Pure pairs. IX. Transversal trees

Alex Scott[1]
Mathematical Institute, University of Oxford, Oxford OX2 6GG, UK

Paul Seymour[2]
Princeton University, Princeton, NJ 08544

Sophie Spirkl[3]
University of Waterloo, Waterloo, Ontario N2L3G1, Canada

September 1, 2020; revised August 4, 2023

Abstract

Fix $k > 0$, and let G be a graph, with vertex set partitioned into k subsets ("blocks") of approximately equal size. An induced subgraph of G is "transversal" (with respect to this partition) if it has exactly one vertex in each block (and therefore it has exactly k vertices). A "pure pair" in G is a pair X, Y of disjoint subsets of $V(G)$ such that either all edges between X, Y are present or none are; and in the present context we are interested in pure pairs (X, Y) where each of X, Y is a subset of one of the blocks, and not the same block. This paper collects several results and open questions concerning how large a pure pair must be present if various types of transversal subgraphs are excluded.

1 Introduction

Graphs in this paper are finite, and without loops or parallel edges. Let $A, B \subseteq V(G)$ be disjoint. We say that A is *complete* to B, or A, B are *complete*, if every vertex in A is adjacent to every vertex in B, and similarly A, B are *anticomplete* if no vertex in A has a neighbour in B. A *pure pair* in G is a pair A, B of disjoint subsets of $V(G)$ such that A, B are complete or anticomplete. The number of vertices of G is denoted by $|G|$. The complement graph of G is denoted by $\overline{G}$. If $X \subseteq V(G)$, $G[X]$ denotes the subgraph induced on X.

A *blockade* $\mathcal{B}$ in a graph G is a sequence $(B_1, \ldots, B_k)$ of pairwise disjoint nonempty subsets of $V(G)$, called the *blocks* of $\mathcal{B}$; and the *width* of $\mathcal{B}$ is the minimum cardinality of its blocks, and its *length* is k. An induced subgraph H of G is $\mathcal{B}$-*rainbow* if each of its vertices belongs to a block of $\mathcal{B}$, and it has at most one vertex in each block; and $\mathcal{B}$-*transversal* if it is $\mathcal{B}$-rainbow and has exactly one vertex in each block (and therefore has exactly k vertices). A *copy* of a graph H in a graph G means an induced subgraph of G that is isomorphic to H.

In earlier papers of this sequence we proved several theorems that say that if we have a blockade $\mathcal{B}$ and there is no $\mathcal{B}$-rainbow copy of some special graph H, then there must be a pure pair X, Y with $|X|, |Y|$ large (in terms of the width of the blockade). For instance, in [1] we proved:

1.1 *For every forest H, there exists $d > 0$, such that, for every graph G with a blockade $\mathcal{B}$ of length at least d, if every vertex of G has degree less than W/d, and there is no anticomplete pair X, Y in G with $|X|, |Y| \geq W/d$ (where W is the width of $\mathcal{B}$), then there is a $\mathcal{B}$-rainbow copy of H in G.*

In this paper we investigate what happens if we ask for a $\mathcal{B}$-transversal copy of H rather than just a $\mathcal{B}$-rainbow copy.

This leads naturally to the question, what if we ask even more? Let H be a graph together with a fixed linear ordering of its vertex set, say $v_1, v_2, \ldots, v_k$. (We call this an "ordered graph".) If $\mathcal{B} = (B_1, \ldots, B_k)$ is a blockade in a graph G, an *ordered $\mathcal{B}$-transversal copy of H* means a $\mathcal{B}$-transversal induced subgraph J of G, such that there is an isomorphism ϕ from H to J, with $\phi(v_i) \in B_i$ for $1 \leq i \leq k$. Erdős, Hajnal and Pach [6] proved:

1.2 *For every ordered graph H, there exist $c, \varepsilon > 0$, such that for every graph G, and every blockade $\mathcal{B} = (B_1, \ldots, B_k)$ in G, where $k = |H|$, either:*

- *there is an ordered $\mathcal{B}$-transversal copy of H in G; or*

- *there are distinct $i, j \in \{1, \ldots, k\}$, and $X \subseteq B_i$ and $Y \subseteq B_j$, such that X, Y is a pure pair, and $|X|, |Y| \geq \varepsilon W^c$, where W is the width of $\mathcal{B}$.*

This is where we will start. We would particularly like to know, for which graphs H can we take $c = 1$ in the second bullet? (Sadly, almost never: only for graphs H with at most two vertices.) And, more promising: for which H can we get $|X| \geq \varepsilon W$ and $|Y| \geq \varepsilon W^c$ in the second bullet?

At the other extreme, we cannot get past the following open question, a variant of a conjecture of Conlon, Fox and Sudakov [2] that is discussed further in [3] (a "triangle" means a copy of the complete graph K_3):

1.3 Question: *Do there exist $c, \varepsilon > 0$ with the following property? Let $\mathcal{B} = (B_1, B_2, B_3)$ be a blockade in a graph G, and let W be its width. If there is no $\mathcal{B}$-transversal triangle, then there exist distinct $i, j \in \{1, 2, 3\}$ and $X \subseteq B_i$ and $Y \subseteq B_j$, such that X, Y is a pure pair, and $|X| \geq \varepsilon W$, and $|Y| \geq \varepsilon W^c$.*

1

Settling 1.3 was our initial goal in this research (although we were not able to do it), and it is easy to see that it can be reduced to the sparse case, when for all distinct $i, j \in \{1, 2, 3\}$, every vertex in B_i has at most $|B_j|/100$ neighbours in B_j (and 100 can be replaced by any other number).

We say a blockade $\mathcal{B} = (B_1, \ldots, B_k)$ has *local degree* λ if $\lambda \geq 0$ is the maximum of the number of neighbours of v in B_j, maximized over all distinct $i, j \in \{1, \ldots, k\}$ and all $v \in B_i$. (We set this to be zero if $k \leq 1$.) All our results concern blockades with local degree at most a small constant times the width. Also, we will look for pure pairs X, Y where X, Y are each a subset of a block, and not the same block; and we will no longer need pure pairs X, Y with X complete to Y. Let us say a blockade $\mathcal{B} = (B_1, \ldots, B_k)$ is (x, y)-*cohesive* if for all distinct $i, j \in \{1, \ldots, k\}$, there do not exist $X \subseteq B_i$ and $Y \subseteq B_j$ such that $|X| \geq x$, and $|Y| \geq y$, and X is anticomplete to Y. (This is true if $k \leq 1$.) Here are our main results, first for unordered copies of H:

1.4 *(Proved in 4.2 and 4.3.) If H is a cycle with $k \geq 4$ vertices, then there exist $\varepsilon, c > 0$ with the following property. Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade in a graph G, with local degree less than εW and $(\varepsilon W, \varepsilon W^c)$-cohesive, where W is its width. Then there is a $\mathcal{B}$-transversal copy of H in G.*

This statement for $k = 3$ is open and equivalent to 1.3.

1.5 *(Proved in 2.2.) For every integer $k \geq 1$, there exists $\varepsilon > 0$ with the following property. Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade in a graph G, with local degree less than εW and $(\varepsilon W, \varepsilon W)$-cohesive, where W is its width. Then there is a $\mathcal{B}$-transversal copy of a k-vertex path in G.*

The statement of 1.5 is also true for the tree obtained from a path with $k - 2$ vertices by adding two extra vertices, both adjacent to the last vertex of the path, provided that $k \geq 5$; and also for the tree obtained from a path with $k - t$ vertices by adding t extra vertices, each adjacent to the last vertex of the path, provided that $k \geq t^2 2^t$. (These are proved in 3.4.) It is *not* true for the tree obtained from a $(k - 6)$-vertex path by adding six extra vertices, three adjacent to the first vertex of the path and three adjacent to the last. (This is 3.1.)

For ordered copies of H, we have:

1.6 *(Proved in 5.1.) If H is an ordered tree with $k \geq 2$ vertices, then there exists $\varepsilon > 0$ with the following property. Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade in a graph G, with local degree less than εW and $(\varepsilon W, \varepsilon W^{1/(k-1)})$-cohesive, where W is its width. Then there is an ordered $\mathcal{B}$-transversal copy of H.*

For caterpillars we can strengthen this. A *caterpillar* is a tree in which all the vertices of degree more than one belong to one path.

1.7 *(Proved in 5.5.) If H is an ordered caterpillar with k vertices, then there exists $\varepsilon > 0$ with the following property. Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade in a graph G, with local degree less than εW, and $(\varepsilon W, \varepsilon W^{1/d})$-cohesive, where W is its width, and d is the maximum degree of H. Then there is an ordered $\mathcal{B}$-transversal copy of H.*

And a counterexample (if $t \geq 1$ is an integer, S_t denotes the star with $t + 1$ vertices, that is, the tree in which one vertex is adjacent to all the others):

1.8 *(Proved in 5.9.) Let $t \geq 3$ be an integer, and let S_t^+ be obtained from S_t by linearly ordering its vertex set. For all $c > 1/t$ and all $\varepsilon > 0$, there is a graph G and a blockade $\mathcal{B} = (B_1, \ldots, B_{t+1})$ in G, with local degree less than εW and $(\varepsilon W, \varepsilon W^c)$-cohesive where W is its width, such that there is no ordered $\mathcal{B}$-transversal copy of S_t^+ in G.*

2   Two easy covering theorems

We say that a graph H has the *strong transversal property* if there exists $\varepsilon > 0$ with the following property: for every graph G, if $\mathcal{B} = (B_1, \ldots, B_{|H|})$ is a blockade in G, with local degree less than εW and $(\varepsilon W, \varepsilon W)$-cohesive, where W is its width, then there is a $\mathcal{B}$-transversal copy of H in G. If this holds we say that ε is an *STP-coefficient*. We start with asking, which graphs have the strong transversal property? A sparse random graph of girth at least $|H| + 1$ shows that every such graph H must be a forest, and one might hope that all forests have the property, extending the results of [1], but this is not true, as we shall see. Nevertheless, some forests have the strong transversal property: here is what we know about them.

2.1 *Let H be a graph.*

- *If H is not a forest then H does not have the strong transversal property.*

- *If every component of H has the strong transversal property then so does H (the converse is false).*

- *If H is a path, then H has the strong transversal property.*

- *If $|H| > 4$ and H is obtained from a path by adding two new vertices adjacent to the last vertex of the path, then H has the strong transversal property.*

- *If $|H| > 2^t(t^2 - t + 1)$ and H is obtained from a path by adding t new vertices adjacent to the last vertex of the path, then H has the strong transversal property.*

- *If H is obtained from a path by adding six new vertices, three adjacent to the first vertex of the path and three adjacent to the last, then H does not have the strong transversal property.*

- *If H has a vertex of degree at least d where $2^{d-1} \geq |H|$, then H does not have the strong transversal property.*

In particular, it is *not* true that if a graph H has the property then so do all its induced subgraphs, or indeed all its components: the graph S_3 does not have the property (by the last bullet of 2.1) but if we add a vertex of degree zero, this five-vertex forest has the property. Indeed, it follows from one of the results of [1] that for any forest, if we add enough vertices of degree zero we will obtain a forest with the strong transversal property.

We will prove the various statements of 2.1 as separate theorems (except for the first two, which we leave to the reader).

If G is a graph and $A, B \subseteq V(G)$ are disjoint, we say A *covers* B if every vertex in B has a neighbour in A. For convenience, let us say a blockade $\mathcal{B} = (B_1, \ldots, B_k)$ in G is ε-*coherent* if for all distinct $i, j \in \{1, \ldots, k\}$:

- each vertex in B_i has fewer than $\varepsilon |B_j|$ neighbours in B_j, and

- there do not exist $X \subseteq B_i$ and $Y \subseteq B_j$ with $|X| \geq \varepsilon |B_i|$ and $|Y| \geq \varepsilon |B_j|$ and X anticomplete to Y.

This is very much like ε-cohesion, but is different when the blocks have different sizes. This new definition is not really needed, but it works nicely and is a little more compact than using cohesion (and we used the same concept in earlier papers). Let us first prove the third statement of 2.1, that is, 1.5, which we restate in a slightly stronger form:

2.2 *Let $k \geq 2$ be an integer, and $0 < \varepsilon \leq 1/(2k-2)$. Let $\mathcal{B} = (B_1, \ldots, B_k)$ be an ε-coherent blockade in a graph G. Then there is a $\mathcal{B}$-transversal k-vertex path in G with an end-vertex in B_1.*

Proof. We define $t_1, \ldots, t_k$ with $\{t_1, \ldots, t_k\} = \{1, \ldots, k\}$, and $A_i \subseteq B_i$ for $1 \leq i \leq k$, as follows. Let $t_1 = 1$. Inductively, let $1 \leq i \leq k$, and suppose that $t_1, \ldots, t_i$ and $A_{t_1}, \ldots, A_{t_{i-1}}$ have been defined, with the properties that

- $\emptyset \neq A_{t_h} \subseteq B_{t_h}$ for $1 \leq h < i$;

- for $1 \leq h < i-1$, A_{t_h} covers $A_{t_{h+1}}$;

- for $1 \leq g < h < i$ with $h - g \geq 2$, there are no edges between A_{t_g} and A_{t_h};

- for each $j \in \{1, \ldots, k\} \setminus \{t_1, \ldots, t_i\}$, at least $(1 - 2(i-1)\varepsilon)|B_j|$ vertices in B_j have no neighbour in $A_{t_1} \cup \cdots \cup A_{t_{i-1}}$; and

- if $i \geq 2$, at least $\varepsilon|B_{t_i}|$ vertices in B_{t_i} have a neighbour in $A_{t_{i-1}}$ and have no neighbour in $A_{t_1} \cup \cdots \cup A_{t_{i-2}}$.

Let $J = \{1, \ldots, k\} \setminus \{t_1, \ldots, t_i\}$. For each $j \in J$, let C_j be the set of vertices in B_j with no neighbour in $A_{t_1} \cup \cdots \cup A_{t_{i-1}}$; thus $|C_j| \geq (1 - 2(i-1)\varepsilon)|B_j|$. If $i = 1$ let $D = B_{t_i}$, and if $i \geq 2$ let D be the set of vertices in B_{t_i} that have a neighbour in $A_{t_{i-1}}$ and have no neighbour in $A_{t_1} \cup \cdots \cup A_{t_{i-2}}$; thus $|D| \geq \varepsilon|B_{t_i}|$. If $i = k$, let $A_{t_i} = D$ and the inductive definition is complete, so we assume that $i \leq k-1$. For each $j \in J$, fewer than $\varepsilon|B_j|$ vertices in B_j have no neighbour in D, since $|D| \geq \varepsilon|B_{t_{i+1}}|$ and the blockade is ε-coherent; and since $|C_j| \geq (1 - 2(i-1)\varepsilon)|B_j| \geq 2\varepsilon|B_j|$, at least $\varepsilon|B_j|$ vertices in C_j have a neighbour in D.

Since $J \neq \emptyset$, there exists $A_{t_i} \subseteq D$ minimal such that for some $j \in J$, at least $\varepsilon|B_j|$ vertices in C_j have a neighbour in A_{t_i}. From the minimality of A_{t_i}, for each $j \in J$ there are fewer than $2\varepsilon|B_j|$ vertices in C_j with a neighbour in A_{t_i}, and hence there are at least $|C_j| - 2\varepsilon|B_j| \geq (1 - 2i\varepsilon)|B_j|$ vertices in C_j with no neighbour in A_{t_i}. Choose $j \in J$ such that at least $\varepsilon|B_j|$ vertices in C_j have a neighbour in A_{t_i}, and define $t_{i+1} = j$. This completes the inductive definition.

Choose $a_{t_k} \in A_{t_k}$. Since $A_{t_{k-1}}$ covers A_{t_k}, there exists $a_{t_{k-1}} \in A_{t_{k-1}}$ adjacent to a_{t_k}; and similarly for $i = k-2, k-3, \ldots, 1$ there exists $a_{t_i} \in A_{t_i}$ adjacent to $a_{t_{i+1}}$. But for $1 \leq i < j \leq k$ with $j \geq i+2$, there are no edges between A_{t_i}, A_{t_j}; so a_{t_1}-a_{t_2}-$\cdots$-a_{t_k} is an induced path. This proves 2.2. ∎

A somewhat similar proof (the proofs of 2.2 and 2.3 are both specializations of the proof of the main theorem of [7]) shows:

2.3 *Let $k \geq 1$ be an integer, let $K = 2^{k-1} + 1$, and let $0 < \varepsilon \leq 3^{-K}$. Let $\mathcal{B} = (B_1, \ldots, B_K)$ be an ε-coherent blockade in a graph G. Then there is a $\mathcal{B}$-rainbow copy of S_k in G.*

4

Proof. If $J \subseteq \{1, \ldots, K\}$, a *star-partition* of $\mathcal{B}$ (see figure 1) with *element set* J is a sequence

$$(h_1, I_1, h_2, I_2, \ldots, h_t, I_t)$$

where $J = \{h_1, \ldots, h_t\} \cup I_1 \cup \cdots \cup I_t$, and a choice of a nonempty subset $A_i \subseteq B_i$ for each $i \in J$, with the following properties:

- $t \geq 1$, and $h_1, \ldots, h_t \in \{1, \ldots, K\}$ are distinct, and the sets $I_1, \ldots, I_t$ are pairwise disjoint subsets of $\{1, \ldots, K\} \setminus \{h_1, h_2, \ldots, h_t\}$;

- for $1 \leq s \leq t$, and all $j \in I_s$, A_j covers A_{h_s};

- for $1 \leq s \leq t$ and all $j \in I_s$, A_j is anticomplete to

  - all the sets $A_{j'}$ for $j' \in I_s \setminus \{j\}$;
  - all the sets $A_{h_{s'}}$ for $s' \in \{1, \ldots, t\} \setminus \{s\}$; and
  - all the sets $A_{j'}$ for $s' \in \{1, \ldots, t\} \setminus \{s\}$ and $j' \in I_{s'}$.
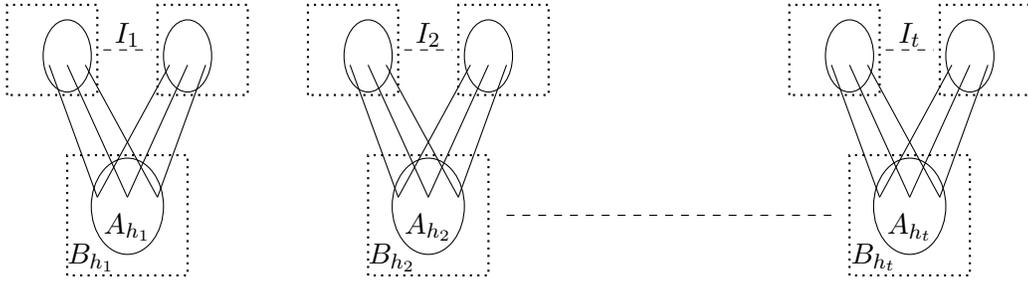

Figure 1: A star-partition.

The *linkage* of a star-partition is the maximum over all distinct $i, j \in \{h_1, \ldots, h_t\}$ of the maximum over $v \in A_i$ of $n/|B_j|$ where n is the number of neighbours of v in A_j (or zero if $t \leq 1$). Its *length* is t, and its *value* is $2^{|I_1|} + \cdots + 2^{|I_t|}$. There is a star-partition with linkage less than ε, length K and value K, since we may set $t = K$ and the sets $I_1, \ldots, I_K$ all empty, and $A_i = B_i$ for $1 \leq i \leq K$. Choose $t \geq 1$ minimum such that there is a star-partition with linkage less than $\varepsilon 3^{K-t}$, length t and value at least K, say

$$(h_1, I_1, h_2, I_2, \ldots, h_t, I_t)$$

and $A_i \subseteq B_i$ for each $i \in J$, where J is its element set.

Suppose that $t \geq 2$. We may assume that $|I_t| \leq |I_1|, \ldots, |I_{t-1}|$. Choose $C_{h_t} \subseteq A_{h_t}$ minimal such that C_{h_t} covers at least one-third of one of the sets $A_{h_1}, \ldots, A_{h_{t-1}}$, say of A_{h_1}. Let C_{h_1} be the set of vertices in A_{h_1} with a neighbour in C_{h_t}, and for $2 \leq s \leq t - 1$, let C_{h_s} be the set of vertices in A_{h_t} with no neighbour in C_{h_t}. Thus $|C_{h_1}| \geq |A_{h_1}|/3$, and from the minimality of C_{h_t}, and since the linkage is less than $\varepsilon 3^{K-t}$, it follows that

$$|C_{h_s}| \geq (2/3 - \varepsilon 3^{K-t})|A_{h_s}| \geq |A_{h_s}|/3$$

5

for $2 \le s \le t - 1$. Let $J' = J \setminus I_t$, and for each $i \in J'$ with $i \notin \{h_1, \ldots, h_{t-1}, h_t\}$ let $C_i = A_i$. Then

$$(h_1, I_1 \cup \{h_t\}, h_2, I_2, \ldots, h_{t-1}, I_{t-1})$$

and the sets C_i $(i \in J')$ form a star-partition with linkage less than $\varepsilon 3^{K-t+1}$, length $t - 1$ and value at least K, contrary to the minimality of t.

Hence $t = 1$, and so $2^{|I_1|} \ge K$, and therefore $|I_1| \ge k$. We may assume that $h_1 = 1$, and $2, \ldots, k + 1 \in I_1$. Choose $u \in A_1$; then for $2 \le i \le k + 1$ there exists $v_i \in A_i$ adjacent to u, and the subgraph induced on $\{u, v_1, \ldots, v_k\}$ is a $\mathcal{B}$-rainbow copy of S_k. This proves 2.3. ∎

Next we show that the expression $2^{k-1} + 1$ in 2.3 is best possible, and hence the final statement of 2.1 holds, because of the following:

2.4 *For every integer $k \ge 2$, and all $\varepsilon > 0$, there is a graph G, and an ε-coherent blockade $\mathcal{B}$ in G of length 2^{k-1}, such that there is no $\mathcal{B}$-rainbow copy of S_k in G.*

The proof needs the following two lemmas, which will also be needed later in the paper. The first is a standard estimate:

2.5 *If $n \ge k \ge 1$ are integers then $\binom{n}{k} \le (en/k)^k$ (where e is Euler's number).*

Proof. By Stirling's formula [8], we have

$$k! \ge \sqrt{2\pi} k^{k+1/2} e^{-k+1/(12k+1)} \ge (k/e)^k$$

for $k \ge 1$, and so

$$\binom{n}{k} \le n^k/k! \le n^k (k/e)^{-k} = (en/k)^k.$$

This proves 2.5. ∎

The second lemma is also a well-known result.

2.6 *Let $\varepsilon > 0$; then there exists $d > 0$ such that for all sufficiently large n, there is a bipartite graph with bipartition (A, B), where $|A| = |B| = n$, such that every vertex has degree less than d, and there do not exist anticomplete sets $A' \subseteq A$ and $B' \subseteq B$ with $|A'|, |B'| \ge \varepsilon n$.*

Proof. Choose $c > 4\varepsilon^{-2} \log 2$, and choose d with $2 \log 2 < d \log(d/(2ce))$. Now let n be sufficiently large, and let A, B be disjoint sets both of cardinality $2n$. For each $u \in A$ and $v \in B$, let u, v be adjacent with probability c/n, independently. The expected number of anticomplete pairs (A', B') with $A' \subseteq A$ and $B' \subseteq B$ and with $|A'|, |B'| \ge \varepsilon n$ is at most

$$2^{4n}(1 - c/n)^{\varepsilon^2 n^2} \le 2^{4n} e^{-c\varepsilon^2 n} \le 1/2$$

(since $c\varepsilon^2 > 4 \log 2$) if n is sufficiently large. The probability that a given vertex $v \in A \cup B$ has degree at least d is at most

$$\binom{2n}{d}(c/n)^d \le (2en/d)^d (c/n)^d = (2ce/d)^d,$$

6

by 2.5. Consequently the probability that at least n vertices in A have degree at least d is at most $2^{2n}(2ce/d)^{nd} < 1/4$ (since $2\log 2 < d\log(d/(2ce))$) if n is sufficiently large. Hence, if n is sufficiently large, with positive probability there is no anticomplete pair (A', B') with $A' \subseteq A$ and $B' \subseteq B$ and with $|A'|, |B'| \geq \varepsilon n$, and at most n vertices in A, and at most n vertices in B have degree at least d. Thus by deleting the n vertices in A with largest degree, and the same for B, we obtain a graph satisfying the theorem. This proves 2.6. ∎

To prove 2.4, for inductive purposes we will prove something a little stronger, the following.

2.7 *Let $\varepsilon > 0$. For every integer $k \geq 2$, and every integer $p \geq 0$, there exists $W(k, p)$ such that for all integers $W \geq W(k, p)$, if H is a graph with $2^{k-1}W$ vertices, and with maximum degree at most p, there is a graph G with the same vertex set and with H as a subgraph, and an ε-coherent blockade $\mathcal{B}$ in G of length 2^{k-1}, such that there is no $\mathcal{B}$-rainbow copy of S_k in G.*

Proof. We prove 2.7 by induction on k. Suppose first that $k = 2$. Choose d and n_0 such that 2.6 holds for all $n \geq n_0$. Let $W(2, p) = \max(n_0, (p + d)/\varepsilon)$; we claim that $W(2, p)$ satisfies the theorem. Let $W \geq W(2, p)$, and let H be a graph with $2W$ vertices, and with maximum degree at most p. Let B_1, B_2 be two disjoint subsets of $V(H)$ both of cardinality W. By 2.6 there is a graph J with bipartition (B_1, B_2), such that every vertex has degree less than d, and there is no anticomplete pair (A', B') with $A' \subseteq B_1$ and $B' \subseteq B_2$ and with $|A'|, |B'| \geq \varepsilon W$, that is, in J the blockade $\mathcal{B} = (B_1, B_2)$ has local degree less than d and is $(\varepsilon W, \varepsilon W)$-cohesive. Let G be the union of H, J; then in G the same blockade $\mathcal{B}$ has local degree less than $p + d$ and is $(\varepsilon W, \varepsilon W)$-cohesive. Since it only has two blocks and therefore there is no $\mathcal{B}$-rainbow copy of S_2, the result holds.

Now we assume inductively that $k \geq 3$ and the theorem holds for $k - 1$. Choose d and n_0 such that 2.6 holds for all $n \geq n_0$, with ε replaced by $2^{2-k}\varepsilon$.

Let
$$W(k, p) = \max\left(n_0, (p + d)/\varepsilon, W(k - 1, p + (p + d)^2)\right).$$

We claim that $W(k, p)$ satisfies the theorem. Let $W \geq W(k, p)$, and let H be a graph with $2^{k-1}W$ vertices, and with maximum degree at most p. Let V_1, V_2 be two disjoint subsets of $V(H)$ both of cardinality $2^{k-2}W$. By 2.6 there is a graph J with bipartition (V_1, V_2), such that every vertex has degree less than d, and there is no anticomplete pair (A', B') with $A' \subseteq V_1$ and $B' \subseteq V_2$ and with

$$|A'|, |B'| \geq (2^{2-k}\varepsilon)|A'| = \varepsilon W.$$

For $i = 1, 2$, let H_i be the graph with vertex set V_i, in which distinct vertices $u, v \in V_i$ are adjacent if and only if either they are H-adjacent, or they have a common $(H \cup J)$-neighbour in $V(H) \setminus V_i$. It follows that H_i has maximum degree at most $p + (p + d)^2$. Since $W(k, p) \geq W(k - 1, p + (p + d)^2)$, the inductive hypothesis implies that for $i = 1, 2$ there is a graph G_i with vertex set V_i and with H_i as a subgraph, and an ε-coherent blockade $\mathcal{B}_i$ in G_i of length 2^{k-2}, such that there is no $\mathcal{B}_i$-rainbow copy of S_{k-1} in G_i. Let G be the union of G_1, G_2 and $H \cup J$, and let $\mathcal{B}$ be the blockade with blocks all the blocks of $\mathcal{B}_1$ and all those of $\mathcal{B}_2$ (in some order). It follows that, in G, the blockade $\mathcal{B}$ is ε-coherent (since $p + d \leq \varepsilon W$). Suppose there is a $\mathcal{B}$-rainbow copy T of S_k in G, and let v be the vertex of T_k that has degree k in T_k. From the symmetry we may assume that $v \in V_1$. Since all neighbours of v in V_2 are pairwise adjacent (since $H_2 \subseteq G_2 \subseteq G$) and T is induced in G, it follows that at most one vertex of T belongs to V_2, and so there is a $\mathcal{B}_1$-rainbow copy of S_{k-1} in G_1, a contradiction. This proves 2.7. ∎

7

3 Brooms

Let P be a path with vertices $p_1, \ldots, p_k$ in order; and let H be obtained from P by adding t new vertices, each adjacent to p_k. We define $B(k,t) = H$; such a graph is a *broom*.

If instead we add $s + t$ new vertices to P, s of them adjacent to p_1 and the other t to p_k, the graph we produce is called a *double broom* and is denoted $B(k, s, t)$. We still have three parts of 2.1 to prove (namely that $B(k, 2)$ has the strong transversal property, and so does $B(k, t)$ if $k \gg t$, and that $B(k, 3, 3)$ does not), and we will do that in this section. We begin with the easiest:

3.1 *For every integer $k \geq 1$, the double broom $B(k, 3, 3)$ does not have the strong transversal property.*

Proof. Suppose that $B(k, 3, 3)$ has the strong transversal property, with STP-coefficient ε. Choose d and n_0 such that 2.6 holds for all $n \geq n_0$. Let

$$W = \max\left(n_0, \left(((k+3)d + 2d + 3(k+3)d^2)^4 + d + 3(k+3)d^2\right)/\varepsilon\right).$$

Let $B_1, \ldots, B_{k+6}$ be pairwise disjoint sets each of cardinality W. Let $\mathcal{B} = (B_1, \ldots, B_{k+6})$, and $V_1 = B_{k+4} \cup B_{k+5} \cup B_{k+6}$, and $V_2 = B_1 \cup \cdots \cup B_{k+3}$. For $1 \leq i < j \leq k+6$, let $J_{i,j}$ be a graph with bipartition (B_i, B_j) with maximum degree less than d, such that there is no anticomplete pair A', B' with $A' \subseteq B_i$ and $B' \subseteq B_j$ and $|A'|, |B'| \geq \varepsilon W$. Let J be the union of all the graphs $J_{i,j}$. Let L be the graph with vertex set V_1 in which distinct u, v are adjacent if there is a $\mathcal{B}$-rainbow path in J with ends u, v, of length one or two and with its interior vertex (if any) in V_2. Let R be the graph with vertex set V_2 in which distinct u, v are adjacent if there is a $\mathcal{B}$-rainbow path of $J \cup L$ with ends u, v and with every internal vertex in V_1. Let $G = J \cup L \cup R$.

Since J is a subgraph of G it follows that $\mathcal{B}$ is $(\varepsilon W, \varepsilon W)$-cohesive in G. The only edges of G between V_1, V_2 are those of J; and L has maximum degree at most $2d + 3(k+3)d^2$, since each vertex in V_1 has degree at most $(k+3)d$ in J, and each of those neighbours has degree at most $3d$ in J. If P is a $\mathcal{B}$-rainbow path of $J \cup L$ with ends in V_2 and with every internal vertex in V_1, then P has at most five vertices; and since $J \cup L$ has maximum degree at most $(k+3)d + 2d + 3(k+3)d^2$, it follows that each vertex in V_2 is an end of at most $((k+3)d + 2d + 3(k+3)d^2)^4$ such paths, and so R has maximum degree at most $((k+3)d + 2d + 3(k+3)d^2)^4$.

Consequently $\mathcal{B}$ has local degree at most

$$(k+3)d + 2d + 3(k+3)d^2)^4 + d + 3(k+3)d^2 < \varepsilon W.$$

From the choice of ε, there is a $\mathcal{B}$-rainbow copy H of $B(k, 3, 3)$ in G. Let it be constructed from a k-vertex path with vertices $p_1, \ldots, p_k$ in order, by adding three new vertices q_1, q_2, q_3 adjacent to p_1 and three new vertices r_1, r_2, r_3 adjacent to p_k. Every path of H with both ends in V_2 has all its internal vertices in V_2, since otherwise there would be an induced $\mathcal{B}$-rainbow path of G with both ends in V_2 and all internal vertices in V_1, with its ends nonadjacent; and this would contradict the construction of R. If one of $p_1, \ldots, p_k$ belongs to V_1 then one of q_1, q_2, q_3 and one of r_1, r_2, r_3 belongs to V_2, contradicting the claim just made; so $p_1, \ldots, p_k \in V_2$. So exactly three of $q_1, q_2, q_3, r_1, r_2, r_3$ belong to V_1; and so two of them have a common neighbour in V_2, contradicting the definition of L. This proves that there is no such ε, and so proves 3.1. ∎

Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade in G, and let J be a digraph with vertex set $\{1, \ldots, k\}$. (Digraphs in this paper do not have loops or parallel edges, but they may have antiparallel edges. Thus, if there is an edge from i to j then it is unique, but there might also be an edge from j to i.) For $\tau > 0$, we say that J is a τ-*covering digraph* for $\mathcal{B}$ if for $1 \le i \le k$ there exists $A_i \subseteq B_i$, and for each edge ij of J there exists $X_{ij} \subseteq B_i$, with the following properties:

- $|A_i| \ge \tau^{|E(J)|}|B_i|$ for $1 \le i \le k$;

- for each edge ij of J, X_{ij} covers A_j and X_{ij} is anticomplete to A_h for all $h \in \{1, \ldots, k\} \setminus \{i, j\}$; and

- for all edges $ij, i'j'$ of J with $i \ne i'$, and $i \ne j'$ and $i' \ne j$, the sets $X_{ij}, X_{i'j'}$ are anticomplete.

We call the sets $(A_1, \ldots, A_k)$ a *core* for J. There is a τ-covering digraph, because we can take J with no edges and $A_i = B_i$ for each i. A τ-covering digraph for $\mathcal{B}$ is *optimal* if no τ-covering digraph for $\mathcal{B}$ has strictly more edges. If $X \subseteq V(G)$ we denote by $N(X)$ or $N_G(X)$ the set of vertices in $V(G) \setminus X$ that have a neighbour in X.

3.2 *Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade in a graph G, let $\tau > 0$, and let J be an optimal τ-covering digraph for $\mathcal{B}$, with core $(A_1, \ldots, A_k)$. Suppose that for all distinct $i, j \in \{1, \ldots, k\}$, every vertex in B_i has fewer than $(1 - 2\tau)\tau^{|E(J)|}|B_j|$ neighbours in B_j. Let $1 \le i \le k$ and $X \subseteq A_i$. Then either*

- *there exists $j \in \{1, \ldots, k\} \setminus \{i\}$ such that $ij \in E(J)$ and $|A_j \setminus N(X)| < \tau^{|E(J)|+1}|B_j|$; or*

- *for every $j \in \{1, \ldots, k\} \setminus \{i\}$ such that $ij \notin E(J)$ we have $|A_j \cap N(X)| < \tau^{|E(J)|+1}|B_j|$.*

Proof. Let $s = |E(J)|$. We assume the second bullet of the theorem is false, so there exists $j \in \{1, \ldots, k\} \setminus \{i\}$ with $ij \notin E(J)$ such that $|A_j \cap N(X)| \ge \tau^{s+1}|B_j|$. Choose $Y \subseteq X$ minimal such that there exists $j \in \{1, \ldots, k\} \setminus \{i\}$ with $ij \notin E(J)$ such that $|A_j \cap N(Y)| \ge \tau^{s+1}|B_j|$. Let $C_j = A_j \cap N(Y)$, let $C_i = A_i$, and for each $h \in \{1, \ldots, k\} \setminus \{i, j\}$ let $C_h = A_h \setminus N(Y)$. Thus $|C_j| \ge \tau^{s+1}|B_j|$. If $|C_h| \ge \tau^{s+1}|B_h|$ for every $h \in \{1, \ldots, k\} \setminus \{i, j\}$, then adding the edge ij to J gives a τ-covering digraph for $\mathcal{B}$ with core $(C_1, \ldots, C_k)$, contrary to the optimality of J. Thus there exists $h \in \{1, \ldots, k\} \setminus \{i, j\}$ with $|C_h| < \tau^{s+1}|B_h|$. If $ih \notin E(J)$, the minimality of Y and the hypothesis about local degree imply that

$$|A_h \setminus C_h| < \left(\tau^{s+1} + ((1 - 2\tau)\tau^s)\right)|B_h| = \left(\tau^s - \tau^{s+1}\right)|B_h|,$$

and so

$$|C_h| \ge \tau^{s+1}|B_h|,$$

a contradiction. Thus $ih \in E(J)$. Since $|C_h| < \tau^{s+1}|B_h|$, and

$$A_h \setminus N(X) \subseteq A_h \setminus N(Y) = C_h,$$

the first bullet of the theorem holds. This proves 3.2. ∎

We deduce:

3.3 *Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade of length at least two in a graph G, and let J be an optimal τ-covering digraph for $\mathcal{B}$, with core $(A_1, \ldots, A_k)$. If $\mathcal{B}$ is $(1 - 2\tau)\tau^{|E(J)|}$-coherent, then every vertex of J has outdegree at least one.*

Proof. Suppose that i has outdegree zero in J; we may assume that $i = 1$. Let $|E(J)| = s$. Since $k \geq 2$ and $\mathcal{B}$ is $(1 - 2\tau)\tau^s$-coherent and $|A_1| \geq \tau^s |B_1|$, it follows that $|A_2 \setminus N(A_1)| < (1 - 2\tau)\tau^s |B_2|$, and so

$$|A_2 \cap N(A_1)| > |A_2| - (1 - 2\tau)\tau^s |B_2| \geq \tau^{s+1} |B_2|.$$

But then both the outcomes of 3.2 (with $i = 1$ and $X = A_1$) are false, a contradiction. This proves 3.3. ∎

We use these results to prove one of the remaining parts of 2.1:

3.4 *Let $k, t \geq 0$ be integers with $t \geq 2$ and $k \geq 2^t(t^2 - t + 1) - t + 1$. Then $B(k, t)$ has the strong transversal property.*

Proof. Let $\tau = 1/6$, and let $\varepsilon = \tau^{(k+t)^2} 3^{-k}$; we will show that $B(k, t)$ has the strong transversal property with STP-coefficient ε.

Thus, let $\mathcal{B} = (B_1, \ldots, B_{k+t})$ be an ε-coherent blockade in a graph G. We must show that there is a $\mathcal{B}$-transversal copy of $B(k, t)$ in G. Let J be an optimal τ-covering digraph for $\mathcal{B}$, with core $(A_1, \ldots, A_{k+t})$. Let $z = |E(J)|$. By 3.3, every vertex of J has outdegree at least one.

Since $|A_i| \geq \tau^z |B_i|$ for $1 \leq i \leq k + t$, and $\tau^z \geq \tau^{(k+t)^2}$, and $\varepsilon \leq \tau^{(k+t)^2} 3^{-k}$, and $\mathcal{B}$ is ε-coherent, it follows that the blockade $(A_1, \ldots, A_k)$ is 3^{-k}-coherent.

(1) *We may assume that every vertex of J has indegree less than t.*

Suppose that, say, k is J-adjacent from each of $k + 1, \ldots, k + t$. Thus for each $j \in \{k + 1, \ldots, k + t\}$, there exists $X_{j,k} \subseteq B_j$ as in the definition of τ-covering digraph. Since the blockade $(A_1, \ldots, A_k)$ is 3^{-k}-coherent, by 2.2 there is an $(A_1, \ldots, A_k)$-transversal k-vertex path in G with an end-vertex in A_k, say with vertices $p_1, \ldots, p_k$ in order, where $p_i \in A_i$ for $1 \leq i \leq k$. For each $j \in \{k + 1, \ldots, k + t\}$, choose $q_j \in X_{j,k}$ adjacent to p_k (this exists, since $X_{j,k}$ covers A_k). Then $q_{k+1}, \ldots, q_{k+t}$ are pairwise nonadjacent, and nonadjacent to $p_1, \ldots, p_{k-1}$, from the properties of the sets $X_{j,k}$. Hence the subgraph induced on $\{p_1, \ldots, p_k, q_1, \ldots, q_t\}$ is a $\mathcal{B}$-transversal copy of $B(k, t)$, as required. This proves (1).

(2) *There is a subset of $2^t + 1$ elements of $\{1, \ldots, k + t\}$, pairwise nonadjacent in J, and such that no two of them have a common out-neighbour in J.*

From (1) and averaging, there is a vertex i of J with outdegree less than t; and so the set of vertices that are either equal to i, J-adjacent to i, J-adjacent from i, or share a J-outneighbour with i, has cardinality at most $1 + 2(t - 1) + (t - 1)(t - 2) = t^2 - t + 1$. By deleting this set from J, we obtain some digraph; again we find a vertex with outdegree at most $t - 1$ in this digraph, and again delete the corresponding set of vertices, and continue. We can repeat this at least $2^t + 1$ times, since

10

$|J| = k + t > 2^t(t^2 - t + 1)$. Thus we construct a set of $2^t + 1$ vertices of J satisfying (2). This proves (2).

From (2) we may assume that $\{1, \ldots, 2^t + 1\}$ is a stable set in J and no two of its members have a common J-outneighbour. If F is a $\mathcal{B}$-rainbow copy of S_{t+1}, then it has a vertex of degree $t + 1$, that belongs to a block B_i say, and $t + 1$ other vertices, in blocks B_j ($j \in I$) say, where $I \subseteq \{1, \ldots, k + t\} \setminus \{i\}$ with $|I| = t + 1$. Let us call (i, I) the *type* of F.

(3) *We may assume that for some $n \geq 0$ there are copies $F_1, \ldots, F_n$ of S_{t+1}, each $(A_1, \ldots, A_{t+2})$-transversal and pairwise disjoint, and all with type $(1, \{2, \ldots, t+2\})$, and there exist $r \in \{1, \ldots, t+2\}$ and $s \in \{t + 3, \ldots, t + k\}$ with $|N(F \cap A_r) \cap A_s| \geq |A_s|/(2t + 4)$, where $F = V(F_1) \cup \cdots \cup V(F_n)$.*

Choose a maximal set $\mathcal{F}$ of pairwise disjoint $(A_1, \ldots, A_{2^t+1})$-rainbow copies of S_{t+1}, and for $1 \leq i \leq 2^t + 1$ let D_i be the set of vertices of A_i that are not in any member of $\mathcal{F}$. There is no $(D_1, \ldots, D_{2^t+1})$-rainbow copy of S_{t+1}; and so by 2.3, the blockade $(D_1, \ldots, D_{2^t+1})$ is not 3^{-2^t-1}-coherent (or one of the sets D_i is empty); and it follows that $|D_r| < \varepsilon 3^{2^t+1}|B_r|$ for some $r \in \{1, \ldots, 2^t + 1\}$. Since $|A_r| \geq \tau^{(k+t)^2}|B_r|$ and $\varepsilon 3^{2^t+1} \leq \tau^{(k+t)^2}/2$, it follows that

$$|A_r \setminus D_r| \geq \left(\tau^{(k+t)^2} - \varepsilon 3^{2^t+1}\right)|B_r| \geq \tau^{(k+t)^2}|B_r|/2.$$

Hence there is a subset $\mathcal{F}' \subseteq \mathcal{F}$ with cardinality at least $\tau^{(k+t)^2}|B_r|/2$, such that each of its members has a vertex in A_r.

There are only at most $(2^t + 1)^{t+2}$ possible types of $\mathcal{A}$-rainbow copies of S_{t+1}; so within our set $\mathcal{F}'$ there is a subset of $\tau^{(k+t)^2}(2^t + 1)^{-t-2}|B_r|/2$ of them all with the same type, and from the symmetry we may assume this common type is $(1, \{2, \ldots, t + 2\})$, and $r \in \{1, \ldots, t + 2\}$.

Thus there are n pairwise disjoint $(A_1, \ldots, A_{t+2})$-rainbow copies of S_{t+1}, all of type $(1, \{2, \ldots, t + 2\})$, say $F_1, \ldots, F_n$, where $n \geq \tau^{(k+t)^2}(2^t + 1)^{-t-2}|B_r|/2$. Let $F = V(F_1) \cup \cdots \cup V(F_n)$. Since $\tau^{(k+t)^2}(2^t + 1)^{-t-2}/2 \geq \varepsilon$, fewer than $\varepsilon|B_{t+3}|$ vertices in B_{t+3} have no neighbour in $F \cap A_r$, and so

$$|N(F \cap A_r) \cap A_{t+3}| \geq |A_{t+3}| - \varepsilon|B_{t+3}| \geq |A_{t+3}|/2 \geq |A_{t+3}|/(2t + 4).$$

We deduce that (3) holds, setting $s = t + 3$. This proves (3).

Let us choose n minimum satisfying (3), and let $F = V(F_1) \cup \cdots \cup V(F_n)$. We recall that $z = |E(J)|$.

(4) *For each $i \in \{1, \ldots, t + 2\}$, there is no $j \in \{1, \ldots, k + t\} \setminus \{i\}$ such that*

$$|A_j \setminus N(F \cap A_i)| \leq \tau^{z+1}|B_j|.$$

Consequently, if $j \in \{1, \ldots, k + t\} \setminus \{i\}$ and $|N(F \cap A_i) \cap A_j| \geq \tau^{z+1}|B_j|$, then j is J-adjacent from i, and so $j \geq t + 3$.

For the first assertion, suppose that such i, j exist. It follows that

$$|A_j \cap N(F \cap A_i)| > |A_j| - \tau^{z+1}|B_j|.$$

11

From the minimality of n, we deduce that

$$|A_j \cap N(F \cap A_i)| \le |A_j|/(2t+4) + \varepsilon|B_j|.$$

Consequently

$$|A_j| - \tau^{z+1}|B_j| < |A_j|/(2t+4) + \varepsilon|B_j|,$$

and so

$$(1 - 1/(2t+4))|A_j| < (\tau^{z+1} + \varepsilon)|B_j|.$$

Since $|A_j| \ge \tau^z|B_j|$, it follows that $(1 - 1/(2t+4))\tau^z < \tau^{z+1} + \varepsilon$, contrary to the choice of ε. This proves the first assertion. The second follows immediately from 3.2, and so this proves (4).

Choose r, s as in (3). Let us say a vertex $v \in A_s$ is *good* if v has a G-neighbour in $F \cap A_r$ and has no G-neighbour in $F \setminus A_r$. Let C_s be the set of all good vertices in A_s.

(5) $|C_s| \ge |A_s|/(4t+8)$.

We have seen that s is J-adjacent from r. Since no two of $1, \ldots, t+2$ have a common J-outneighbour, s is not J-adjacent from any element in $\{1, \ldots, t+2\} \setminus \{r\}$. From the second assertion of (4), $|N(F \cap A_i) \cap A_s| < \tau^{z+1}|B_s|$ for each $i \in \{1, \ldots, t+2\} \setminus \{r\}$. Hence at most $(t+1)\tau^{z+1}|B_s|$ vertices in B_s have a G-neighbour in $F \setminus A_r$; and so at least $|A_s|/(2t+4) - (t+1)\tau^{z+1}|B_s|$ vertices in A_s are good. Since

$$(t+1)\tau^{z+1}|B_s| \le \tau^z|B_s|/(4t+8) \le |A_s|/(4t+8)$$

we deduce that $|C_s| \ge |A_s|/(4t+8)$. This proves (5).

We know that $r \in \{1, \ldots, t+2\}$, but the argument to come depends on whether $r = 1$ or not. If $r = 1$ let $t' = t+1$ and $F' = F \setminus A_{t+2}$, and otherwise let $t' = t+2$ and $F' = F$. For $1 \le m \le n$ let $F'_m = F_m \setminus A_{t+2}$ if $r = 1$, and otherwise let $F'_m = F_m$. Thus if $r = 1$, F'_m is a copy of S_t, and otherwise F'_m is a copy of S_{t+1}. For each $j \in \{t'+1, \ldots, k+t\}$ with $j \ne s$, let C_j be the set of all vertices in A_j with no neighbour in F'.

(6) $|C_j| \ge |A_j|/(4t+8)$ *for each* $j \in \{t'+1, \ldots, k+t\}$.

Let $j \in \{t'+1, \ldots, k+t\}$. By (5) we may assume that $j \ne s$. From the minimality of n, for $1 \le i \le t'$ at most $|A_j|/(2t+4) + \varepsilon|B_j|$ vertices in A_j have a neighbour in $F' \cap A_i$; and so at most $(t+2)(|A_j|/(2t+4) + \varepsilon|B_j|)$ have a neighbour in F. Thus

$$|C_j| \ge |A_j| - |A_j|/2 - (t+2)\varepsilon|B_j| = |A_j|/2 - (t+2)\varepsilon|B_j| \ge |A_j|/(4t+8).$$

This proves (6).

From (6) the blockade $(C_{t'+1}, \ldots, C_{k+t})$ is $(4t+8)\varepsilon\tau^{-z}$-coherent, and since $(4t+8)\varepsilon\tau^{-z} \le 3^{1-(k+t)+t'}$ (because $z \le (k+t)(k+t-1)$ and $t'-t \le 2$) 2.2 implies that there is a $(C_{t'+1}, \ldots, C_{k+t})$-transversal path P with end-vertex in C_s. Let its end in C_s be v. From the definition of C_s, there exists $u \in F' \cap A_r$ adjacent to v. Let $u \in F'_m$ say, where $1 \le m \le n$. Since $v \in C_s$, v is good and so has no neighbour in F'_m except u; and since every other vertex of P belongs to some set C_j where $j \in \{t'+1, \ldots, k+t\} \setminus \{s\}$, the edge uv is the only edge of G between $V(P)$ and $V(F'_m)$. Thus the union of P, F_m and the edge uv is a $\mathcal{B}$-rainbow copy of $B(k,t)$. This proves 3.4. ∎

Thus, if $t \leq \log_2 k - (2 + o(1)) \log \log_2 k$, 3.4 tells us that $B(k, t)$ has the strong transversal property, and for $t > (1 + o(1)) \log_2 k$, 2.7 tells us that it does not. More exactly, for $t \geq 2$, if $k \geq 2^t(t^2 - t + 1) - t + 1$ then $B(k, t)$ has the property, and if $2 \leq k \leq 2^t - t$ then it does not. We have not decided the values of k in the middle, except when $t = 2$. In that case 3.4 tells us that $B(k, 2)$ has the property when $k \geq 11$, but this can be improved to:

3.5 *If $k = 1$ or $k \geq 3$ then $B(k, 2)$ has the strong transversal property, and if $k = 2$ it does not.*

Proof. We just sketch the proof, since it is similar to that of 3.4. The claims for $k = 1$ and $k = 2$ follow from 1.5 and 2.4 respectively, so we assume that $k \geq 3$. With an appropriate choice of τ and ε, we choose an optimal τ-covering digraph J; and we may assume no vertex has indegree more than one in J, as in the proof of 3.4, and every vertex has outdegree at least one, by 3.3. Consequently J is a disjoint union of directed cycles. Next we use a lemma (we omit the proof here), that for any τ-covering digraph J and for every directed cycle of J, some edge of the cycle is in a directed cycle of length two (this is true in general, not just in the present context). Consequently J is the disjoint union of directed cycles of length two, and in particular, $|J|$ is even. Thus there are three pairwise nonadjacent vertices of J, say $1, 2, 3$; and as in the proof of 3.4 we find many pairwise disjoint copies of S_2, all (A_1, A_2, A_3)-rainbow and all with middle vertex in A_1. (Note that we look for copies of S_2, not for copies of S_3, which is what setting $t = 2$ in the proof of 3.4 would suggest.) We may assume that $4, 5$ and 6 are J-adjacent to and from $1, 2, 3$ respectively. Now the proof is finished more-or-less as in 3.4; with notation as in 3.4, if $r = 1$ we follow the proof of 3.4, that is, we delete from $A_4, \ldots, A_{k+2}$ the small number of vertices with the wrong adjacency to F, and then apply 2.2 to the resulting blockade $(C_4, C_5, \ldots, C_{k+2})$, finding a path with first vertex in C_4. If $r = 2$ say, we do the same, but apply 2.2 to the blockade $(C_5, \ldots, C_{k+2})$, finding a path with first vertex in C_5, and then turning this into a copy of $B(2, k)$ by adding F_m and a vertex of $X_{4,1}$. We omit further details. ∎

4 The cycle

Our remaining results all concern looking for an anticomplete pair of sets that have polynomial size rather than linear size. So, we are not working with the strong transversal property any more, nor with ε-coherence.

In this section we prove 1.4. We handle the cases $k = 4$ and $k \geq 5$ separately. Both proofs are related to the proofs of theorems in [4]. (These theorems have recently been superceded by theorems in [5], but the proof methods of the latter are quite different.)

We will need the following lemma:

4.1 *Let $0 < \varepsilon \leq 1/2$, and let $\mathcal{B} = (B_1, B_2)$ be a blockade in a graph G, with local degree less than εW and $(\varepsilon W, \varepsilon W^c)$-cohesive where W is its width. Then $\varepsilon W^c > 1$; and if $X \subseteq B_1$ with $|X| \geq 2\varepsilon W$, there are fewer than εW^c vertices in B_2 that have at most $W^{1-c}/2$ neighbours in X.*

Proof. Let $v \in B_1$. Since $\mathcal{B}$ has local degree less than εW, v has at most εW neighbours in B_2; and so has at least $(1 - \varepsilon)W \geq \varepsilon W$ non-neighbours in B_2. Thus $\mathcal{B}$ is not $(\varepsilon W, 1)$-cohesive, and so $\varepsilon W^c > 1$. This proves the first assertion.

Suppose the second assertion is false; then there exists $Y \subseteq B_2$ with $|Y| = \lceil \varepsilon W^c \rceil$, such that every vertex in Y has at most $W^{1-c}/2$ neighbours in B_2. Since $\varepsilon W^c > 1$, it follows that $|Y| \le 2\varepsilon W^c$. Hence at most $|Y|W^{1-c}/2 \le \varepsilon W$ vertices in X have a neighbour in Y, and since $|X| \ge 2\varepsilon W$, X has a subset of cardinality at least εW that is anticomplete to Y, a contradiction. This proves 4.1. ∎

First we show the following, which is a strengthening of the $k = 4$ case of 1.4:

4.2 *Let $\varepsilon = 1/4$, and let $\mathcal{B} = (B_1, \ldots, B_4)$ be a blockade in a graph G, with local degree less than εW and $(\varepsilon W, \varepsilon W^{1/3})$-cohesive where W is its width. Then there is a $\mathcal{B}$-transversal copy of a cycle of length four in G.*

Proof. Let $\mathcal{B} = (B_1, \ldots, B_4)$ be a blockade in a graph G, with local degree less than εW and $(\varepsilon W, \varepsilon W^{1/3})$-cohesive where W is its width. From 4.1:

(1) $\varepsilon W^{1/3} > 1$; and if $i, j \in \{1, \ldots, 4\}$ are distinct, and $X \subseteq B_i$ with $|X| \ge 2\varepsilon W$, and Y is a set of vertices in B_j each with at most $W^{2/3}/2$ neighbours in X, then $|Y| < \varepsilon W^{1/3}$.

Let $v_3 \in B_3$ and $v_4 \in B_4$ be adjacent. We say the edge $v_3 v_4$ is

- *1-good* if v_3 has at least $W^{2/3}/2$ neighbours in B_4;

- *2-good* if it is 1-good and v_4 has at least $W^{2/3}/2$ neighbours in B_2 that are nonadjacent to v_3.

We claim:

(2) *More than half the edges between B_3, B_4 are 2-good.*

By (1), fewer than $\varepsilon W^{1/3}$ vertices in B_3 have at most $W^{2/3}/2$ neighbours in B_4, so at most $\varepsilon W^{1/3}W^{2/3}/2 = \varepsilon W/2$ edges between B_3, B_4 are not 1-good. Now let $v_3 \in B_3$ and let N_4 be the set of its neighbours in B_4, and let N_2 be the set of its neighbours in B_2. Thus $|N_2| < \varepsilon|B_2|$, and so $|B_2 \setminus N_2| \ge (1-\varepsilon)|B_2| \ge 2\varepsilon W$; so by (1), fewer than $\varepsilon W^{1/3}$ vertices in N_4 have fewer than $W^{2/3}/2$ neighbours in $B_2 \setminus N_2$. Consequently at most $\varepsilon W^{1/3}$ of the edges between v_3 and N_4 are 1-good and not 2-good. Since this holds for every choice of $v_3 \in B_3$, it follows that at most $\varepsilon W^{1/3}|B_3|$ edges between B_3, B_4 are 1-good and not 2-good. Hence in total, at most $\varepsilon W^{1/3}|B_3| + \varepsilon W/2$ edges between B_3, B_4 are not 2-good. But at least $|B_3| - \varepsilon W^{1/3}$ vertices in B_3 have at least $W^{2/3}/2$ neighbours in B_4, so there are at least
$$(|B_3| - \varepsilon W^{1/3})W^{2/3}/2$$
edges between B_3, B_4. Since
$$\varepsilon W^{1/3}|B_3| + \varepsilon W/2 < (|B_3| - \varepsilon W^{1/3})W^{2/3}/4$$
because, for instance,
$$(W^{2/3}/4 - \varepsilon W^{1/3})|B_3| \ge (W^{2/3}/4 - \varepsilon W^{1/3})W > \varepsilon W/2 + \varepsilon W/4,$$
this proves (2).

(3) *We may assume that there exist anticomplete subsets $C_1 \subseteq B_1$ and $C_2 \subseteq B_2$, with $|C_1|, |C_2| \geq W^{2/3}/2$.*

By (2), and the same statement with B_3, B_4 exchanged and B_1, B_2 exchanged, it follows that there is an edge $v_3 v_4$ with $v_3 \in B_3$ and $v_4 \in B_4$, such that v_4 has at least $W^{2/3}/2$ neighbours in B_2 that are nonadjacent to v_3, and v_3 has at least $W^{2/3}/2$ neighbours in B_1 that are nonadjacent to v_4. Let C_1 be the set of vertices in B_1 that are adjacent to v_3 and not to v_4, and define $C_2 \subseteq B_2$ similarly. If there is an edge between C_1, C_2, then adding v_3, v_4 makes a $\mathcal{B}$-transversal cycle of length four; so we may assume there is no such edge. This proves (3).

Choose C_1, C_2 as in (3).

(4) *There exists $v_3 \in B_3$ with at least $\varepsilon W^{1/3}$ neighbours in C_1 and at least $\varepsilon W^{1/3}$ neighbours in C_2.*

Suppose that there is a set $A_3 \subseteq B_3$ with cardinality $\lceil 2\varepsilon W \rceil$, such that each of its members has fewer than $\varepsilon W^{1/3}$ neighbours in C_1. By (1), fewer than $\varepsilon W^{1/3}$ vertices in C_1 have fewer than $W^{2/3}/2$ neighbours in A_3; and so there are at least

$$(|C_1| - \varepsilon W^{1/3})W^{2/3}/2$$

edges between C_1, A_3. But from the definition of A_3, there are at most $|A_3|\varepsilon W^{1/3}$ such edges; so

$$(|C_1| - \varepsilon W^{1/3})(W^{2/3}/2) < |A_3|(\varepsilon W^{1/3}).$$

Since $|C_1| \geq W^{2/3}/2$, and $|A_3| \leq 3\varepsilon W$ (since $\varepsilon = 1/4$ and $W \geq 64$), it follows that

$$(W^{2/3}/2 - \varepsilon W^{1/3})(W^{2/3}/2) < 3\varepsilon W(\varepsilon W^{1/3}),$$

which (since $\varepsilon = 1/4$) simplifies to $W^{1/3} < 2$, a contradiction since $W^{1/3} \geq \varepsilon^{-1} = 4$. Thus there are fewer than $2\varepsilon W$ vertices in B_3 with fewer than $\varepsilon W^{1/3}$ neighbours in C_1; and fewer than $2\varepsilon W$ vertices in B_3 with fewer than $\varepsilon W^{1/3}$ neighbours in C_2, similarly. Since $4\varepsilon W = W \leq |B_3|$, this proves (4).

Choose v_3 as in (4), and for $i = 1, 2$ let A_i be the set of neighbours of v_3 in C_1 and in C_2 respectively. Since $\mathcal{B}$ is $(\varepsilon W, \varepsilon W^{1/3})$-cohesive, fewer than εW vertices in B_4 have no neighbour in A_1; fewer than εW vertices in B_4 have no neighbour in A_2; and fewer than εW vertices in B_4 are adjacent to v_3. Since $3\varepsilon W < W \leq |B_4|$, there exists $v_4 \in B_4$ with a neighbour $v_1 \in A_1$, and a neighbour $v_2 \in A_2$, and non-adjacent to v_3. But then there is a $\mathcal{B}$-transversal 4-cycle induced on $\{v_1, v_2, v_3, v_4\}$. This proves 4.2. ∎

To complete the proof of 1.4, next we prove the following, a strengthening of 1.4 when $k \geq 5$:

4.3 *Let $k \geq 5$ be an integer, and let $\varepsilon = 1/(3k)$. Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade in a graph G, with local degree less than εW and $(\varepsilon W, \varepsilon W^{1/2})$-cohesive where W is its width. Then there is a $\mathcal{B}$-transversal copy of a cycle of length k in G.*

Proof. Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade in a graph G, with local degree less than εW and $(\varepsilon W, \varepsilon W^{1/2})$-cohesive where W is its width. From 4.1:

(1) $\varepsilon W^{1/2} > 1$; and if $i, j \in \{1, \ldots, k\}$ are distinct, and $X \subseteq B_i$ with $|X| \geq 2\varepsilon W$, and Y is a set of vertices in B_j each with fewer than $W^{1/2}/2$ neighbours in X, then $|Y| < \varepsilon W^{1/2}$.

Next we prove the following:

(2) Let $i_1, i_2, \ldots, i_s \in \{1, \ldots, k\}$ be distinct, and let $D_{i_r} \subseteq B_{i_r}$ for $1 \leq r \leq s$, such that $|D_{i_1}| \geq \varepsilon W^{1/2}$, and $|D_{i_r}| \geq \varepsilon s |B_{i_r}|$ for $2 \leq r \leq s$. There there is an induced path of G with vertices $v_{i_1} \text{-} \cdots \text{-} v_{i_s}$, where $v_{i_r} \in D_{i_r}$ for $1 \leq r \leq s$.

We proceed by induction on s; if $s = 1$ the result is trivial, so we assume that $s \geq 2$ and the result holds for $s - 1$. Let $i_1, i_2, \ldots, i_s$ and $D_{i_r} \subseteq B_{i_r}$ for $1 \leq r \leq s$ as above. From the symmetry we may assume that $i_r = r$ for $1 \leq r \leq s$. Since $s \geq 2$, and consequently $|D_2| \geq s\varepsilon|B_2| \geq 2\varepsilon|B_2|$, (1) implies that there exists $v_1 \in D_1$ with at least $W^{1/2}/2$ neighbours in D_2. Let E_2 be the set of these neighbours, and for $3 \leq r \leq s$ let E_r be the set of vertices in D_r nonadjacent to v_1. Since v_1 has fewer than $\varepsilon|B_r|$ neighbours in B_r, it follows that $|E_r| \geq |D_r| - \varepsilon|B_r| \geq \varepsilon(s-1)|B_r|$. Hence from the inductive hypothesis applied to $2, \ldots, s$ and the sets $E_2, \ldots, E_s$, there is an induced path of G with vertices $v_2 \text{-} \cdots \text{-} v_s$, where $v_r \in E_r$ for $2 \leq r \leq s$. Adding v_1 and the edge $v_1 v_2$ gives a path satisfying (2). This proves (2).

From (1), all vertices in B_1 except at most $\varepsilon W^{1/2}$ have at least $W^{1/2}/2$ neighbours in B_2, and the same for B_3; so there exists $v_1 \in B_1$ with at least $W^{1/2}/2$ neighbours in B_2 and at least $W^{1/2}/2$ neighbours in B_3. For $i = 2, 3$ let A_i be the set of neighbours of v_1 in B_i, and for $4 \leq i \leq k$ let A_i be the set of vertices in B_i that are nonadjacent to v_1. Thus $|A_2|, |A_3| \geq W^{1/2}/2$. Since v_1 has at most $\varepsilon|B_j|$ neighbours in B_j, it follows that $|A_j| \geq (1 - \varepsilon)|B_j|$ for $4 \leq j \leq k$.

All except at most εW vertices in A_4 have a neighbour in A_2, so we may choose $C_2 \subseteq A_2$ minimal such that for some $j \in \{4, \ldots, k\}$, at least $|B_j|/3$ vertices in A_j have a neighbour in C_2. Choose some such j; and from the symmetry we may assume that $j = 4$. Let C_4 be the set of vertices in A_4 that have a neighbour in C_2. For $5 \leq i \leq k$, let C_i be the set of vertices in A_i with no neighbour in C_2. Thus $|C_4| \geq |B_4|/3$; and from the minimality of C_2, it follows that fewer than $|B_i|/3 + \varepsilon|B_i|$ have a neighbour in C_2, for each $i \in \{5, \ldots, k\}$. Hence for each $i \in \{5, \ldots, k\}$,

$$|C_i| \geq |A_i| - |B_i|/3 - \varepsilon|B_i| \geq (1 - \varepsilon)|B_i| - |B_i|/3 - \varepsilon|B_i| = (2/3 - 2\varepsilon)|B_i| \geq |B_i|/3.$$

By (1), at most $\varepsilon W^{1/2}$ vertices in A_3 have fewer than $W^{1/2}/2$ neighbours in C_5 (this is where we use $k \geq 5$); and since $|A_3| \geq W^{1/2}/2 > \varepsilon W^{1/2}$, there exists $v_3 \in A_3$ with at least $W^{1/2}/2$ neighbours in C_5.

Let D_5 be the set of neighbours of v_3 in C_5; let $C_1 = B_1$, and for $i \in \{1, 4\} \cup \{6, \ldots, k\}$, let D_h be the set of vertices in C_h nonadjacent to v_3. Hence $|D_5| \geq W^{1/2}/2$, and

$$|D_h| \geq |C_h| - \varepsilon|B_h| \geq (2/3 - 3\varepsilon)|B_i|$$

for $i \in \{1, 4\} \cup \{6, \ldots, k\}$.

16

Every vertex in D_4 has a neighbour in C_2, which may or may not be adjacent to v_3. So either at least $|D_4|/2$ vertices in D_4 have a neighbour in C_2 nonadjacent to v_3, or at least $|D_4|/2$ vertices in D_4 have a neighbour in C_2 adjacent to v_3. We handle these two cases separately.

First, assume that at least $|D_4|/2$ vertices in D_4 have a neighbour in C_2 nonadjacent to v_3; let D_2 be the set of vertices in C_2 nonadjacent to v_3, and let D_4' be the set of vertices in D_4 with a neighbour in D_2. Thus

$$|D_4'| \geq |D_4|/2 \geq (2/3 - 3\varepsilon)|B_4|/2 \geq (k-3)\varepsilon|B_4|,$$

since $\varepsilon \leq 1/(3k)$. By (2), there is an induced path P of G v_5-v_6-$\cdots$-v_k-v_4, where $v_r \in D_r$ for $5 \leq r \leq k$ and $v_4 \in D_4'$. Choose $v_2 \in D_2$ adjacent to v_4; then the union of P and the path v_4-v_2-v_1-v_3-v_5 gives a $\mathcal{B}$-transversal cycle satisfying the theorem.

Now we assume that at least $|D_4|/2$ vertices in D_4 have a neighbour in C_2 adjacent to v_3; let D_2 be a subset of C_2, all adjacent to v_3, minimal such that either at least $|D_4|/2$ vertices in D_4, or at least $|D_1|/2$ vertices in D_1, have a neighbour in D_2.

Suppose there is a set $D_4' \subseteq D_4$ with cardinality at least $|D_4|/2$, all with a neighbour in D_2. From the minimality of D_2, at most $|D_1|/2 + \varepsilon|B_1|$ vertices in D_1 have a neighbour in D_2, and so there is a subset $D_1' \subseteq D_1$ with cardinality at least $|D_1|/2 - \varepsilon|B_1| \geq (k-2)\varepsilon|B_1|$, anticomplete to D_2. By (2), there is an induced path P of G with vertices v_5-$\cdots$-v_k-v_1'-v_4, where $v_r \in D_r$ for $5 \leq r \leq k$ and $v_1' \in D_1'$ and $v_4 \in D_4'$. Choose $v_2 \in D_2$ adjacent to v_4; then the union of P and the path with vertices v_4-v_2-v_3-v_5 gives a $\mathcal{B}$-transversal cycle satisfying the theorem.

Finally we may assume that there is a set $D_1' \subseteq D_1$ with cardinality at least $|D_1|/2$, all with a neighbour in D_2. From the minimality of D_2, at most $|D_4|/2 + \varepsilon|B_4|$ vertices in D_4 have a neighbour in D_2, and so there is a subset $D_4' \subseteq D_4$ with cardinality at least $|D_4|/2 - \varepsilon|B_4| \geq (k-2)\varepsilon|B_4|$, anticomplete to D_2. By (2), there is an induced path P of G with vertices v_5-$\cdots$-v_k-v_4-v_1', where $v_r \in D_r$ for $5 \leq r \leq k$ and $v_4 \in D_4'$ and $v_1' \in D_1'$. Choose $v_2 \in D_2$ adjacent to v_1'; then the union of P and the path with vertices v_1'-v_2-v_3-v_5 gives a $\mathcal{B}$-transversal cycle satisfying the theorem. This proves 4.3. ∎

We do not know whether the exponents of $1/3$ (in 4.2) and $1/2$ (in 4.3) are best possible.

5 Ordered transversal subgraphs

Now we turn to excluding ordered graphs. We begin with 1.6, which we restate:

5.1 *If H is an ordered tree with $k \geq 1$ vertices, then there exists $\varepsilon > 0$ with the following property. Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade in a graph G, with local degree less than εW and $(\varepsilon W, \varepsilon W^{1/(k-1)})$-cohesive where W is its width. Then there is an ordered $\mathcal{B}$-transversal copy of H.*

To prove this we need to prove a strengthening (which implies 5.1 by setting $c = 1/(k-1)$):

5.2 *Let H be an ordered tree with $k \geq 1$ vertices, and let $c > 0$ with $(k-1)c \leq 1$. Let $\varepsilon = 4^{1-k}$. Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade in a graph G, with local degree less than εW and $(\varepsilon W, \varepsilon W^c)$-cohesive where W is its width. Then there are at least $4^{1-k}W^{k-(k-1)c}$ ordered $\mathcal{B}$-transversal copies of H.*

Proof. We proceed by induction on k. The result is trivial for $k = 1$, so we assume that $k \geq 2$ and the result holds for $k - 1$. Let the ordering of H be $p_1, \ldots, p_k$. We may assume that p_k has degree one in H, and p_{k-1} is its unique neighbour. Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade in a graph G, with local degree less than εW and $(\varepsilon W, \varepsilon W^c)$-cohesive where W is its width. We may assume that $|B_i| = W$ for $1 \leq i \leq k$. From 4.1:

(1) *$\varepsilon W^c > 1$; and if $X \subseteq B_k$ with $|X| \geq 2\varepsilon W$, there are fewer than εW^c vertices in B_{k-1} that have at most $W^{1-c}/2$ neighbours in X.*

In particular, there are at least $(W - \varepsilon W^c)W^{1-c}/2 \geq W^{2-c}/4$ edges between B_{k-1} and B_k, so if $k = 2$ the result is true. Thus we may assume that $k \geq 3$.

Let H' be the ordered tree obtained from H by deleting p_k (with ordering $p_1, \ldots, p_{k-1}$), and similarly let H'' be the ordered forest obtained by deleting both p_k, p_{k-1}. Let $\mathcal{B}' = (B_1, \ldots, B_{k-1})$, and $\mathcal{B}'' = (B_1, \ldots, B_{k-2})$. Let $\mathcal{H}$ be the set of all ordered $\mathcal{B}$-transversal copies of H, let $\mathcal{H}'$ be the set of all ordered $\mathcal{B}'$-transversal copies of H', and let $\mathcal{H}''$ be the set of all ordered $\mathcal{B}''$-transversal copies of H''.

For each $F \in \mathcal{H}''$, let $n(F)$ be the number of vertices $v \in B_{k-1}$ such that the subgraph induced on $V(F) \cup \{v\}$ is an ordered $\mathcal{B}'$-transversal copy of H'. Let $m(F)$ be the number of edges uv with $u \in B_k$ and $v \in B_{k-1}$ such that the subgraph induced on $V(F) \cup \{u, v\}$ is an ordered $\mathcal{B}$-transversal copy of H.

(2) *For each $F \in \mathcal{H}''$, $m(F) \geq (n(F) - \varepsilon W^c)W^{1-c}/2$.*

Let $F \in \mathcal{H}''$ and let N be the set of vertices $v \in B_{k-1}$ such that the subgraph induced on $V(F) \cup \{v\}$ is an ordered $\mathcal{B}'$-transversal copy of H'. Let X be the set of vertices in B_k with no neighbours in $V(F)$. Thus

$$|X| \geq |B_1| - (k - 2)\varepsilon|B_1| \geq 2\varepsilon W,$$

(since $\varepsilon \leq 1/k$), and so by (1), there are fewer than εW^c vertices in N that have at most $W^{1-c}/2$ neighbours in X. All the others have more than $W^{1-c}/2$ neighbours in X, and every such edge contributes to $m(F)$. This proves (2).

Summing $n(F), m(F)$ and 1 over all $F \in \mathcal{H}''$ gives $|\mathcal{H}|, |\mathcal{H}'|$ and $|\mathcal{H}''|$ respectively, so by summing the inequality of (2) over all $F \in \mathcal{H}''$, we deduce that

$$2|\mathcal{H}| \geq W^{1-c}|\mathcal{H}'| - \varepsilon W|\mathcal{H}''| \geq W^{1-c}|\mathcal{H}'| - \varepsilon W^{k-1}.$$

But from the inductive hypothesis,

$$|\mathcal{H}'| \geq 4^{2-k}W^{k-1-(k-2)c} \geq 2\varepsilon W^{k-2+c}$$

(the latter since $c \leq 1/(k-1)$ and $\varepsilon = 4^{1-k}$). Consequently

$$2|\mathcal{H}| \geq W^{1-c}|\mathcal{H}'|/2 \geq 4^{2-k}W^{k-1-(k-2)c}W^{1-c}/2 = 2 \cdot 4^{1-k}W^{k-(k-1)c}.$$

This proves 5.2. ∎

The exponent of $1/(k-1)$ in 5.1 is best possible in the sense that for the tree S_{k-1}, the exponent cannot be replaced by any larger constant, as we shall see. But perhaps it can be replaced by $1/d$ where d is the maximum degree of the tree? We propose:

5.3 Conjecture: *If H is an ordered tree with $k \geq 1$ vertices and maximum degree d, then there exists $\varepsilon > 0$ with the following property. Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade in a graph G, with local degree less than εW and $(\varepsilon W, \varepsilon W^{1/d})$-cohesive where W is its width. Then there is an ordered $\mathcal{B}$-transversal copy of H.*

The next result shows that this is true for caterpillars:

5.4 *Let H be an ordered caterpillar with $k \geq 1$ vertices and maximum degree d, and let $\varepsilon = 4^{-d}/k$. Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade in a graph G, with local degree less than εW and $(\varepsilon W, \varepsilon W^{1/d})$-cohesive where W is its width. Then there is an ordered $\mathcal{B}$-transversal copy of H.*

For inductive purposes it is helpful to prove something stronger. If H is a caterpillar, there is a path of H containing all vertices of H with degree more than one. If there is such a path with one end v we call v a *head* of the caterpillar. (Thus, the head is not necessarily unique.) We will show:

5.5 *Let H be an ordered caterpillar with $k \geq 1$ vertices, with ordering $v_1, \ldots, v_k$ where v_1 is a head. Let v_1 have degree d_1, and let every vertex of H have degree at most d, and let $\varepsilon = 4^{-d}/k$. Let $\mathcal{B} = (B_1, \ldots, B_k)$ be a blockade in a graph G, with local degree less than εW and $(\varepsilon W, \varepsilon W^{1/d})$-cohesive where W is its width. Let $C_1 \subseteq B_1$, where $|C_1| \geq 4^{d_1 - d} W^{d_1/d}$. Then there is an ordered $(C_1, B_2, \ldots, B_k)$-transversal copy of H.*

Proof. Let G, $\mathcal{B} = (B_1, \ldots, B_k)$ and C_1 be as in the theorem. We may assume that $|B_1|, \ldots, |B_k| = W$. If $k = 1$ the result is trivial; and if $k = 2$, then $d_1 = d = 1$, and so $|C_1| \geq W$, and therefore there is an edge between C_1, B_2 and the claim holds. So we may assume that $k \geq 3$, and proceed by induction on k.

Suppose first that $d_1 \geq 2$, and let v_k say be a neighbour of v_1 that has degree one in H. Let H' be obtained from H by deleting v_k. By 4.1, there are fewer than $\varepsilon W^{1/d}$ vertices in C_1 that have at most $W^{1-1/d}/2$ neighbours in B_k. Hence there are at least $(|C_1| - \varepsilon W^{1/d}) W^{1-1/d}/2 \geq |C_1| W^{1-1/d}/4$ edges between C_1 and B_k, since $|C_1| \geq W^{d_1/d} \geq 2\varepsilon W^{1/d}$. Consequently some vertex $u_k \in B_k$ has at least $|C_1| W^{-1/d}/4$ neighbours in C_1. Let C_1' be the set of these neighbours; then

$$|C_1'| \geq |C_1| W^{-1/d}/4 \geq 4^{d_1 - d} W^{d_1/d} W^{-1/d}/4 \geq 4^{d_1 - 1 - d} W^{(d_1 - 1)/d}.$$

For $2 \leq i \leq k - 1$, let B_i' be the set of vertices in B_i nonadjacent to u_k; so $|B_i'| \geq (1 - \varepsilon)W$. Let the blockade $(B_1, B_2', \ldots, B_{k-1}')$ have width W' say; then $W' \geq (1 - \varepsilon)W$. Let $\varepsilon' = 4^{-d}/(k-1)$; then $\varepsilon' W' \geq \varepsilon W$, and so $(B_1, B_2', \ldots, B_{k-1}')$ has local degree less than $\varepsilon' W'$ and is $(\varepsilon' W', \varepsilon' (W')^{1/d})$-cohesive. (Note that $\varepsilon'(W')^{1/d} \geq \varepsilon W^{1/d}$.) From the inductive hypothesis, there is an ordered $(C_1', B_2', \ldots, B_{k-1}')$-transversal copy of $H \setminus \{v_k\}$; and adding u_k gives an ordered $\mathcal{B}$-transversal copy of H containing a vertex of C_1.

So we may assume that $d_1 = 1$; let v_2 be the unique neighbour of v_1 in H, and let H' be obtained from H by deleting v_1. Thus v_2 is a head of H'.

By 4.1, since $|C_1| \geq 4^{1-d} W^{1/d} > \varepsilon W^{1/d}$, there is a vertex $u_1 \in C_1$ with at least $W^{1-1/d}/2$ neighbours in B_2; let C_2 be the set of these neighbours. For $3 \leq i \leq k$, let B_i' be the set of vertices

19

in B_i nonadjacent to u_1, so $|B_i'| \geq (1 - \varepsilon)W$. Let the blockade $(B_2, B_3', \ldots, B_k')$ have width W' say. Then $W' \geq (1 - \varepsilon)W$, and as before, it has local degree less than $\varepsilon'W'$ and is $(\varepsilon'W', \varepsilon'(W')^{1/d})$-cohesive where $\varepsilon' = 4^{-d}/(k-1)$. From the inductive hypothesis, there is an ordered $(C_2, B_3', \ldots, B_k')$-transversal copy of $H \setminus \{v_1\}$, and adding u_1 gives a an ordered $\mathcal{B}$-transversal copy of H containing a vertex of C_1. This proves 5.5. ∎

Finally, let us see that the exponents in 5.4 and 5.5 cannot be replaced by any larger constant. We need the following three lemmas:

5.6 *Let $t \geq 3$ be an integer. Let $0 < \varepsilon < 1$ be rational, and let $c > d > 1/t$, where $d - 1/t < (c - 1/t)/(t - 1)$, and $d < 2/t$, and c, d are rational. Let $K > \varepsilon \log(e/\varepsilon)/(-\log(1 - \varepsilon))$. Let n be an integer such that $n^c, n^d, n^{1-d}, n^{1/t}, \varepsilon n$ are all integers. If n is sufficiently large, there is a graph with bipartition A, B, where $|A| = n$ and $|B| = n^{2/t}$, such that*

- *every vertex in A has degree at most $t - 1$;*

- *for every $X \subseteq A$ with $|X| \geq \varepsilon n^c$, there are at least Kn^d vertices in B that have a neighbour in X;*

- *for every $X \subseteq A$ with $|X| \geq \varepsilon n$, there are at least $(\varepsilon/e)|B|$ vertices in B with a neighbour in X; and*

- *every vertex in B has less than n^{1-d} neighbours in A.*

Proof. Let A, B be disjoint sets of cardinalities $n, n^{2/t}$ respectively. For each $v \in A$, choose $v_1, \ldots, v_{t-1}$ in B uniformly and independently at random (and therefore not necessarily distinct), and add edges to make v adjacent to $v_1, \ldots, v_{t-1}$. Let G be the graph this constructs. We claim that if n is sufficiently large then with high probability G satisfies the theorem.

Let $X \subseteq A$ with $|X| = \varepsilon n^c$, and let $Y \subseteq B$ with $|Y| = Kn^d$. The probability that for every vertex in X, all its neighbours are in Y, is

$$\left(\left(Kn^d\right)/|B|\right)^{\varepsilon(t-1)n^c} = \left(Kn^{d-2/t}\right)^{\varepsilon(t-1)n^c}.$$

By 2.5 there are at most $((e/\varepsilon)n^{1-c})^{n^c}$ choices of X, and at most $((e/K)n^{2/t-d})^{Kn^d}$ choices of Y. Thus the probability that there is a choice of X, Y such that for every vertex in X, all its neighbours are in Y, is at most the product of these, that is

$$\left(Kn^{d-2/t}\right)^{\varepsilon(t-1)n^c} \left((e/\varepsilon)n^{1-c}\right)^{\varepsilon n^c} \left((e/K)n^{2/t-d}\right)^{Kn^d}.$$

The logarithm of this (L say) is

$$\varepsilon(t-1)n^c \left(\log K + (d - 2/t)\log n\right) + \varepsilon n^c \left(\log(e/\varepsilon) + (1 - c)\log n\right) + Kn^d \left(\log(e/K) + (2/t - d)\log n\right).$$

Since $c > d, 1/t$, for sufficiently large n the two terms containing $n^c \log n$ are much larger than the others, and the sum of their coefficients is $\varepsilon(t - 1)(d - 2/t) + \varepsilon(1 - c)$. This is negative, since $d - 1/t < (c - 1/t)/(t - 1)$; and so for sufficiently large n, L is large and negative, and therefore with high probability, the second bullet of the theorem holds.

20

Now let $X \subseteq A$ and $Y \subseteq B$, with $|X| = \varepsilon n$ and $|Y| = \lfloor (\varepsilon/e)|B| \rfloor$. The probability that for every vertex in X, all its neighbours are in Y, is at most $(\varepsilon/e)^{(t-1)\varepsilon n}$. By 2.5 there are at most $(e/\varepsilon)^{\varepsilon n}$ choices of X, and at most $2^{n^{2/t}}$ choices of Y. Thus the probability that there is a choice of X, Y such that for every vertex in X, all its neighbours are in Y, is at most

$$(\varepsilon/e)^{(t-1)\varepsilon n}(e/\varepsilon)^{\varepsilon n}2^{n^{2/t}}.$$

The logarithm of this (L say) is

$$(t-1)\varepsilon n \log(\varepsilon/e) + \varepsilon n \log(e/\varepsilon) + n^{2/t}\log 2.$$

The two terms linear in n dominate for large n, and the sum of their coefficients is

$$(t-1)\varepsilon \log(\varepsilon/e) + \varepsilon \log(e/\varepsilon) = -(t-2)\varepsilon \log(e/\varepsilon) < 0,$$

so for sufficiently large n, L is large and negative, and therefore with high probability, the third bullet of the theorem holds.

Finally, let $v \in B$ and let $X \subseteq A$ with $|X| = n^{1-d}$. The probability that v is adjacent to every vertex in X is at most $\left((t-1)n^{-2/t}\right)^{n^{1-d}}$. The number of choices of X is at most $\left(en^d\right)^{n^{1-d}}$, so the probability that some vertex in B has degree at least n^{1-d} is at most

$$\left((t-1)n^{-2/t}\right)^{n^{1-d}}\left(en^d\right)^{n^{1-d}}n.$$

The logarithm of this is

$$n^{1-d}\left(\log(t-1) - (2/t)\log n\right) + n^{1-d}(1 + d\log n) + \log n.$$

The $n^{1-d}\log n$ terms dominate, for large n, and the sum of their coefficients is $-2/t + d$; and this is negative since $d < 2/t$. Consequently with high probability, the fourth bullet of the theorem holds. This proves 5.6. ∎

5.7 *Let $t \geq 3$ be an integer. Let $0 < \varepsilon < 1$ be rational, and let $c > d > 1/t$, where $d - 1/t < (c - 1/t)/(t - 1)$ and $d < 2/t$, and c, d are rational. Let $K > \log(e/\varepsilon)$. Let n be an integer such that $n^c, \varepsilon n^c, n^d, n^{1/t}, \varepsilon n, n^{1-d}, (\varepsilon/2)n^d$ are all integers. If n is sufficiently large, there is a graph with bipartition B, C, where $|B| = n^{2/t}$ and $|C| = n$, such that*

- *every vertex in B has degree at most n^{1-d};*

- *for every $X \subseteq B$ with $|X| \geq Kn^d$, there are more than $(1 - \varepsilon)n$ vertices in C that have a neighbour in X; and*

- *for every $X \subseteq B$ with $|X| \geq (\varepsilon/e)|B|$, there are more than $n - \varepsilon n^c$ vertices in C with a neighbour in X.*

- *every vertex in C has degree at most $(\varepsilon/2)n^d$.*

Proof. Let B, C be disjoint sets of cardinalities $n^{2/t}, n$ respectively. For each $v \in B$, choose n^{1-d} vertices in B uniformly and independent at random (and therefore not necessarily distinct), and add edges to make v adjacent to them. Let G be the graph this constructs. We claim that if n is sufficiently large then with high probability G satisfies the theorem.

Let $X \subseteq B$ with $|X| = Kn^d$, and let $Y \subseteq C$ with $|Y| = (1 - \varepsilon)n$. The probability that for every vertex in X, all its neighbours are in Y, is

$$(1 - \varepsilon)^{Kn} \le e^{-\varepsilon K n}.$$

By 2.5 there are at most $\left((e/K)n^{2/t-d}\right)^{Kn^d}$ choices of X, and at most $(e/\varepsilon)^{\varepsilon n}$ choices of Y. Thus the probability that there is a choice of X, Y such that for every vertex in X, all its neighbours are in Y, is at most

$$e^{-\varepsilon K n}\left((e/K)n^{2/t-d}\right)^{Kn^d}(e/\varepsilon)^{\varepsilon n}.$$

The logarithm of this is

$$-\varepsilon K n + K n^d \left(\log(e/K) + (2/t - d)\log n\right) + \varepsilon n \log(e/\varepsilon).$$

For sufficiently large n the terms linear in n dominate, and the sum of their coefficients is $-\varepsilon K + \varepsilon \log(e/\varepsilon)$; and this is negative since $K > \log(e/\varepsilon)$, so with high probability, the second bullet of the theorem holds.

Now let $X \subseteq B$ and $Y \subseteq C$, with $|X| = \lceil (\varepsilon/e)|B| \rceil$ and $|Y| = n - \varepsilon n^c$. The probability that for every vertex in X, all its neighbours are in Y, is at most

$$\left(1 - \varepsilon n^{c-1}\right)^{(\varepsilon/e)n^{2/t}n^{1-d}} = \left(1 - \varepsilon n^{c-1}\right)^{(\varepsilon/e)n^{1+2/t-d}} \le e^{-\varepsilon n^{c-1}(\varepsilon/e)n^{1+2/t-d}} = e^{-(\varepsilon^2/e)n^{c+2/t-d}}.$$

By 2.5 there are at most $2^{n^{2/t}}$ choices of X, and at most $\left((e/\varepsilon)n^{1-c}\right)^{\varepsilon n^c}$ choices of Y. Thus the probability that there is a choice of X, Y such that for every vertex in X, all its neighbours are in Y, is at most

$$e^{-(\varepsilon^2/e)n^{c+2/t-d}}2^{n^{2/t}}\left((e/\varepsilon)n^{1-c}\right)^{\varepsilon n^c}.$$

The logarithm of this is

$$-(\varepsilon^2/e)n^{c+2/t-d} + n^{2/t}\log 2 + \varepsilon n^c \left(\log(e/\varepsilon) + (1-c)\log n\right).$$

Since $c + 2/t - d > \max(2/t, c)$, the first term dominates if n is sufficiently large, and so with high probability, the third bullet of the theorem holds.

Let $v \in C$ and let $X \subseteq B$ with $|X| = (\varepsilon/2)n^d$. The probability that v is adjacent to every vertex in X is $\left(n^{-d}\right)^{(\varepsilon/2)n^d} = n^{-(\varepsilon/2)dn^d}$. There are at most $\left((2e/\varepsilon)n^{2/t-d}\right)^{(\varepsilon/2)n^d}$ choices of X by 2.5, and n choices of v, so the probability that there is a choice of v, X such that v is adjacent to every vertex in X, is at most

$$n^{-(\varepsilon/2)dn^d}\left((2e/\varepsilon)n^{2/t-d}\right)(\varepsilon/2)n^d n.$$

The logarithm of this is

$$-(\varepsilon/2)dn^d \log n + (\varepsilon/2)n^d \left(\log(2e/\varepsilon) + (2/t - d)\log n\right) + \log n.$$

The terms in $n^d \log n$ dominate for large n, and the sum of their coefficients is

$$-(\varepsilon/2)d + (\varepsilon/2)(2/t - d) = \varepsilon(1/t - d) < 0.$$

Consequently with high probability the fourth bullet holds. This proves 5.7. ∎

5.8 *Let* $0 < c \leq 1$. *If* n *is sufficiently large, and* εn, εn^c *are both integers, there is a graph with bipartition* A, B, *where* $|A| = |B| = n$, *such that every vertex has degree at most* $(2/\varepsilon^2)n^{1-c}$, *and the blockade* (A, B) *is* $(\varepsilon n, \varepsilon n^c)$-*cohesive.*

We leave the proof to the reader; it is like that of 2.6. The three preceding lemmas are used for the following:

5.9 *Let* $t \geq 3$, *and let* S_t^+ *be obtained from* S_t *by ordering its vertex set. For all* $c > 1/t$ *and all* $\varepsilon > 0$, *there is a graph* G, *and a blockade* $(B_1, \ldots, B_{t+1})$ *in* G, *with local degree less than* εW *and* $(\varepsilon W, \varepsilon W^c)$-*cohesive where* W *is its width, such that there is no ordered* $\mathcal{B}$-*rainbow copy of* S_t^+ *in* G.

Proof. We may assume that c, ε are rational, by slightly decreasing them if necessary. We call the vertex of S_t of degree t its *centre*. From the symmetry we may assume that the centre is the last in the ordering of S_t^+. Choose d such that $c > d > 1/t$, where $d - 1/t < (c - 1/t)/(t - 1)$ and $d < 2/t$, and d is rational. Choose an integer n such that $n^c, n^d, n^{1/t}, \varepsilon n, n^{1-d}, \varepsilon n^d$ are all integers, and n is large enough to satisfy each of 5.6, 5.7 and 5.8. Choose K as in 5.6 and 5.7.

Take $t + 2$ pairwise disjoint sets $B_0, B_1, \ldots, B_{t+1}$, where $|B_0| = n^{2/t}$ and $B_1, \ldots, B_{t+1}$ all have cardinality n. We attach bipartite graph onto various pairs of the sets $B_0, \ldots, B_{t+1}$ as follows:

- Let $J_{t+1,0}$ be a copy of the graph of 5.6 with bipartition B_{t+1}, B_0.

- For $1 \leq i \leq t$, let $J_{0,i}$ be a copy of the graph of 5.7 with bipartition (B_0, B_i).

- For $1 \leq i < j \leq t$, let $J_{i,j}$ be a copy of the graph of 5.8 with bipartition B_i, B_j.

Now for $1 \leq i < j \leq t+1$, and all $u \in B_i$ and $v \in B_j$, add an edge between u, v if they have a common neighbour in B_0. Finally, delete B_0; this defines a graph G, with a blockade $\mathcal{B} = (B_1, \ldots, B_{t+1})$ of width $W = n$, and we claim it satisfies the theorem.

Suppose first that there is an ordered $\mathcal{B}$-rainbow copy of S_t^+ in G. Thus there exists $v_{t+1} \in B_{t+1}$ adjacent in G to some $v_i \in B_i$ for $1 \leq i \leq t$, such that $v_1, \ldots, v_t$ are pairwise G-nonadjacent. From the construction, for $1 \leq i \leq t$ every vertex of B_i G-adjacent to v_{t+1} is $J_{0,i}$-adjacent to a vertex $w \in B_0$ that is $J_{t+1,0}$-adjacent to v_{t+1}. There are only $t - 1$ such vertices, because of the properties of $J_{t+1,0}$; so there exist distinct $i, j \in \{1, \ldots, t\}$ and $w \in B_0$, such that v_i, v_j, v_{t+1} are all adjacent (in $J_{0,i}, J_{0,j}, J_{t+1,0}$ respectively) to w. But then v_i is G-adjacent to v_j, a contradiction. This proves that there is no ordered $\mathcal{B}$-rainbow copy of S_t^+ in G.

To check the local degree of $\mathcal{B}$: each vertex in B_{t+1} is $J_{t+1,0}$-adjacent to at most $t - 1$ vertices in B_0; and each of these neighbours has at most n^{1-d} $J_{0,i}$-neighbours in B_i; so each vertex in B_{t+1} has at most $(t - 1)n^{1-d}$ G-neighbours in B_i, for $1 \leq i \leq t$, and $(t - 1)n^{1-d} < \varepsilon n$ if n is large enough.

Each vertex in B_i is $J_{0,i}$-adjacent to at most $(\varepsilon/2)n^d$ vertices in B_0, and they have degree less than n^{1-d} in $J_{t+1,0}$; so each vertex in B_i has at most $\varepsilon n/2$ G-neighbours in B_{t+1}, for $1 \leq i \leq t$.

For $1 \leq i < j \leq t$, each vertex in B_i has at most $(2/\varepsilon^2)n^{1-c}$ $J_{i,j}$-neighbours in B_j; and in addition, it is $J_{0,i}$-adjacent to most $(\varepsilon/2)n^d$ vertices in B_0, and they have degree at most n^{1-d} in $J_{0,j}$. Consequently each vertex in B_i is G-adjacent to at most $(2/\varepsilon^2)n^{1-c} + (\varepsilon/2)n^d n^{1-d}$ vertices in B_j. For large n, this is less than εn. The same holds for $j > i$. Consequently $\mathcal{B}$ has local degree less than εn.

To check that it is $(\varepsilon n, \varepsilon n^c)$-coherent: first, let $X \subseteq B_{t+1}$ and $Y \subseteq B_i$ where $1 \leq i \leq t$, with $|X| \geq \varepsilon n$ and $|Y| \geq \varepsilon n^c$. From the properties of $J_{t+1,0}$ there is a set $Z \subseteq B_0$ with $|Z| \geq (\varepsilon/e)n^{2/t}$,

23

and all its members have a $J_{t+1,0}$-neighbour in X; and from the properties of $J_{0,i}$, more than $n - \varepsilon n^c$ vertices in B_i have a $J_{0,i}$-neighbour in Z; and consequently some vertex in Y is G-adjacent to a vertex in X.

Next, let $X \subseteq B_{t+1}$ and $Y \subseteq B_i$ where $1 \leq i \leq t$, with $|X| \geq \varepsilon n^c$ and $|Y| \geq \varepsilon n$. By a similar argument it follows that X, Y are not anticomplete in G.

Finally, let $1 \leq i < j \leq t$, and let $X \subseteq B_i$ and $Y \subseteq B_j$, with $|X| \geq \varepsilon n$ and $|Y| \geq \varepsilon n^c$. From the properties of $J_{i,j}$, X, Y are not anticomplete in $J_{i,j}$ and hence not in G. This proves 5.9. ∎

References

[1] M. Chudnovsky, A. Scott, P. Seymour and S. Spirkl, "Pure pairs. I. Trees and linear anticomplete pairs", *Advances in Math.*, **375** (2020), 107396, `arXiv:1809.00919`.

[2] D. Conlon, J. Fox and B. Sudakov, "Recent developments in graph Ramsey theory", in: *Surveys in Combinatorics 2015*, Cambridge University Press, 2015, 49–118.

[3] M. Chudnovsky, J. Fox, A. Scott, P. Seymour and S. Spirkl, "Pure pairs. III. Sparse graphs with no polynomial-sized anticomplete pairs", *J. Graph Theory*, **95** (2020), 315–340, `arXiv:1810.00058`.

[4] M. Chudnovsky, J. Fox, A. Scott, P. Seymour and S. Spirkl, "Towards Erdős-Hajnal for graphs with no 5-hole", *Combinatorica* **39** (2019), 983–991, `arXiv:1803.03588`.

[5] M. Chudnovsky, A. Scott, P. Seymour and S. Spirkl, "Erdős-Hajnal for graphs with no five-hole", *Proc. London Math. Soc.*, to appear, `arXiv:2102.04994`.

[6] P. Erdős, A. Hajnal and J. Pach, "A Ramsey-type theorem for bipartite graphs", *Geombinatorics* **10** (2000), 64–68.

[7] A. Liebenau, M. Pilipczuk, P. Seymour and S. Spirkl, "Caterpillars in Erdős-Hajnal", *J. Combinatorial Theory, Ser. B*, **136** (2019), 33–43, `arXiv:1810.00811`.

[8] H. Robbins, "A remark on Stirling's formula", *Amer. Math. Monthly* **62** (1955), 26–29.

[9] V. Rödl, "On universality of graphs with uniformly distributed edges", *Discrete Math.* **59** (1986), 125–134.

[10] A. Scott, P. Seymour and S. Spirkl, "Pure pairs. VIII. Excluding a sparse graph", submitted for publication, `arXiv:2201.04062`.