Combinatorics, Probability and Computing (2009) 18, 247–258.
doi:10.1017/S0963548307008875 Printed in the United Kingdom

Constrained Ramsey Numbers

PO-SHEN LOH[1†] and BENNY SUDAKOV[2‡]

[1]Department of Mathematics, Princeton University, Princeton, NJ 08544, USA
(e-mail: ploh@math.princeton.edu)

[2]Department of Mathematics, UCLA, Los Angeles, CA 90095, USA
(e-mail: bsudakov@math.ucla.edu)

Received 20 December 2006; revised 14 October 2007; first published online 10 December 2007

For two graphs S and T, the constrained Ramsey number $f(S, T)$ is the minimum n such that every edge colouring of the complete graph on n vertices (with any number of colours) has a monochromatic subgraph isomorphic to S or a rainbow subgraph isomorphic to T. Here, a subgraph is said to be rainbow if all of its edges have different colours. It is an immediate consequence of the Erdős–Rado Canonical Ramsey Theorem that $f(S, T)$ exists if and only if S is a star or T is acyclic. Much work has been done to determine the rate of growth of $f(S, T)$ for various types of parameters. When S and T are both trees having s and t edges respectively, Jamison, Jiang and Ling showed that $f(S, T) \leqslant O(st^2)$ and conjectured that it is always at most $O(st)$. They also mentioned that one of the most interesting open special cases is when T is a path. In this paper, we study this case and show that $f(S, P_t) = O(st \log t)$, which differs only by a logarithmic factor from the conjecture. This substantially improves the previous bounds for most values of s and t.

1. Introduction

The Erdős–Rado Canonical Ramsey Theorem [6] guarantees that for any m, there is some n such that any edge colouring of the complete graph on the vertex set $\{1, \dots, n\}$, with arbitrarily many colours, has a complete subgraph of size m whose colouring is one of the following three types: monochromatic, rainbow, or lexical. Here, a subgraph is rainbow if all edges receive distinct colours, and it is lexical when there is a total order of its vertices such that two edges have the same colour if and only if they share the same larger endpoint.

† Research supported in part by a Fannie and John Hertz Foundation Fellowship, an NSF Graduate Research Fellowship, and a Princeton Centennial Fellowship.
‡ Research supported in part by NSF CAREER award DMS-0546523, NSF grant DMS-0355497, USA–Israeli BSF grant, and by an Alfred P. Sloan fellowship.

Since the the first two types of colourings are somewhat more natural, it is interesting to study the cases when we can guarantee the existence of either monochromatic or rainbow subgraphs. This motivates the notion of constrained Ramsey number $f(S,T)$, which is defined to be the minimum n such that every edge colouring of the complete graph on n vertices (with any number of colours) has a monochromatic subgraph isomorphic to S or a rainbow subgraph isomorphic to T. It is an immediate consequence of the Canonical Ramsey Theorem that this number exists if and only if S is a star or T is acyclic, because stars are the only graphs that admit a simultaneously lexical and monochromatic colouring, and forests are the only graphs that admit a simultaneously lexical and rainbow colouring.

The constrained Ramsey number has been studied by many researchers [1, 3, 4, 7, 8, 10, 13, 14, 18], and the bipartite case in [2]. In the special case when $H = K_{1,k+1}$ is a star with $k+1$ edges, colourings with no rainbow H have the property that every vertex is incident to edges of at most k different colours, and such colourings are called k-local. Hence $f(S, K_{1,k+1})$ corresponds precisely to the local k-Ramsey numbers, $r^k_{\mathrm{loc}}(S)$, which were introduced and studied by Gyárfás, Lehel, Schelp and Tuza in [11]. These numbers were shown to be within a constant factor (depending only on k) of the classical k-coloured Ramsey numbers $r(S;k)$, by Truszczyński and Tuza [17].

When S and T are both trees having s and t edges respectively, Jamison, Jiang and Ling [13] conjectured that $f(S,T) = O(st)$, and provided a construction which showed that the conjecture, if true, is best-possible up to a multiplicative constant. Here is a variant of such a construction, which we present for the sake of completeness, which shows that in general the upper bound on $f(S,T)$ cannot be brought below $(1+o(1))st$. For a prime power t let $\mathbb{F}_t$ be the finite field with t elements. Consider the complete graph with vertex set equal to the affine plane $\mathbb{F}_t \times \mathbb{F}_t$, and colour each edge based on the slope of the line between the corresponding vertices in the affine plane. The number of different slopes (hence colours) is $t+1$, so there is no rainbow graph with $t+2$ edges. Also, monochromatic connected components are cliques of order t, corresponding to affine lines. Therefore, if $\Omega(\log t) < s < t$, we can take a random subset of the construction (taking each vertex independently with probability s/t) to obtain a colouring of the complete graph of order $(1+o(1))st$ with $t+1$ colours in which all monochromatic connected components have size at most $(1+o(1))s$.

Although Jamison, Jiang and Ling were unable to prove their conjecture, they showed that $f(S,T) = O(st \cdot d_T) \leqslant O(st^2)$, where d_T is the diameter of T. Since this bound clearly gets weaker as the diameter of T grows, they asked whether a pair of paths maximizes $f(S,T)$, over all trees with s and t edges, respectively. This generated much interest in the special case when T is a path P_t. In [18], Wagner proved that $f(S,P_t) \leqslant O(s^2 t)$. This bound grows linearly in t when s is fixed but still has order of magnitude t^3 for trees of similar size. Although Gyárfás, Lehel and Schelp [10] recently showed that, for small t (less than 6), paths are not the extremal example, they remain one of the most interesting cases of the constrained Ramsey problem.

In this paper we prove the following theorem, which agrees with the conjecture, up to a logarithmic factor and the fact that T is a path. It significantly improves the previous bounds for most values of s and t, and in particular gives the first sub-cubic bound for the case when the monochromatic tree and rainbow path are of comparable size.

Theorem 1.1. *Let S be any tree with s edges, and let t be a positive integer. Then, for any $n \geqslant 3600st\log_2 t$, every colouring of the edges of the complete graph K_n (with any number of colours) contains a monochromatic copy of S or a rainbow t-edge path.*

This supports the conjectured upper bound of $O(st)$ for the constrained Ramsey number of a pair of trees. With Oleg Pikhurko, the second author obtained another result which provides further evidence for the conjecture. This result studies a natural relaxation of the above problem, in which one wants to find either a monochromatic copy of a tree S or a properly coloured copy of a tree T. It appears that in this case the logarithmic factor can be removed, giving an $O(st)$ upper bound. We view this result as complementary to our main theorem, and therefore have included its short proof in the Appendix to our paper.

We close this section by comparing our approach to Wagner's, as the two proofs share some similarities. This will also lead us to introduce one of the the main tools that we will use later. Both proofs find a structured subgraph $G' \subset G$ in which one may direct some edges in such a way that directed paths correspond to rainbow paths. Wagner's approach imposes more structure on G', which simplifies the task of finding directed paths, but this comes at the cost of substantially reducing $|G'|$. In particular, his $|G'|$ is s times smaller than $|G|$, which contributes a factor of s to his ultimate bound $O(s^2t)$. We instead construct a subgraph with weaker properties, but of order which is a constant fraction of $|G|$ (hence saving a factor of s in the bound). This complicates the problem of finding the appropriate directed paths, which we overcome by using the following notion of *median order*.

Definition. Let G be a graph, some of whose edges are directed. Given a linear ordering $\sigma = (v_1, \dots, v_n)$ of the vertex set, a directed edge $\overrightarrow{v_iv_j}$ is said to be *forward* if $i < j$, and *backward* if $i > j$. If σ maximizes the number of forward edges, it is called a *median order*.

Median orders were originally studied for their own sake; for example, finding a median order for a general digraph is known to be NP-hard. More recently, Havet and Thomassé [12] discovered that they are a powerful tool for inductively building directed paths in tournaments (complete graphs with all edges directed). Their paper used this method to produce a short proof of Dean's conjecture (see [5]) that every tournament has a vertex whose second neighbourhood is at least as large as the first. Havet and Thomassé also used a median order to attack Sumner's conjecture (see [19]) that every tournament of order $2n-2$ contains every oriented tree of order n. They succeeded in proving this conjecture precisely for arborescences (oriented trees where every vertex except the root has in-degree one) and within a factor-2 approximation for general oriented trees.

The only property that they used is the so-called *feedback property*: if $\sigma = (v_1, \dots, v_n)$ is a median order, then for any pair $i < k$, the number of forward edges $\overrightarrow{v_iv_j}$ with $i < j \leqslant k$ is at least the number of backward edges $\overleftarrow{v_iv_j}$ with $i < j \leqslant k$. This property is easily seen to be true by comparing σ to the linear order

$$\sigma' = (v_1, v_2, \dots, v_{i-1}, v_{i+1}, v_{i+2}, \dots, v_k, v_i, v_{k+1}, v_{k+2}, \dots, v_n),$$

which was obtained from σ by moving v_i to the position between v_k and v_{k+1}. As an illustration of the simple power of this property, consider the following well-known result, which we will in fact use later in our proof.

Claim. Every tournament has a directed Hamiltonian path.

Proof. Let $\sigma = (v_1, \dots, v_n)$ be a median order. For each i, the edge $v_i v_{i+1}$ is directed in some way because we have a tournament, and so the feedback property applied with $k = i + 1$ implies that it is in fact a forward edge $\overrightarrow{v_i v_{i+1}}$. Therefore, $(v_1, \dots, v_n)$ is already a directed path, so we are done. □

2. Proof of Theorem 1.1

Let us assume for the sake of contradiction that $n \geqslant 3600 st \log_2 t$, but there is no monochromatic copy of S and no rainbow t-edge path. In earlier papers on the constrained Ramsey numbers of trees [13, 18], and in this work, the following well-known crude lemma is the only method used to exploit the non-existence of a monochromatic S. Its proof follows from the observation that every graph with average degree $\geqslant 2s$ has an induced subgraph with minimum degree $\geqslant s$.

Lemma 2.1. *Let S be a tree with s edges, and let $G = (V, E)$ be a simple graph, edge-coloured with k colours, with no monochromatic subgraph isomorphic to S. Then $|E| < ks|V|$.*

The rest of the proof of our main theorem roughly separates into two main steps. First, we find a structured subgraph $G' \subset G$ whose order is within a constant factor of $|G|$. We aim to arrive at a contradiction by using G' to construct a rainbow t-edge path. The structure of G' allows us to direct many of its edges in such a way that certain directed paths are automatically rainbow. In the second step, we use the median order's feedback property to find many directed paths, which we then connect into a single long rainbow path using the structure of G'.

2.1. Passing to a directed graph

In this section, we show how to find a nicely structured subset of our original graph, at a cost of a constant factor reduction of the size of our vertex set. We then show how the search for a rainbow path reduces to a search for a particular collection of directed paths.

Lemma 2.2. *Let S be a tree with s edges and t be a positive integer. Let G be a complete graph on $n \geqslant 310st$ vertices whose edges are coloured (in any number of colours) in such a way that G has no monochromatic copy of S and no rainbow t-edge path.*

Then there exists a set R of 'rogue colours', a subset $U \subset V(G)$ with a partition $U = U_1 \cup \cdots \cup U_r$, an association of a distinct colour $c_i \notin R$ to each U_i, and an orientation of some of the edges of the induced subgraph $G[U]$, which satisfy the following properties.

(i) $|U| > \frac{n}{10}$, *$|R| < t$, and each $|U_i| < 2s$.*

(ii) *For any edge between vertices $x \in U_i$ and $y \in U_j$ with $i \neq j$, if it is directed $\overrightarrow{xy}$, its colour is c_i, if it is directed $\overrightarrow{yx}$, its colour is c_j, and if it is undirected, its colour is in R.*

(iii) *For any pair of vertices $x \in U_i$ and $y \in U_j$ (where i may equal j), there exist at least t vertices $z \notin U$ such that the colour of the edge xz is c_i and the colour of yz is c_j.*

Proof. Let us say that a vertex v is *t-robust* if, for every set F of t colours, there are at least $n/5$ edges adjacent to v that are not in any of the colours in F. Let $V_1 \subset V$ be the set of t-robust vertices. We will need a lower bound on $|V_1|$, but this is just a special case of Lemma A.2 (whose short proof appears in the Appendix). Substituting the values $a = n/5$ and $b = t$ into this lemma gives $|V(G) \setminus V_1| \leqslant 2(ts + n/5) < 4n/5$ and so $|V_1| \geqslant n/5$.

Now, let P be a rainbow path of maximal length in G such that at least one of its endpoints is in V_1, and let R be the set of colours of the edges of P. Thus $|R| < t$, by the assumption that G contains no rainbow t-edge path. Let B be the set of vertices that have at least $n/15$ adjacent edges in a colour in R. Then G contains at least $|B|n/30$ edges with colours in R. On the other hand, by applying Lemma 2.1 to the subgraph of G determined by taking only the edges with colours in R, we see that the total number of edges in G with colour in R is less than $|R|sn < tsn$, and so $|B| < 30st$.

Let v be an endpoint of P which is in V_1. Define the sets U_i as follows. Let $\{c_1, \ldots, c_r\}$ be the non-R colours that appear on edges adjacent to v. For each such c_i, let U_i be the set of vertices that are not in B or P, and are adjacent to v via an edge of colour c_i. Set $U = U_1 \cup \cdots \cup U_r$. We claim that these designations will satisfy the desired properties.

Consider arbitrary vertices $x \in U_i$ and $y \in U_j$, where i may equal j. Since $n \geqslant 30t$, we have $|V_1 \setminus P| \geqslant (2/15)n + t$, so $x, y \notin B$ implies that there are at least t choices for $z \in V_1 \setminus P$ such that both of the edges xz and yz have colours not in R. Each such xz must be in colour c_i, or else the extension of P by the path vxz would contradict maximality of P, and similarly each yz must be in colour c_j. Finally, $U \cap V_1 = \emptyset$, because any $w \in U \cap V_1$ would allow us to extend P by the edge vw. Therefore, we have property (iii).

For property (ii), let $x \in U_i$ and $y \in U_j$, with $i \neq j$. By property (iii), there exists some vertex $z \in V_1 \setminus P$ such that yz is in colour c_j. Then the colour of the edge xy must be in $\{c_i, c_j\} \cup R$, or else the extension of P by the path $vxyz$ would contradict its maximality. Therefore, we can leave it undirected if the colour is in R, and direct it according to property (ii) otherwise.

It remains to show property (i). We have already established that $|R| < t$ and we can obtain the first inequality from the construction of V_1 as follows. Since $v \in V_1$, it is t-robust and so is adjacent to at least $n/5$ edges in non-R colours. Therefore, using that $n \geqslant 310st$ we get

$$|U| \geqslant n/5 - |B| - |P| > n/5 - 30st - t \geqslant n/10.$$

For the last part, assume for the sake of contradiction that $|U_i| \geqslant 2s$. Arbitrarily select a subset $U_i' \subset U_i$ of size $2s$, and consider the subgraph G' formed by the edges of colour

c_i among vertices in $U_i' \cup V_1$. By the argument that showed property (iii), every edge between U_i' and V_1 has colour in $R \cup \{c_i\}$. So, since $U_i \cap B = \emptyset$, every $x \in U_i'$ is adjacent to at least $|V_1| - n/15 \geqslant (2/3)|V_1|$ vertices in V_1 via edges of colour c_i. Therefore, using that $|V_1|/3 \geqslant 2s = |U_i'|$, we have

$$e(G') \geqslant |U_i'| \cdot (2/3)|V_1| = (4/3)s|V_1| = s(|V_1| + (1/3)|V_1|) \geqslant s \cdot v(G').$$

Then Lemma 2.1 implies that G' has a copy of tree S, which is monochromatic by construction of G'. This contradiction completes the proof of the last part of property (i), and the proof of the lemma. □

The partially directed subgraph of Lemma 2.2 allows us to find rainbow paths by looking for certain types of directed paths. For example, if Lemma 2.2 produces $U = U_1 \cup \cdots \cup U_m$, and we have found a directed path $\overrightarrow{v_1 \cdots v_t}$ with each v_i from a distinct U_j, then it must be rainbow by property (ii) of the construction of U. Unfortunately, the following simple construction of a set, with no monochromatic S that satisfies the structure conditions of Lemma 2.2, shows that we cannot hope to obtain our rainbow path by searching for a single (long) directed path: re-index $\{U_i\}$ with ordered pairs as $\{U_{ij}\}_{i=1,j=1}^{h,t-1}$, let all $|U_{ij}| = s/3$, for all $1 \leqslant i < j \leqslant h$ direct all edges between any $U_{i,*}$ and $U_{j,*}$ in the direction $U_{i,*} \to U_{j,*}$, and for all $1 \leqslant i \leqslant j < t$ and $1 \leqslant k \leqslant h$ colour all edges between $U_{k,i}$ and $U_{k,j}$ in colour r_i, where $R = \{r_1, \ldots, r_{t-1}\}$. Although it is clear that this construction has no directed paths longer than $h = O\left(\frac{|U|}{st}\right)$, it is also clear that one could build a long rainbow path by combining undirected edges and directed paths. The following lemma makes this precise.

Lemma 2.3. *Let $U = U_1 \cup \cdots \cup U_m$ be a subset of $V(G)$ satisfying the structural conditions of Lemma 2.2, and let R be the associated set of rogue colours. Suppose we have a collection of $r < t$ edges $\{u_i v_i\}_{i=1}^r$ in $G[U]$ whose colours are distinct members of R, and a collection of directed paths $\{P_i\}_{i=0}^r$, with P_i starting at v_i for $i \geqslant 1$. Then, as long as all of the vertices in $\{u_1, \ldots, u_r\} \cup P_0 \cup \cdots \cup P_r$ belong to distinct sets U_j, there exists a rainbow path in G that contains all of the paths P_i and all of the edges $u_i v_i$. In short, one can link all of the fragments together into a single rainbow path.*

Proof. For each i, let w_i be the final vertex in the directed path P_i. For a vertex $v \in U$, let $c(v)$ denote the colour associated with the set U_i that contains v. Since $r < t$, by property (iii) of Lemma 2.2, for each $0 \leqslant i < r$, there exists a distinct vertex $x_i \notin U$ such that the colour of the edge $w_i x_i$ is $c(w_i)$ and the colour of the edge $x_i u_{i+1}$ is $c(u_{i+1})$. These vertices x_i together with paths P_i form a path P of distinct vertices, which we will now prove is rainbow.

Note that our linking process only adds edges with non-rogue colours. Since we assumed that the $u_i v_i$ have distinct colours, and the edges of the P_i are directed paths (hence with non-rogue colours), it is immediate that P has no duplicate rogue colours. Also note that among all directed edges in $\{P_i\}$, no pair of edges has initial endpoint in the same U_j by assumption. Therefore, they all have distinct colours by property (ii) of Lemma 2.2. Furthermore, none of these directed edges originates from any point in any U_j that intersects $\{u_1, \ldots, u_r, w_1, \ldots, w_r\}$, so they share no colours with

$C' = \{c(u_1), \ldots, c(u_r), c(w_1), \ldots, c(w_r)\}$; finally the colours in C' are themselves distinct because of our assumption that all vertices in $\{u_1, \ldots, u_r\} \cup P_0 \cup \cdots \cup P_r$ come from distinct U_j. This proves that P is a rainbow path. □

2.2. Finding directed paths

Now apply Lemma 2.2, and let us focus on $U = U_1 \cup \cdots \cup U_m$, which is of size at least $n/10 \geqslant 360st \log_2 t$. Let us call the edges which have colours in R 'rogue edges'. Note that if all edges were directed (*i.e.*, we have a tournament), then the existence of a long directed path follows from the fact that every tournament has a Hamiltonian path. The main issue is the presence of undirected edges. We treat these by observing that each undirected edge must have one of $|R| < t$ rogue colours. Then, we use the machinery of median orders to repeatedly halve the number of rogue colours, at the expense of losing only $O(st)$ vertices each time. This is roughly the source of the $\log_2 t$ factor in our final bound.

Now we provide the details to make the above outline rigorous. Applying Lemma 2.1 to the subgraph consisting of all rogue edges, we see that the average rogue degree (number of adjacent rogue edges) in $G[U]$ is at most $2s|R| \leqslant 2st$. So, we can delete all vertices in U with rogue degree at least $4st$ at a cost of reducing $|U|$ by at most half. Let us also delete all edges within each U_i for the sake of clarity of presentation. Note that the reduced U still has size at least $180st \log_2 t$. Let σ be a median order for this partially directed graph induced by U. We will use the feedback property to find directed paths (and this is the only property of median orders that we will use).

We wish to apply Lemma 2.3, so let us inductively build a matching of distinct rogue colours, and accumulate a bad set that we call B and which we will maintain and update through the entire proof in this section. Let v_1 be the first vertex according to σ, and start with $B = U_\ell$, where $U_\ell \ni v_1$. Proceed through the rest of the vertices in the order of σ. For the first stage, stop when we first encounter a vertex not in B that is adjacent to a rogue edge (possibly several) whose other endpoint is also not in B, and call the vertex v_2. Arbitrarily select one of those rogue edges adjacent to v_2, call it e_2, and call its colour r_2. Since we deleted all edges inside U_i, e_2 links two distinct U_i and U_j. Add all vertices of U_i and U_j to B. In general, if we have already considered all vertices up to v_k, continue along the median order (starting from the vertex immediately after v_k) until we encounter a vertex not in B that is adjacent to an edge of a new rogue colour which is not in $\{r_2, \ldots, r_k\}$, again with the other endpoint also not in B. Call that vertex v_{k+1}, the edge e_{k+1}, and its colour r_{k+1}. Add to B all the vertices in the two sets U_i which contain the endpoints of e_{k+1}. Repeat this procedure until we have gone through all of the vertices in order. Suppose that this process produces vertices $v_1, v_2, \ldots, v_f$. Then, to simplify the statements of our lemmas, also let $v_{f+1}, v_{f+2}, \ldots, v_{2f}$ refer to the final vertex in the median order. Our goal will be to find directed paths from $\{v_i\}_{i=1}^{f}$, which via Lemma 2.3 will then extend to a rainbow path.

Note that if $|B| \geqslant 2st$, then the number of vertices in $\{v_1\} \cup e_2 \cup \cdots \cup e_f$ is at least t by property (i). Thus, applying Lemma 2.3 with $P_i = \{v_i\}$, we can produce a rainbow path with at least t edges. Therefore, we may assume for the rest of this proof that $|B| < 2st$. Also observe that this argument implies that $f \leqslant t/2$.

The following technical lemma will help us to build the directed paths $\{P_i\}$.

Lemma 2.4. *Let v be a vertex in U, and let B be a set of size at most $2st$. Then, among the $8st$ vertices immediately following v in the median order, there is always some $w \notin B$ such that there is a directed edge from v to w.*

Proof. First, note that since we deleted all vertices with rogue degree at least $4st$, more than $4st$ of the $8st$ vertices immediately after v are connected to v by a directed edge. Since we have a median order, the feedback property implies that only at most half of those edges can be directed back towards v; therefore, there are more than $2st$ vertices there that have a directed edge from v. Since $|B| < 2st$, at least one of these vertices will serve as our w. $\square$

Consider the vertices $v_1, v_2, v_4, \dots, v_{2^{\lfloor \log_2 2f \rfloor}}$. Since we have already established that $f \leqslant t/2$, this is a list of at most $1 + \log_2 t$ vertices, the first and last of which are also the first and last vertices in the median order. Since U still has at least $180st \log_2 t$ vertices, the pigeonhole principle implies that there must be some pair of vertices $\{v_\ell, v_{2\ell}\}$ in that list such that the number of vertices between them in the median order is at least $180st - 2$. Thus, the following lemma will provide the desired contradiction.

Lemma 2.5. *If there is any $1 \leqslant \ell \leqslant f$ such that there are at least $176st$ vertices between v_ℓ and $v_{2\ell}$ in the median order, then G has a rainbow t-edge path.*

Proof. Suppose we have an ℓ that satisfies the conditions of the lemma. Let S_1 be the first $8st$ vertices immediately following v_ℓ in the median order, and let S_2 be the next $168st$ vertices in the median order.

Let us first build for every $i \leqslant \ell$ a directed path P_i from v_i to S_1 by repeatedly applying Lemma 2.4. Start with each such $P_i = \{v_i\}$, and as long as one of those P_i does not reach S_1, apply the lemma to extend it forward to a new vertex w, and add the set U_k containing w to the set of bad vertices B. If at any stage we have $|B| \geqslant 2st$, we can immediately apply Lemma 2.3 to find a rainbow path with at least t edges, just as in the argument directly preceding the statement of Lemma 2.4. So, suppose that does not happen, and let $\{w_i\}_1^\ell \subset S_1$ be the endpoints of these paths. We will show that we can further extend these paths into S_2 by a total amount of at least t, in such a way that we never use two vertices from the same set U_k. This will complete our proof because Lemma 2.3 can link them into a rainbow path with at least t edges.

Recall that all of the sets U_i had size at most $2s$. Therefore, we can partition S_2 into disjoint sets U'_j with $2s \leqslant |U'_j| \leqslant 4s$, where each U'_j is obtained by taking a union of some sets $U_i \cap S_2$. We will design our path extension process such that it uses at most one vertex from each U'_j, and hence it will also intersect each U_k at most once. We use the probabilistic method to accomplish this.

Perform the following randomized algorithm, which will build a collection of sets $\{T_i\}_{i=1}^\ell$. First, activate each U'_j with probability $1/8$. Next, for each activated U'_j, select one of its vertices uniformly at random, and assign it to one of the T_i, again uniformly at random. For each $i \leqslant \ell$, let T'_i be obtained from T_i by deleting every vertex in B, and every vertex that is not pointed to by a directed edge from v_i. Finally, let T''_i be derived from T_i' by (arbitrarily) deleting one vertex from every rogue edge with both endpoints in T_i'. Observe that now each T_i'' spans a tournament, so as we saw at the end of the Introduction, it contains a directed Hamiltonian path P_i'. Since w_i has a directed edge to every vertex in T_i'', this P_i' can be used to extend P_i. Therefore, if we can construct sets T_i'' such that $|T_1''| + \cdots + |T_\ell''| \geqslant t$, we will be done.

Fix an $i \leqslant \ell$, and let us compute $\mathbb{E}[|T_i'|]$. By the feedback property of a median order, the number of (backward) directed edges from S_2 to $\{w_i\}$ is at most half of the number of directed edges between w_i and the vertices in $S_1 \cup S_2$ which follow it in the median order. Since the latter number is bounded by $|S_1 \cup S_2| = 176st$, the number of directed edges from S_2 to w_i is at most $88st$. Also, the number of rogue edges between S_2 and $\{w_i\}$ is at most $4st$ because by construction all rogue degrees are bounded by $4st$. Therefore, the number of (forward) directed edges from w_i to S_2 is at least $168st - 88st - 4st = 76st$. Since we will delete up to $2st$ vertices which were from B, the number of directed edges from w_i to vertices in $S_2 \setminus B$ is at least $74st$. Suppose $\overrightarrow{w_i x}$ is one of these directed edges, and suppose that $x \in U_k'$. The probability that x is selected for T_i is precisely

$$\frac{1}{8 \cdot \ell \cdot |U_k'|} \geqslant \frac{1}{8 \cdot \ell \cdot 4s},$$

and by construction of x, we know that if it is selected for T_i, it will also remain in T_i'. Therefore, by linearity of expectation,

$$\mathbb{E}[|T_i'|] \geqslant 74st \cdot \frac{1}{8 \cdot \ell \cdot 4s} = \frac{37}{16} \frac{t}{\ell}.$$

To bound $\mathbb{E}[|T_i'| - |T_i''|]$, observe that the number of rogue colours in the graph spanned by $S_2 \setminus B$ is less than 2ℓ, by construction of the sequence $\{v_i\}$. Therefore, Lemma 2.1 implies that there are fewer than $2\ell \cdot s \cdot 168st$ rogue edges spanned by $S_2 \setminus B$. Consider one of these rogue edges xy. If we select both of its endpoints for T_i, it will contribute at most 1 (possibly 0) to $|T_i'| - |T_i''|$; otherwise it will contribute 0. Above, we have already explained that the probability that the vertex $x \in U_j'$ is selected for T_i is precisely

$$\frac{1}{8 \cdot \ell \cdot |U_k'|} \leqslant \frac{1}{8 \cdot \ell \cdot 2s}.$$

If x and y come from distinct U_j', then the probabilities that they were both selected for T_i are independent, and otherwise it is impossible that they both were selected. Hence

$$\mathbb{E}[|T_i'| - |T_i''|] \leqslant 2\ell \cdot s \cdot 168st \cdot \left(\frac{1}{8 \cdot \ell \cdot 2s}\right)^2 = \frac{21}{16} \frac{t}{\ell}.$$

Therefore, by linearity of expectation, $\mathbb{E}[|T_i''|] \geqslant t/\ell$, and thus $\mathbb{E}[|T_1''| + \cdots + |T_\ell''|] \geqslant t$. This implies that there exists an instance of our random procedure for which $|T_1''| + \cdots + |T_\ell''| \geqslant t$, so we are done. □

3. Concluding remarks

In our proof, we apply Lemma 2.2 to produce a structured set $U = U_1 \cup \cdots \cup U_m$ of size $\Omega(st \log t)$. The argument in Section 2.2 is quite wasteful because, in particular, Lemma 2.5

attempts to build a collection of directed paths with total length $\geqslant t$, but essentially using only the vertices in the median order between v_ℓ and $v_{2\ell}$. This dissection of the vertex set into dyadic chunks incurs the logarithmic factor in our bound. We believe that with a better argument, one might be able to complete the proof using a structured set $U = U_1 \cup \cdots \cup U_m$ of size only $\Omega(st)$. If this were indeed possible, then Lemma 2.2 would immediately imply that $f(S, P_t) = O(st)$, because one loses only a constant factor in passing from $V(G)$ to U.

It would be very interesting to obtain a better bound on $f(S, T)$ for general trees T. Our approach, based on the median order, seems particularly promising here since it might be combined with the following result of Havet and Thomassé [12] on Sumner's conjecture: every tournament of order $4n$ contains every directed tree of order n as a subgraph.

Appendix (by Oleg Pikhurko and Benny Sudakov)

Consider the following variant of the constrained Ramsey number. Let $g(S, T)$ be the minimum integer n such that every colouring of the edges of the complete graph K_n contains either a monochromatic copy of S or a properly coloured copy of T. (In contrast, recall that the definition of $f(S, T)$ requires T to be rainbow.) As for constrained Ramsey numbers, it is easy to see that $g(S, T)$ exists (*i.e.*, it is finite) if and only if S is a star or T is acyclic. Although there has been little success bounding $f(S, T)$ by $O(st)$, it turns out that we can prove a quadratic upper bound for $g(S, T)$, which is of course no larger than $f(S, T)$.

Theorem A.1. *Let S and T be two trees with s and t edges, respectively. Then $g(S, T) \leqslant 2st + t^2$.*

The following construction shows that the upper bound is tight up to a constant factor. Let S be a path with $s+1$ edges and let T be a star with $t+1$ edges. Then let $V_1, \ldots, V_t$ be disjoint sets of size $\lfloor s/2 \rfloor$ each. Colour all edges inside V_i and from V_i to V_j with $j > i$ by colour i. This produces a graph on $t\lfloor s/2 \rfloor$ vertices with no monochromatic S and no properly coloured T.

To prove Theorem A.1, we first need the following lemma.

Lemma A.2. *Consider an edge colouring of the complete graph which contains no monochromatic copy of a fixed tree S with s edges. Let U be the set of vertices such that, for every $u \in U$, one can delete at most a edges from the graph such that the remaining edges which connect u to the rest of the graph have at most b colours. Then $|U| \leqslant 2(bs + a)$.*

Proof. Focus on the subgraph induced by U. Now we can delete at most a edges at every vertex so that the remaining edges at that vertex have at most b colours. Let G be the graph obtained after all of these deletions. If $m = |U|$, then the number of edges of G is at least $\binom{m}{2} - am$. For every remaining colour c, let G_c be the subgraph of all edges of colour c. By Lemma 2.1, we have $e(G_c) < s \cdot v(G_c)$ for each c. Also, since every vertex of G is incident with edges of at most b colours, we have that $\sum_c v(G_c) \leqslant bm$. Combining all these inequalities, we have

$$\binom{m}{2} - am \leqslant \sum_c e(G_c) < \sum_c s \cdot v(G_c) \leqslant sbm.$$

This implies that $m < 2(bs + a) + 1$. □

Proof of Theorem A.1. The proof is by induction on t. The statement is trivial for $t = 1$ because any edge will give us a properly coloured T. Now suppose that T is a tree with $t > 1$ edges, and we have a colouring of $G = K_{2st+t^2}$ with no monochromatic copy of S. It suffices to show that we can find a properly coloured copy of T. Select an edge (u, v) of T such that all neighbours of v except u are leaves $v_1, \dots, v_k$. Delete $v_1, \dots, v_k$ from T and call the new tree T_1. The number of edges in T_1 is $t_1 = t - k$.

Let U be the set of vertices of G such that, for every $u \in U$, one can delete at most t_1 edges from G such that the edges which connect u to the rest of the graph have at most k colours. By the previous lemma $|U| \leqslant 2(ks + t_1)$, and let $W = V(G) \setminus U$. Then we have that

$$|W| = 2st + t^2 - |U| \geqslant 2st_1 + t^2 - 2t_1 > 2st_1 + t_1^2.$$

Therefore, by induction we can find a properly coloured copy of the tree T_1 inside W. Let u', v' be the images in this copy of the vertices of u, v of T_1. By definition of W, the vertex v' has edges of at least $k + 1$ colours connecting it with vertices outside this copy of T_1. At least k of these colours are different from that of the edge (u', v'), so we can extend the tree to a properly coloured copy of T.

Using a more careful analysis in the above proof, which we omit, one can slightly improve the term t^2 in Theorem A.1.

References

[1] Alon, N., Jiang, T., Miller, Z. and Pritikin, D. (2003) Properly colored subgraphs and rainbow subgraphs in edge-colorings with local constraints. *Random Struct. Alg.* **23** 409–433.

[2] Balister, P., Gyárfás, A. Lehel, J. and Schelp, R. (2006) Mono-multi bipartite Ramsey numbers, designs, and matrices, *J. Combin. Theory Ser. A* **113** 101–112.

[3] Bialostocki, A. and Voxman, W. (2003) On monochromatic-rainbow generalizations of two Ramsey type theorems. *Ars Combinatoria* **68** 131–142.

[4] Chen, G., Schelp, R. and Wei, B. (2001) Monochromatic-rainbow Ramsey numbers. Unpublished work in *14th Cumberland Conference Abstracts*.

[5] Dean, N. and Latka, B. (1995) Squaring the tournament: An open problem. *Congr. Numer.* **109** 73–80.

[6] Erdős, P. and Rado, R. (1950) A combinatorial theorem. *J. London Math. Soc.* **25** 249–255.

[7] Eroh, L. (2004) Constrained Ramsey numbers of matchings. *J. Combin. Math. Combin. Comput.* **51** 175–190.

[8] Eroh, L. (2004) Rainbow Ramsey numbers of stars and matchings. *Bull. Inst. Combin. Appl.* **40** 91–99.

[9] Gyárfás, A., Lehel, J., Nešetřil, J., Rödl, V., Schelp, R. and Tuza, Z. (1987) Local k-colorings of graphs and hypergraphs. *J. Combin. Theory Ser. B* **43** 127–139.

[10] Gyárfás, A., Lehel, J. and Schelp, R. (2007) Finding a monochromatic subgraph or a rainbow path. *J. Graph Theory* **54** 1–12.
[11] Gyárfás, A., Lehel, J., Schelp, R. and Tuza, Z. (1987) Ramsey numbers for local colorings. *Graphs Combin.* **3** 267–277.
[12] Havet, F. and Thomassé, S. (2000) Median orders of tournaments: A tool for the second neighborhood problem and Sumner's conjecture. *J. Graph Theory* **35** 244–256.
[13] Jamison, R., Jiang, T. and Ling, A. C. H. (2003) Constrained Ramsey numbers of graphs. *J. Graph Theory* **42** 1–16.
[14] Jamison, R. and West, D. (2004) On pattern Ramsey numbers of graphs. *Graphs Combin.* **20** 333–339.
[15] Schelp, R. (1997) Local and mean k-Ramsey numbers for complete graphs. *J. Graph Theory* **24** 201–203.
[16] Thomason, A. and Wagner, P. (2007) Complete graphs with no rainbow path, *J. Graph Theory* **54** 261–266.
[17] Truszczyński, M. and Tuza, Z. (1987) Linear upper bounds for local Ramsey numbers. *Graphs Combin.* **3** 67–73.
[18] Wagner, P. (2006) An upper bound for constrained Ramsey numbers. *Combin. Probab. Comput.* **15** 619–626.
[19] Wormald, N. (1983) Subtrees of large tournaments. In *Combinatorial Mathematics X*, Vol. 1036 of *Lecture Notes in Mathematics*, Springer, pp. 417–419.