J. Eur. Math. Soc. 26, 275–287 (2024) DOI 10.4171/JEMS/1362

István Tomon

String graphs have the Erdős–Hajnal property

Received February 28, 2020; revised May 23, 2022

Abstract. The following Ramsey-type question is one of the central problems in combinatorial geometry. Given a collection of certain geometric objects in the plane (e.g. segments, rectangles, convex sets, arcwise connected sets) of size n, what is the size of the largest subcollection in which either any two elements have a nonempty intersection, or any two elements are disjoint? We prove that there exists an absolute constant $c > 0$ such that if $\mathcal{C}$ is a collection of n curves in the plane, then $\mathcal{C}$ contains at least n^c elements that are pairwise intersecting, or n^c elements that are pairwise disjoint. This resolves a problem raised by Alon, Pach, Pinchasi, Radoičić and Sharir, and Fox and Pach. Furthermore, as any geometric object can be arbitrarily closely approximated by a curve, this shows that the answer to the aforementioned question is at least n^c for any collection of n geometric objects.

Keywords. String graphs, Ramsey theory

1. Introduction

Erdős–Hajnal property

One of the starting points of Ramsey theory is the classical result of Erdős and Szekeres [7] that if G is a graph on n vertices, then G contains either a clique or an independent set of size at least $\frac{1}{2}\log_2 n$. This bound is optimal up to a constant factor: using probabilistic techniques, Erdős [5] proved the existence of graphs with no clique or independent set of size larger than $2\log_2 n$. However, Erdős and Hajnal [5] noticed that the lower bound can be significantly improved if we restrict our attention to a nontrivial hereditary family of graphs. Here, a family of graphs is *hereditary* if it is closed under taking induced subgraphs, and it is *nontrivial* if it is not the family of all graphs. Indeed, they proved that if $\mathcal{G}$ is such a family, then there exists a constant $c = c(\mathcal{G}) > 0$ such that each $G \in \mathcal{G}$ with n vertices contains a clique or an independent set of size at least $e^{c\sqrt{\log n}}$. The celebrated Erdős–Hajnal conjecture asks whether this bound can be improved to n^c. This motivates the following definitions.

István Tomon: Department of Mathematics and Mathematical Statistics, Umeå University, 901 87 Umeå, Sweden; istvantomon@gmail.com

Mathematics Subject Classification (2020): Primary 05C55; Secondary 52C10

Definition 1. A family $\mathcal{G}$ of graphs has the *Erdős–Hajnal property* if there exists $c = c(\mathcal{G}) > 0$ such that every $G \in \mathcal{G}$ contains a clique or an independent set of size at least $|V(G)|^c$.

Also, a family $\mathcal{G}$ of graphs has the *strong Erdős–Hajnal property* if there exists $b = b(\mathcal{G}) > 0$ such that the following holds for every $G \in \mathcal{G}$. There exist two disjoint sets $A, B \subset V(G)$ such that $|A| = |B| \geq b|V(G)|$, and either there are no edges between A and B, or every $a \in A$ is joined to every $b \in B$ by an edge.

It was shown in [1] that the strong Erdős–Hajnal property implies the Erdős–Hajnal property in hereditary families. However, not every nontrivial hereditary family has the strong Erdős–Hajnal property. Despite the considerable attention the Erdős–Hajnal conjecture received, it is mostly wide open; see the survey of Chudnovsky [3] for a general reference.

Erdős–Hajnal type questions are also extensively studied in geometric settings. The *intersection graph* of a family $\mathcal{C}$ of geometric objects is the graph whose vertices correspond to the elements of $\mathcal{C}$, and two vertices are joined by an edge if the corresponding sets have a nonempty intersection.

Classical intersection graphs

Perhaps one of the first geometric Erdős–Hajnal type results is the classical folklore result that if G is the intersection graph of a family of n intervals, then G is perfect, so G contains either a clique or an independent set of size at least $n^{1/2}$. Larman et al. [20] proved that if G is the intersection graph of axis-parallel rectangles, then G contains a clique or an independent set of size $\Omega(\sqrt{n/\log n})$, and it is open whether this bound can be improved to $\Omega(\sqrt{n})$. In the same paper, they also proved that if $\mathcal{C}$ is a collection of n convex sets, then $\mathcal{C}$ contains a subcollection of size $n^{1/5}$ such that either any two elements have a nonempty intersection, or any two elements are disjoint. On the other hand, Kynčl [19] constructed a collection of n segments whose intersection graph has no clique or independent set of size larger than $n^{\log 8/\log 169} \approx n^{0.405}$. Kostochka [17] proved that the intersection graph of n chords of a cycle contains either a clique or an independent set of size at least $\Omega(\sqrt{n/\log n})$, and he also showed that this bound is best possible up to a constant factor. Pach and Solymosi [24] proved that the family of intersection graphs of line segments has the strong Erdős–Hajnal property.

Modern geometric graphs

While intersection graphs of intervals, segments, convex sets, etc. were the subjects of study in the mid and late 20th century, in the past two decades the focus shifted to more general geometric graphs. One such generalization is semi-algebraic graphs, that is, graphs whose vertices correspond to points in $\mathbb{R}^d$, and the edges are defined by polynomial relations (for precise definitions, see [1]). Indeed, intersection graphs of intervals, segments, and rectangles are special instances of semi-algebraic graphs. However, intersection graphs of convex sets are not. In general, it was proved by Alon et al. [1] that if $\mathcal{G}$ is a family of semi-algebraic graphs of bounded complexity, then $\mathcal{G}$ has the strong Erdős–Hajnal property.

Another generalization, which is the main interest of our paper, is string graphs. A *curve* (or a *string*) is the image of a continuous function $\phi : [0, 1] \to \mathbb{R}^2$. A *string graph* is the intersection graph of a family of curves. String graphs were introduced by Benzer [2] in 1959 to study topological properties of genetic structures, and later Sinden [26] considered such graphs to model printed circuits. Since then, combinatorial and computational properties of string graphs have been extensively studied. Note that, in a certain sense, curves are the most general geometric objects on the plane one can consider: indeed, when talking about geometric objects, one of the weakest geometric properties one should require is arcwise connectedness, and any arcwise connected set on the plane can be approximated arbitrarily closely by curves. In particular, all of the aforementioned intersection graphs are string graphs as well.

The question whether the family of string graphs has the Erdős–Hajnal property became one of the central problems in the area, and was settled in a number of interesting special cases. This question first appeared in the paper of Alon et al. [1], and later in papers of Fox and Pach [9] and Fox, Pach, and Tóth [14]. In particular, [9] is a nice survey on the topic. In the aforementioned paper of Larman et al. [20] it is also established that if G is the intersection graph of n x-monotone curves (a curve is *x-monotone* if every vertical line intersects it in at most one point), then G contains a clique or an independent set of size at least $n^{1/5}$. Also, it was proved by Fox, Pach and Tóth [14] that for every fixed k, the family of intersection graphs of curves, where any two curves intersect in at most k points, has the strong Erdős–Hajnal property. Here, let us remark that a surprising construction of Kratochvíl and Matoušek [18] shows the existence of string graphs on n vertices such that in any realization of these graphs with curves there are two curves that intersect in at least $2^{\Omega(n)}$ points. In general, Fox and Pach [13] proved that if G is a string graph on n vertices, then G contains either a clique or an independent set of size $n^{\Omega(1/\log\log n)}$.

Our results

The main result of our paper is that the family of string graphs has the Erdős–Hajnal property, which implies the Erdős–Hajnal property of all of the aforementioned families of intersection graphs.

Theorem 1. *There exists an absolute constant $c > 0$ such that for every positive integer n, if G is a string graph on n vertices, then G contains either a clique or an independent set of size at least n^c.*

Let us remark that Theorem 1 has no analogue in higher dimensions. Indeed, Tietze [27] proved that any graph can be realized as the intersection graph of convex sets in $\mathbb{R}^3$.

Our proof of Theorem 1 closely follows the path laid out by the works of Fox and Pach [11–13]. In the next subsection, we discuss their ideas and outline our proof strategy. We also introduce our notation, which is mostly conventional.

2. Overview of the proof

Given a graph G and two subsets A and B of $V(G)$, say that A *is complete to* B if every $a \in A$ and is joined by an edge to every $b \in B$. A *spanning subgraph* of G is a subgraph G' with $V(G) = V(G')$.

A graph G is a *comparability graph* if there exists a partial ordering $\prec$ on $V(G)$ such that for any $v, w \in V(G)$, we have $v \prec w$ or $w \prec v$ if and only if vw is an edge of G. Also, G is an *incomparability graph* if it is the complement of a comparability graph.

Previous approach

It turns out that string graphs and incomparability graphs are closely related. Indeed, it was proved by Lovász [22] and re-proved in [15, 25] that every incomparability graph is a string graph. On the other hand, Fox and Pach [12] proved that every dense string graph contains a dense incomparability graph as a spanning subgraph. (Here, by a *dense graph* we refer to a sequence of graphs G_i with $|V(G_i)| \to \infty$ and $\liminf |E(G_i)|/|V(G_i)|^2 > 0$.) This result is going to be one of the main ingredients of our proof of Theorem 1; see Section 5 for more details.

One approach to proving Theorem 1 would be to show that the family of string graphs has the strong Erdős–Hajnal property. Indeed, with the exception of intersection graphs of x-monotone curves, every family of intersection graphs where the Erdős–Hajnal property is known also has the strong Erdős–Hajnal property.

Unfortunately, the family of string graphs does not have the strong Erdős–Hajnal property, but it has something close to it. Let G be a string graph on n vertices. A separator theorem of Lee [21] shows that if G is sufficiently sparse (meaning that $|E(G)| \le \lambda n^2$ for some small constant λ), then $V(G)$ contains two linear sized subsets with no edges between them. This result is going to be another important ingredient in our proof; see Section 5 for more details. On the other hand, by a result of Fox [8], every dense incomparability graph on n vertices contains two disjoint sets A and B of size $\Omega(n/\log n)$ such that A is complete to B, and this bound is the best possible up to a constant factor. But then remembering that every dense string graph contains a dense incomparability graph, we find that if G is dense, then G contains two disjoint sets A and B of size $\Omega(n/\log n)$ such that A is complete to B. Therefore, one can conclude the following "almost-strong Erdős–Hajnal property":

Theorem 2 ([12, 21]). *If G is a string graph on n vertices, then $V(G)$ contains two disjoint sets A and B such that either*

(1) $|A| = |B| = \Omega(n)$ *and there are no edges between A and B, or*

(2) $|A| = |B| = \Omega(n/\log n)$ *and A is complete to B.*

Again, these bounds are the best possible up to a constant factor. Then, by a recursive argument this theorem implies that if G is a string graph on n vertices, then G contains a clique or an independent set of size $n^{\Omega(1/\log\log n)}$ (see [13]).

New ideas

In order to improve this bound, we do the following. Instead of proving the strong Erdős–Hajnal property, we prove something slightly weaker which we call the *quasi-Erdős–Hajnal property*. Roughly, a family of graphs $\mathcal{G}$ has this property if for every $G \in \mathcal{G}$ there exist some $t \geq 2$ and t disjoint subsets $X_1, \ldots, X_t$ of $V(G)$ such that $|X_1|, \ldots, |X_t|$ are "large" with respect to t and $|V(G)|$, and either there are no edges between X_i and X_j for $1 \leq i < j \leq t$, or X_i is complete to X_j for $1 \leq i < j \leq t$. It turns out that in hereditary families, the quasi-Erdős–Hajnal property is equivalent to the Erdős–Hajnal property; see Section 4 for more details. Then, our main contribution to the proof of Theorem 1 is that in every dense incomparability graph G, there exist $t \geq 2$ and t disjoint subsets $X_1, \ldots, X_t$ such that $|X_1|, \ldots, |X_t|$ are "large" with respect to t and $|V(G)|$, and X_i is complete to X_j for $1 \leq i < j \leq t$. This can be found in Section 3. But then, together with the aforementioned results of Lee [21] and Fox and Pach [12], this implies that the family of string graphs has the quasi-Erdős–Hajnal property.

Notation

In the rest of our paper, we use the following standard graph-theoretic notations. If G is a graph, $\Delta(G)$ denotes the maximum degree of G, and if $v \in V(G)$, then $N(v) = \{w \in V(G) : vw \in E(G)\}$ is the neighborhood of v. If U is a subset of the vertex set, then $G[U]$ is the subgraph of G induced on U. Given a poset P with partial ordering $\prec$, a total ordering $<_l$ is a *linear extension* of $\prec$ if $x \prec y$ implies $x <_l y$ for all $x, y \in P$. It is well known that every partial ordering has a linear extension (which might not be unique). Also, if $A, B \subset P$, then we write $A <_l B$ if $a <_l b$ for all $a \in A$ and $b \in B$. We omit floors and ceilings whenever they are not crucial.

3. Incomparability graphs

Our main contribution to the proof of Theorem 1 is the following result about partial orders, which might be of independent interest.

Theorem 3. *For every $\alpha > 0$ there exists $c > 0$ such that the following holds. Let G be an incomparability graph with n vertices and at least $\alpha\binom{n}{2}$ edges. Then there exist $t \geq 2$ and t disjoint subsets $X_1, \ldots, X_t$ of $V(G)$ such that X_i is complete to X_j for $1 \leq i < j \leq t$, and $(n/|X_i|)^c < t$ for $i = 1, \ldots, t$.*

We would like to emphasize that t depends on the incomparability graph G. In order to prove this theorem, it is slightly more convenient to work with comparability graphs instead of incomparability graphs, so we prove the following equivalent statement instead.

Theorem 4. *For every $\alpha > 0$ there exists $c > 0$ such that the following holds. Let P be a comparability graph with n vertices and at most $(1-\alpha)\binom{n}{2}$ edges. Then there exist $t \geq 2$ and t disjoint subsets $X_1, \ldots, X_t$ of $V(P)$ such that there are no edges between X_i and X_j for $1 \leq i < j \leq t$, and $(n/|X_i|)^c < t$ for $i = 1, \ldots, t$.*

Let us prove this theorem. Instead of working with very dense comparability graphs, we would like to work with sparse ones. With the help of a technical lemma, we show that if P satisfies the conditions of the previous theorem, then P contains two linear sized subsets A and B such that the bipartite graph induced between A and B is sufficiently sparse. In order to show this, we make use of the following simple result.

Lemma 5. *For every* $0<\alpha<1$ *there exists* $\alpha_1>0$ *such that the following holds. If* G *is a graph with* $n\geq 2$ *vertices and at most* $(1-\alpha)\binom{n}{2}$ *edges, then* G *contains an induced subgraph* G' *such that* $|V(G')|\geq\alpha_1 n$ *and* $\Delta(G')\leq(1-\alpha_1)|V(G')|$.

Proof. We show that $\alpha_1=\alpha/8$ suffices. Let H be the complement of G. Then H has at least $\alpha\binom{n}{2}$ edges. Keep removing vertices of H as long as H has a vertex of degree less than $\frac{\alpha}{8}n$, and let H' be the resulting induced subgraph of H. In total, we remove at most $\frac{\alpha}{8}n^2$ edges, so $|E(H')|\geq\frac{\alpha}{2}\binom{n}{2}$. Let $n'=|V(H')|$. Then H' is nonempty and has minimum degree at least $\frac{\alpha}{8}n\geq\alpha_1 n'$. Also, as $\binom{n'}{2}\geq|E(H')|$, we get $n'\geq\frac{\sqrt{\alpha}}{4}n>\alpha_1 n$. Setting G' to be the complement of H' gives the desired induced subgraph of G. ∎

Lemma 6. *For every* $\alpha,\varepsilon>0$ *there exists* $\beta>0$ *such that the following holds. Let* P *be a comparability graph with* n *vertices and at most* $(1-\alpha)\binom{n}{2}$ *edges, and let* $<_l$ *be a linear extension of the underlying partial order. Then there exist two disjoint subsets* A *and* B *of* $V(P)$ *such that* $A<_l B$, $|A|=|B|\geq\beta n$, $|N(v)\cap B|\leq\varepsilon|B|$ *for every* $v\in A$, *and* $|N(w)\cap A|\leq\varepsilon|A|$ *for every* $w\in B$.

Proof. By Lemma 5, there exists $\alpha_1>0$ (depending only on α) such that P contains an induced subgraph P' with $n'=|V(P')|\geq\alpha_1 n$ and $\Delta(P')\leq(1-\alpha_1)n'$. We show that $\beta=\alpha_1^2\min\{\alpha_1,\varepsilon\}/36$ suffices. Suppose that there exists no pair (A,B) satisfying the desired conditions.

Let $\prec$ be the partial ordering of the underlying poset of P. Let T be the $\frac{\alpha_1}{6}n'$ largest elements of P' with respect to the linear extension $<_l$, let $S=P'\setminus T$, and cut T into two equal sized parts, X and Y, with $X<_l Y$. Note that $|X|=|Y|=\frac{\alpha_1}{12}n'\geq\frac{\alpha_1^2}{12}n$.

Say that a vertex $v\in S$ is *heavy* if $|N(v)\cap X|\geq\beta n$. Observe that for every $v\in S$, setting $A_0=N(v)\cap X$ and $B_0=Y\setminus N(v)$, there are no edges between A_0 and B_0. Indeed, otherwise, if $x\in A_0$ and $y\in B_0$ are joined by an edge, then $x<_l y$ implies $x\prec y$, and as $v\prec x$, we find that $v\prec y$, contradicting $y\notin N(v)$. But if v is heavy, then $|A_0|\geq\beta n$, so we must have $|B_0|<\beta n\leq\frac{\alpha_1}{3}|Y|$. Otherwise, taking A and B to be arbitrary βn-element subsets of A_0 and B_0, respectively, the pair (A,B) satisfies the required conditions. Thus, $|N(v)\cap Y|>(1-\alpha_1/3)|Y|$ for every heavy vertex v.

If $H\subset S$ is the set of heavy vertices, then the number of edges between H and Y is at least $(1-\alpha_1/3)|Y|\,|H|$, which implies that there exists a vertex in Y of degree at least $(1-\alpha_1/3)|H|$. Therefore, by the maximum degree condition, we can write $(1-\alpha_1/3)|H|\leq(1-\alpha_1)n'$, which gives

$$|H|\leq\frac{1-\alpha_1}{1-\alpha_1/3}n'<\left(1-\frac{\alpha_1}{3}\right)n'.$$

But then $|S \setminus H| = n' - |T| - |H| > \frac{\alpha_1}{6}n'$, or in other words, there are at least $\frac{\alpha_1}{6}n'$ vertices $v \in S$ such that $|N(v) \cap X| < \beta n$. Let A be an arbitrary set of $\frac{\alpha_1}{24}n'$ such vertices, where we remark that $|A| > \beta n$ is satisfied. The number of edges between A and X is at most $\beta n|A|$, so the number of vertices $v \in X$ such that $|N(v) \cap A| \geq \varepsilon|A|$ is at most $\frac{\beta}{\varepsilon}n < |X|/2$. Delete all such vertices from X, and perhaps some more, to get a set B of size $|A| = |X|/2 = \frac{\alpha_1}{24}n'$. Then we have $|A| = |B| > \beta n$, $|N(v) \cap B| \leq \beta n < \varepsilon|B|$ for every $v \in A$, and $|N(v) \cap A| \leq \varepsilon|A|$ for every $v \in B$. Therefore, the pair (A, B) satisfies the desired conditions, a contradiction. ∎

Most of the work needed to prove Theorem 4 is put into the following lemma.

Lemma 7. *There exist positive real numbers ε and δ such that the following holds. Let P be a comparability graph on $2n$ vertices, and let $<_l$ be a linear extension of the underlying poset. Let A be the n smallest elements of P with respect to $<_l$, let $B = P \setminus A$, and suppose that $|N(v) \cap B| \leq \varepsilon n$ for every $v \in A$ and $|N(w) \cap A| \leq \varepsilon n$ for every $w \in B$. Then there exist $t \geq 2$ and t disjoint sets $X_1, \ldots, X_t \subset V(P)$ such that there are no edges between X_i and X_j for $1 \leq i < j \leq t$, and $\delta(n/|X_i|)^{1/2} < t$ for $i = 1, \ldots, t$.*

Proof. We prove that we can choose $\varepsilon = \frac{1}{500}$ and $\delta = \frac{1}{100}$. Let $\prec$ be the partial ordering of the underlying poset of P.

Let $J = J_0 = \lfloor \log_2 \varepsilon n \rfloor + 1$. For $j = 1, \ldots, J_0$, let $t_j = n^{1/2}2^{j/2}$. Then

$$\sum_{i=1}^{J_0} t_i = \sum_{i=1}^{J_0} n^{1/2}2^{i/2} \leq 2n\varepsilon^{1/2}\frac{1}{1-2^{-1/2}} < \frac{n}{4}. \tag{1}$$

Also, let $A' = \emptyset$ and $B' = \emptyset$. In what follows, we define an algorithm, which we shall refer to as the *main algorithm*, which will find and output the desired t and the t sets $X_1, \ldots, X_t$. During each step of the algorithm, we will make the following changes: we will move certain elements of A into A', move certain elements of B into B', and decrease J. We think of the elements of A' and B' as "leftovers". We will make sure that at the end of each step of the algorithm, the following properties are satisfied:

(a) $|A| + |A'| = |B| + |B'| = n$,

(b) $|A'|, |B'| \leq 2\sum_{i=J+1}^{J_0} t_i$,

(c) for every $v \in B$, $|N(v) \cap A| < 2^J$.

Note that by (1) and properties (a) and (b), we have $|A|, |B| \geq n/2$. Also, these properties are certainly satisfied at the beginning of the algorithm. Now let us describe a general step of our main algorithm.

Main algorithm. If $J = 0$, then stop the main algorithm, and output $t = 2$ and $X_1 = A$, $X_2 = B$. Note that in this case there is no edge between A and B by property (c), and $|A|, |B| \geq n/2$. By the choice of δ, this output has the desired properties.

Suppose that $J \geq 1$. For $i = 1, \ldots, J$, let V_i be the set of vertices $v \in B$ such that $2^{i-1} \leq |N(v) \cap A| < 2^i$, and let V_0 be the set of vertices $v \in B$ such that $N(v) \cap A = \emptyset$. Then by property (c), $B = \bigcup_{i=0}^{J} V_i$.

Let $1 \le k \le J$ be maximal such that $t_k < |V_k|$. First, consider the case where there exists no such k. Then

$$n - 2\sum_{i=J+1}^{J_0} t_i - |V_0| \le n - |B'| - |V_0| = |B| - |V_0| = \sum_{i=1}^{J} |V_i| \le \sum_{i=1}^{J} t_i,$$

where the first inequality follows from property (b), and the first equality is a consequence of property (a). Comparing the left and right hand sides, and using (1), we get $|V_0| \ge n/2$. In this case, stop the algorithm and output $t = 2$, $X_1 = V_0$ and $X_2 = A$. Note that $\delta(n/|X_i|)^{1/2} < t$ is satisfied for $i = 1, 2$.

Now suppose that there exists such a k. Remove the elements of V_i for $i > k$ from B, and add them to B'. Thus we add at most $\sum_{i=k+1}^{J} t_i$ elements to B'. Set $J := k$. Then properties (a)–(c) are still satisfied (we note that J may not have decreased yet).

Now we shall run a *subalgorithm*. Let $W_0 = V_k$. Then with the help of the subalgorithm we construct a sequence $W_0 \supset \dots \supset W_r$ satisfying the following properties. During each step of the subalgorithm, we either find our desired t and t sets $X_1, \dots, X_t$, or move certain elements of A to A'. At the end of the l-th step of the subalgorithm, W_l be the set of vertices in B that still have at least 2^{k-1} neighbors in A. We stop the algorithm if W_l is too small.

Subalgorithm. Suppose that W_l is already defined. If $|W_l| < 2t_k$, then let $r = l$, stop the subalgorithm, remove the elements of W_l from B and add them to B'. Update $J := k - 1$, and move to the next step of the main algorithm. Note that B' has property (b). Later, we will see that all the other properties are satisfied.

On the other hand, if $|W_l| \ge 2t_k$, we define W_{l+1} as follows. Let $x_l = |W_l|/t_k$. Say that a vertex $v \in A$ is *heavy* if

$$|N(v) \cap W_l| \ge \frac{x_l 2^k}{t_k}|W_l| = \left(\frac{|W_l|}{t_k}\right)^2 2^k = \frac{|W_l|^2}{n} =: \Delta_l,$$

and let H_l be the set of heavy vertices. Counting the number f of edges between H_l and W_l in two ways, we can write

$$|H_l|\Delta_l \le f < |W_l|2^k,$$

which gives $|H_l| < t_k/x_l$. Remove the elements of H_l from A and add them to A'. Examine how the degrees of the vertices in W_l have changed, and consider the following two cases:

Case 1: At least $|W_l|/2$ vertices in W_l have at least 2^{k-1} neighbors in A. Let T be the set of vertices in W_l that have at least 2^{k-1} neighbors in A, so $|T| \ge |W_l|/2$. Pick each element of A with probability $p = 2^{-k}$, and let S be the set of selected vertices. Say that $v \in T$ is *good* if $|N(v) \cap S| = 1$, and let Y be the set of good vertices. Then

$$\mathbb{P}(v \text{ is good}) = |N(v) \cap A|p(1-p)^{|N(v)\cap A|-1} \ge \tfrac{1}{2}(1 - 2^{-k})^{2^k} \ge \tfrac{1}{6},$$

so $\mathbb{E}(|Y|) \geq |T|/6 \geq |W_l|/12$. Therefore, there exists a choice for S such that $|Y| \geq |W_l|/12$; let us fix such an S. For each $v \in S$, let Y_v be the set of elements $w \in Y$ such that $N(w) \cap S = \{v\}$. An important observation is that if $v, v' \in S$ and $v \neq v'$, then there is no edge between Y_v and $Y_{v'}$. Indeed, otherwise, if $w \in Y_v$ and $w' \in Y_{v'}$ are such that $w \prec w'$, then $v \prec w \prec w'$, which means that $\{v, v'\} \in N(w') \cap S$, contradicting the assumption that w' is good. Also, note that

$$|Y_v| \leq |N(v) \cap W_l| \leq \min\{\varepsilon n, \Delta_l\} =: \Delta_l'.$$

In other words, the sets Y_v for $v \in S$ partition Y into sets of size at most Δ_l'. Here, we have

$$\frac{|Y|}{\Delta_l'} \geq \frac{|W_l|}{12\Delta_l'} \geq \max\left\{\frac{n}{12|W_l|}, \frac{|W_l|}{\varepsilon n}\right\}.$$

By the choice of ε, the right hand side is always at least 6. But then we can partition S into $t \geq \frac{|Y|}{3\Delta_l'} \geq 2$ parts $S_1, \dots, S_t$ such that the sets $X_i = \bigcup_{v \in S_i} Y_v$ have size at least Δ_l' for $i = 1, \dots, t$. The resulting sets $X_1, \dots, X_t$ are such that there are no edges between X_i and X_j for $1 \leq i < j \leq t$ and

$$t \geq \frac{|Y|}{3\Delta_l'} \geq \frac{n}{36|W_l|} \geq \frac{1}{36}\left(\frac{n}{\Delta_l}\right)^{1/2} \geq \frac{1}{36}\left(\frac{n}{|X_i|}\right)^{1/2}.$$

Stop the main algorithm, and output t and $X_1, \dots, X_t$. By the choice of δ, this output has the desired properties.

Case 2: At most $|W_l|/2$ vertices in W_l have at least 2^{k-1} neighbors in A. In this case, define W_{l+1} as the set of elements of W_l with at least 2^{k-1} neighbors in A (then W_{l+1} is the set of all elements in B with at least 2^{k-1} neighbors in A as well). Also, move to the next step of the subalgorithm.

Let us check that if the main algorithm is not terminated, then at the end of the subalgorithm, properties (a)–(c) are still satisfied. Indeed, (a) and (c) are clearly true, and (b) holds for B'. It remains to show that (b) holds for A' as well. Note that as $|W_{l+1}| \leq |W_l|/2$ for $l = 0, \dots, r-1$, and $|W_{r-1}| \geq 2t_k$, we have $|W_l| \geq 2^{r-l}t_k$ and $x_l \geq 2^{r-l}$. Compared to the first step of the subalgorithm, $|A'|$ increased by

$$\sum_{l=0}^{r-1} |H_l| \leq \sum_{l=0}^{r-1} \frac{t_k}{x_l} \leq \sum_{l=0}^{r-1} \frac{t_k}{2^{r-l}} < t_k.$$

Therefore, property (b) also holds.

Note that in every step of the main algorithm, J decreases by at least 1, so the main algorithm will stop in a finite number of steps, and it will output the desired t and t sets $X_1, \dots, X_t$. This finishes the proof. ∎

Now we are ready to prove the main theorem of this section.

Proof of Theorem 4. Let $\varepsilon, \delta > 0$ be the constants given by Lemma 7. By Lemma 6, there exists $\beta > 0$ such that the following holds. Let $<_l$ be a linear extension of the underlying partial order of P. Then there exist two disjoint subsets A and B of P such that $A <_l B$, $|A| = |B| \geq \beta n$, $|N(v) \cap B| \leq \varepsilon|B|$ for every $v \in A$, and $|N(w) \cap A| \leq \varepsilon|A|$ for every $w \in B$.

Apply Lemma 7 to the comparability graph $P' = P[A \cup B]$. We conclude that there exist $t \geq 2$ and t disjoint subsets $X_1, \dots, X_t$ of P' such that there are no edges between X_i and X_j for $1 \leq i < j \leq t$, and $\delta(|A|/|X_i|)^{1/2} < t$ for $i = 1, \dots, t$. Here, $\delta(|A|/|X_i|)^{1/2} \geq \delta\beta^{1/2}(n/|X_i|)^{1/2}$. Choose $c > 0$ small enough such that $2 > (\frac{4}{\delta^2\beta})^c$. This choice guarantees that if $\delta\beta^{1/2}(n/|X_i|)^{1/2} \geq 2$, then $(n/|X_i|)^c < \delta\beta^{1/2}(n/|X_i|)^{1/2} < t$. On the other hand, if $\delta\beta^{1/2}(n/|X_i|)^{1/2} < 2$, then $n/|X_i| \leq \frac{4}{\delta^2\beta}$, so $(n/|X_i|)^c < 2 \leq t$. Therefore, c suffices. ∎

4. The quasi-Erdős–Hajnal property

Say that a family $\mathcal{G}$ of graphs has the *quasi-Erdős–Hajnal property* if there exists a constant $c = c(\mathcal{G}) > 0$ such that the following holds for every $G \in \mathcal{G}$ with at least two vertices: there exist $t \geq 2$ and t disjoint subsets $X_1, \dots, X_t$ of $V(G)$ such that $t \geq (|V(G)|/|X_i|)^c$ for $i = 1, \dots, t$, and either

(i) X_i is complete to X_j for $1 \leq i < j \leq t$, or

(ii) there is no edge between X_i and X_j for $1 \leq i < j \leq t$.

We show that for hereditary graph families, the Erdős–Hajnal property is actually equivalent to the quasi-Erdős–Hajnal property.

Lemma 8. *If $\mathcal{G}$ is a hereditary family of graphs, then $\mathcal{G}$ has the Erdős–Hajnal property if and only if it has the quasi-Erdős–Hajnal property.*

Proof. If $\mathcal{G}$ has the Erdős–Hajnal property, then there exists $c > 0$ such that every $G \in \mathcal{G}$ contains a clique or an independent set of size at least $|V(G)|^c$. But then setting $t = |V(G)|^c$ and defining $X_1, \dots, X_t$ to be the single element sets formed by the vertices of such a clique or independent set shows that $\mathcal{G}$ also has the quasi-Erdős–Hajnal property. It remains to show the other direction.

Suppose that $\mathcal{G}$ has the quasi-Erdős–Hajnal property with a constant $c > 0$. Let $G \in \mathcal{G}$ be a graph on n vertices. Let $\mathcal{X} = \{V(G)\}$ and let H be the graph with vertex set $\mathcal{X}$ (that is, H has exactly one vertex, namely $V(G)$). We repeat the following procedure until every element of $\mathcal{X}$ has only one vertex. If $\mathcal{X}$ contains a set of size at least 2, say $X \in \mathcal{X}$, then consider the induced subgraph $G[X] \in \mathcal{G}$. Then there exist $t \geq 2$ and t disjoint subsets $X_1, \dots, X_t$ of X such that $t \geq (|X|/|X_i|)^c$ for $i = 1, \dots, t$, and either

(i) X_i is complete to X_j for $1 \leq i < j \leq t$, or

(ii) there is no edge between X_i and X_j for $1 \leq i < j \leq t$.

Remove the set X from $\mathcal{X}$ and add the sets $X_1, \dots, X_t$. Also, if (i) happens, replace the vertex X in H with a clique on $\{X_1, \dots, X_t\}$, otherwise, replace X in H with an independent set on $\{X_1, \dots, X_t\}$. More precisely, X_i has the same neighborhood as X had outside of $\{X_1, \dots, X_t\}$, and $\{X_1, \dots, X_t\}$ induces either a clique or an independent set depending on whether (i) or (ii) holds, respectively.

Note that $\sum_{i=1}^{t} |X_i|^c \geq |X|^c$, therefore the sum $\sum_{Y \in \mathcal{X}} |Y|^c$ did not decrease after the change. Thus, we have $\sum_{Y \in \mathcal{X}} |Y|^c \geq n^c$ in each step of the procedure. This implies that at the end of the procedure, that is, when every element of $\mathcal{X}$ is a single vertex set, we have $|\mathcal{X}| \geq n^c$.

Moreover, at each step of the procedure, the graph H is a cograph. It is well known that cographs are perfect, therefore, at the end of the procedure, either H or its complement contains a clique of size at least $n^{c/2}$. This clique corresponds to a clique or an independent set of size at least $n^{c/2}$ in G. As this is true for every $G \in \mathcal{G}$, $\mathcal{G}$ has the Erdős–Hajnal property with constant $c/2$. ■

5. String graphs

In this section, we put all the ingredients together to prove Theorem 1.

A *separator* in a graph G is a subset S of the vertices such that after the removal of S, every connected component of G has size at most $2|V(G)|/3$. It was proved by Fox and Pach [10] that if G is the intersection graph of a family of n curves and g is the total number of crossings between the curves, then G contains a separator of size $O(\sqrt{g})$. Later, Fox and Pach [11] showed that if G is a string graph with m edges, then it contains a separator of size $O(m^{3/4}\sqrt{\log m})$, and proposed the conjecture that one can also find a separator of size $O(\sqrt{m})$, which is then optimal up to a constant factor. In [11, 13], Fox and Pach also provide a number of applications of the existence of small separators. The size of the smallest separator was improved to $O(\sqrt{m}\log m)$ by Matoušek [23], and recently Lee [21] completely settled the aforementioned conjecture of Fox and Pach. The result of Lee immediately implies the following lemma, which will be the first key ingredient in our proof.

Lemma 9. *There exists a constant $\lambda > 0$ such that the following holds. If G is a string graph with n vertices and at most λn^2 edges, then there exist two disjoint subsets X_1 and X_2 of vertices such that there are no edges between X_1 and X_2, and*

$$|X_1| = |X_2| \geq \lambda n.$$

Let us remark that this lemma also follows from a recent graph-theoretic result of Chudnovsky et al. [4]. Also, the present author [28] proved the following sharpening of Lemma 9: If the edge density of a string graph is below $1/4$, then one can find two linear sized sets of vertices with no edges between them. However, there are string graphs with edge density arbitrarily close to $1/4$ which contain only logarithmic sized such sets.

The final ingredient we need for our proof is the following result of Fox and Pach [12], which tells us that every dense string graph contains a dense incomparability graph as a spanning subgraph.

Lemma 10. *For every $\lambda > 0$ there exist $\varepsilon > 0$ such that the following holds. If G is a string graph with n vertices and at least λn^2 edges, then G contains a spanning subgraph G' such that G' is an incomparability graph with at least εn^2 edges.*

By Lemma 8, in order to prove Theorem 1, it is enough to show that the family of string graphs has the quasi-Erdős–Hajnal property. This almost immediately follows from a combination of the results discussed in this paper.

Theorem 11. *The family of string graphs has the quasi-Erdős–Hajnal property.*

Proof. Let λ be the constant given by Lemma 9, and let ε be the constant given by Lemma 10 (with respect to λ). Also, let c_0 be the constant c given by Theorem 3 with respect to $\alpha = 2\varepsilon$. We show that the family of string graphs has the quasi-Erdős–Hajnal property with exponent

$$c = \min\left\{c_0, \frac{1}{\log_2(1/\lambda)}\right\}.$$

Let G be a string graph with n vertices. If G has at most λn^2 edges, then G contains two disjoint subsets X_1 and X_2 with no edges between them such that $|X_1| = |X_2| \geq \lambda n$. Setting $t = 2$, we have $t \geq (1/\lambda)^c \geq (n/|X_i|)^c$ for $i = 1, 2$.

Now suppose that G has more than λn^2 edges. Then G contains an incomparability graph G' with at least εn^2 edges. Then, by Theorem 3, there exist $t \geq 2$ and t disjoint subsets $X_1, \ldots, X_t$ of G' such that X_i is complete to X_j for $1 \leq i < j \leq t$, and $(n/|X_i|)^c < t$ for $i = 1, \ldots, t$. This finishes the proof. ∎

Acknowledgments. We would like to thank the anonymous referee for useful comments and suggestions, and Jacob Fox and János Pach for their valuable remarks on the presentation of this paper. Also, we would like to thank Adam Zsolt Wagner for valuable discussions.

Funding. Most of this manuscript was prepared while the author was employed at ETH Zurich, where the research was supported by the SNSF grant 200021-175573.

References

[1] Alon, N., Pach, J., Pinchasi, R., Radoičić, R., Sharir, M.: Crossing patterns of semi-algebraic sets. J. Combin. Theory Ser. A **111**, 310–326 (2005) Zbl 1099.14048 MR 2156215

[2] Benzer, S.: On the topology of the genetic fine structure. Proc. Nat. Acad. Sci. USA **45**, 1607–1620 (1959)

[3] Chudnovsky, M.: The Erdös-Hajnal conjecture—a survey. J. Graph Theory **75**, 178–190 (2014) Zbl 1280.05086 MR 3150572

[4] Chudnovsky, M., Scott, A., Seymour, P., Spirkl, S.: Pure pairs. II. Excluding all subdivisions of a graph. Combinatorica **41**, 379–405 (2021) Zbl 1488.05427 MR 4282718

[5] Erdős, P.: Some remarks on the theory of graphs. Bull. Amer. Math. Soc. **53**, 292–294 (1947) Zbl 0032.19203 MR 19911

[6] Erdős, P., Hajnal, A.: Ramsey-type theorems. Discrete Appl. Math. **25**, 37–52 (1989) Zbl 0715.05052 MR 1031262

[7] Erdős, P., Szekeres, G.: A combinatorial problem in geometry. Compos. Math. **2**, 463–470 (1935) Zbl 0012.27010 MR 1556929

[8] Fox, J.: A bipartite analogue of Dilworth's theorem. Order **23**, 197–209 (2006) Zbl 1108.06002 MR 2308906

[9] Fox, J., Pach, J.: Erdős–Hajnal-type results on intersection patterns of geometric objects. In: Horizons of combinatorics, Bolyai Society Mathematical Studies 17, Springer, Berlin, 79–103 (2008) Zbl 1170.05327 MR 2432528

[10] Fox, J., Pach, J.: Separator theorems and Turán-type results for planar intersection graphs. Adv. Math. **219**, 1070–1080 (2008) Zbl 1148.05050 MR 2442062

[11] Fox, J., Pach, J.: A separator theorem for string graphs and its applications. Combin. Probab. Comput. **19**, 371–390 (2010) Zbl 1230.05186 MR 2607373

[12] Fox, J., Pach, J.: String graphs and incomparability graphs. Adv. Math. **230**, 1381–1401 (2012) Zbl 1244.05073 MR 2921183

[13] Fox, J., Pach, J.: Applications of a new separator theorem for string graphs. Combin. Probab. Comput. **23**, 66–74 (2014) Zbl 1287.05066 MR 3197967

[14] Fox, J., Pach, J., Tóth, C. D.: Intersection patterns of curves. J. London Math. Soc. (2) **83**, 389–406 (2011) Zbl 1238.05269 MR 2776643

[15] Golumbic, M. C., Rotem, D., Urrutia, J.: Comparability graphs and intersection graphs. Discrete Math. **43**, 37–46 (1983) Zbl 0502.05050 MR 680302

[16] Kostochka, A.: Coloring intersection graphs of geometric figures with a given clique number. In: Towards a theory of geometric graphs, Contemporary Mathematics 342, American Mathematical Society, Providence, RI, 127–138 (2004) Zbl 1064.05059 MR 2065259

[17] Kostochka, A. V.: Upper bounds on the chromatic number of graphs. Trudy Inst. Mat. (Novosibirsk) **10**, 204–226, 265 (1988) (in Russian) Zbl 0677.05027 MR 945704

[18] Kratochvíl, J., Matoušek, J.: String graphs requiring exponential representations. J. Combin. Theory Ser. B **53**, 1–4 (1991) Zbl 0675.05059 MR 1122293

[19] Kynčl, J.: Ramsey-type constructions for arrangements of segments. Eur. J. Combin. **33**, 336–339 (2012) Zbl 1230.05207 MR 2864419

[20] Larman, D., Matoušek, J., Pach, J., Törőcsik, J.: A Ramsey-type result for convex sets. Bull. London Math. Soc. **26**, 132–136 (1994) Zbl 0809.05090 MR 1272297

[21] Lee, J. R.: Separators in region intersection graphs. In: 8th Innovations in Theoretical Computer Science Conference, LIPIcs–Leibniz International Proceedings in Informatics 67, Schloss Dagstuhl, Leibniz-Zentrum für Informatik, Wadern, art. 1, 8 pp. (2017) Zbl 1402.05151 MR 3754925

[22] Lovász, L.: Perfect graphs. In: Selected topics in graph theory, 2, Academic Press, London, 55–87 (1983) Zbl 0536.05055 MR 797249

[23] Matoušek, J.: Near-optimal separators in string graphs. Combin. Probab. Comput. **23**, 135–139 (2014) Zbl 1287.05094 MR 3197972

[24] Pach, J., Solymosi, J.: Crossing patterns of segments. J. Combin. Theory Ser. A **96**, 316–325 (2001) Zbl 0989.05031 MR 1864126

[25] Pach, J., Tóth, G.: Comment on Fox news. Geombinatorics **15**, 150–154 (2006) Zbl 1145.05020 MR 2300647

[26] Sinden, F. W.: Topology of thin film RC-circuits. Bell System Technological J. 1639–1662 (1966) Zbl 0144.45601

[27] Tietze, H.: Über das Problem der Nachbargebiete im Raum. Monatsh. Math. Phys. **16**, 211–216 (1905) Zbl 36.0530.04 MR 1547413

[28] Tomon, I.: A sharp threshold phenomenon in string graphs. Discrete Comput. Geom. **67**, 338–359 (2022) Zbl 1480.05038 MR 4356254