+1}}$ be the set of vertices that have a blue in-neighbour in X_j.

We now show that $|X_{j+1}| \geqslant |Y_{j+1}|$. Let $v \in Y_{j+1} \subseteq S_{i_{j+1}}$ and $u \in E_{i_{j+1}}$ be its corresponding end-vertex of a path Q of order ℓ. Since $v \in Y_{j+1}$, there exists a vertex w in X_j such that the edge wv is blue in G. By definition of X_j, w is the end-vertex of a path P of order at least $j \cdot \ell$. Then $PwvQ$ is a path of order at least $(j+1)\cdot \ell$ in $\bigcup_{r\in[j+1]} V_{i_r}$, hence $u \in X_{j+1}$. This shows that $|X_{j+1}| \geqslant |Y_{j+1}|$.

As there are no blue edges between X_j and $S_{i_{j+1}} \setminus Y_{j+1}$ and since $|X_j| \geqslant k/16$ (by induction) we have $|S_{i_{j+1}} \setminus Y_{j+1}| < k/16$, by the definition of the auxiliary graph K. This implies that $|Y_{j+1}| > k/8 - k/16 \geqslant k/16$, so $|X_{j+1}| \geqslant |Y_{j+1}| \geqslant k/16$, as required. □

By applying Claim 4.16 with $j = t/4$, we find a blue path of order $t/4 \cdot n/(m^{1-3\alpha})$ in G. By inequality (1) this blue path has order at least n, as desired.

There is a large red–red matching in K. Now we consider the case where there is a red–red matching M, which covers $t/4$ vertices of K. Let us denote the edges of M by

$$(V_{i_1}, V_{j_1}), (V_{i_2}, V_{j_2}), \ldots, (V_{i_{t/8}}, V_{j_{t/8}}),$$

where $i_s < j_s$ for every $s \in [t/8]$ and $i_1 < \cdots < i_{t/8}$. By definition of the auxiliary graph K, there are subsets $A_s \subseteq V_{i_s}$ and $B_s \subseteq V_{j_s}$ of size $k/16$ each, such that all edges from B_s to A_s are red; fix such subsets. The vertices $i_1, \ldots, i_{t/8}$ form a red transitive tournament that respects this ordering, *i.e.* $i_s i_r$ is an edge if $s < r$. By Observation 3.7, we may find within this tournament a copy of a $(3,\alpha)$-tree-split T' of T, since T' is an out-directed tree of size smaller than $2m^{1-3\alpha} \leqslant t/8$ (by inequality (2)).

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

It remains to find appropriate candidate sets D_i for each vertex i from the tree-split so that we can find a red copy of T in G, by Claim 4.9. We will construct D_i such that they have size at least $k/32$. As before, we start with leaf-trees and work our way up the tree, in such a way that when we are about to define a candidate set D_i, the candidate sets of subtrees corresponding to out-neighbours of the vertex i in the tree-split are already defined.

Let $j_1, \ldots, j_h$ be the out-neighbours in the tree-split of a vertex i. We assume that $D_{j_s} \subseteq A_{j_s}$ has been defined and has size at least $k/32$. Note that $h \leqslant m^\alpha$ due to the bound on the number of leaves of T, and possibly $h = 0$ if i corresponds to a leaf-tree. Let X_s be the set of vertices in A_i which have at least one red out-neighbour in D_{i_s}, for $s \in [h]$. Let Y be the set of vertices in B_i which have at least $|T_i| + \sigma \log N$ red out-neighbours in X_s for every $s \in [h]$; if $h = 0$ we define Y to be the set of vertices in B_i that have at least $|T_i| + \sigma \log N$ red out-neighbours in A_i. Finally, let D_i be the set of vertices in A_i that have at least $|T_i|$ red out-neighbours in Y.

Claim 4.17. $|D_i| \geqslant |A_i| - \sigma \log N$.

Proof. Firstly, note that every vertex in A_i has at most $am^{2\alpha}$ blue out-neighbours in A_{j_s} (as j_s corresponds to a set that appears later in the ordering $V_1, \ldots, V_{t/8}$ than the set that contains A_i). It follows from pseudorandomness that all but at most $\sigma \log N \leqslant k/32$ vertices in A_i have at least $\varepsilon|A_{j_s}| > am^{2\alpha}$ out-neighbours in A_{j_s}, at least one of which is red. In particular, $|X_s| \geqslant k/32$.

Secondly, again by pseudorandomness and by the fact that all edges from B_i to A_i are red, all but at most $\sigma \log N$ vertices in B_i have at least $\varepsilon|X_s| \geqslant |T_i| + \sigma \log N$ red neighbours in X_s. It follows that $|Y| \geqslant |B_i| - h \cdot \sigma \log N \geqslant k/32$. If $h = 0$, then, similarly, $|Y| \geqslant k/32$.

Finally, recall that the vertices in A_i have at most $am^{2\alpha}$ blue out-neighbours in B_i. Hence, by pseudorandomness, all but at most $\sigma \log N$ vertices in B_i have at least $\varepsilon|Y|$ out-neighbours in Y, at least $\varepsilon|Y| - am^{2\alpha} \geqslant |T_i|$ of which are red. It follows that $|D_i| \geqslant |A_i| - \sigma \log N$, as required. □

By Claim 4.9 we may find a red copy of T in G. This completes the proof of Theorem 4.7. □

4.3 General trees

We are now ready to prove Theorem 4.1, without the constraint on the number of leaves. Our proof strategy is to consider the $m^{1/6}$-core-split $F_1, \ldots, F_\ell$ of a tree on m vertices. Then each tree in the split has at most $m^{1/6}$ leaves, so we can use the intermediate result, Theorem 4.7, to find it in the right neighbourhood.

Proof of Theorem 4.1. Without loss of generality, we assume that T is out-directed, as otherwise we can look at in-neighbourhoods instead of out-neighbourhoods in G. Suppose that G, together with a fixed 2-colouring, has no blue copy of $\overrightarrow{P_n}$. Let c_1 be the constant from Theorem 4.7 for parameters ε and 2σ. Define $\delta = \varepsilon^2/(32 \cdot 6)$. We assume that $c \geqslant \max(2c_1\delta^{-1}, 4\delta^{-2})$.

Claim 4.18. *Let $U \subseteq V(G)$ be a set of size at least $\delta N - m$ and let T' be an out-directed tree on at most m vertices with at most $m^{1/6}$ leaves. Then $G[U]$ contains a red copy of T'.*

Proof. Firstly, we claim that $G[U]$ is $(\varepsilon, 2\sigma \log M)$-pseudorandom, where $M = |U|$. Indeed, note that

$$M \geqslant \delta N - m \geqslant (\delta - 1/c)N \geqslant \frac{\delta}{2}N \geqslant \frac{\delta\sqrt{c}}{2}\sqrt{N} \geqslant \sqrt{N},$$

using $N \geqslant c$ and $c \geqslant 4/\delta^2$. Thus $2\sigma \log M \geqslant \sigma \log N$, so $G[U]$ is $(\varepsilon, 2\sigma \log M)$-pseudorandom, using $(\varepsilon, \sigma \log N)$-pseudorandomness of G. Next, note that

$$n, m \leqslant \frac{N}{c} \leqslant \frac{2}{\delta c}M \leqslant \frac{M}{c_1},$$

and

$$nm \leqslant \frac{N^2}{c^2 \log N} \leqslant \frac{4}{c^2\delta^2}\frac{M^2}{\log M} \leqslant \frac{M^2}{c_1^2 \log M},$$

as $c \geqslant 2c_1/\delta$. Hence, by Theorem 4.7, U contains either a red T' or a blue $\overrightarrow{P_n}$; by assumption it follows that U contains a red T', as required. □

Let $F_1, \dots, F_\ell$ be the $m^{1/6}$-core-split of T. By Proposition 4.4, $\ell \leqslant 6$ and each tree in a forest F_i has at most $m^{1/6}$ leaves.

Define $U_0 = V(G)$, and for $i \leqslant 5$ let U_i be the set of vertices in $V(G)$ that have at least δN red out-neighbours in U_{i-1}.

Claim 4.19. *$|U_i| \geqslant N/6$ for $i \leqslant 5$.*

Proof. We prove by induction on i that $|U_i| \geqslant (1 - i/6)N$. This holds trivially for $i = 0$, as $U_0 = V(G)$. Now let $1 \leqslant i \leqslant 5$, and suppose that the statement holds for $i - 1$. Consider the set $W := U_{i-1} \setminus U_i$. Suppose that $|U_i| < (1 - i/6)N$, then by induction $|W| \geqslant N/6$. Also, by the definition of U_i, the number of red edges in W is at most $|W| \cdot \delta N \leqslant (\varepsilon^2/32)|W|^2$ (recall that $\delta = \varepsilon^2/(32 \cdot 6)$). It follows from Lemma 3.3 that W contains a blue $\overrightarrow{P_n}$, as $(\varepsilon/4)|W| \geqslant \max(\sigma \log N, n)$, a contradiction. Hence, $|U_i| \geqslant (1 - i/6)N \geqslant N/6$, as required. □

We now show how to find a red copy of T. We first find a red copy of F_1 in U_5; this is possible due to Claim 4.18 and the fact that F_1 is an out-directed tree on at most m vertices with at most $m^{1/6}$ leaves. Suppose that we found a red copy of $T \setminus (V(F_\ell) \cup \cdots \cup V(F_i))$ for some $2 \leqslant i \leqslant \ell$, such that the vertices corresponding to F_{i-1} are in U_{7-i}. We embed the trees in F_i one by one. Let T' be one such tree, and let u be the vertex in U_{7-i} that corresponds to the parent of T' in T. Let W be the set of red out-neighbours of u in U_{6-i} that are still available. By choice of U_i, $|W| \geqslant \delta N - m$, so by Claim 4.18 there is a red T' in W. Continuing in this way, we find a copy of $T \setminus (V(F_\ell) \cup \cdots \cup V(F_{i+1}))$ such that the vertices corresponding to F_i are in U_{6-i}. Doing this until $\ell = 6$, we find a red copy of T. This completes the proof of Theorem 4.1. □

5. Tree versus tree

In this section we extend Theorem 4.1 to the case of two general (*i.e.* not necessarily directed) trees. We start by proving it for a directed tree with few leaves versus any directed tree (see Theorem 5.1). We then remove the assumption that one of the trees has few leaves (Theorem 5.5). Finally, we also remove the assumption that the trees are directed (Theorem 5.8). We will often start by embedding a subtree T' of a tree T, and then attempt to embed the trees in $T \setminus V(T')$ in the neighbourhood of a suitable vertex in T'.

5.1 Directed tree versus directed tree with few leaves

Our first goal is to prove the following theorem.

Theorem 5.1. *Given $0 < \varepsilon < 1/2$ and $k, \sigma > 0$ there exists a constant $c > 0$ such that the following holds. Let G be a tournament on N vertices which is $(\varepsilon, \sigma \log N)$-pseudorandom, let S be a directed tree on n vertices, and let T be a directed trees on m vertices with at most k leaves, where $m, n \leqslant N/c$ and $nm \leqslant N^2/(c^2 \log N)$. Then $G \rightarrow (S, T)$.*

Before turning to the proof, we give a definition. Let T be an out-directed tree. The *disjoint paths layer* of T, denoted $\mathrm{L}(T)$ (see Figure 5), is the collection of paths of T that end at a non-root

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

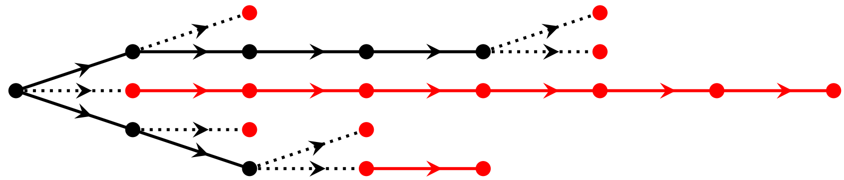

Figure 5. The disjoint paths layer of a tree.

leaf u and start one vertex after the last branching vertex, or root, between the root and u; in the case where T consists of a single vertex (which is the root), we instead define $\mathrm{L}(T) = T$. In particular, the vertices in $L(T)$, except for the leaves of T, have degree exactly 2 in T.

Proposition 5.2. *The following properties hold for every out-directed tree T:*

(i) $\mathrm{L}(T)$ *is a union of pairwise vertex-disjoint directed paths of T,*
(ii) $T \setminus V(\mathrm{L}(T))$ *is an out-directed tree,*
(iii) *the number of non-root leaves in $T \setminus V(\mathrm{L}(T))$ is at most half the number of non-root leaves in T.*

Proof. The first two properties are immediate from the definition. Property (iii) follows as each non-root leaf in $T \setminus V(\mathrm{L}(T))$ sends at least two edges to paths of $\mathrm{L}(T)$. □

Proof of Theorem 5.1. Without loss of generality, suppose that T is out-directed. We assume that G has no blue S. Let c_1 be the constant from Theorem 4.1 with parameters ε and 2σ, set $\delta := \varepsilon^2/(32(\log k + 2))$, and pick c such that $c \geqslant \max\{2c_1/\delta, 4(\log k + 2)/\varepsilon\}$. We use the following claim.

Claim 5.3. *Let U be a set of at least $\delta N - m$ vertices. Then U contains a red $\overrightarrow{P_m}$.*

Proof. Let $M := |U| \geqslant \delta N - m$. Then, using $c \geqslant 2c_1/\delta \geqslant 2/\delta$,

$$M \geqslant \delta N - \frac{N}{c} \geqslant \frac{\delta}{2} \cdot N \geqslant \sqrt{N}.$$

Since G, and thus $G[U]$, is $(\varepsilon, \sigma \log N)$-pseudorandom, $G[U]$ is $(\varepsilon, 2\sigma \log M)$-pseudorandom. Using $c \geqslant 2c_1/\delta$, we have $M \geqslant (\delta/2)N \geqslant (c_1/c)N$. Thus, by the assumptions on n and m,

$$n, m \leqslant \frac{N}{c} \leqslant \frac{M}{c_1} \quad \text{and} \quad nm \leqslant \frac{N^2}{c^2 \log N} \leqslant \frac{M^2}{c_1^2 \log M}.$$

Hence, by definition of c_1 (according to Theorem 4.1), U contains a red $\overrightarrow{P_m}$ or a blue S. Since we assumed that the latter does not hold, U contains a red $\overrightarrow{P_m}$, as required. □

Our plan is to embed a red copy of T layer by layer. To this end, define $T_0 := T$ and, for $i \geqslant 1$, $T_i := T_{i-1} \setminus V(\mathrm{L}(T))$, and let h be the largest i such that T_i is non-empty. Note that T_h is a singleton (as the root is not removed unless the root is the only vertex) and, by Proposition 5.2(iii), T_i has at most $k \cdot 2^{-i}$ non-root leaves; in particular, $h \leqslant \log k + 1$.

Define $U_0 := V(G)$, and for $1 \leqslant i \leqslant h$ let U_i be the set of vertices in U_{i-1} whose red out-degree in U_{i-1} is at least δN. We shall need the following claim.

Claim 5.4. *$U_h \neq \emptyset$.*

Proof. The proof is essentially identical to that of Claim 4.19. We prove by induction that

$$|U_i| \geqslant \left(1 - \frac{i}{h+1}\right)N$$

for $0 \leqslant i \leqslant h$. This is trivial for $i = 0$, as $U_0 = V(G)$. Let $0 < i \leqslant h$, and suppose that the statement holds for $i-1$, that is,

$$|U_{i-1}| \geqslant \left(1 - \frac{i-1}{h+1}\right)N.$$

Set $W := U_{i-1} \setminus U_i$. Suppose that

$$|U_i| < \left(1 - \frac{i}{h+1}\right)N,$$

so $|W| \geqslant N/(h+1)$. We now wish to apply Lemma 3.3. To do so, note that, by definition of U_i, the number of red edges in W is at most

$$|W| \cdot \delta N = |W| \cdot \frac{\varepsilon^2}{32(\log k + 2)} \cdot N \leqslant \frac{\varepsilon^2}{32}|W| \cdot \frac{N}{h+1} \leqslant \frac{\varepsilon^2}{32}|W|^2,$$

using the definition of δ and the bounds $h \leqslant \log k + 1$ and $|W| \geqslant N/(h+1)$. We also have $(\varepsilon/4)|W| \geqslant \sigma \log N$ (since $N \geqslant c$ and we take c large enough, in terms of σ, ε, k). Thus, by Lemma 3.3, G contains any blue tree on at most $(\varepsilon/4)|W|$ vertices. Since

$$m \leqslant \frac{N}{c} \leqslant \frac{\varepsilon}{4} \cdot \frac{N}{h+1} \leqslant \frac{\varepsilon}{4}|W|$$

(using $c \geqslant 4(\log k + 2)/\varepsilon \geqslant 4(h+1)/\varepsilon$), it follows that G contains a blue copy of T, a contradiction. □

We now show that there is a red copy of T_i in U_i, by induction on $0 \leqslant i \leqslant h$. Since T_h is a singleton and U_h is non-empty, there is indeed a red copy of T_h in U_h. Now suppose that for some $0 \leqslant i < h$, there is a red copy of T_{i+1} in U_{i+1}. Recall that $T_{i+1} = T_i \setminus V(\mathrm{L}(T_i))$; hence it suffices to show that the paths in $\mathrm{L}(T_i)$ can be embedded in the red out-neighbourhoods of the corresponding vertices in T_{i+1}. We embed the paths in $\mathrm{L}(T_i)$ one by one. Let P be a path in $\mathrm{L}(T_i)$ of order ℓ, let v be its start-vertex and let u be the vertex in U_{i+1} that corresponds to the parent of v in T. Let W denote the red out-neighbours of u in U_i which are still available. Then, since u is in U_{i+1} and at most m vertices are used, $|W| \geqslant \delta N - m$. By Claim 5.3, W contains a red P, as required. We are thus able to embed each of the paths in $\mathrm{L}(T_i)$ in U_i so as to obtain a red copy of T_i in U_i. In particular, by taking $i = 0$, we see that G has a red copy of T, as required for the proof of Theorem 5.1. □

5.2 Directed trees

With the next theorem we further generalize the result to the case of any directed trees S and T. We once again obtain a reduction to the previous result, Theorem 5.1. This time we make use of k-cores, which we have already encountered in the proof of Theorem 4.1 (see Definition 3.4).

Theorem 5.5. *Given $0 < \varepsilon < 1/2$ and $\sigma > 0$, there exists a constant $c > 0$ such that the following holds. Let G be a tournament on N vertices which is $(\varepsilon, \sigma \log N)$-pseudorandom. Then $G \rightarrow (S, T)$ for any directed trees S and T on n and m vertices, respectively, where $n, m \leqslant N/c$ and $nm \leqslant N^2/(c^2 \log N)$.*

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

Proof. Our goal is to reduce the statement of this theorem to the case when one of the trees has a constant number of leaves. We iteratively make the trees S and T smaller, using Theorem 5.1, until one of them becomes a singleton.

Define $\delta = \varepsilon^2/64$, $\ell := 8/\delta^2$, $k := \ell^2$, let c_1 be the constant from Theorem 5.1 with parameters ε, 3σ and k, and let $c := \max\{c_1\ell, 8\ell/\varepsilon\}$. Set $h := \lceil \log_k N \rceil$ for $0 \leqslant i \leqslant h$, and write $n_i := n \cdot k^{-i}$, $m_i := m \cdot k^{-i}$ and $N_i := N \cdot \ell^{-i}$. We shall use the following proposition.

Proposition 5.6. *The following properties hold.*

(i) *Let U be a set of at least N_{i+1} vertices, let S be a directed tree on n_i vertices, and let T be a directed tree on m_i vertices with at most k leaves. Then U contains a blue S or a red T.*

(ii) *Let U be a set of at least N_{i+1} vertices. Then either it contains a blue copy of any tree on n_i vertices, or the set of vertices in U whose red out-degree in U is at least $\delta|U|$ has size at least $|U|/2$.*

Proof. Firstly, note that for every $0 \leqslant i \leqslant h$

$$N_i \geqslant N \cdot \ell^{-\log_k N - 1} = N \cdot k^{-\frac{1}{2}\log_k N} \cdot \frac{1}{\ell} = \frac{\sqrt{N}}{\ell} \geqslant N^{1/3}.$$

It follows that every subset $U \subseteq V(G)$ of size at least N_i, where $0 \leqslant i \leqslant h$, is $(\varepsilon, 3\sigma \log |U|)$-pseudorandom.

Note that

$$n_i = nk^{-i} \leqslant n\ell^{-i} \leqslant (N/c)\ell^{-i} = N_{i+1}\ell/c \leqslant N_{i+1}/c_1$$

(using $c \geqslant c_1\ell$). Similarly, $m_i \leqslant N_{i+1}/c_1$ and $n_i m_i \leqslant N_{i+1}^2/c_1^2 \log N_{i+1}$. Property (i) thus follows from the definition of c_1 (via Theorem 5.1).

Property (ii) can be deduced from Lemma 3.3 as follows. Suppose that the set X of vertices in U whose red out-degree is smaller than $\delta|U|$ has size at least $|U|/2$. Then the number of red edges spanned by X is at most $|X| \cdot \delta|U| \leqslant (\varepsilon^2/32)|X|^2$. Thus, by Lemma 3.3, $G[X]$ contains a blue copy of any tree on at most $(\varepsilon/4)|X| \geqslant n_i$ vertices, as required, where we used the inequalities

$$\frac{\varepsilon}{4}|X| \geqslant \frac{\varepsilon}{8}N_{i+1} \geqslant \frac{\varepsilon c}{8\ell}n_i \geqslant n_i$$

(using $n_i \leqslant N_{i+1}\ell/c$ and $c \geqslant 8\ell/\varepsilon$) and

$$\frac{\varepsilon}{4}|X| \geqslant \frac{\varepsilon}{8}N_{i+1} \geqslant \sigma \log N. \qquad \square$$

We complete the proof with the following claim.

Claim 5.7. *Let $U \subseteq V(G)$ be a set of size at least N_i, where $0 \leqslant i \leqslant h$, and let S and T be directed trees of order n_i and m_i, respectively. Then U contains a blue S or a red T.*

Proof. We prove the claim by induction on i. Note that when $i = h$ the claim holds trivially as $n_h, m_h \leqslant 1$ and $N_h \geqslant 1$. Now suppose that $0 \leqslant i < h$ and the claim holds for $i + 1$.

Suppose that U does not contain a blue S or a red T. For convenience, we assume that S and T are out-directed; the remaining cases follow similarly. Let S' and T' be the k-cores of S and T, respectively. Then S' and T' have at most k leaves, $S \setminus V(S')$ is a forest of trees of order at most n_{i+1}, and $T \setminus V(T')$ is a forest of trees of order at most m_{i+1}.

Let X be the set of vertices in U whose red out-degree in U is at least $\delta|U|$. Then, by Proposition 5.6(ii) and the assumption that U does not contain a blue S, we have $|X| \geqslant |U|/2 \geqslant N_i/2 \geqslant N_{i+1}$. By Proposition 5.6(i) and the assumption that U does not have a blue S, X contains

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

a red T'. We attempt to extend the copy of T' to a red T in U by attaching, one at a time, copies of the trees in $T \setminus V(T')$. As U does not have a red copy of T, at some point we fail. Let T'' be the tree in $T \setminus V(T')$ that we fail to embed (while T' and some of $T \setminus V(T')$ is already embedded). Denote the root of T'' by u, and let u' be the vertex in X in which we embedded the parent of u in T.

Let Y denote the set of red out-neighbours of u' in U which have not been used yet, so by the failure to embed T'', Y does not have a red T''. Let Y' be the set of vertices in Y whose blue out-degree in Y is at least $\delta|Y|$. Then

$$|Y| \geqslant \delta|U| - m_i \geqslant \frac{\delta}{2}|U| \geqslant \frac{\delta}{4}N_i \geqslant N_{i+1}$$

and hence, by Proposition 5.6(ii), with red and blue swapped,

$$|Y'| \geqslant |Y|/2 \geqslant \frac{\delta}{8}N_i \geqslant N_{i+1}.$$

As Y, and thus Y', does not contain a red T'', it follows from Proposition 5.6(i) that Y' contains a blue S'. Again, we try to extend this copy of S' to a blue copy of S in Y', by attaching one tree of $S \setminus V(S')$ at a time. As there is no blue copy of S in Y', at some point we fail; let S'' denote the tree that we fail to embed. Let v be the root of S'', and let v' be the vertex in Y' where we embedded the parent of v in S.

Let Z be the set of blue out-neighbours of v' in Y' which are not used. Then

$$|Z| \geqslant \delta|Y| - n_i \geqslant \frac{\delta}{2}|Y| \geqslant \frac{\delta^2}{8}N_i = N_{i+1}$$

and Z does not have a red T'' or a blue S'', contrary to the induction hypothesis. It follows that U contains a red T or a blue S, as required. □

The proof of Theorem 5.5 follows immediately from Claim 5.7 by taking $i = 0$. □

5.3 General trees

Our final aim is to generalize Theorem 5.5 to arbitrary oriented trees, as follows.

Theorem 5.8. *Given $0 < \varepsilon < 1/2$ and $\sigma > 0$, there exists a constant $c > 0$ such that the following holds. Let G be a tournament on N vertices which is $(\varepsilon, \sigma \log N)$-pseudorandom, and let S and T be trees of orders n and m, respectively, where $m, n \leqslant N/c$ and $nm \leqslant N^2/(c^2 \log N)$. Then $G \to (S, T)$.*

We will use the next definition and lemma in the proof.

Definition 5.9. Let G be an oriented graph and k a positive constant. We call a pair of disjoint subsets $(A, B) \subseteq V(G)^2$ a *k-mindegree pair* if every vertex in A has at least k out-neighbours in B and every vertex in B has at least k in-neighbours in A.

Lemma 5.10. *Let $0 < \delta < 1/4$. In every oriented graph G with at least $\delta|G|^2$ edges, there is a $(\delta/4)|G|$-mindegree pair.*

Proof. Let us define a partition (X, Y) of $V(G)$ by putting each vertex independently with probability $1/2$ either in X or in Y. Note that the expectation of $e(X, Y)$ is at least $e(G)/4 \geqslant (\delta/4)|G|^2$. Thus, there exist disjoint sets X and Y with $e(X, Y) \geqslant (\delta/4)|G|^2$.

Now we consider the underlying subgraph of G whose edges are those going from X to Y. We remove one by one all vertices with degree less than $(\delta/4)|G|$ in this underlying graph. Let $A \subseteq X$ and $B \subseteq Y$ be the sets of remaining vertices. Note that both A and B are non-empty, since

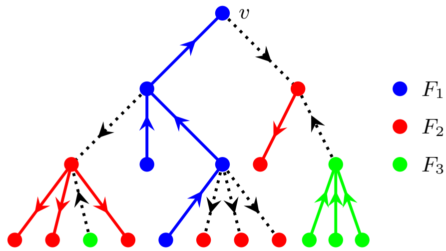

Figure 6. The in–out split of a tree.

otherwise all vertices would be removed by this process, each contributing less than $(\delta/4)|G|$ edges. This would imply $e(X, Y) < (\delta/4)|G|^2$, a contradiction. Therefore, (A, B) is a $(\delta/4)|G|$-mindegree pair in G. □

We shall use Theorem 5.5 in our proof of Theorem 5.8. For this we need a suitable split of T.

Let v be the root of T. Let F_1 be the induced subtree of T containing v and all vertices of T, which can be reached from v by following in-edges (it is possible that F_1 contains only v). Let U_2 be the set of roots of the trees in the forest $T \setminus V(F_1)$. We define F_2 to be the forest of induced subtrees of T consisting of the vertices in U_2 and all vertices in $T \setminus V(F_1)$ that can be reached from U_2 by following out-edges. We continue this procedure and eventually we obtain a split of T into layers of in- and out-forests $F_1, \ldots, F_\ell$, such that the forest F_i consists of in-directed trees for odd $i \in [\ell]$ and out-directed trees for even $i \in [\ell]$. Moreover, all edges in T are either contained in a forest F_i or are between consecutive layers F_i and F_{i+1}, and they are directed from F_i to F_{i+1} if i is odd, and are directed from F_{i+1} to F_i if i is even. We call this split the *in–out split* of T (see Figure 6).

Proof of Theorem 5.8. We first prove the theorem under the additional assumption that S is directed, using Theorem 5.5, and then we use this to prove the theorem in full generality. In order to avoid repeating the arguments, we use Proposition 5.11 below.

Let $\delta = \varepsilon^2/32$. Let c_1 be the constant from Theorem 5.5 with parameters ε and 2σ. Without loss of generality $c_1 \geqslant 8/\delta$, and let $c = c_1^3$. Let N, n, m be fixed (such that the inequalities in the statement of the theorem hold), and let G be a 2-coloured tournament on N vertices.

Proposition 5.11. *Let $N/c_1 \leqslant M \leqslant N$. Let $U \subseteq V(G)$ be a set of size M, and suppose that every subset of U of size at least M/c_1 contains a red copy of every* directed *tree of order at most m. Then U contains a blue copy of every tree (not necessarily directed) of order n, or a red copy of every tree of order m.*

The same holds with the roles of red and blue, and the roles of n and m, swapped.

Proof. As we have done already several times, by Lemma 3.3 we can assume that U spans at least δM^2 red edges, since otherwise U contains a blue copy of every tree of order $(\varepsilon/4)M \geqslant n$. Then by Lemma 5.10 there exist disjoint sets $A, B \subseteq U$ such that (A, B) is a $(\delta/4)M$-mindegree pair in the red subgraph of G. Let T be a tree on m vertices, consider its in–out split $F_1, \ldots, F_\ell$, and let $V_1, \ldots, V_\ell$ denote the corresponding partition of vertices of the tree T; recall that F_1 is an in-directed subtree of T. We will embed every in-directed tree of the in–out split inside A and every out-directed tree in B.

Claim 5.12. *For every $i \in [\ell]$ there is a red copy of $T[V_1 \cup \cdots \cup V_i]$ such that V_i is embedded in A if i is odd and in B if i is even.*

Proof. We prove this by induction. For the basis, note that F_1 is a single in-directed tree and $|A| \geqslant (\delta/4)M \geqslant M/c_1$, thus by assumption there is a red copy of F_1 inside A.

https://doi.org/10.1017/S0963548319000373 Published online by Cambridge University Press

Now let us assume that the claim holds for $1 \leqslant i-1 < \ell$. For convenience we assume that i is even; the case where i is odd follows similarly. So, we have found a red copy of $T[V_1 \cup \cdots \cup V_{i-1}]$ such that V_{i-1} is embedded in A. Now we need to show how to embed the trees of the forest F_i. Let T' be one of the trees in the forest F_i and let $v \in A$ be the vertex corresponding to the parent of the roof of T' in $T[V_1 \cup \cdots \cup V_{i-1}]$. Since (A, B) is a $(\delta/4)M$-mindegree pair, v has at least $(\delta/4)M$ red out-neighbours in B. So far we embedded at most n vertices of the tree T, so the number of available vertices in the neighbourhood is at least $(\delta/4)M - n \geqslant (\delta/4)M - M/c_1 \geqslant M/c_1$. Therefore, by assumption, there is a red copy of T' in B rooted at some vertex w, such that edge vw is red.

This way we can embed all the trees in F_i and extend the red copy of $T[V_1 \cup \cdots \cup V_{i-1}]$ to a red copy of $T[V_1 \cup \cdots \cup V_i]$ satisfying the conditions of the claim. □

By Claim 5.12 with $i = \ell$, U contains a red T. As T was an arbitrary tree on m vertices, the proof is complete. An analogous argument can be used to prove the statement of the proposition with the roles of red and blue, and of m and n, swapped. □

We now show how to complete the proof of Theorem 5.8 using Proposition 5.11. Suppose that there exists a subset $U \subseteq V(G)$ of size at least N/c_1, whose subsets of size at least $|U|/c_1$ all contain a red copy of every directed tree of order m. Then, by Proposition 5.11, U contains a red copy of every tree of order m or a blue copy of every tree of order n, and we are done. Thus we may assume that every subset $U \subseteq V(G)$ has a subset W_U of size at least $|U|/c_1$ such that W_U does not contain a red T_U, for some directed tree T_U of order m. But then, by Theorem 5.5, every such W_U contains a blue copy of every directed tree on n vertices (using the definition of c_1, and the inequalities $n, m \leqslant N/c = N/c_1^3 \leqslant |W_U|/c_1$ and $nm \leqslant N^2/c^2 \log N \leqslant |W_U|^2/c_1^2 \log|W_U|$). In particular, every set $U \subseteq V(G)$ of size at least N/c_1 contains a blue copy of every directed tree on n vertices. By Proposition 5.11 again (with the roles of red and blue and n and m swapped), either G contains a red copy of every tree on m vertices, or a blue copy of every tree on n vertices, as required. □

6. Concluding remarks and open problems

In this paper we have proved that, with high probability, in every 2-edge-colouring of a random tournament on $Cn\sqrt{\log n}$ vertices there exists a monochromatic copy of any tree of order n.

Bucić, Letzter and Sudakov [4] proved tight results for both oriented and directed Ramsey numbers of trees for the case of more than two colours as well. It seems that the methods used in their proofs do not extend directly to the random tournament setting, so it could be very interesting to extend our result to k-colours. In the case of paths they showed in [5] that, with high probability, in any k-edge colouring of a random tournament on $\Omega(n^{k-1}\sqrt{\log n})$ vertices, there is a monochromatic path of length n. Moreover, an example by Ben-Eliezer, Krivelevich and Sudakov [2] shows that there is a k-edge colouring of any tournament on $cn^{k-1}(\log n)^{1/k}$ vertices with no monochromatic paths of length n, for some constant $c > 0$. We believe the upper bound should be tight, for random tournaments, but the k-colour case is still open, even for directed paths.

Burr and Erdős [6] initiated the study of Ramsey numbers of bounded degree graphs in 1975. They conjectured that the Ramsey number of bounded degree graphs is linear in their size. This was subsequently proved by Chvátal, Rödl, Szemerédi and Trotter [7]. The dependence of the constant factor on the maximum degree in this bound was later improved, first by Eaton [9], then by Graham, Rödl and Ruciński [11] and the current best bound is due to Conlon, Fox and Sudakov [8]. Bucić, Letzter and Sudakov [4] pose an interesting analogous problem in the oriented and directed Ramsey settings. They ask if, for every d, there is a constant $c = c(d)$ such that any tournament on cn vertices contains any acyclic graph on at most n vertices with maximum degree at most d. This can be thought of as the one-colour version of the more general question of determining the r-colour oriented Ramsey number of bounded degree acyclic graphs. A similar question arises naturally in the random setting. Here the one-colour version is a simple consequence of the directed version of Szemerédi's Regularity Lemma [18] due to Alon and Shapira [1]. However, the question of the two colours is open and interesting, and it seems likely that a result in any setting could also help with the other settings.

Theorem 1.1 is tight up to a constant factor, as long as the only information we are given on the tree is its order. However, it is not tight for every tree of order n. For example, if the tree in question T is a star of order n, then it is not hard to see that the random tournament G is only required to have order $\Omega(n)$ in order to satisfy $G \to T$, as opposed to a bound of $\Omega(n\sqrt{\log n})$ which is needed for a directed path on n vertices, or for trees which contain directed subpaths of order $\Omega(n)$. With this in mind, it is natural to ask if the tight bound for a tree T depends only on the order of the tree and the length of its longest directed subpath, denoted by $\ell(T)$. More precisely, Bucić, Letzter and Sudakov [4] ask if the directed Ramsey number of a tree is $O(|T| \cdot \ell(T))$; if this holds, it can readily be seen to be tight. They prove that this holds for oriented paths. It would also be interesting to tackle this question in the random tournament setting.

Acknowledgements

We would like to thank the anonymous referee for helpful comments.

References

[1] Alon, N. and Shapira, A. (2004) Testing subgraphs in directed graphs. *J. Comput. Syst. Sci.* **69** 354–382.
[2] Ben-Eliezer, I., Krivelevich, M. and Sudakov, B. (2012) The size Ramsey number of a directed path. *J. Combin. Theory Ser. B* **102** 743–755.
[3] Bermond, J.-C. (1974) Some Ramsey numbers for directed graphs. *Discrete Math.* **9** 313–321.
[4] Bucić, M., Letzter, S. and Sudakov, B. (2019) Directed Ramsey number for trees. *J. Combin. Theory Ser. B* **137** 145–177.
[5] Bucić, M., Letzter, S. and Sudakov, B. (2019) Monochromatic paths in random tournaments. *Random Struct. Alg.* **54** 69–81.
[6] Burr, S. A. and Erdős, P. (1975) On the magnitude of generalized Ramsey numbers for graphs. *Colloq. Math. Soc. János Bolyai* **10** 215–240.
[7] Chvatál, V., Rödl, V., Szemerédi, E. and Trotter, W. T. (1983) The Ramsey number of a graph with bounded maximum degree. *J. Combin. Theory Ser. B* **34** 239–243.
[8] Conlon, D., Fox, J. and Sudakov, B. (2012) On two problems in graph Ramsey theory. *Combinatorica* **32** 513–535.
[9] Eaton, N. (1998) Ramsey numbers for sparse graphs. *Discrete Math.* **185** 63–75.
[10] Gallai, T. (1968) On directed paths and circuits. In *Theory of Graphs (Proc. Colloq., Tihany, 1966)* (P. Erdős and G. Katona, eds), Academic Press, pp. 115–118.
[11] Graham, R. L., Rödl, V. and Ruciński, A. (2000) On graphs with linear Ramsey numbers. *J. Graph Theory* **35** 176–192.
[12] Gyárfás, A. and Lehel, J. (1973) A Ramsey-type problem in directed and bipartite graphs. *Period. Math. Hungar.* **3** 299–304.
[13] Harary, F. and Hell, P. (1974) Generalized Ramsey theory for graphs, V: The Ramsey number of a digraph. *Bull. London Math. Soc.* **6** 175–182.
[14] Hasse, M. (1965) Zur algebraischen Begründung der Graphentheorie, I. *Math. Nachr.* **28** 275–290.
[15] Ramsey, F. P. (1930) On a problem of formal logic. *Proc. London Math. Soc.* **30** 264–285.
[16] Raynaud, H. (1973) Sur le circuit hamiltonien bi-coloré dans les graphes orientés. *Period. Math. Hungar.* **3** 289–297.
[17] Roy, B. (1967) Nombre chromatique et plus longs chemins d'un graphe. *Rev. Française Informat. Recherche Opérationnelle* **1** 129–132.
[18] Szemerédi, E. (1978) Regular partitions of graphs. In *Problèmes combinatoires et théorie des graphes*, Vol. 260 of Proc. Colloq. Internat. CNRS, pp. 399–401.
[19] Vitaver, L. M. (1962) Determination of minimal coloring of vertices of a graph by means of Boolean powers of the incidence matrix. *Dokl. Akad. Nauk SSSR* **147** 758–759.
[20] Williamson, J. E. (1973) A Ramsey type problem for paths in digraphs. *Math. Ann.* **203** 117–118.

Cite this article: Bucić M, Heberle S, Letzter S and Sudakov B (2020). Monochromatic trees in random tournaments. *Combinatorics, Probability and Computing* **29**, 318–345. https://doi.org/10.1017/S0963548319000373