Size-Ramsey numbers of powers of hypergraph paths

Shoham Letzter* Alexey Pokrovskiy † Liana Yepremyan‡

Abstract

The *s-colour size-Ramsey number* of a hypergraph H is the minimum number of edges in a hypergraph G whose every s-edge-colouring contains a monochromatic copy of H. We show that the s-colour size-Ramsey number of the r-uniform tight path on n vertices is linear in n, for every fixed r, s, thereby answering a question of Dudek, La Fleur, Mubayi and Rödl (2017).

1 Introduction

For two hypergraphs $\mathcal{G}$ and $\mathcal{H}$ and an integer $s \geq 2$ write $\mathcal{G} \xrightarrow{s} \mathcal{H}$ if in every s-edge-colouring of $\mathcal{G}$ there is a monochromatic copy of $\mathcal{H}$. The s-colour *size-Ramsey number* of a hypergraph $\mathcal{H}$, denoted $\hat{r}_s(\mathcal{H})$, is the minimum number of edges in a hypergraph $\mathcal{G}$ satisfying $\mathcal{G} \xrightarrow{s} \mathcal{H}$. Namely,

$$\hat{r}_s(\mathcal{H}) = \min\{e(\mathcal{G}) : \mathcal{G} \xrightarrow{s} \mathcal{H}\}.$$

When $s = 2$, we often omit the subscript 2, and refer to the 2-colour size-Ramsey number of $\mathcal{H}$ as, simply, the *size-Ramsey number* of $\mathcal{H}$.

The notion of size-Ramsey numbers of graphs was introduced by Erdős, Faudree, Rousseau and Schelp [11] in 1978. It is an interesting and well-studied variant of the classical s-colour *Ramsey number* of $\mathcal{H}$, denoted $r_s(\mathcal{H})$, which can be defined as

$$r_s(\mathcal{H}) = \min\{|\mathcal{G}| : \mathcal{G} \xrightarrow{s} \mathcal{H}\},$$

i.e. it is the minimum number of vertices in a hypergraph $\mathcal{G}$ such that $\mathcal{G} \xrightarrow{s} \mathcal{H}$. The study of Ramsey numbers, especially for $s = 2$, is one of the most central and well-studied topics in combinatorics. In this paper we study size-Ramsey numbers of graphs and hypergraphs, an active field of research in recent years. One of the earliest results regarding size-Ramsey numbers of graphs, obtained by Beck [2] in 1983, asserts that the size-Ramsey number of a path is linear in its length; more precisely, Beck showed that $\hat{r}(P_n) \leq 900n$ for large n. The problem of determining the size-Ramsey number of a path has been the subject of many papers [1, 2, 4, 5, 8, 9, 17], and the currently best-known bounds are as follows:

$$(3.75 + o(1))n \leq \hat{r}(P_n) \leq 74n,$$

*Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK. Email: s.letzter@ucl.ac.uk. Research supported by the Royal Society.

†Department of Mathematics, University College London, Gower Street, London WC1E 6BT, UK. Email: a.pokrovskiy@ucl.ac.uk.

‡Deparment of Mathematics, London School of Economics, London WC2A 2AE, UK. Email: l.yepremyan@lse.ac.uk. Research supported by Marie Sklodowska Curie Global Fellowship, H2020-MSCA-IF-2018:846304.

where the lower bound is due to Bal and DeBiasio [1] and the upper bound is by Dudek and Prałat [9]. One can easily generalise Beck's arguments to the multicolour setting, showing that $\hat{r}_s(P_n) = O_s(P_n)$. This multicolour variant has received its own fair share of attention [1, 9, 10, 16]; and currently the best-known bounds are as follows, where the lower bound is due to Dudek and Prałat [9] and the upper bound is by Krivelevich [16].

$$\hat{r}_s(P_n) = \Omega(s^2 n), \qquad \hat{r}_s(P_n) = O(s^2 \log s \cdot n).$$

In contrast, the study of size-Ramsey numbers of hypergraphs was initiated only recently (in 2017) by Dudek, La Fleur, Mubayi and Rödl [7]. One of the first problems proposed in their paper is the following generalisation of Beck's result [2] regarding size-Ramsey numbers of paths. The *r-uniform tight path* on n vertices, denoted $P_n^{(r)}$, is the r-uniform hypergraph on vertex set $[n]$ whose edges are all sets of r consecutive elements in $[n]$. Observe that $P_n^{(2)}$ is the path P_n on n vertices. Dudek, La Fleur, Mubayi and Rödl [7] asked if, similarly to the graph case, $\hat{r}\big(P_n^{(r)}\big) = O(n)$. This was answered affirmatively for $r = 3$ by Han, Kohayakawa, Letzter, Mota and Parczyk [13]. For $r \geq 4$, the best-known bound prior to our work was $\hat{r}(P_n^{(r)}) = O\big((n \log n)^{r/2}\big)$, due to Lu and Wang [19]. In our main result we settle the problem, and, in fact, prove a stronger statement where the number of colours is arbitrary (rather than 2).

Theorem 1.1. *Fix integers $r, s \geq 1$. Then $\hat{r}_s\big(P_n^{(r)}\big) = O(n)$.*

Our methods are also powerful enough to show that Ramsey numbers of other hypergraphs are linear; these hypergraphs include powers of tight paths, tight hypergraph trees and their powers, and tight subdivisions of small hypergraphs. The proofs for these more general results are significantly more technical (though they follow exactly the same proof approach as Theorem 1.1) and can be found in the unpublished manuscript [18].

2 Proof overview

In a graph G, denote by $d_G(x, y)$ the distance between x and y in G, i.e. the length of the shortest path from x to y.

Definition 2.1 (Powers)**.** *For a graph G, let G^c denote the graph formed by connecting every pair of vertices with $d_G(x, y) \leq c$. We define $\mathcal{G}^c_{(r)}$ to be the r-uniform hypergraph on $V(G)$ whose hyperedges are all r-sets of vertices that are within distance at most c from each other in G (i.e. $E(\mathcal{G}^c_{(r)}) = \{\{v_1, \ldots, v_r\} \subseteq V(G) : d_G(v_i, v_j) \leq c$ for all $i, j \in [r]\}$).*

Throughout the paper r will be a fixed integer which is at least 3, so we abbreviate $\mathcal{G}^c_{(r)}$ as $\mathcal{G}^c$. To give a linear upper bound on $\hat{r}_s\big(P_n^{(r)}\big)$ we need to come up with a hypergraph G with $O(n)$ edges that satisfies $G \xrightarrow{s} P_n^{(r)}$. The host graph we use is not new — we take a bounded degree expanding graph G and show that for some large constant p, we have $\mathcal{G}^p \xrightarrow{s} P_n^{(r)}$. This idea dates back to [6], and has been used in many papers since [12, 13, 14, 15]. There is really only one property of G that we use, namely that

$$G^c \longrightarrow \Big(\overbrace{P_{n/c}, \ldots, P_{n/c}}^{t}, K_d\Big). \tag{1}$$

(This means that for every $(t+1)$-colouring of G^c there is either a monochromatic $P_{n/c}$ in one of the first s colours, or a monochromatic K_d in the last colour; here c is sufficiently large in terms of d and t). This is proved in Lemma 3.8 — though again this is not really a new result and similar lemmas have underpinned a lot of recent work about size-Ramsey numbers. The basic structure of our proof is the following.

We consider an s-colouring of $\mathcal{G}^p$, and suppose for contradiction that there are no monochromatic tight paths of length $\Omega(n)$. To each vertex $v \in V(G)$ we associate a set $T(v)$, denoting $F = \{T(v) : v \in V(G)\}$. Then we repeat the following procedure:

- ❖ We consider a carefully defined auxiliary colouring φ of the graph G^c, based on F. By (1) (with t an appropriate function of s and r), there are two possibilities:
  - (a) If φ has a monochromatic path of length $\Omega(n)$ in one of the first t colours, then this will produce a tight path of length $\Omega(n)$ in $\mathcal{G}^p$ (contradicting our initial assumption).
  - (b) If instead φ has many disjoint monochromatic K_d's in the last colour, then we use their presence to deduce that the original colouring had a certain amount of "structure" to it. We use this structure to define a new family F', and return to step 1 with this family F'.

By repeating this process, we gradually find more and more "structure" in our original colouring. This eventually leads to a contradiction, because we also prove that there is an absolute limit to how much "structure" a colouring can have.

The big thing missing from the above overview is an indication of what "structure" we find in our colourings. The definition of this is new and quite complicated — in fact, almost all the details of this paper are there to formally define this structure and prove lemmas to work with it. In the remainder of this section we give some informal indication of how things work.

First, the sets $T(v)$ we associate with each vertex v above are not just sets, but rather *trees* (in the graph G^k for suitable k). Given a set S of leaves of such a tree we define type(S) to be the isomorphism class of the subtree of $T(v)$ whose leaves are S. The auxiliary colouring of G^k in step 1 roughly joins $x, y \in V(G)$ by an edge whenever there exist two $(r-1)$-sets $A, B \subseteq V(T(v))$ of the same type and a short monochromatic tight path P_{xy} between them, and colour xy by (τ, c) where τ is the type of A and c is the colour of P_{xy}. It can be shown that long paths in this colouring of G^k can be used to build long tight paths in $\mathcal{G}^p$ (though this is far from obvious, and requires a lot of additional machinery based around concatenating short tight paths). Thus, if step (a) fails, then we obtain disjoint sets $\{T(v_1), \ldots, T(v_d)\}$ with the property that there is no short monochromatic tight path from a subset $A \subseteq V(T(v_i))$ to a subset $B \subseteq V(T(v_j))$, for any A, B of the same type and any $i \neq j$. This allows us to find a new family $F' = \{T'(v) : v \in V(G)\}$, where the sets $T'(v)$ are trees which are taller than the trees $T(v)$, and with the following property: "for any x, y with $d_G(x, y) \leq k$, there is no short tight path between any pair of subsets of $F'(x)$ and $F'(y)$ of the same type". We call such a colouring "k-disconnected" — and this is exactly the "structure" we build throughout the proof. One of our lemmas (Lemma 4.25) says that there do not exist k-disconnected families of trees of arbitrarily large height — this is what we meant by "there is an absolute limit to how much structure a colouring can have", and this is what gives us our final contradiction.

The structure of the paper is as follows. In the next section we define expanders and prove that they satisfy (1). In Section 4.1 we introduce tree families F, and give a formal definition of what we mean by "type". In Section 4.2, we prove some Ramsey-type lemmas about tree families (which are needed for concatenating short monochromatic tight paths later in the proof, and also for proving Lemma 4.25). In Section 4.3 we formally define k-disconnectedness, and prove that there do not exist k-disconnected families of trees of arbitrary height (Lemma 4.25). In Section 4.4, we define an operation called "forest augmentation" which is how the forest F changes between different iterations of ❖. In Section 4.5 we introduce and study things called "versatile sets of vertices", which are another tool we need for concatenating short tight paths. In Section 5 we prove Theorem 1.1.

3 Ramsey properties of expanders

The Ramsey graphs we construct are based around expanders.

Definition 3.1 (Expander). *A graph G on n vertices is an ε-*expander *if for any two sets of vertices A, B of size at least εn, there is an edge starting in A and ending B.*

It is a standard result that there exist arbitrarily large ε-expanders of bounded degree (e.g. $G(n, C/n)$ satisfies this after deleting high degree vertices. See [18], Proposition 3.2).

Lemma 3.2. *For all $\varepsilon > 0$ and large n, there exist connected ε-expanders G on n vertices with $\Delta(G) \le \varepsilon^{-2}$.*

Our goal in this section is to prove that powers of expanders have a certain Ramsey property (Lemma 3.9). We will need the following three well-known bipartite Ramsey results. Here the notation $G \longrightarrow (H_1, \dots, H_k)$ means that in every k-edge-colouring of G there is a copy of H_i in colour i, for some $i \in [k]$.

Lemma 3.3. *$K_{3n,3n} \longrightarrow (P_n, K_{n,n})$.*

Proof. Let $K_{3n,3n}$ be red/blue coloured. Let the parts of $K_{3n,3n}$ be X, Y. It is known that one can partition the vertices of a graph into a path P and two disjoint sets of vertices A, B with $|A| = |B|$, such that there are no edges between A and B (see [3]). Apply this to the subgraph of red edges to get a red path P and corresponding sets A, B. If $|P| \ge n$ we are done, so suppose that $|P| < n$. Since $|A| = |B|$, the number of vertices outside P is even and so $|P|$ is even. Since vertices of P alternate between X and Y, this tells us that $|V(P) \cap X| = |V(P) \cap Y| \le n$, or, equivalently, $|A \cap X| + |B \cap X| = |A \cap Y| + |B \cap Y| \ge 2n$. Without loss of generality, $|A \cap X| \ge |B \cap X|$, which gives $|A \cap X| \ge n$. Adding $|A \cap X| + |A \cap Y| = |A| = |B| = |B \cap X| + |B \cap Y|$ to $|A \cap X| + |B \cap X| = |A \cap Y| + |B \cap Y|$, we get $|A \cap X| = |B \cap Y|$. Now, we have two sets $A' = A \cap X$ and $B' = B \cap Y$ of size at least n with no red edges between them (since $A' \subseteq A, B' \subseteq B$), and also no non-edges between them (since $A' \subseteq X, B' \subseteq Y$) i.e. they form a blue complete bipartite graph $K_{n,n}$. □

Lemma 3.4. *$K_{3^r n, 3^r n} \longrightarrow \Big(\overbrace{P_n, \dots, P_n}^{r \text{ times}}, K_{n,n}\Big)$ for every integer $r \ge 1$.*

Proof. We proceed by induction on r. The initial case $r = 1$ is exactly Lemma 3.3. Consider an $(r+1)$-coloured $K_{3^r n, 3^r n}$. By Lemma 3.3, either there is a P_n coloured 1 or a $K_{3^{r-1}n, 3^{r-1}n}$ coloured by $\{2, \dots, r+1\}$. In the former case, we are done immediately. In the latter case, we are done by induction. □

We write K_n^s for the complete s-partite graph with parts of size n.

Lemma 3.5. *$K_{6^{rs}} \to \Big(\overbrace{P_n, \dots, P_n}^{r \text{ times}}, K_n^s\Big)$ for every integers $r \ge 1$ and $s \ge 2$.*

Proof. Fix $r \geq 1$ and write $a_s = (2 \cdot 3^r)^{s-1}$. We will prove that the following holds for $s \geq 2$.

$$K_{a_s n, a_s n} \longrightarrow \Big(\overbrace{P_n, \ldots, P_n}^{r \text{ times}}, K_n^s\Big).$$

Notice that this will prove the lemma as $6^{rs} \geq a_s$ for every $s \geq 2$. We proceed by induction on s. The initial case $s = 2$ follows from Lemma 3.4 (because $a_2 = 2 \cdot 3^r$ and so $K_{a_2 n}$ contains a copy of $K_{3^r n, 3^r n}$). Suppose then that $s \geq 3$, and that the lemma holds for smaller s. Consider an $(r+1)$-coloured $K_{a_s n}$, and suppose that there is no monochromatic P_n with a colour in $[r]$. By Lemma 3.4, there is a copy of $K_{a_{s-1}n, a_{s-1}n}$ coloured $r+1$ (because $a_s = 2 \cdot 3^r \cdot a_{s-1}$). Let A, B be the parts of this $K_{a_{s-1}n, a_{s-1}n}$. By induction applied to $K_n[A]$ and $K_n[B]$, the parts A and B each either contain a monochromatic P_n with a colour in $[r]$, in which case we are done, or a K_n^{s-1} coloured $r+1$. The edges between these K_n^{s-1}'s are all colour $r+1$, giving a K_n^{2s-2} in this colour (and this contains a colour $r+1$ copy of K_n^s, since $2s - 2 \geq s$ for $s \geq 3$). $\square$

The following is a basic property of expanders.

Lemma 3.6. *Let $S_1, \ldots, S_r$ be disjoint sets of size at least $r\varepsilon n$ in an ε-expander G on n vertices. Then there is a path $v_1 \ldots v_r$ in G with $v_i \in S_i$.*

Proof. Notice that between any two disjoint sets S and T in an ε-expander there is a matching of size at least $\min\{|S|, |T|\} - \varepsilon n$ (if M is a maximal S-T matching, then we must have $|S \setminus V(M)| = |T \setminus V(M)| \leq \varepsilon n$ as otherwise ε-expansion would give another S-T edge from $S \setminus V(M)$ to $T \setminus V(M)$ disjoint from M). Set M_1 to be a maximal S_1-S_2 matching and write $m = r\varepsilon n$; so $|M_1| \geq m - \varepsilon n$. Build matchings $M_2, \ldots, M_{r-1}$ one by one by taking M_i to be a maximal $(S_i \cap V(M_{i-1}))$-S_{i+1} matching. By the first sentence of the proof (applied with $S = S_i \cap V(M_{i-1})$ and $T = S_{i+1}$), we have

$$|M_i| \geq \min\{|S_i \cap V(M_{i-1})|, |S_{i+1}|\} - \varepsilon n \geq \min\{|M_{i-1}|, m\} - \varepsilon n \geq m - i\varepsilon n,$$

where the last inequality follows from iterating the inequality $|M_i| \geq \min\{|M_{i-1}|, m\} - \varepsilon n$ and using $|M_1| \geq m - \varepsilon n$. In particular, $|M_{r-1}| \geq m - (r-1)\varepsilon n > 0$, i.e. $M_{r-1} \neq \emptyset$. Also, by construction, each edge of M_i touches an edge of M_{i-1}, for every $i \in [r-1]$. These two facts together give a sequence $v_1 \ldots v_r$ with $v_{i-1}v_i \in M_i$. This is thus a path satisfying the lemma. $\square$

We use $N_G(X)$ to denote the neighbourhood of X, namely the set of vertices outside of X with an edge in G into X, and omit the subscript G when it is clear from context.

Lemma 3.7. *Let G be a 0.01-expander of order n. Then there is subgraph H of order at least $0.97n$ with H^k a $\frac{1}{5 \cdot 2^{k/2-2}}$-expander.*

Proof. Set $r = \lfloor k/2 - 1 \rfloor$ and $\varepsilon = \frac{1}{5 \cdot 2^{k/2-2}} \geq \frac{1}{5 \cdot 2^r}$. Let $X \subseteq V(G)$ be a maximal subset with $|X| \leq n/4$ and $|N(X)| \leq 2|X|$. Then $|X| \leq n/100$, since otherwise $|X \cup N(X)| \geq 99n/100$ (if not, we would have $|V(G) \setminus (X \cup N(X))|, |X| \geq 0.01n$. Then the definition of 0.01-expander applies to give an edge from X to $V(G) \setminus (X \cup N(X))$, which is impossible) so $|N(X)| \geq 99n/100 - |X| \geq 74n/100 > 2|X|$, a contradiction. Set $H := G \setminus (X \cup N(X))$ (i.e. G with the vertices in $X \cup N(X)$ deleted) to get a graph of order at least $n - 3|X| \geq 97n/100$. Note that all $Y \subseteq V(H)$ with $|Y| \leq n/5$ have $|N_H(Y)| > 2|Y|$ (since otherwise $X \cup Y$ is a set with $|N(X \cup Y)| \leq 2|X \cup Y|$ and $|X \cup Y| \leq n/100 + n/5 \leq n/4$, contradicting maximality of X).

Now consider $A, B \subseteq V(H^r) = V(H)$ of size εn. Write A_i for the set of vertices at distance at most i from A in H, and define B_i analogously with respect to B. Then $|A_i| \geq |N(A_{i-1})| \geq$

$\min\{2|A_{i-1}|, n/5\}$ for all i, giving $|A_r| \geq \min(2^r \varepsilon n, n/5) = n/5$. Similarly, $|B_r| \geq n/5$. Using 0.01-expansion, this gives an edge between A_r and B_r in G (and hence also in H) i.e. $d_H(A, B) \leq 2r+1$. Thus there is an A-B edge in $H^{2r+1} \subseteq H^k$, as required. $\square$

The following is a Ramsey property of expanders.

Lemma 3.8. *Let $c \gg d$ ($c \geq 10 \cdot 6^{d^2}$ works). Let G be a 0.01-expander on n vertices. Then*

$$G^c \to \Big(\overbrace{P_{n/c}, \ldots, P_{n/c}}^{d \text{ times}}, K_d\Big).$$

Proof. Write $\varepsilon = \frac{1}{5 \cdot 2^{c/2-2}}$. Pick c so that $0.97/6^{d^2} \geq \max\{d\varepsilon, 1/c\}$. Let G^c be $(d+1)$-coloured. Apply Lemma 3.7 to find a subgraph $H \subseteq G$ of order $0.97n$ with H^c an ε-expander. Think of this as an $(d+2)$-colouring of $K_{0.97n}$ with non-edges of H^c having colour 0. Apply Lemma 3.5 with $r = s = d$, thinking of $\{0, d+1\}$ as a single colour. This either gives a monochromatic path of length $m := 0.97n/6^{d^2}$ coloured by $[d]$ or a $\{0, d+1\}$-coloured K_m^d. In the former case, we are done (since $m \geq n/c$). In the latter case, let $S_1, \ldots, S_d$ be the parts of the $\{0, d+1\}$-coloured K_m^d. Using $m \geq d\varepsilon n$, apply Lemma 3.6 to get a path $v_1 \ldots v_d$ in H with $v_i \in S_i$. Then $H^c[\{v_1, \ldots, v_d\}]$ is a K_d coloured $d+1$ (it is complete because the vertices are within distance $d \leq c$ in H. It is colour $d+1$ because all its edges go between different parts S_i, S_j, and the only colour permitted there is $d+1$). $\square$

The following is the main result of this section.

Lemma 3.9 (Ramsey of expander powers)**.** *Let $c \gg d, \varepsilon^{-1}$, and $d \geq s$. Let G be an ε-expander on n vertices. Then for every $U \subseteq V(G)$ with $|U| \geq n/2$, we have*

$$G[U]^c \to \Big(\overbrace{P_{n/c}, \ldots, P_{n/c}}^{s \text{ times}}, \overbrace{K_d \,\dot\cup\, \ldots \,\dot\cup\, K_d}^{\text{covering at least } (1-200\varepsilon)|U| \text{ vertices}}\Big).$$

Proof. Without loss of generality, $s = d$ (if $s < d$, just think of any s-colouring as a d-colouring, with some of the colours not appearing). Let $c = c'/(100\varepsilon)$ where c' is the constant from Lemma 3.8. Consider a maximum collection of pairwise vertex-disjoint K_d's in U coloured $s+1$, which we denote by $K^1, \ldots, K^m$. Suppose, towards contradiction, that $m \leq (1-200\varepsilon)|U|/d$, and define $H = G[U] \setminus V(K^1 \cup \cdots \cup K^m)$ to get an induced subgraph H of size at least $200\varepsilon|U| \geq 100\varepsilon n$, noting that H is a 0.01-expander. By Lemma 3.8 we either get a colour 1 copy of K_d (contradicting the maximality of our collection), or we get a monochromatic path in one of the other colours of order at least $100\varepsilon n/c' = n/c$, as required. $\square$

4 Ordered forests

4.1 Basic definitions

In the paper we work exclusively with trees that are both rooted and ordered.

Here "rooted" means that the tree has a designated vertex called the *root* and all edges are directed away from the root. A *rooted forest* is a forest whose components are rooted trees. A *leaf* in a rooted forest is defined as a vertex with out-degree 0. The *level* of a vertex is the distance it has to the root in its component. The *height* of a forest is the maximum level of a vertex in the forest. We emphasise that the root of a tree of height at least 1 is never a leaf (even if it has degree 1), whereas in a height 0 tree the single vertex it contains is both a root and a leaf. A forest is *balanced* if all its leaves are on the same level. We remark that, aside from the beginning of Section 4.5, all forests everywhere will be balanced. The out-neighbours of a vertex v are called *children of* v, and the in-neighbour is called the *parent of* v. Vertices to which there is a directed path from v are called *descendants of* v, while vertices from which there is a directed path to v are called *ancestors of* v. We remark that v is both a descendent and ancestor of itself. We call two vertices *comparable* if one is a descendent of the other. A d-ary forest is a rooted forest where all non-leaf vertices have out-degree d (and in-degree 1, aside from the root which has in-degree 0).

Ordered trees are defined as follows.

Definition 4.1 (Ordered trees, forests, tree families)**.** *An* ordered forest *is a rooted forest F (i.e. one in which every tree of F is rooted), together with an ordering of $V(F)$ such that*

(F1) *If v' is a descendant of v then $v < v'$.*

(F2) *For every incomparable u, v (namely, u is neither an ancestor nor a predecessor of v), with $u < v$, all descendants u' of u and v' of v have $u' < v'$.*

Throughout the paper all trees/forests will be ordered and so we sometimes suppress the word "ordered" for clarity. The *root set* of F, denoted $\mathrm{root}(F)$, is the (ordered) set of roots of F. We say "F is rooted on V" if $\mathrm{root}(F) = V$. For ordered forests F, F', we write $F' \leq F$ if F' is an ordered subforest of F with the same height and root set.

Notation 4.2. *There is a natural way of identifying height h ordered forests with subsets of $\mathbb{N}^i$: we identify vertices on level i with vectors of the form $(x_1, \dots, x_i, 0, \dots, 0)$ and do this level by level. If a vertex $v = (x_1, \dots, x_i, 0, \dots, 0)$ has children $v_1, \dots, v_k$ ordered $v_1 < \dots < v_k$, then we identify v_i with $(x_1, \dots, x_i, i, \dots, 0)$ (see Figure 1). Note that the ordering of the vertices is exactly the lexicographic ordering on corresponding vectors (i.e. $(a_1, \dots, a_h) < (b_1, \dots, b_h)$ if the smallest coordinate with $a_t \neq b_t$ has $a_t < b_t$).*

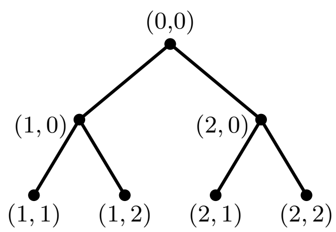

Figure 1: A height 3 ordered tree with the natural labelling in $\mathbb{N}^3$

Definition 4.3 ($L(F)$ and $A^F(S)$)**.** *Given an ordered forest F, we write $L(F)$ for the (ordered) set of leaves of F (i.e. vertices with out-degree 0).*

Given a set of vertices $S \subseteq V(F)$, we let $A^F(S)$ be the induced subforest of F consisting of all ancestors of vertices of S (see Figure 2).

Most of the time it will be clear what F is from context and we abbreviate $A(S) = A^F(S)$. It is useful to note that the above two definitions are inverses to each other, in the sense that for any set of leaves S in a forest F we have $L\left(A^F(S)\right) = S$ and for any $F' \leq F$, we have $A^F(L(F')) = F'$ (for this we use that F and F' have the same root set).

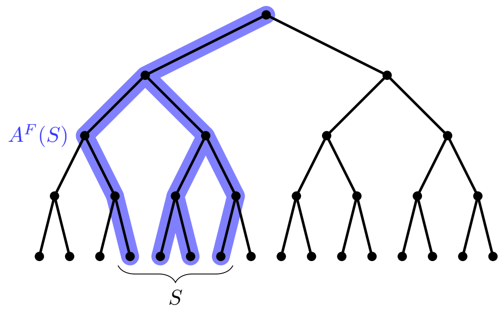

Figure 2: A set of leaves S and the corresponding $A^F(S)$

It is also useful to note that $A^F(A^F(S)) = A^F(S)$ (since "is an ancestor of" is a transitive relation). A consequence of this is that every root of $A^F(S)$ is a root of F (let x be a root of $A^F(S)$. If it is not a root of F, then it has an ancestor y in $F \setminus A^F(S)$. But then y is in $A^F(A^F(S))$, but not in $A^F(S)$).

We use the following notions of homomorphism/monomorphism/isomorphism.

Definition 4.4 (Homomorphisms, monomorphisms and isomorphisms)**.** *A homomorphism φ of ordered forests, from F to F', denoted $\varphi : F \to F'$, is a function from $V(F)$ to $V(F')$, which maps edges to edges and preserves order (i.e. for every $u, v \in V(F)$, if $uv \in E(F)$, then $\varphi(u)\varphi(v) \in E(F')$, and if $u < v$, then $\varphi(u) < \varphi(v)$). An injective homomorphism is called a* monomorphism *and a bijective homomorphism is called an* isomorphism*.*

A *copy* of F in F' is an image of F in F' under some monomorphism $\varphi : F \to F'$.

The following two lemmas show that levels/ancestors are preserved by monomorphisms.

Lemma 4.5. *Let F, F' be two balanced forests of the same height, and $\varphi : F \to F'$ a monomorphism. For any $x \in V(F)$, the vertex x is in the same level in F as $\varphi(x)$ is in F'.*

Proof. Denote the height of F and F' by h. Notice that an edge uv in F, with u being the parent of v, is mapped to an edge $\varphi(u)\varphi(v)$, with $\varphi(u)$ being the parent of $\varphi(v)$, due to φ being order-preserving. Thus, if $v_0 \dots v_h$ is a path in F with $\text{level}(v_i) = i$, then $\varphi(v_0) \dots \varphi(v_h)$ is a path in F' with $\varphi(v_{i-1})$ being the parent of $\varphi(v_i)$. Due to F and F' both being height h trees, this implies that $\text{level}(\varphi(v_i)) = i = \text{level}(v_i)$. Since every vertex in F is in a path $v_0 \dots v_h$ with v_i having level i, it follows that $\text{level}(\varphi(v)) = \text{level}(v)$ for every vertex v in F. $\square$

Lemma 4.6. *Let $\varphi : F \to F'$ be a monomorphism between two balanced forests of the same height. For every $X \subseteq V(F)$, we have $\varphi(A^F(X)) = A^{F'}(\varphi(X))$.*

Proof. We claim that "x is an ancestor of y" if and only if "$\varphi(x)$ is an ancestor of $\varphi(y)$". Indeed, if x an ancestor of y then there is a path $x = v_0 \dots v_k = y$ where v_{i-1} is the parent of v_i, for $i \in [k]$. But then $\varphi(x) = \varphi(v_0) \dots \varphi(v_k) = y$ is a path in F' where $\varphi(v_{i-1})$ is the parent of $\varphi(v_i)$, showing that $\varphi(x)$ is an ancestor of $\varphi(y)$. This shows the "if" direction. For the "only if" part, since by Lemma 4.5 we have $\text{level}(x) = \text{level}(\varphi(x))$, the number of ancestors of x in F is the same as the number of ancestors of $\varphi(x)$ in F' (and both equal $\text{level}(x) + 1$). By definition of A^F, we have

$$\begin{aligned} V\left(A^{F'}(\varphi(X))\right) &= \{y \in V(F') : y \text{ is an ancestor of } \varphi(x) \text{ for some } x \in X\} \\ &= \{\varphi(y) : y \in V(F) \text{ and } \varphi(y) \text{ is an ancestor of } \varphi(x) \text{ for some } x \in X\} \\ &= \varphi\left(\{y \in V(F) : y \text{ is an ancestor of some } x \in X\}\right) = V\left(\varphi\left(A^F(X)\right)\right), \end{aligned}$$

completing the proof. $\square$

We can now introduce the crucial definition of the "type" of a subset.

Definition 4.7 (Type)**.** *For a forest F and subset $e \subseteq L(F)$, the* type *of e in F, denoted* $\mathrm{type}_F(e)$*, is the isomorphism class of $A^F(e)$ (as an ordered forest).*

Let $\mathrm{Types}(h, r)$ *be the set of different types of subsets of size at most r inside ordered trees of height at most h. Note that* $|\mathrm{Types}(h, r)|$ *is finite since there are only finitely many rooted trees of height at most h with at most r leaves.*

For an ordered forest F, let F^- denote F, but with all roots deleted (see Figure 3). If τ is the type of F, then define τ^- to be the type of F^- (noting that this depends only on the type τ, and not on F).

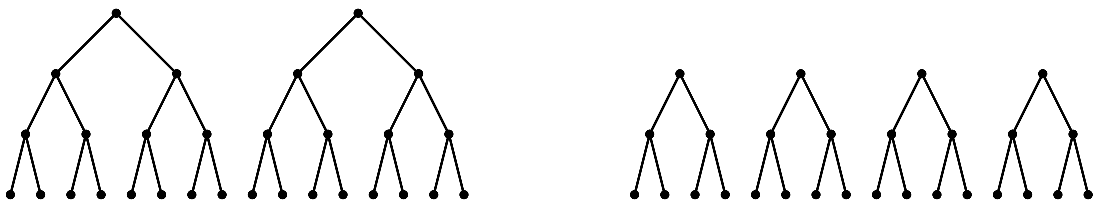

Figure 3: A forest F (on the left) and the forest F^- (on the right)

Lemma 4.8. *For $X \subseteq L(F)$, we have $A^{F^-}(X) = (A^F(X))^-$.*

Proof. By definition,

$$
\begin{aligned}
V\left(A^{F^-}(X)\right) &= \{y \in V(F^-) : y \text{ is an ancestor of some } x \in X\} \\
&= \{y \in V(F) \setminus \mathrm{root}(F) : y \text{ is an ancestor of some } x \in X\} \\
&= \{y \in V(F) : y \text{ is an ancestor of some } x \in X\} \setminus \mathrm{root}(F) \\
&= V\left(A^F(X)\right) \setminus \mathrm{root}(F) = V\left(A^F(X)\right) \setminus \mathrm{root}(A^F(X)) = V\left((A^F(X))^-\right). \quad \square
\end{aligned}
$$

In the next two definitions we show how to relate rooted forests to hypergraphs.

Definition 4.9. *For a set S, an S-*forest *is a balanced, ordered forest F with $V(F) \subseteq S \times S$ so that all edges are of the form $(y, s)(y', s)$, and all roots are of the form (s, s).*

For a set S, an element $s \in S$, and an S-forest F, define $F(s)$ to be the subgraph of F induced on $\{(y, s) \in V(F) : y \in S\}$.

Observation 4.10. *For a set S, an element $s \in S$, and an S-forest F, the subgraph $F(s)$ is a tree.*

Proof. For some $(y, s) \in V(F(s))$, let P be the path in F from a root r to (y, s) (P exists since we are in a rooted forest). Write P as $r = (v_1, s_1) \dots (v_k, s_k) = (y, s)$. Then $s_1 = \dots = s_k = s$, because all edges in F are of form $(y, s')(y', s')$, for some $s' \in S$, and $s_k = s$. Thus $r = (s, s)$, because roots have form (s', s'), with $s' \in S$. This shows that P is a path from (y, s) to (s, s) in $F(s)$. Therefore, $F(s)$ is connected and thus a tree. $\square$

We will call the tree $F(s)$ the "subtree of F rooted at (s, s)". We will use S-forests as a way to represent a collection of balanced trees of the same height, with roots and vertices in S, where each element of S is the root of at most one tree, but can be a vertex in many of these trees.

For an S-forest F, a vertex $v \in V(F)$, and $i \leq \text{level}(v)$, we let $\pi_i^F(v)$ be the first coordinate of the ancestor of v at level i. In particular, $\pi^F_{\text{level}(y,s)}\big((y, s)\big) = y$ and $\pi^F_0\big((y, s)\big) = s$. We abbreviate $\pi(v) := \pi^F_{\text{level}(v)}(v)$ and $\pi_0(v) := \pi^F_0(v)$ (noting that these do not depend on F, since they are just projections on the first/second coordinates of v).

The following is how we relate hypergraphs to forests.

Definition 4.11 ($\mathcal{H} \otimes_r F$)**.** *For a hypergraph $\mathcal{H}$, a $V(\mathcal{H})$-forest F, and an integer $r \geq 1$, let $\mathcal{H} \otimes_r F$ be the r-uniform hypergraph with vertex set $L(F)$, where an r-tuple e is an edge wherever $\pi_0(e) \subseteq f$ for some edge f of $\mathcal{H}$. Equivalently, $\{(v_1, x_1), \ldots, (v_r, x_r)\} \in E(\mathcal{H} \otimes_r F)$ whenever $\{x_1, \ldots, x_r\}$ is contained in some edge in $\mathcal{H}$.*

When r is known from context or its value is not relevant for the arguments, we often abbreviate $\mathcal{H} \otimes_r F$ to $\mathcal{H} \otimes F$.

4.2 Clean and separated tree assignments

The goal of this section is to prove that every $V(\mathcal{H})$-forest, where $\mathcal{H}$ is a hypergraph with bounded maximum degree, has a large subforest with the same root set that interacts nicely with $\mathcal{H}$.

Definition 4.12 (Clean hypergraphs)**.** *For a hypergraph $\mathcal{H}$ and a $V(\mathcal{H})$-forest F, we say that a colouring of the hypergraph $\mathcal{H} \otimes F$ is* clean *if any two edges e, f satisfying* $\text{root}(e) = \text{root}(f)$ *and* $\text{type}(e) = \text{type}(f)$ *(i.e.* $\text{type}\big(e \cap V(F(v))\big) = \text{type}\big(f \cap V(F(v))\big)$ *for every $v \in$* $\text{root}(f)$*) have the same colour.*

This is equivalent to asking that, for any edge $e \in \mathcal{H} \otimes F$ and monomorphism $\varphi : A(e) \to F$ which preserves roots, $\varphi(e)$ has the same colour as e.

Definition 4.13 (Separated tree assignments)**.** *Let G be a graph and F a $V(G)$-forest. We say that F is d-*separated *on G if for $u, v \in V(G)$ with $d_G(u, v) \leq d$ (namely the distance between u and v on G is at most d), we have $\pi(F(u)) \cap \pi(F(v)) = \emptyset$.*

Note that this is equivalent to "for $u, v \in V(F)$ with $d_G(\pi_0(u), \pi_0(v)) \leq d$ we have $\pi(u) \neq \pi(v)$". Informally, thinking of a $V(G)$-forest as a collection of trees on $V(G)$, where each vertex is the root of at most one tree, this means that trees rooted at vertices that are at distance at most d from each other in G are vertex-disjoint.

The goal of the section is to prove Lemma 4.17 which shows that under certain assumptions on $\mathcal{H}$, and for $d \gg d'$, any forest of d-ary trees rooted on $V(\mathcal{H})$ contains a d'-ary subforest with the same root set that is both clean and separated. We start with establishing "separated".

Lemma 4.14. *There is a function $D(h, d)$, such that the following holds for d_1, d_2, d_1', d_2' with $d_1 \geq D(h, d_1')$ and $d_2 \geq D(h, d_2')$. Let T_1 and T_2 be balanced, ordered d_1- and d_2-ary trees of height h, whose vertex sets may intersect, but whose roots are distinct. Then there are balanced d_1'- and d_2'-ary, height h subtrees $T_1' \subseteq T_1$ and $T_2' \subseteq T_2$ with $V(T_1')$, $V(T_2')$ disjoint.*

Proof. Assign to each non-root vertex u of $V(T_1) \cup V(T_2)$ a random variable $\tau(u)$ which is chosen to be either 1 or 2 uniformly at random, independently of other elements. We claim that, with positive probability, for every non-leaf in T_i, at least $d_i/3$ of its children are assigned i, for $i \in [2]$. Indeed, the probability that this fails for a particular non-leaf in T_i is at most $e^{-d_i/36}$, by Chernoff's bounds, and thus by a union bound the probability that some non-leaf in T_i has fewer than $d_i/3$ children that were assigned i is at most $d_i^h e^{-d_i/36} < 1/2$ (using $d_i \ge D(h, d'_i)$), as claimed. It follows that there exist subtrees $T'_i \subseteq T_i$, for $i \in [2]$, such that T'_i is a d'_i-ary tree of height h and its non-root vertices were assigned i (and are not the root of T_{3-i}; using $d_i \ge D(h, d'_i)$). Then the trees T'_1, T'_2 are vertex-disjoint (using that the roots of T_1 and T_2 are distinct), as required. $\square$

The above can be used to show that F always contains a separated subforest. In the proof, for a function f, we write f^i to denote $\overbrace{f \circ \cdots \circ f}^{i \text{ times}}$. Recall that $F' \le F$ means that F' is a subforest of F with the same root set.

Lemma 4.15 (Separating forests)**.** *Let $d \gg b, h, \Delta$. Let G be a graph with maximum degree at most Δ and F be a $V(G)$-forest of balanced, d-ary, height h trees. Then there is a forest $F' \le F$ of balanced, b-ary, height h trees such that F' is b-separated on G.*

Proof. Fix b, h, Δ, and let $f(d') := D(h, d')$, where D is the function from Lemma 4.14. Let $g(d) = \max\{d' : d \ge f(d')\}$, and note that $g(f(d)) \ge d$ for every d. Set $d = f^{\Delta^b}(b)$. Then $g^i(d) \ge b$ for every $i \le \Delta^b$.

We modify the forest F, by doing the following for each edge xy in G^b, one at a time. Suppose that $F(x)$ and $F(y)$ are d_x- and d_y-ary trees of height h. Apply Lemma 4.14 to get $g(d_x)$- and $g(d_y)$-ary subtrees $F''(x) \le F(x)$ and $F''(y) \le F(y)$, which are balanced height h forests and are vertex-disjoint, and replace $F(x)$ by $F''(x)$ and $F(y)$ by $F''(y)$. Notice that each subtree $F(x)$ is modified at most Δ^b times, and so it is $g^i(d)$-ary for some $i \le \Delta^b$. Since $g^i(d) \ge b$, we can choose $F'(x)$ to be a balanced b-ary subtree of height h of the tree $F(x)$ at the end of the process. Taking F' to be the union of trees $F'(x)$ completes the proof. $\square$

Next we focus on establishing cleanliness.

Lemma 4.16. *There is a function $D(d, r, h, k, s)$ such that the following is true. Let d, r, h, k, s be positive integers and set $D = D(d, r, h, k, s)$. Let F be a balanced, D-ary ordered forest of height h with k components. Let S be a balanced, ordered forest of height h with r leaves and k components. Given any s-colouring of the copies of S in F, there is a balanced ordered d-ary subforest $F' \le F$ of height h with k components, such that all copies of S in F' have the same colour.*

Proof. We prove the lemma by induction on h and k. The initial case is $h = k = 1$, namely F and S are both stars, with D and r leaves, respectively, and so copies of S correspond to r-sets of leaves of F. This case thus follows from Ramsey's theorem.

Now fix $(h, k) \ne (1, 1)$ and suppose that the lemma holds for all (h', k') with either $h' = h, k' < k$ or with $h' < h, k' \ge 1$. Let χ be an s-colouring of the copies of S in F.

Suppose first that $k = 1$ and $h \ge 2$, so F and S are trees of height h. Denote the root of F by r_F and the root of S by r_S. Let $F' = F \setminus \{r_F\}$, let $S' = S \setminus \{r_S\}$, and let ℓ be the number of components in S'; so $1 \le \ell \le r$ as S has r leaves. Consider the s-colouring χ' of copies of S' in F' defined as follows: if S'' is a copy of S' in F', consider the copy of S obtained by joining

r_F to the roots of the components of S' (since both S' and F' have height $h-1$, the roots of components of S' are roots of components of F', so the copy of S formed in this way is indeed a subtree of F), and colour S'' by the colour of this copy of S in F according to χ.

Let α be such that $d, r, h, s \ll \alpha \ll D$, and set $t = \binom{\alpha}{\ell}$. By choice of parameters, there is a sequence $d_0, \dots, d_t$ such that $d_0 = D$, $d_t = d$, and $r, h, s, d_t \ll d_{t-1} \ll \dots \ll d_0$. Let A be a set of α children of r_F in F, and let $A_1, \dots, A_t$ be any enumeration of its ℓ-subsets. For $a \in A$, let T_a be the subtree of F rooted at a. We claim that there exist sequences $T_a = T_a^{(0)} \supseteq T_a^{(1)} \supseteq \dots \supseteq T_a^{(t)}$ for $a \in A$, where $T_a^{(i)}$ is a d_i-ary ordered tree of height $h-1$ with the following property: all copies of S' in the forest $\bigcup_{a \in A_i} T_a^{(i)}$ have the same colour. To see this, apply the induction hypothesis with $h' = h-1$ and $k' = \ell$, to $\bigcup_{a \in A_i} T_a^{(i-1)}$, letting $\bigcup_{a \in A_i} T_a^{(i)}$ be the resulting subforest, and take $T_a^{(i)} \subseteq T_a^{(i-1)}$ to be an arbitrary d_i-ary ordered subforest of height $h-1$ for $a \in A \setminus A_i$. The subtrees $T_a^{(i)}$ satisfy the requirements. Denote by c_i the colour of any copy of S' in $\bigcup_{a \in A_i} T_a^{(i)}$ (this does not depend on the particular choice of a copy of S').

Consider the auxiliary colouring of the complete ℓ-graph on A, where the edge A_i has colour c_i. Then by Ramsey's theorem and choice of α there is a monochromatic subset $A' \subseteq A$ of size d; say the common colour is c. Let F'' be the d-ary tree of height h obtained by reattaching the root of F to the subtrees $T_a^{(t)}$ with $a \in A'$. Then $F'' \subseteq F$ is an ordered d-ary tree of height h whose copies of S all have colour c, as required.

Now suppose $k \geq 2$. Let T be the first tree in F, let $F' = F \setminus V(T)$, let R be the first tree in S and let $S' = S \setminus V(R)$. Choose d' satisfying $d, r, h, k, s \ll d' \ll D$, and let T' be any d'-ary subtree of T of height h. Enumerate the copies of R in T' by $R_1, \dots, R_t$. By choice of d' and t there exist $d_0, \dots, d_t$ such that $d_0 = D$, $d_t = d$, and $r, s, h, k, d_t \ll d_{t-1} \ll \dots \ll d_0$. By induction, there exist subforests $F' \supseteq F_1 \supseteq \dots \supseteq F_t$, such that F_i is a d_i-ary forest of height h with $k-1$ components, whose copies of S in $T' \cup F_i$ that contain R_i all have the same colour, denoted c_i. Now consider the colouring of copies of R in T', where R_i is coloured c_i. By induction, T' contains a d-ary subtree T'' of height h whose copies of R all have the same colour, say red. Then the forest $T'' \cup F_t$ is a d-ary forest of height h with k components whose copies of S are red. $\square$

The above can be used to show that $\mathcal{H} \otimes F$ always contains a cleanly coloured hypergraph.

Lemma 4.17 (Cleaning forests)**.** *Let $d \gg b, h, s, r, \Delta$. Let $\mathcal{H}$ be an r-uniform hypergraph with maximum degree at most Δ and F a $V(\mathcal{H})$-forest of balanced d-ary trees of height h.*

Then for any s-colouring of $\mathcal{H} \otimes F$, there is some ordered subforest $F' \leq F$ of balanced, b-ary, height h trees such that the subhypergraph $\mathcal{H} \otimes F'$ is cleanly coloured.

Proof. Fix b, h, s, r, Δ, and let $f(d) = \max\{D(d, r, h, k, s) : k \in [r]\}$, where D is the function from Lemma 4.16. Let $g(d) = \max\{d' : d \geq f(d')\}$, noting that $g(f(d)) \geq d$ for every $d \geq 1$. Set $\sigma = 2^r |\operatorname{Types}(h, r)|$, $\rho = r\Delta + 1$, $d = f^{\sigma\rho}(b)$, and observe that $g^i(d) \geq b$ for $i \leq \rho\sigma$.

Let $\mathcal{H}'$ be the subhypergraph of $\mathcal{H}$, induced on vertices $v \in V(\mathcal{H})$ such that $F(v) \neq \emptyset$. We claim that the edges of $\mathcal{H}'$ can be partitioned into at most ρ matchings (which are collection of pairwise vertex-disjoint edges). Indeed, consider the line graph $L(\mathcal{H}')$ of $\mathcal{H}'$, which is the graph on $E(\mathcal{H}')$ where ef is an edge whenever e and f intersect. This graph has maximum degree at most $r\Delta = \rho - 1$, and so its chromatic number is at most ρ. In other words, there is a proper colouring of $L(\mathcal{H}')$ with ρ colours, and this corresponds to a partition of $E(\mathcal{H}')$ into at most ρ matchings. Denote the matchings involved in such a partition by $M_1, \dots, M_\rho$.

Write $F_0 = F$. We define a sequence $F_1, \dots, F_\rho$, as follows, so that $F_i \leq F_{i-1}$ and F_i is a balanced $g^{\sigma i}(d)$-ary forest of height h for $i \in [k]$. Let $i \in [k]$ and suppose that $F_0, \dots, F_{i-1}$ are defined and satisfy the requirements. For each edge $e \in M_i$ we modify the trees $F_{i-1}(v)$ for $v \in e$ (notice that each such tree is a balanced $g^{\sigma(i-1)}(d)$-ary tree of height h) in at most σ rounds, once for each type $\tau \in \mathrm{Types}(h, r)$ and subset $e' \subseteq e$ of size $|\mathrm{root}(\tau)|$. We make sure that after the j-th round each $F_{i-1}(v)$ with $v \in e$ is a balanced $g^{\sigma(i-1)+j}(d)$-ary tree of height h. Say we are in the j-th round, the corresponding type is τ and the corresponding subset of e is e'. Apply Lemma 4.16 to the ordered forest $\bigcup_{v \in e'} F_{i-1}(v)$ to obtain balanced $g^{(i-1)\sigma+j}(d)$-ary subtrees $F'(v) \leq F_{i-1}(v)$ of height h, for $v \in e'$, such that all type τ subforests of $\bigcup_{v \in e'} F'(v)$ have the same colour. Replace $F_{i-1}(v)$ by $F'(v)$ for every $v \in e'$, and replace each $F_{i-1}(v)$ with $v \in e \setminus e'$ by a balanced $g^{(i-1)\sigma+j}(d)$-ary subtree of height h, chosen arbitrarily. After these at most σ rounds, each $F_{i-1}(v)$ with $v \in e$ is a balanced $g^k(d)$-ary tree of height h, with $k \leq i\sigma$, and for every $\tau \in \mathrm{Types}(h, r)$, all type τ subforests of $\bigcup_{v \in e} F_{i-1}(v)$ with the same root set have the same colour. For each $v \in V(\mathcal{H}')$, take $F_i(v)$ to be a balanced $g^{i\sigma}(d)$-ary subtree of $F_{i-1}(v)$ of height h, chosen arbitrarily, and set $F_i = \bigcup_{v \in V(\mathcal{H}')} F_i(v)$.

Notice that F_ρ is a balanced b-ary forest (using $b = g^{\sigma\tau}(d)$) of height h, satisfying that $F_\rho \leq F$ and $\mathcal{H} \otimes F_\rho$ is cleanly coloured. $\square$

4.3 Distances, short trees, and disconnected tree assignments

Given a graph G, a $V(G)$-forest F, and a hypergraph $\mathcal{H}$ on $L(F)$ we will need to track what sorts of paths/walks one can find in $\mathcal{H}$ between different trees of F. How long/short a path is will be judged relative to distances in G via the following definition.

Definition 4.18 ($|| * ||_{F,G}$)**.** *Let F be a $V(G)$-forest for a graph G. For a set $e \subseteq L(F)$, we define $||e||_G = \max_{x,y \in e} d_G(\pi_0(x), \pi_0(y))$.*

We remark that if P, Q have the same root sets, then $||P||_G = ||Q||_G$. This definition gives a concise way of defining $\mathcal{G}^t \otimes F$ (recall that $\mathcal{G}^t$ is a shorthand for the r-uniform hypergraph $G^t_{(r)}$, which is the r-uniform hypergraph on $V(G)$ whose edges are r-sets of vertices that are pairwise at distance at most t from each other in G).

Observation 4.19. *For a graph G and forest F rooted on $V(G)$, the edge set of the hypergraph $\mathcal{G}^t \otimes_r F$ consists of exactly the r-sets $e \subseteq L(F)$ which have $||e||_G \leq t$.*

If $\mathcal{H}$ is a hypergraph whose vertex set is ordered, a *tight walk* in $\mathcal{H}$ is a sequence of vertices $P = v_1 \dots v_k$ so that $\{v_i, v_{i+1}, \dots, v_{i+r-1}\} \in E(\mathcal{H})$ for each $i \in [k - r + 1]$, and also $v_1 < \dots < v_{r-1}$ and $v_{k-r+1} < \dots < v_k$. We call $\{v_1, \dots, v_{r-1}\}$ and $\{v_{k-r+1}, \dots, v_k\}$ the *start* and *end* of the tight walk. This definition has the consequence that a subsequence $v_i v_{i+1} \dots v_j$ of P need not be a tight walk. However, it has the useful property that given two tight walks P, Q such that P is from S to T and Q is from T to R, we can concatenate P and Q to get a walk PQ from S to R.

We will need the following.

Definition 4.20 (Independent sets of leaves)**.** *For a forest F we say that two sets of leaves $e, f \subseteq L(F)$ are* independent *if there are vertices $v_e, v_f \in V(F)$ such that the vertices in e are descendants of v_e, the vertices in f are descendants of v_f, and v_e, v_f are incomparable.*

Observation 4.21. *If e, f are independent sets of leaves in an S-forest, then $|\pi_0(e)| = |\pi_0(f)| = 1$.*

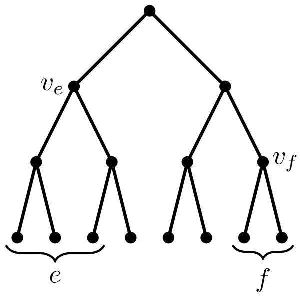

Figure 4: Two independent sets of leaves e, f and the corresponding ancestors v_e, v_f

Proof. Let all of e descend from $v_e = (x, y)$. By definition of edges in S-forests, all descendants v of (x, y) have $\pi_0(v) = y$, showing that $|\pi_0(e)| = |\{y\}| = 1$. The same works for f. $\square$

A crucial definition in the paper is the following one — it is the structure we find in our coloured graph without monochromatic tight paths.

Definition 4.22 (k-disconnected)**.** *Let F be a $V(G)$-forest for a graph G, and let $\mathcal{H}$ be a coloured r-uniform hypergraph with $V(\mathcal{H}) = L(F)$. We say that a colouring of $\mathcal{H}$ is k-*disconnected *on (G, F) if for any $v \in V(G)$, any independent $(r-1)$-sets $S, T \subseteq L(F(v))$ of the same type, and any monochromatic tight walk P of length at most $3r$ from S to T in $\mathcal{H}$, we have $||V(P)||_G \geq k$.*

Height 1 families are always disconnected.

Observation 4.23. *Let G be a graph and F be a $V(G)$-forest of height 1 trees (i.e. stars). Then every hypergraph $\mathcal{H}$ of uniformity at least 3 with $V(\mathcal{H}) = L(F)$ is k-disconnected on (G, F) for every k.*

Proof. Note that any $S, T \subseteq L(F(v))$ of size at least 2 are not independent (since $|S| \geq 2$, the only common ancestor of the vertices in S is the root v, and similarly for T. Thus the only choice of v_S, v_T for the definition of "independent" could be $v_S = v, v_T = v$, which does not satisfy "v_S, v_T incomparable"). Therefore, the definition of "k-disconnected" holds vacuously since there are no pairs S, T for which the definition needs to be checked. $\square$

In contrast to this there do not exist 1-disconnected families of large height. This will be proved via the following two lemmas.

Lemma 4.24. *Let $r, s \ll h \ll d$ and let $\ell \leq r$. Let T be a balanced d-ary ordered tree of height h, and let $\mathcal{H}$ be the r-uniform complete graph on $L(T)$. Then for every s-colouring of $\mathcal{H}$ there exist disjoint sets $X, Y, Z \subseteq L(T)$ of size ℓ, such that X, Z are independent and of the same type, and $X \cup Y$ and $Y \cup Z$ are monochromatic cliques of the same colour.*

Proof. Fix an s-colouring of all r-subsets of $L(T)$. By Lemma 4.17 applied to T, there is a balanced binary subtree $T' \leq T$ of height h, such that r-sets of leaves of the same type have the same colour. Let v_0 be the first vertex in $L(T')$ (according to the ordering of T', inherited from T), and let $v_0 \dots v_h$ be the path from v_0 to the root of T'. Let u_i be the last leaf in the tree rooted at v_i for $i \in \{0, \dots, h\}$. (See Figure 5 for an illustration of the vertices v_i and u_i.) By Ramsey's theorem there is a subset $A \subseteq \{u_0, \dots, u_h\}$ of size $2\ell + 1$ whose r-subsets all have the same colour, say red. Let $w_1, \dots, w_{2\ell+1}$ be the vertices in A, in order.

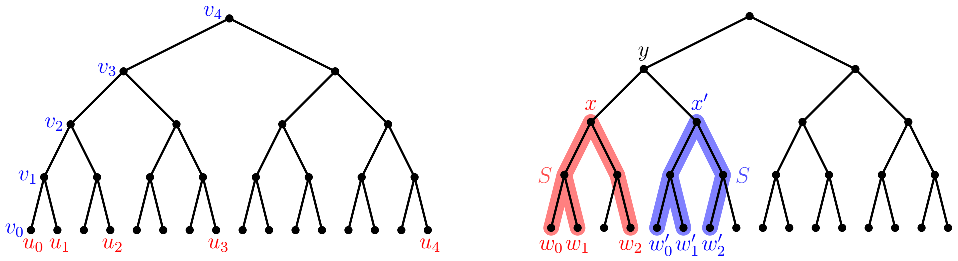

Figure 5: An illustration of the sets $\{v_0, \dots, v_h\}$, $\{u_0, \dots, u_h\}$, $\{w_1, \dots, w_\ell\}$ and $\{w'_1, \dots, w'_h\}$; here the leaves are ordered from left to right.

Let S be the minimal ordered subtree of T' whose leaves are $\{w_1, \dots, w_\ell\}$. Denote its root by x and denote the father of x by y. Let x' be the child of y which is not x, let S' be a copy of S rooted at x', and denote the leaves of S' by $\{w'_1, \dots, w'_\ell\}$, noting that $\{w_1, \dots, w_\ell\}$ and $\{w'_1, \dots, w'_\ell\}$ are independent (See Figure 5 for an illustration of the vertices w_i and w'_i.)

Notice that $\{w_1, \dots, w_\ell, w_{\ell+2}, \dots, w_{2\ell+1}\}$ and $\{w'_1, \dots, w'_\ell, w_{\ell+2}, \dots, w_{2\ell+1}\}$ have the same type (we omitted $w_{\ell+1}$ as it may be a descendant of y, whereas $w_{\ell+1}, \dots, w_{2\ell+1}$ are not), so by choice of $w_1, \dots, w_{2\ell+1}$ and T' both sets are red cliques. Take $X = \{w_1, \dots, w_\ell\}$, $Y = \{w_{\ell+2}, \dots, w_{2\ell+1}\}$ and $Z = \{w'_1, \dots, w'_\ell\}$. These sets satisfy the requirements of the lemma. $\square$

The following lemma is what gives a contradiction in our main proof.

Lemma 4.25 (Disconnection is bounded)**.** *Let $r \le s \ll h \ll d$. Let $t \ge 0$ be an integer, let G be a graph, and let F be a $V(G)$-forest consisting of balanced d-ary trees of height at least h. Then there is no s-colouring of $\mathcal{G}^t \otimes_r F$ which is 1-disconnected on (G, F).*

Proof. Suppose, for contradiction, that we have a 1-disconnected, s-colouring of $\mathcal{G}^t \otimes F$. Consider some $v \in V(G)$ and the corresponding tree $F(v) \in F$. By Lemma 4.24, there are disjoint $(r-1)$-sets $X = \{x_1, \dots, x_{r-1}\}$, $Y = \{y_1, \dots, y_{r-1}\}$, $Z = \{z_1, \dots, z_{r-1}\} \subseteq L(F(v))$ such that X, Z are independent and of the same type, and $X \cup Y$, $Y \cup Z$ induce monochromatic complete hypergraphs. Now $P = x_1 \dots x_{r-1} y_1 \dots y_{r-1} z_1 \dots z_{r-1}$ is a monochromatic tight path in $\mathcal{G}^t \otimes F$ from X to Z. Notice that for all $x \in V(P)$ we have $\pi_0(x) = v$. This gives $||e||_G = 0$ for every set of r consecutive vertices in P, showing that P is indeed a tight path in $\mathcal{G}^t \otimes F$. It also gives $||V(P)||_G = 0$, showing that $\mathcal{G}^t \otimes F$ is not 1-disconnected on (F, G). $\square$

The following allows us to keep track of distances when we modify one forest into another.

Definition 4.26 (Short trees)**.** *For a graph G and a $V(G)$-forest F, we say that level i is k-short in G if for all its edges xy between level $i-1$ and level i, we have $d_G(\pi(x), \pi(y)) \le k$.*

Note that this is equivalent to "for all $v \in L(F)$, we have $d_G(\pi^F_{i-1}(v), \pi^F_i(v)) \le k$". Short forests have the property that for two vertices x, y in them it is easy to estimate $d_G(x, y)$.

Lemma 4.27. *Let G be a graph, and F a $V(G)$-forest whose $1, \dots, \max\{i, j\}$ levels are k-short. For any set $W \subseteq L(F)$ and vertices $u \in \pi^F_i(W)$ and $v \in \pi^F_j(W)$, we have $d_G(u, v) \le ||W||_G + ik + jk$.*

Proof. Let $P = (u_0, x)(u_1, x), \dots, (u_h, x)$ be a root-leaf path in F such that $u_i = u$ and $(u_h, x) \in W$ (this exists by assumption on u). Since levels $1, \dots i$ are k-short, we have that $d_G(u_{t-1}, u_t) \le k$ for all $t \le i$, which gives $d_G(x, u) = d_G(u_0, u_i) \le ik$. Similarly, there is a root-leaf path $Q = (v_0, y)(v_1, y), \dots, (v_h, y)$ such that $v_j = v$ and $(v_h, y) \in Q$, so $d_G(y, v) \le jk$.

Also, since $(u_h, x), (v_h, y) \in W$, we have $d_G(x, y) = d_G\big(\pi_0(u_h, x), \pi_0(v_h, y)\big) \le ||W||_G$. Using the triangle inequality, we get $d_G(u, v) \le d_G(u, x) + d_G(x, y) + d_G(y, v) \le ||W||_G + ik + jk$. $\square$

The hypergraph we really care about is $\mathcal{G}^p$. The following is how we relate it to the hypergraphs $\mathcal{G}^t \otimes_r F$ for which all our machinery works.

Lemma 4.28. *Let $p \gg t, h, k, r$. Let G be a graph and F a $V(G)$-forest of height at most h that is t-separated on G and whose levels are all k-short on G. Then $\pi : \mathcal{G}^t \otimes_r F \to \mathcal{G}^p$ maps edges to edges.*

Proof. Let $\{v_1, \dots, v_r\} \in E(\mathcal{G}^t \otimes F)$. By Observation 4.19, we have $||\{v_1, \dots, v_r\}||_G \le t$, or, equivalently, $d_G(\pi_0(v_i), \pi_0(v_j)) \le t$ for every $i, j \in [r]$. Lemma 4.27 implies $d_G(\pi(v_i), \pi(v_j)) \le t + 2hk \le p$ for all i, j. Also, by F being t-separated on G, we have $\pi(v_i) \ne \pi(v_j)$ for all distinct i, j (indeed, write $u = \pi_0(v_i)$ and $v = \pi_0(v_j)$. If $u = v$ then $\pi(v_i) \ne \pi(v_j)$ because otherwise $v_i = v_j$. If $u \ne v$, since $d_G(u, v) \le t$ and F is t-separated on G, then $\pi(F(u)) \cap \pi(F(v)) = \emptyset$). We have shown that $\{\pi(v_1), \dots, \pi(v_r)\}$ is a set of r vertices in G within distance p of each other in G, i.e. an edge of $\mathcal{G}^p$. $\square$

4.4 Augmentations

In this section we take a balanced forest F of height h, and build a balanced forest of height $h + 1$.

The following definition takes an S-forest and yields another S-forest with the roots of the former removed.

Definition 4.29. *For an S-forest F the* root-deletion function *is the function $\mathrm{rd}^F : V(F) \setminus \{(x, x) : x \in S\} \to S \times S$ by $\mathrm{rd}^F : (x, y) \to \big(x, \pi_1^F((x, y))\big)$.*

Observation 4.30. *If $F \le F'$ are two forests, then restricting $\mathrm{rd}^{F'}$ to F gives the function rd^F, i.e. $\mathrm{rd}^{F'}|_F = \mathrm{rd}^F$.*

The following is how we build trees of larger heights. Informally, what the next definition does is take disjoint trees $T(u_1), \dots, T(u_k)$, and adds one extra vertex v that is connected to all their roots, to get a new tree rooted at v.

Definition 4.31 (Augmentation of a single tree)**.** *Let $u_1, \dots, u_k, v$ be distinct elements in a set S, and let $T_1, \dots, T_k$ be trees, such that $\mathrm{root}(T_i) = \{u_i\}$ for $i \in [k]$, and the sets $\pi\big(V(T_1)\big), \dots, \pi\big(V(T_k)\big), \{v\}$ are pairwise disjoint. Define an S-tree, denoted by $(v; T_1; \dots; T_k)$, as follows:*

- *First, for each i, define T_i^* by*

$$V(T_i^*) = \{(x, v) : (x, u_i) \in V(T_i)\}$$
$$E(T_i^*) = \{(x, v)(y, v) : (x, u_i)(y, u_i) \in E(T_i)\},$$

  and order each $V(T_i^*)$ by $(x, v) < (y, v) \iff (x, u_i) < (y, u_i)$.

- *Then, order $\bigcup_{i \in [k]} V(T_i^*)$ by $V(T_1^*) < \dots < V(T_k^*)$ thereby making the union $\bigcup_{i \in [k]} T_i^*$ an ordered forest.*

- *Finally, define $(v; T_1; \ldots; T_k)$ to have vertex set V and edge set E, defined as follows.*

$$V = \{(v, v)\} \cup \bigcup_{i \in [k]} V(T_i^*),$$
$$E = \{(v, v)(u_i, v) : i \in [k]\} \cup \bigcup_{i \in [k]} E(T_i^*),$$

and extend the ordering of $\bigcup_{i \in [k]} V(T_i^*)$ to an ordering of V by letting (v, v) precede all other vertices.

It is perhaps not obvious that the above produces an ordered tree — this is proved in the following lemma. Recall that F^- denotes the subforest of F with all roots deleted.

Lemma 4.32. *Let F be an S-forest, and let $u_1, \ldots, u_k \in \pi_0(V(F))$ be such that the sets $\pi\big(V(F(u_1))\big), \ldots, \pi\big(V(F(u_k))\big), \{v\}$ are pairwise disjoint. Set $T = (v; F(u_1); \ldots; F(u_k))$. Then T is an S-tree and $\mathrm{rd}^T|_{T^-} : T^- \to F$ is a monomorphism.*

Proof. With $T_i := F(u_i)$ and T_i^* as in Definition 4.31, note that $\varphi : \bigcup_{i \in [k]} F(u_i) \to \bigcup_{i \in [k]} T_i^*$ defined by $\varphi((x, u_i)) = (x, v)$ is an isomorphism. This shows that $\bigcup_{i \in [k]} T_i^*$ is an ordered forest with roots $(u_1, v), \ldots, (u_k, v)$. The vertex (v, v) is joined to exactly these vertices making T a rooted tree. To see that the vertices of T are ordered correctly we need to check that (F1) and (F2) hold when one of the vertices involved is (v, v) (since we have already established that $T \setminus \{v\}$ is an ordered forest). Part (F1) holds for (v, v) since v is an ancestor of everything and also ordered to be before everything. Part (F2) does not need to be checked for (v, v) since it concerns pairs of incomparable vertices — but (v, v) is comparable with everything.

For the second part of this lemma, note that for $(x, v) \in V(T_i^*)$, we have that $\pi_1^T((x, v)) = u_i$ (the vertex (u_i, v) is on level 1 of T because it is a child of the root (v, v), also (u_i, v) is an ancestor of (x, v) since $(x, v) \in T_i^*$). This shows that for $(x, v) \in V(T_i^*)$ we have $\mathrm{rd}^T((x, v)) = (x, u_i) = \varphi^{-1}(x, v)$ — and hence the function from $\bigcup_{i \in [k]} T_i^*$ to $\bigcup_{i \in [k]} F(u_i)$ given by rd^T is an isomorphism. Since $T^- = \bigcup_{i \in [k]} T_i^*$, we get that $\mathrm{rd}^T|_{T^-} : T^- \to F$ is a monomorphism. □

The following lemma works out how distances change with the above operation.

Lemma 4.33. *For a graph G, let F be a balanced $V(G)$-forest of height h, and let $u_1, \ldots, u_k \in \pi_0(V(F))$ be chosen so that $\pi\big(V(F(u_1))\big), \ldots, \pi\big(V(F(u_k))\big), \{v\}$ are pairwise disjoint. Set $T = (v; F(u_1); \ldots; F(u_k))$. For $i \in [h]$, if level i of each $F(u_1), \ldots, F(u_k)$ is k-short, then level $i + 1$ of T is k-short. Moreover, if $d(v, u_1), \ldots, d(v, u_k) \le k$, then level 1 of T is k-short.*

Proof. For the first part, note that from Definition 4.31, with $T_i = F(u_i)$ and T_i^* as in Definition 4.31, we get that vertices of T on level i are the vertices of $\bigcup_{i \in [k]} T_i^*$ on level $i - 1$. Let $\varphi : \bigcup_{i \in [k]} F(u_i) \to \bigcup_{i \in [k]} T_i^*$ be the function from the proof of the previous lemma, which was shown to be an isomorphism. For a vertex x in $\bigcup_{i \in [k]} T_i^*$, we have $\varphi^{-1}(x)$ is a vertex of $\bigcup_{i \in [k]} F(u_i)$ with $\mathrm{level}(x) = \mathrm{level}(\varphi^{-1}(x))$ (by Lemma 4.5). Thus

$$\begin{aligned} &\max\{d_G\big(\pi(x), \pi(y)\big) : xy \in E(T_i^*), \text{ with } x \text{ on level } i - 1 \text{ and } y \text{ on level } i\} \\ = &\max\{d_G\big(\pi(x), \pi(y)\big) : xy \in E(F(u_i)), \text{ with } x \text{ on level } i - 1 \text{ and } y \text{ on level } i\} \le k. \end{aligned}$$

This shows that level i in $\bigcup_{i \in [k]} T_i^*$ is k-short, implying that level $i + 1$ in T is k-short, for every $i \in [h]$.

For the second part note that from Definition 4.31, the only edges of T from level 0 to level 1 are the edges $\{(v, v)(u_i, v) : i \in [k]\}$, and $d_G(v, u_i) \le k$, showing that level 1 is k-short. □

Roughly speaking, we think of one forest F' is an augmentation of another forest F if $F' \leq F''$ for some forest F'', whose trees are constructed from trees of F via Definition 4.31. This informal definition is cumbersome to work with, so we instead work with the following definition via the rd function.

Definition 4.34 (Augmentation of tree families)**.** *Let F, F' be S-forests. Say F' augments F if every $v \in S$ satisfies $\mathrm{rd}^{F'}(F'(v)^-) \subseteq V(F)$ and moreover $\mathrm{rd}^{F'}|_{F'(v)^-} : F'(v)^- \to F$ is a monomorphism (see Figure 6).*

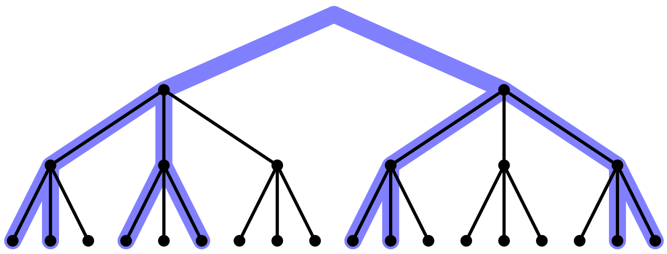

Figure 6: An illustration of a forest F' (in blue) which is an augmentation of a forest F (in black)

The following lemma gives us the properties of augmentation we use.

Lemma 4.35 (Properties of augmentation)**.** *Let G be a graph and F, F' be $V(G)$-forests with F' augmenting F. The following are true.*

(i) $\pi^F \circ \mathrm{rd}^{F'} = \pi^{F'}$ *and* $\pi_1^{F'}|_{F'-} = \pi_0^F \circ \mathrm{rd}^{F'}$.

(ii) *If level 1 of F' is k-short on G, then for any $S \subseteq L(F)$, we have $||S||_G \geq ||\,\mathrm{rd}^{F'}(S)||_G - 2k$.*

(iii) *If $t \geq s + 2k$, F' is s-separated on G, and its level 1 is k-short on G, then the function $\mathrm{rd}^{F'} : \mathcal{G}^s \otimes_r F' \to \mathcal{G}^t \otimes_r F$ maps edges to edges.*

Proof.

(i) The equality $\pi^F \circ \mathrm{rd}^{F'} = \pi^{F'}$ is immediate from the fact that π is always just a projection on the first coordinate and since rd does not change the first coordinate. Meanwhile, $\pi_1^{F'}|_{F'-} = \pi_0^F \circ \mathrm{rd}^{F'}$ is immediate from the definition of $\mathrm{rd}^{F'}$.

(ii) By definition, $\tau = \mathrm{type}_{F'}(X)$ is the isomorphism class of $A^{F'}(X)$, and $\mathrm{type}_F(\mathrm{rd}^{F'}(X))$ is the isomorphism class of $A^F(\mathrm{rd}^{F'}(X))$. By Lemma 4.6, since $\mathrm{rd}^{F'}|_{F'(v)^-} : F'(v)^- \to F$ is a monomorphism, we have $A^F(\mathrm{rd}^{F'}(X)) = \mathrm{rd}^{F'}(A^{F'(v)^-}(X)) = \mathrm{rd}^{F'}(A^{F'-}(X))$, so $\mathrm{type}_F(\mathrm{rd}^{F'}(X))$ is the isomorphism class of $A^{F'-}(X)$. By Lemma 4.8, we also have $A^{F'-}(X) = A^{F'}(X)^-$, and by definition of τ^-, the isomorphism class of $A^{F'}(X)^-$ is τ^-. Altogether, $\mathrm{type}_F(\mathrm{rd}^{F'}(X)) = \tau^-$, as required.

(iii) Recall that edges of $\mathcal{G}^s \otimes F'$ are the r-sets $e \subseteq L(F')$ which have $||e||_G \leq s$. Part (ii) shows that $||\,\mathrm{rd}^{F'}(e)||_G \leq s + 2k \leq t$ — therefore $\mathrm{rd}^{F'}(e)$ is contained in an edge of $\mathcal{G}^t \otimes F$. Moreover, since F' is s-separated, every edge $e \in \mathcal{G}^s \otimes F'$ satisfies $|\pi(e)| = |e| = r$. Indeed, consider $u, v \in e$ distinct. If $\pi_0(u) = \pi_0(v)$ then $\pi(u) \neq \pi(v)$ by distinctness of u, v; if instead $\pi_0(u) \neq \pi_0(v)$ then $\pi\big(F(\pi_0(u))\big) \cap \pi\big(F(\pi_0(v))\big) = \emptyset$, showing $\pi(u) \neq \pi(v)$. Using (i), we have $|\,\mathrm{rd}^{F'}(e)| \geq |\pi(\mathrm{rd}^{F'}(e))| = |\pi(e)| = r$, which tells us that $|\,\mathrm{rd}^{F'}(e)| = r$, and so $\mathrm{rd}^{F'}(e)$ is an edge in $\mathcal{G}^s \otimes F'$. $\square$

4.5 Versatile sets

The goal in this section is to set up a framework under which we can concatenate tight paths/walks. This is based around the following concept.

Definition 4.36 (Versatile sets)**.** *For a balanced, height h ordered tree T, a subset $e \subseteq L(T)$ is* t-versatile *in T if for every balanced, height h ordered tree T' with at most t leaves, and monomorphism $\varphi : A^T(e) \to T'$, there is a monomorphism $\psi : T' \to T$ with $\psi \circ \varphi = \mathrm{id}$.*

We will prove that versatile sets exist inside all balanced d-ary trees with large enough d. This will involve an inductive argument where trees are built one vertex at a time via adding leaves. We first set up the theory of extending ordered trees like this. We remark that, unlike the rest of the paper, in this section we do not require trees to be balanced.

In a tree S, we call a path $P = v_1 \dots v_k$ *extendible* if $v_1 < \cdots < v_k$ (i.e. v_{i-1} is the parent of v_i for $i \in \{2, \dots, k\}$) and there is no vertex x satisfying $x > v_k$ which is comparable with v_1 (see Figure 7). It is useful to note that isomorphisms map extendible paths to extendible paths. I changed the definition from: $P = v_1 \dots v_k$ is extendible if there is no $x > v_k$ which is comparable with some v_i, with $i \in [2, k-1]$. It seems that this is what we use, and I find it a bit easier to understand

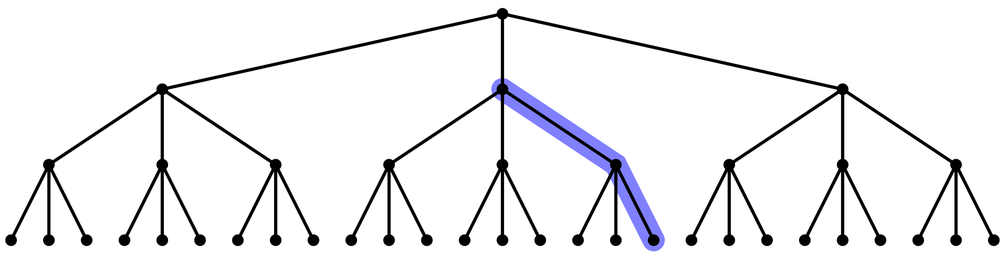

Figure 7: An extendible path in an ordered trees (with ordering of vertices in each level from left to right).

Observation 4.37. *A path $P = v_1 \dots v_k$ in an ordered tree S is extendible if and only if v_i is the latest child of v_{i-1} in the ordering of $V(S)$, for $i \in \{2, \dots, k\}$, and v_k is a leaf.*

Proof. For the "if" part, by definition we have $v_1 < \cdots < v_k$, showing that v_i is a child of v_{i-1}, for $i \in \{2, \dots, k\}$. Let u be the last descendent of v_1 according to the ordering of $V(S)$. Then u is a leaf, and $v_k = u$, implying that v_i is the latest child of v_{i-1} for every i, by (F2).

For the "only if" part, suppose that P has the property given in the observation, and suppose that P is not extendible. This means that there is a vertex x which satisfies $x > v_k$ and x is comparable with v_1. If x is an ancestor of v_1 then $x < v_1 < v_k$, a contradiction. If x is a descendant of v_1 then $x \leq v_k$, because v_k is the last descendant of v_1, again a contradiction. $\square$

Let $S \leq T$ be ordered trees and P an extendible path in S. We call T a P-extension of S if T has one extra vertex y, one extra edge $v_1 y$, and y is the successor of v_k in T (i.e. y is immediately after v_k in the ordering on $V(T)$).

Lemma 4.38. *Let $S \leq T$ be ordered trees sharing a root so that T has one extra vertex. Then T is a P-extension of S for some extendible path P in S.*

Proof. Let y be the vertex of $V(T) \setminus V(S)$, noting that it must be a leaf. Let $v_1 y$ be the edge of $E(T) \setminus E(S)$, and let v_0 be the predecessor of y in the ordering of $V(T)$. We claim that v_0, v_1 are comparable. Indeed, if not, then by part (F1) of the definition of ordered trees, we have $y > v_1$, and by choice of v_0 we have $v_1 < v_0 < y$. Thus, by part (F2) of the same definition, with $u = v_1, u' = y, v = v' = v_0$, we get $y < v_0$, a contradiction. So v_0, v_1 are comparable, which in fact implies that v_0 is a descendant of v_1 (because otherwise $v_0 < v_1 < y$, contrary to the choice of v_0 as y's predecessor). Thus, there is a path $P = v_1 \dots v_k = v_0$ in S, with $v_1 < \dots < v_k$.

We claim that P is extendible. To see this, consider a vertex x satisfying $x > v_k$, and suppose for contradiction that x, v_1 are comparable. If x is an ancestor of v_1, then $x < v_1 < v_k$, a contradiction. If x is a descendant of v_1, then v_1, y are incomparable (since y is a leaf, it is only comparable with v_1 and its ancestors). Then (F2) with $u = v_1, u' = x, v = v' = y$ gives $x < y$, a contradiction to v_k being the predecessor of y. $\square$

Repeatedly applying the above lemma gives the following.

Corollary 4.39. *Let $S \le T$ be ordered trees sharing a root. Then T can be constructed from S via a sequence of P-extensions.*

Extensions can be used to build monomorphisms as follows.

Observation 4.40. *Let S, T, S', T' be ordered trees, and P_S, P_T paths in S, T, such that S' is a P_S-extension of S, and T' is a P_T-extension of T. If $f : S \to T$ an isomorphism with $f(P_S) = P_T$, then f can be extended to an isomorphism $f : S' \to T'$.*

Proof. Let s'' be the vertex of $V(S') \setminus V(S)$, and t'' the vertex of $V(T') \setminus V(T)$. Extend f by defining $f(s'') = t''$. This is clearly a bijection, maps edges to edges (for this notice that all vertices in P_S precede s'' in S', and so if P_S has ends a, b with $a < b$, then b is the predecessor of s'', showing that s'' is adjacent to a. A similar argument holds for P_T, and then the claim follows by the fact that f is an isomorphism which maps P_S to P_T), and maps the predecessor of s'' to the predecessor of t'', which implies that f is order preserving, i.e. it is an isomorphism. $\square$

Our main building block for versatile sets is the following lemma.

Lemma 4.41. *Let $d \gg h, k, s$. Let T_0 be a balanced d-ary tree of height h. There is a sequence of trees $T_0 \ge \dots \ge T_k$ such that T_k is a balanced s-ary tree of height h and we have the following property:*

($*$) *For any $S \le T_i$ and extendible path P in S, there is a P-extension of S contained in T_{i-1}.*

Proof. Without loss of generality, we can assume that $d = 2^k(s+1) - 1$ (otherwise pass to a subtree which is $(2^k(s+1)-1)$-ary and continue the proof there). For each $i \in [k]$, set $d_i = 2^{k-i}(s+1) - 1$, noting that $d_k = s$ and $d_0 = d$. Label $V(T)$ by $([d] \cup \{0\})^h$ as explained in Notation 4.2. Define

$$X_i = 2^i \cdot [d_i] = \{2^i j : j \in [d_i]\},$$

noting that $[d] = X_0 \supseteq \dots \supseteq X_k$. Let T_i be the subtree of T consisting of vertices in $(X_i \cup \{0\})^h$, noting that T_i is a d_i-ary tree, and we have $T = T_0 \ge \dots \ge T_k$. For property ($*$), consider some $S \le T_i$ and extendible path P in S. Then P can be written as $v_0 \dots v_\ell$, with $v_0 < \dots < v_\ell$, where

$$\begin{aligned} v_0 &= (2^i a_1, \dots, 2^i a_t, 0, \dots, 0) \\ v_\ell &= (2^i a_1, \dots, 2^i a_t, 2^i a_{t+1}, \dots, 2^i a_{t+\ell}, 0, \dots, 0), \end{aligned}$$

for some $t, \ell \in [h]$ and $a_1, \ldots, a_{t+\ell} \in [d_i]$. Define

$$y = (2^i a_1, \ldots, 2^i a_t, 2^i a_{t+1} + 2^{i-1}, 0, \ldots, 0).$$

Notice that $y \in V(T_{i-1}) \setminus V(T_i)$ (as $2^i[d_i] + 2^{i-1} \subseteq 2^{i-1}[d_{i-1}]$, using $d_{i-1} \geq 2d_i + 1$) and that y is a child of v_0 in T_{i-1}, and add y to S to get a P-extension. To see that this is indeed a P-extension, we need to show that y is the successor of v_ℓ in S. First, observe that $v_\ell < y$ (as the first t coordinates of v_ℓ and y are the same, and the $(t+1)$-st coordinate of v_ℓ is smaller than that of y). It thus suffices to show that there is no vertex $x \in V(S)$ satisfying $v_\ell < x < y$. Suppose there is such $x = (x_1, \ldots, x_h)$. The first coordinate where x disagrees with v_ℓ and y must be $t+1$ — otherwise we would have $x < v_\ell$ or $x > y$. On the $(t+1)$-st coordinate we would need $2^i a_t < x_{t+1} < 2^i a_t + 2^{i-1}$ — but then $2^{-(i-1)} x_{t+1}$, which is an integer by definition of X_{i-1}, satisfies $2a_t < 2^{-(i-1)} x_{t+1} < 2a_t + 1$, a contradiction. □

The following is one of the key results in this section.

Lemma 4.42 (Existence of versatile sets)**.** *Let $d \gg h, s, t$. Let T be a balanced d-ary tree of height h and let $\tau \in \mathrm{Types}(h, s)$. There is subset $e \subseteq L(T)$ of type τ that is t-versatile.*

Proof. Consider a sequence $T = T_0 \geq \cdots \geq T_{ht}$, as given by Lemma 4.41. Let S be a type τ subtree of T_{ht} (which exists because T_{ht} is a balanced s-ary tree of height h), and set $e = L(S)$ (so $S = A^T(e)$ and e is type τ). Consider a height h tree T' with at most t leaves and a monomorphism $\varphi : S \to T'$ as in Definition 4.36 of "t-versatility". We need to find a monomorphism $\psi : T' \to T$ such that $\psi \circ \varphi = \mathrm{id}_S$.

Set $k := |T'| - |S|$, noting $k \leq ht$. Use Lemma 4.39 to get a sequence of trees $\varphi(S) = A'_0 \leq A'_1 \leq \cdots \leq A'_k = T'$ and paths $P'_i \subseteq A'_i$ such that each A'_{i+1} is constructed from A'_i via a P'_i-extension.

We build trees $A_i \subseteq T_{ht-i}$ and isomorphisms $\psi_i : A'_i \to A_i$, for $i \in \{0, \ldots, k\}$, as follows. Start with $A_0 := S$ and $\psi_0 = (\varphi|_{\varphi(S)})^{-1}$. Now assume that we have defined $A_0, \ldots, A_i$ and $\psi_0, \ldots, \psi_i$ with the desired properties. Since A'_{i+1} is built from A'_i via a P'_i-extension, $\psi_i(P'_i)$ is extendible in A_i. Thus, using $(*)$ and $A_i \leq T_{ht-i}$, there is a $\psi_i(P'_i)$-extension $A_{i+1} \leq T_{ht-i-1}$ of A_i. Use Observation 4.40 to extend ψ_i to an isomorphism $\psi_{i+1} : A'_{i+1} \to A_{i+1}$.

We claim that ψ_k satisfies the requirements, namely it is a function from T' to T (this holds by construction) and $\psi \circ \varphi = \mathrm{id}_S$. For the latter bit, by construction, we have $\psi_k|_{\varphi(S)} = \psi_k|_{A'_0} = \psi_0 = (\varphi|_{\varphi(S)})^{-1}$. Therefore $\psi_k(\varphi(x)) = \varphi^{-1}(\varphi(x)) = x$ for all $x \in V(S)$, as required. □

The following lemma is how we use versatile sets. Given an arbitrary tight walk $\mathcal{P}$, we get another one that goes between two versatile sets — without changing the types of edges in $\mathcal{P}$.

Lemma 4.43. *Let $\mathcal{H}$ be an r-uniform hypergraph, let F be a balanced $V(\mathcal{H})$-forest and $\mathcal{P}$ a tight walk in $\mathcal{H} \otimes F$ with $V(\mathcal{P}) \subseteq L(F)$ and write $\mathcal{P} = v_1 \ldots v_k$. Write $a = (v_1, \ldots, v_{r-1})$ and $b = (v_{k-r+2}, \ldots, v_k)$, and suppose that u_a, u_b are distinct roots of F such that $V(a) \subseteq L(F(u_a))$ and $V(b) \subseteq L(F(u_b))$.*

Suppose that $a' \subseteq L(F(u_a))$ and $b' \subseteq L(F(u_b))$ are k-versatile in $F(u_a)$ and $F(u_b)$, respectively, and satisfy $\mathrm{type}_F(a') = \mathrm{type}_F(a)$ and $\mathrm{type}_F(b') = \mathrm{type}_F(b)$. Then there is a tight walk $\mathcal{P}'$ in $\mathcal{H} \otimes F$ from a' to b' of order k with the property that the i-th edge of $\mathcal{P}'$ has the same root set and type as the i-th edge of $\mathcal{P}$.

Proof. Let $A = A^F\big(V(\mathcal{P}) \cap V(F(u_a))\big)$ and $B = A^F\big(V(\mathcal{P}) \cap V(F(u_b))\big)$, and note that A, B are trees with at most k leaves. Let $\varphi_a : A^F(a') \to A^F(a)$ and $\varphi_b : A^F(b') \to A^F(b)$ be isomorphisms (which exist because a, a' and b, b' have the same types). Noting that $A^F(a) \subseteq A$ and $A^F(b) \subseteq B$, we can use Definition 4.36 of "versatility" to get monomorphisms $\psi_a : A \to F(u_a)$ and $\psi_b : B \to F(u_b)$ with $\psi_a \circ \varphi_a = \mathrm{id}_{A^F(a')}$ and $\psi_b \circ \varphi_b = \mathrm{id}_{A^F(b')}$.

Let $\psi : A^F(V(\mathcal{P})) \to V(F)$ be the function which is ψ_a on A, ψ_b on B, and the identity everywhere else (this is well defined as A and B are vertex-disjoint by assumption on u_a, u_b). Note that ψ is the identity on roots and a monomorphism (if ψ fixes roots, and $\psi|_{F(r)}$ is a monomorphism for every r then it is a monomorphism). Denote by e_i the i-th edge of $\mathcal{P}$, and write $e_i' = \psi(e_i)$. We claim that e_i' is an edge in $\mathcal{H} \otimes F$. Indeed, because ψ is injective we have $|e_i'| = |\psi(e_i)| = |e_i| = r$. Moreover, $\pi_0(e_i') = \pi_0(\psi(e_i)) = \pi_0(e_i)$, because $\pi_0(\psi(x)) = \pi_0(x)$ for every vertex x in the range of ψ (as this holds for ψ_a, ψ_b, and the identity). By definition of $\mathcal{H} \otimes F$, since e_i is an edge we have $\pi_0(e_i)$ is a subset of an edge in $\mathcal{H}$, and thus the same holds for e_i', implying that e_i' is an edge in $\mathcal{H} \otimes F$. Define $\mathcal{P}' = \psi(\mathcal{P})$. Then, by e_i' being an edge in $\mathcal{H} \otimes F$ for every i, we have that $\mathcal{P}'$ is a tight path in $\mathcal{H} \otimes F$, whose i-th edge is e_i'.

Lemma 4.6 gives $\psi(A^F(e_i)) = A^F(\psi(e_i)) = A^F(e_i')$, and hence e_i, e_i' have the same type. We have already seen that the root set of e_i' has the same root set as e_i. Also, $\psi(a) = \psi(\varphi_a(a')) = \psi_a(\varphi_a(a')) = a'$ and $\psi(b) = \psi(\varphi_b(b')) = \psi_b(\varphi_b(b')) = b'$, by construction, showing that $\mathcal{P}'$ satisfies the requirements of the lemma. $\square$

5 Proof of the main theorem

The following is the main technical result of the paper — Theorem 1.1 will follow from it via a short argument. Here a *tight walk* of order k in an r-uniform hypergraph $\mathcal{H}$ is a sequence $v_1 \dots v_k$ of vertices in $\mathcal{H}$ that need not be distinct, such that $(v_i, \dots, v_{i+r-1})$ is an edge in $\mathcal{H}$ for $i \in [k - r + 1]$. The *degree* of a vertex v in a hypergraph is the number of edges containing v.

Theorem 5.1. *For all integers $r \geq 3$ and $s \geq 2$, there exists an integer p such that: for all sufficiently large n, there exists an r-uniform hypergraph $\mathcal{H}$ on n vertices, with maximum degree at most p, such that in every s-colouring of $\mathcal{H}$ there is a monochromatic tight walk of order at least n/p in which every vertex is used at most p times.*

Proof. Use Lemma 4.25 to pick a number $h := h(r,s)$ satisfying that lemma (i.e. if F is a balanced d-ary $V(G)$-forest of height h then no s-colouring of $\mathcal{G}^t \otimes_r F$ is 1-disconnected, for every non-negative integer t). Pick numbers $p, p', \varepsilon, a_1, b_1, c_1, d_1, \dots, a_h, b_h, c_h, d_h$, as follows.

$$p \gg p' \gg d_1 \gg b_1 \gg c_1 \gg a_1 \gg d_2 \gg b_2 \gg c_2 \gg a_2 \gg \dots$$
$$\dots \gg d_h \gg b_h \gg c_h \gg a_h \gg \varepsilon^{-1} \gg h, r, s.$$

Use Lemma 3.2 to obtain a connected graph G on n vertices which is an ε-expander with maximum degree at most $\Delta := \varepsilon^{-2}$. Order $V(G)$ arbitrarily. Let $\mathcal{H}$ be the r-uniform hypergraph $\mathcal{G}^{p'}$, noting that $\Delta(\mathcal{H}) \leq \binom{(\Delta+1)^{p'}}{r} \leq p$. Fix an s-colouring of $\mathcal{H}$, and suppose for contradiction that there is no monochromatic tight walk of length at least n/p in $\mathcal{H}$ in which every vertex is used at most p times.

Throughout the proof, all graphs of the form $\mathcal{G}^t \otimes F$ we consider satisfy $t \leq d_1$, and F being balanced, t-separated, of height at most h, and having all levels d_1-short. For such hypergraphs, Lemma 4.28 tells us that $\pi : \mathcal{G}^t \otimes F \to \mathcal{G}^{p'}$ is a hypergraph homomorphism (i.e. maps edges to edges). This gives a natural s-colouring on $\mathcal{G}^t \otimes F$ where we colour $e \in \mathcal{G}^t \otimes F$ by the colour of $\pi(e)$ in $\mathcal{H} = \mathcal{G}^{p'}$. Note that, this way, the following holds.

✧ If W is a monochromatic tight walk in $\mathcal{G}^t \otimes F$, with $t \leq d_1$, then $\pi(W)$ is a monochromatic tight walk in $\mathcal{H}$.

We will construct subsets $V(G) = U_1 \supseteq \cdots \supseteq U_h$, and $V(G)$-forests $F_1, \ldots, F_h$, with the following properties holding for every $i \in [h]$.

(F1) $|U_i| \geq (1 - 200\varepsilon)^i n$ (which in particular implies $|U_i| \geq (1 - 200h\varepsilon)n \geq n/2$).

(F2) Trees in F_i are balanced c_i-ary of height i with $\pi_0(F_i) = U_i$.

(F3) All levels in F_i are d_1-short on G, and F_i is b_i-separated on G.

(F4) The hypergraph $\mathcal{G}^{c_i} \otimes F_i$ is cleanly coloured.

(F5) The colouring on $\mathcal{G}^{c_i} \otimes F_i$ is c_i-disconnected on (G, F_i).

(F6) F_i is an augmentation of F_{i-1}, for $i \geq 2$.

We will show that under the assumption that $\mathcal{H}$ has no monochromatic tight walk of length at least n/p where every vertex is used at most p times, then sequences as above can indeed be constructed. In particular, we have a forest F_h which is a balanced d_h-ary $V(G)$-forest of height h, such that $\mathcal{G}^{d_h} \otimes F_h$ is 1-disconnected on (G, F_h). This is a contradiction to the choice of h, proving the theorem.

We begin by constructing U_1, F_1 satisfying (F1) – (F5). Set $U_1 = V(G)$ (ensuring (F1)). Noting that "G connected $\implies \delta(G^{d_1}) \geq d_1$", for each $v \in V(G)$ pick distinct neighbours $u_1^v, \ldots, u_{d_1}^v$ of v in G^{d_1}. Define $F_1''(v)$ by $V(F_1''(v)) = \{(v,v), (v, u_1^v), \ldots, (v, u_{d_1}^v)\}$, with (v,v) connected to all other vertices, making it an ordered $V(G)$-star (with the ordering of the leaves inherited from $V(G)$, and (v,v) preceding all leaves). Taking the union of these (and ordering them according to the ordering of the roots, inherited from the ordering of $V(G)$), we get a $V(G)$-forest F_1'' of (balanced) d_1-ary trees of height 1. Note that each $F_1''(v)$ has all levels d_1-short in G since it is a subgraph of G^{d_1}, implying that F_1'' has all levels d_1-short. Apply Lemma 4.15 (with $(F, G, d, b, h, \Delta)_{4.15} = (F_1'', G, d_1, b_1, 1, \Delta)$) to get a subforest $F_1' \leq F_1''$ of balanced b_1-ary trees of height 1 which is b_1-separated on G. Now apply Lemma 4.17 (with $(F, \mathcal{H}, d, b, h, r, s, \Delta)_{4.17} = (F_1', \mathcal{G}^{c_1}, b_1, c_1, 1, r, s, \binom{(\Delta+1)^{c_1}}{r})$) to get a c_1-ary subforest F_1 satisfying (F1) to (F4). Property (F5) holds by Observation 4.23.

Now suppose that for some $i \in [h-1]$, we have constructed U_i, F_i satisfying (F1) to (F5). We will show how to construct U_{i+1} and F_{i+1} satisfying (F1) to (F6), thereby proving the theorem. Consider an auxiliary colouring on the graph $G^{a_i}[U_i]$ with colour set $([s] \times \mathrm{Types}(i, r-1)) \cup \{\text{grey}\}$, such that each edge uv is coloured as follows.

- Colour uv with colour $(c, \tau) \in [s] \times \mathrm{Types}(i, r-1)$ if there are $(r-1)$-sets $U \subseteq L(F_i(u))$ and $V \subseteq L(F_i(v))$, both of type τ, and a colour c tight walk P in $\mathcal{G}^{c_i} \otimes F_i$ from U to V of length at most $3r$, which satisfies $||P||_G \leq c_i$.
- If the previous item fails for all $(c, \tau) \in [s] \times \mathrm{Types}(h, r-1)$, colour uv grey.

(Notice that edges uv could have more than one colour in $[s] \times \mathrm{Types}(i, r-1)$.) By Lemma 3.9 (applied to G with $(U, c, d, \varepsilon, s)_{3.9} = (U_i, a_i, d_{i+1} + 1, \varepsilon, s \cdot |\mathrm{Types}(i, r-1)|)$), either we get a non-grey monochromatic path of length n/a_i, or we can cover at least $(1 - 200\varepsilon)|U_i|$ vertices by vertex-disjoint grey $K_{d_{i+1}+1}$'s.

Claim 5.2. *There is no non-grey monochromatic path of length at most n/a_i.*

Proof. Suppose to the contrary that $P = p_1 \dots p_m$ is a monochromatic non-grey path of length at least n/a_i in the above colouring. Let $(c, \tau) \in [s] \times \text{Types}(i, r-1)$ be the colour of P. For each $j \in [m]$, use Lemma 4.42 (with $(T, \tau, d, h, s, t)_{4.42} = (F_i(p_j), \tau, c_i, i, r-1, 3r))$ to pick a type τ set $A_j \subseteq L(F_i(p_j))$ which is $3r$-versatile in $F_i(p_j)$. For each $j \in [m-1]$, since $p_j p_{j+1}$ is a colour (c, τ) edge, there is a colour c tight walk P_j in $\mathcal{G}^{c_i} \otimes F_i$ from a type τ subset $U_i \subseteq L(F_i(p_j))$ to a type τ subset $V_j \subseteq L(F_i(p_{j+1}))$, whose length is at most $3r$ and which satisfies $||P_j||_G \leq c_i$. By Lemma 4.43 (with $(F, \mathcal{H}, \mathcal{P}, u_a, u_b, a, b, a', b', k, r)_{4.43} = (F_i, \mathcal{G}^{c_i}, P_j, p_j, p_{j+1}, U_j, V_j, A_j, A_{j+1}, 3r, r))$, there is a tight walk Q_j from A_j to A_{j+1}, of the same length as P_j, and whose edges have the same types and roots as corresponding edges of P_j. Since the colouring of $G^{c_i} \otimes F_i$ is clean (by (F4)), we get that Q_j is also colour c. Thus $W = Q_1 \dots Q_{m-1}$ is a colour c tight walk in $\mathcal{G}^{c_i} \otimes F_i$.

Recall that, by $\diamondsuit$, the image $\pi(W)$ of W in $\mathcal{H}$ is a monochromatic tight walk, of length at least $n/a_i \geq n/p$. We claim that each vertex appears at most p times in $\pi(W)$. This would contradict the assumption on $\mathcal{H}$, proving the claim. To show that no vertex appears too much in $\pi(W)$, notice that any vertex u can appear in $\pi(Q_j)$ only for indices j with $d_G(u, p_j) \leq c_i + hd_1$ (suppose $u \in \pi(Q_j)$. We have $\pi_0(P_j) = \pi_0(Q_j)$, since P_j and Q_j have the same root set. It follows that $p_j \in \pi_0(P_j) = \pi_0(Q_j)$, and $||Q_j||_G = ||P_j||_G$. Using Lemma 4.27 gives $d_G(u, p_j) \leq ||Q_j||_G + hd_1 = ||P_j||_G + hd_1 \leq c_i + hd_1$). Therefore, every vertex is used at most $3r(\Delta+1)^{c_i+hd_1} \ll p$ times in $\pi(W)$, as claimed. $\square$

By the last claim, the first outcome does not hold, and so the second one does, meaning that we can cover at least $(1-200\varepsilon)|U_i|$ vertices by vertex-disjoint grey $K_{d_{i+1}+1}$'s. Let U_{i+1} be the set of vertices covered by by these cliques (which ensures that (F1) holds for U_{i+1}). For a $K_{d_{i+1}+1}$ with vertices $\{v_0, \dots, v_{d_{i+1}}\}$ define trees $F''_{i+1}(v_j)$ for $j \in [0, d_{i+1}]$, as follows.

$$F''_{i+1}(v_j) := \big(v_j; F_i(v_0); \dots; F_i(v_{j-1}); F_i(v_{j+1}); \dots; F_i(v_{d_{i+1}})\big)$$

(This is well defined due to $\pi(F_i(v_0)), \dots, \pi(F_i(v_{d_{i+1}}))$ being vertex-disjoint, which is true because (F3) tells us that F_i is b_i-separated in G, and because $d_G(v_s, v_t) \leq a_i \leq b_i$ for all i, as each $v_j v_{j'}$ is an edge in G^{a_i}.) Note that since $d_G(v_j, v_{j'}) \leq a_i$ for all j, j', level 1 of each $F''_{i+1}(v_j)$ is a_i-short, and all other levels are d_1-short (from Lemma 4.33 and (F3)). Define $F''_{i+1} = \bigcup_{v \in U_{i+1}} F''(v)$ (ordering so that $F''(u) < F''(v)$ if and only if $u < v$ in $V(G)$). Then F''_{i+1} is a $V(G)$-forest with $\pi_0(F''_{i+1}) = U_{i+1}$. By Lemma 4.32 and Observation 4.30, $\text{rd}^{F''_{i+1}}|_{F''_{i+1}(v)^-} : F''_{i+1}(v)^- \to F_i$ is an monomorphism for all $v \in U_{i+1}$, i.e. F''_{i+1} is an augmentation of F_i.

Apply Lemma 4.15 (with $(F, G, d, b, h, \Delta)_{4.15} = (F''_{i+1}, G, d_{i+1}, b_{i+1}, i+1, \Delta))$ to get a subforest $F'_{i+1} \leq F''_{i+1}$ of b_{i+1}-ary trees which is b_{i+1}-separated on G. Apply Lemma 4.17 (with $(F, \mathcal{H}, d, b, h, s, r, \Delta)_{4.17} = (F'_{i+1}, \mathcal{G}^{c_{i+1}}, b_{i+1}, c_{i+1}, i+1, s, r, \binom{(\Delta+1)^{c_{i+1}}}{r}))$ to get a c_{i+1}-ary subforest $F_{i+1} \leq F'_{i+1}$ satisfying (F1) to (F4) and (F6). It remains to show (F5).

Claim 5.3. *$\mathcal{G}^{c_{i+1}} \otimes F_{i+1}$ is c_{i+1}-disconnected on (F_{i+1}, G).*

Proof. Fix some $v \in v(G)$, as in the definition of "c_{i+1}-disconnected". Let $X, Y \subseteq L(F_{i+1}(v))$ be independent $(r-1)$-sets of the same type, and let P be a length at most $3r$ monochromatic tight walk from X to Y in $\mathcal{G}^{c_{i+1}} \otimes F_{i+1}$. Let c be the colour of P and τ be the type of X and Y. By Lemma 4.35 (iii) (applied with $(F', F, G, s, t, k)_{4.35} = (F_{i+1}, F_i, c_{i+1}, c_i, a_i))$, we have that $\text{rd}^{F_{i+1}}(P)$ is also a colour c tight walk in $\mathcal{G}^{c_i} \otimes F_i$ from $\text{rd}^{F_{i+1}}(X)$ to $\text{rd}^{F_{i+1}}(Y)$, and $\text{rd}^{F_{i+1}}(X)$

and $\mathrm{rd}^{F_{i+1}}(Y)$ are independent and are both of type τ^- (to see that $\mathrm{rd}^{F_{i+1}}(P)$ and P have the same colour, recall that for any edge $e \in \mathcal{G}^t \otimes F$ its colour is the colour of $\pi(e)$. By 4.35 (i), we have $\pi(\mathrm{rd}^{F_{i+1}}(e)) = \pi(e)$ showing that e and $\mathrm{rd}^{F_{i+1}}(e)$ have the same colour).

Let $\{v, u_1, \ldots, u_{d_{i+1}}\}$ be the grey clique containing v, noting that children of (v, v) in F_{i+1} are contained in $\{(u_1, v), \ldots, (u_{d_{i+1}}, v)\}$. This shows that for $\ell \in L(F(v))$, we have $\pi_0(\mathrm{rd}^{F_{i+1}}(\ell)) = \pi_1^{F_{i+1}}(\ell) \subseteq \{u_1, \ldots, u_{d_{i+1}}\}$. This tells us that $\pi_0(\mathrm{rd}^{F_{i+1}}(X)), \pi_0(\mathrm{rd}^{F_{i+1}}(Y)) \subseteq \{u_1, \ldots, u_{d_{i+1}}\}$, which, together with Observation 4.21, tells us that $\pi_0(\mathrm{rd}^{F_{i+1}}(X)) = x$ and $\pi_0(\mathrm{rd}^{F_{i+1}}(Y)) = y$ for some $x, y \in \{u_1, \ldots, u_{d_{i+1}}\}$ (possibly with $x = y$), or, equivalently, $\mathrm{rd}^{F_{i+1}}(X) \subseteq F_{i+1}(x)$ and $\mathrm{rd}^{F_{i+1}}(Y) \subseteq F_{i+1}(y)$.

We claim that $||\,\mathrm{rd}^{F_{i+1}}(P)||_G \geq c_i$. Indeed, if $x \neq y$, since the edge xy is not coloured (c, τ) (it is coloured grey; in particular, it is indeed an edge in G^{a_i}), we have $||\,\mathrm{rd}^{F_{i+1}}(P)||_G \geq c_i$. Otherwise, $x = y$ and then by c_i-disconnectedness of F_i, we again have $||\,\mathrm{rd}^{F_{i+1}}(P)||_G \geq c_i$. By Lemma 4.35 (ii), we have $||P||_G > ||\,\mathrm{rd}^{F_{i+1}}(P)||_G - 2a_i \geq c_i - 2a_i \geq c_{i+1}$, as required for showing c_{i+1}-disconnectedness. □

We have proved that F_{i+1} satisfies (F1) to (F6), completing the proof of the theorem. □

We use $\mathcal{H}[t]$ to denote the t-blow-up of a hypergraph $\mathcal{H}$ — the hypergraph formed by replacing each vertex v by t copies $v[1], \ldots, v[t]$ and letting $E(\mathcal{H}[t])$ be the set of r-sets e of form $\{u_1[i_1], \ldots, u_r[i_r]\}$ where $(u_1, \ldots, u_r) \in E(\mathcal{H})$ and $i_1, \ldots, i_r \in [t]$. Now we deduce the main theorem of the paper.

Proof of Theorem 1.1. Let $d \gg p, r, s$. Use Theorem 5.1 to find an r-uniform hypergraph $\mathcal{H}$ with n vertices, maximum degree at most p, and such that "in each s-colouring of $\mathcal{H}$ there is a tight walk of length at least n/p in which every vertex is used at most p times". To prove Theorem 1.1, it is sufficient to show that $\mathcal{H}[d] \xrightarrow{s} \mathcal{P}_{n/p}$ (since $e(\mathcal{H}[d]) \leq e(H)\binom{pr}{r} \leq n \cdot p\binom{pr}{r}$). Consider an s-colouring of $\mathcal{H}[d]$. Let S be a $V(\mathcal{H})$-forest formed by assigning an (arbitrary) size d star to each vertex, and define $\varphi : \mathcal{H}[d] \to \mathcal{H} \otimes S$, by mapping each of $v[1], \ldots, v[d]$ to a different leaf of the star $S(v)$, for every $v \in V(\mathcal{H})$. Notice that φ is an isomorphism, and colour $\mathcal{H} \otimes S$ according to φ, namely by colouring each edge e by the colour of $\varphi^{-1}(e)$.

Lemma 4.17 (applied with $(F, \mathcal{H}, d, b, h, s, r, \Delta)_{4.17} = (S, \mathcal{H}, d, p, 1, s, r, p)$), gives us a p-ary subforest $S' \leq S$ with $\mathcal{H} \otimes S'$ cleanly coloured. By possibly relabelling the indices of the vertices $v[1], \ldots, v[d]$ for $v \in V(\mathcal{H})$, we may assume that $\varphi^{-1}(L(S'(v))) = \{v[1], \ldots, v[p]\}$, for every $v \in V(\mathcal{H})$. Let $\mathcal{H}'$ be the subgraph of $\mathcal{H}[p]$ on $\{v[1] : v \in V(\mathcal{H})\}$, and notice that $\mathcal{H}'$ is isomorphic to $\mathcal{H}$. The property of Theorem 5.1 thus gives us a monochromatic, say colour c, tight walk $\mathcal{W}$ in $\mathcal{H}'$ of order at least n/p, that uses each vertex at most p times. Then $\mathcal{W}$ can be written as $w_1[1] \ldots w_\ell[1]$, where $w_i \in V(\mathcal{H})$ and $\ell \geq n/p$. Let m_i be the number of repetitions of the vertex w_i in the subsequence $w_1, \ldots, w_i$, noting that $m_i \leq p$ always. Now take $\mathcal{W}' = w_1[m_1] \ldots w_\ell[m_\ell]$.

We claim that $\mathcal{W}'$ is a colour c tight path of order ℓ in $\mathcal{H}[p]$. Indeed, first notice that $w_i[m_i]$ is a vertex in $\mathcal{H}[p]$ for every $i \in [\ell]$, as $m_i \leq p$. Second, notice that the vertices $w_i[m_i]$ are distinct, by definition of m_i. Third, since $(w_i[1], \ldots, w_{i+r-1}[1])$ is an edge in $\mathcal{H}'$, we have that $\{w_i, \ldots, w_{i+r-1}\}$ is an edge in $\mathcal{H}$, showing that $(w_i[m_i], \ldots, w_{i+r-1}[m_{i+r-1}])$ is an edge in $\mathcal{H}[p]$. Finally, by cleanliness we have that the colours of $(w_i[1], \ldots, w_{i+r-1}[1])$ and $(w_i[m_i], \ldots, w_{i+r-1}[m_{i+r-1}])$ are the same, and thus by choice of $\mathcal{W}$ they are both c. Therefore, indeed, $\mathcal{W}'$ is a tight path of order $\ell \geq n/p$ in $\mathcal{H}[p]$, proving the theorem. □

References

[1] D. Bal and L. DeBiasio, *New lower bounds on the size-Ramsey number of a path*, arXiv:1909.06354 (2019). 1, 2

[2] J. Beck, *On size Ramsey number of paths, trees, and circuits. I*, J. Graph Theory **7** (1983), 115–129. 1, 2

[3] I. Ben-Eliezer, M. Krivelevich, and B. Sudakov, *Long cycles in subgraphs of (pseudo)random directed graphs*, J. Graph Theory **70** (2012), 284–296. 4

[4] B. Bollobás, *Extremal graph theory with emphasis on probabilistic methods*, no. 62, American Math. Soc., 1986. 1

[5] ______, *Random graphs*, no. 73, Cambridge university press, 2001. 1

[6] D. Clemens, M. Jenssen, Y. Kohayakawa, N. Morrison, G. O. Mota, D. Reding, and B. Roberts, *The size-Ramsey number of powers of paths*, J. Graph Theory **91** (2019), 290–299. 2

[7] A. Dudek, S. La Fleur, D. Mubayi, and V. Rödl, *On the size-Ramsey number of hypergraphs*, J. Graph Theory **86** (2017), 104–121. 2

[8] A. Dudek and P. Prałat, *An alternative proof of the linearity of the size-Ramsey number of paths*, Combin. Probab. Comput. **24** (2015), 551–555. 1

[9] ______, *On some multicolor Ramsey properties of random graphs*, SIAM J. Discr. Math. **31** (2017), 2079–2092. 1, 2

[10] ______, *Note on the multicolour size-Ramsey number for paths*, Electron. J. Combin. **25** (2018), P3.35. 2

[11] P. Erdős, R. J. Faudree, C. C. Rousseau, and R. H. Schelp, *The size Ramsey number*, Periodica Mathematica Hungarica **9** (1978), 145–161. 1

[12] J. Han, M. Jenssen, Y. Kohayakawa, G. O. Mota, and B. Roberts, *The multicolour size-Ramsey number of powers of paths*, J. Combin. Theory, Ser. B **145** (2020), 359–375. 2

[13] J. Han, Y. Kohayakawa, S. Letzter, G. O. Mota, and O. Parczyk, *The size-Ramsey number of 3-uniform tight paths*, Adv. Combin. (2021), to appear. 2

[14] N. Kamčev, A. Liebenau, D. Wood, and L. Yepremyan, *The size Ramsey number of graphs with bounded treewidth*, arXiv:1906.09185 (2019). 2

[15] Y. Kohayakawa, T. Retter, and V. Rödl, *The size Ramsey number of short subdivisions of bounded degree graphs*, Random Struct Alg. **54** (2019), 304–339. 2

[16] M. Krivelevich, *Long cycles in locally expanding graphs, with applications*, Combinatorica **39** (2019), 135–151. 2

[17] S. Letzter, *Path Ramsey number for random graphs*, Combin. Probab. Comput. **25** (2016), 612–622. 1

[18] S. Letzter, A. Pokrovskiy, and L. Yepremyan, *Size-Ramsey numbers of powers of hypergraph trees and long subdivisions*, arXiv:2103.01942 (2021). 2, 4

[19] L. Lu and Z. Wang, *On the size-Ramsey number of tight paths*, SIAM J. Discr. Math. **32** (2018), 2172–2179. 2