Erdős-Hajnal conjecture for graphs with bounded VC-dimension

Jacob Fox * János Pach† Andrew Suk‡

Abstract

The *Vapnik-Chervonenkis dimension* (in short, VC-dimension) of a *graph* is defined as the VC-dimension of the set system induced by the neighborhoods of its vertices. We show that every n-vertex graph with bounded VC-dimension contains a clique or an independent set of size at least $e^{(\log n)^{1-o(1)}}$. The dependence on the VC-dimension is hidden in the $o(1)$ term. This improves the general lower bound, $e^{c\sqrt{\log n}}$, due to Erdős and Hajnal, which is valid in the class of graphs satisfying any fixed nontrivial hereditary property. Our result is almost optimal and nearly matches the celebrated Erdős-Hajnal conjecture, according to which one can always find a clique or an independent set of size at least $e^{\Omega(\log n)}$. Our results partially explain why most geometric intersection graphs arising in discrete and computational geometry have exceptionally favorable Ramsey-type properties.

Our main tool is a partitioning result found by Lovász-Szegedy and Alon-Fischer-Newman, which is called the "ultra-strong regularity lemma" for graphs with bounded VC-dimension. We extend this lemma to k-uniform hypergraphs, and prove that the number of parts in the partition can be taken to be $(1/\varepsilon)^{O(d)}$, improving the original bound of $(1/\varepsilon)^{O(d^2)}$ in the graph setting. We show that this bound is tight up to an absolute constant factor in the exponent. Moreover, we give an $O(n^k)$-time algorithm for finding a partition meeting the requirements. Finally, we establish tight bounds on Ramsey-Turán numbers for graphs with bounded VC-dimension.

1 Introduction

During the relatively short history of computational geometry, there were many breakthroughs that originated from results in extremal combinatorics [27]. Range searching turned out to be closely related to discrepancy theory [10], linear programming to McMullen's Upper Bound theorem and to properties of the facial structure of simplicial complexes [46], motion planning to the theory of Davenport-Schinzel sequences and to a wide variety of other forbidden configuration results [40], graph drawing and VLSI design to the crossing lemma, to the Szemerédi-Trotter theorem, and to flag algebras [48]. A particularly significant example that found many applications in discrete and computational geometry, was the discovery of Haussler and Welzl [30], according to which many geometrically defined set systems have bounded Vapnik-Chervonenkis dimension. Erdős's "Probabilistic Method" [5] or "Random Sampling" techniques, as they are often referred to in computational context, had been observed to be "unreasonably effective" in discrete geometry

*Stanford University, Stanford, CA. Supported by a Packard Fellowship, by NSF CAREER award DMS-1352121, and by an Alfred P. Sloan Fellowship. Email: jacobfox@stanford.edu.

†EPFL, Lausanne and Rényi Institute, Budapest. Research partially supported by Swiss National Science Foundation grants 200020-144531, 200021-137574, and 200021-162884. Email: pach@cims.nyu.edu.

‡Department of Mathematics, University of California at San Diego, La Jolla, CA, 92093 USA. Supported by NSF grant DMS-1500153, an NSF CAREER award, and an Alfred Sloan Fellowship. Email: asuk@ucsd.edu.

and geometric approximation algorithms [28]. Haussler and Welzl offered an explanation and a tool: set systems of bounded Vapnik-Chervonenkis dimension admit much smaller hitting sets and "epsilon-nets" than other set systems with similar parameters.

It was also observed a long time ago that geometrically defined graphs and set systems have unusually strong Ramsey-type properties. According to the quantitative version of Ramsey's theorem, due to Erdős and Szekeres [22], every graph on n vertices contains a clique or an independent set of size at least $\frac{1}{2}\log n$. In [15], Erdős proved that this bound is tight up to a constant factor. However, every intersection graph of n segments in the plane, say, has a much larger clique or an independent set, whose size is at least n^{ε} for some $\varepsilon > 0$ [34]. The proof extends to intersection graphs of many other geometric objects [3]. Interestingly, most classes of graphs and hypergraphs in which a similar phenomenon has been observed turned out to have (again!) bounded Vapnik-Chervonenkis dimension. (We will discuss this fact in a little more detail at the end of the Introduction.)

The problem can be viewed as a special case of a celebrated conjecture of Erdős and Hajnal [16], which is one of the most challenging open problems in Ramsey theory. Let P be a *hereditary* property of finite graphs, that is, if G has property P, then so do all of its induced subgraphs. Erdős and Hajnal conjectured that for every hereditary property P which is not satisfied by all graphs, there exists a constant $\varepsilon(P) > 0$ such that every graph of n vertices with property P has a clique or an independent set of size at least $n^{\varepsilon(P)}$. They proved the weaker lower bound $e^{\varepsilon(P)\sqrt{\log n}}$. According to the discovery of Haussler and Welzl mentioned above, the Vapnik-Chervonenkis dimension of most classes of "naturally" defined graphs arising in geometry is bounded from above by a constant d. The property that the Vapnik-Chervonenkis dimension of a graph is at most d, is hereditary.

The aim of this paper is to investigate whether the observation that the Erdős-Hajnal conjecture tends to hold for geometrically defined graphs can be ascribed to the fact that they have bounded VC-dimension. Our first theorem (Theorem 1 below) shows that the answer to this question is likely to be positive. To continue, we need to agree on the basic definitions and terminology.

Let $\mathcal{F}$ be a set system on a ground set V. The *Vapnik-Chervonenkis dimension* (*VC-dimension*, for short) of $\mathcal{F}$ is the *largest* integer d for which there exists a d-element set $S \subset V$ such that for every subset $B \subset S$, one can find a member $A \in \mathcal{F}$ with $A \cap S = B$. Given a graph $G = (V, E)$, for any vertex $v \in V$, let $N(v)$ denote the neighborhood of v in G, that is, the set of vertices in V that are connected to v. We note that v itself is not in $N(v)$. Then we say that G has *VC-dimension* d, if the set system induced by the neighborhoods in G, i.e. $\mathcal{F} = \{N(v) \subset V : v \in V\}$, has VC-dimension d. Let us remark that although the edges of G also form a 2-uniform set system $\mathcal{F}' = \{e \in E(G)\}$, the VC-dimension of G defined above is usually different from the VC-dimension of $\mathcal{F}'$.

The VC-dimension of a set system is one of the most useful combinatorial parameters that measures its complexity, and, apart from its geometric applications, it has proved to be relevant in many other branches of pure and applied mathematics, such as statistics, logic, learning theory, and real algebraic geometry. The notion was introduced by Vapnik and Chervonenkis [49] in 1971, as a tool in mathematical statistics. Kranakis et al. [33] observed that the VC-dimension of a graph can be determined in quasi-polynomial time and, for bounded degree graphs, in quadratic time. Schaefer [39], addressing a question of Linial, proved that determining the VC-dimension of a set system is Σ_3^p-complete. For each positive integer d, Anthony, Brightwell, and Cooper [6] determined the threshold for the Erdős-Rényi random graph $G(n, p)$ to have VC-dimension d (see also [32]). Given a bipartite graph F, its *closure* is defined as the set of all graphs that can be obtained from F by adding edges between two vertices in the same part. It is known (see [35]) that a class of graphs has bounded VC-dimension if and only if none of its members contains any induced subgraph that belongs to the closure of some fixed bipartite graph F.

Our first result states that the Erdős-Hajnal conjecture "almost holds" for graphs of bounded VC-dimension.

Theorem 1.1. *Let d be a fixed positive integer. If G is an n-vertex graph with VC-dimension at most d, then G contains a clique or independent set of size $e^{(\log n)^{1-o(1)}}$.*

Note that the dependence of the bound on d is hidden in the $o(1)$-notation.

There has been a long history of studying off-diagonal Ramsey numbers, where one is interested in finding the maximum size of an independent set guaranteed in a K_s-free graph on n vertices with s fixed. An old result of Ajtai, Komlós, and Szemerédi [1] states that all such graphs contain independent sets of size $cn^{\frac{1}{s-1}}(\log n)^{\frac{s-2}{s-1}}$. In the other direction, Spencer [44] used the Lovász Local Lemma to show that there are K_s-free graphs on n vertices and with no independent set of size $c'n^{\frac{2}{s+1}}\log n$. This bound was later improved by Bohman and Keevash [8] to $c'n^{\frac{2}{s+1}}(\log n)^{1-\frac{2}{(s+1)(s-2)}}$. In Section 4, we give a simple proof, extending Spencer's argument, showing that there are K_s-free graphs with bounded VC-dimension and with no large independent sets.

Theorem 1.2. *For fixed $s \geq 3$ and $d \geq 5$ such that $d \geq s+2$, there exists a K_s-free graph on n vertices with VC-dimension at most d and no independent set of size $cn^{\frac{2}{s+1}}\log n$, where $c = c(d)$.*

For large s ($s > d$), a result of Fox and Sudakov (Theorem 1.9 in [26]) implies that all n-vertex K_s-free graphs G with VC-dimension d contain an independent set of size $n^{\frac{1}{c\log s}}$ where $c = c(d)$.

Regularity lemma for hypergraphs with bounded VC-dimension. First, we generalize the definition of VC-dimension for graphs to *hypergraphs*. Given a k-uniform hypergraph $H = (V, E)$, for any $(k-1)$-tuple of distinct vertices $v_1, \ldots, v_{k-1} \in V$, let

$$N(v_1, \ldots, v_{k-1}) = \{u \in V : \{v_1, \ldots, v_{k-1}, u\} \in E(H)\}.$$

Then we say that H *has VC-dimension* d, if the set system

$$\mathcal{F} = \{N(v_1, \ldots, v_{k-1}) \subset V : v_1, \ldots, v_{k-1} \in V\}$$

has VC-dimension d. Of course, the hyperedges of H form a set system, but the VC-dimension of this set system is usually *different* from the VC-dimension of H defined above. The latter one is defined as the VC-dimension of the set system $\mathcal{F}$ induced by the neighborhoods of the vertices of H, rather than by the hyperedges.

The *dual* of the set system $(V, \mathcal{F})$ on the ground set V is the set system obtained by interchanging the roles of V and $\mathcal{F}$. That is, it is the set system $(\mathcal{F}, \mathcal{F}^*)$, where the ground set is $\mathcal{F}$ and

$$\mathcal{F}^* = \{\{A \in \mathcal{F} : v \in A\} : v \in V\}.$$

In other words, $\mathcal{F}^*$ is isomorphic to the set system whose ground set is $\binom{V}{k-1}$, and each set is a maximal collection of $(k-1)$-tuples $\{S_1, \ldots, S_p\}$ such that for all i, $v \cup S_i \in E(H)$ for some fixed v. Hence, we have $(\mathcal{F}^*)^* = \mathcal{F}$, and it is known that if $\mathcal{F}$ has VC-dimension d, then $\mathcal{F}^*$ has VC-dimension at most $2^d + 1$. We say that $H = (V, E)$ has *dual VC-dimension* d if $\mathcal{F}^*$ has VC-dimension d.

The main tool used to prove Theorem 1.1 is an ultra-strong regularity lemma for graphs with bounded VC-dimension obtained by Lovász and Szegedy [35] and Alon, Fischer, and Newman [2]. Here, we extend the ultra-strong regularity lemma to uniform hypergraphs.

Given k vertex subsets $V_1, \ldots, V_k$ of a k-uniform hypergraph H, we write $E(V_1, \ldots, V_k)$ to be the set of edges going across $V_1, \ldots, V_k$, that is, the set of edges with exactly one vertex in each V_i. The *density* across $V_1, \ldots, V_k$ is defined as $\frac{|E(V_1,\ldots,V_k)|}{|V_1|\cdots|V_k|}$. We say that the k-tuple $(V_1, \ldots, V_k)$ is *ε-homogeneous* if the density across it is less than ε or greater than $1-\varepsilon$. A partition is called *equitable* if any two parts differ in size by at most one.

In [35], Lovász and Szegedy established an *ultra-strong* regularity lemma for graphs ($k = 2$) with bounded VC-dimension, which states that for any $\varepsilon > 0$, there is a (least) $K = K(\varepsilon)$ such that the vertex set V of a graph with VC-dimension d has a partition into at most $K \leq (1/\varepsilon)^{O(d^2)}$ parts such that all but at most an ε-fraction of the pairs of parts are ε-homogeneous. A better bound was obtained by Alon, Fischer, and Newman [2] for bipartite graphs with bounded VC-dimension, who showed that the number of parts in the partition can be taken to be $(d/\varepsilon)^{O(d)}$. Since the VC-dimension of a graph G is equivalent to the dual VC-dimension of G, we generalize their result to hypergraphs with the following result.

Theorem 1.3. *Let $\varepsilon \in (0, 1/4)$ and let $H = (V, E)$ be an n-vertex k-uniform hypergraph with dual VC-dimension d. Then V has an equitable partition $V = V_1 \cup \cdots \cup V_K$ with $8/\varepsilon \leq K \leq c(1/\varepsilon)^{2d+1}$ parts such that all but an ε-fraction of the k-tuples of parts are ε-homogeneous. Here $c = c(d, k)$ is a constant depending only on d and k. Moreover, there is an $O(n^k)$ time algorithm for computing such a partition.*

Our next result shows that the partition size in the theorem above is tight up to an absolute constant factor in the exponent.

Theorem 1.4. *For $d \geq 16$ and $\varepsilon \in (0, 1/100)$, there is a graph G with VC-dimension d such that any equitable vertex partition on G with the property that all but an ε-fraction of the pairs of parts are ε-homogeneous, requires at least $(5\varepsilon)^{-d/4}$ parts.*

Ramsey-Turán numbers. Let F be a fixed graph. The Ramsey-Turán number $\mathbf{RT}(n, F, o(n))$ is the maximum number of edges an n-vertex graph G can have without containing F as a subgraph and having independence number $o(n)$. Ramsey-Turán numbers were introduced by Sós [43], motivated by the classical theorems of Ramsey and Turán and their connections to geometry, analysis, and number theory. One of the earliest results in Ramsey-Turán theory appeared in [21]. It states that for $p \geq 2$, we have

$$\mathbf{RT}(n, K_{2p-1}, o(n)) = \frac{1}{2}\left(1 - \frac{1}{p-1}\right)n^2 + o(n^2).$$

For the case when the excluded clique has an even number of vertices, Szemerédi [47] applied the graph regularity lemma to show that

$$\mathbf{RT}(n, K_4, o(n)) \leq \frac{1}{8}n^2 + o(n^2),$$

and several years later, Bollobás–Erdős [9] gave a surprising geometric construction which shows that this bound is tight. For larger cliques, a result of Erdős, Hajnal, Sós, and Szemerédi [18] states that

$$\mathbf{RT}(n, K_{2p}, o(n)) = \frac{1}{2}\left(1 - \frac{3}{3p-2}\right)n^2 + o(n^2)$$

holds for every $p \geq 2$. For more results in Ramsey-Turán theory, see the survey of Simonovits and Sós [42].

Here we give tight bounds on Ramsey-Turán numbers for graphs with bounded VC-dimension, showing that the densities for K_{2p} and for K_{2p-1} are the same in this setting, and are different from what we have in the classical setting in the even case.

Let $\mathbf{RT}_d(n, K_p, o(n))$ be the maximum number of edges that an n-vertex K_p-free graph of VC-dimension at most d can have if its independence number is $o(n)$.

Theorem 1.5. *For fixed integers $d \geq 4$ and $p \geq 3$, we have*

$$\mathbf{RT}_d(n, K_{2p-1}, o(n)) = \mathbf{RT}_d(n, K_{2p}, o(n)) = \frac{1}{2}\left(1 - \frac{1}{p-1}\right)n^2 + o(n^2).$$

Semi-algebraic graphs vs. graphs with bounded VC-dimension. A *semi-algebraic* graph G, is a graph whose vertices are points in $\mathbb{R}^d$ and edges are pairs of points that satisfy a semi-algebraic relation of constant complexity.[1] In a sequence of recent works [3, 12, 25], several authors have shown that classical Ramsey and Turán-type results in combinatorics can be significantly improved for semi-algebraic graphs.

It follows from the Milnor-Thom theorem (see [36]) that semi-algebraic graphs of bounded complexity have bounded VC-dimension. Therefore, all results in this paper on properties of graphs of bounded VC-dimension apply to semi-algebraic graphs of bounded description complexity. However, a graph being semi-algebraic of bounded complexity is a much more restrictive condition than having bounded VC-dimension. In particular, it is known (it follows, e.g., from [6]) that for each $\varepsilon > 0$ there is a positive integer $d = d(\varepsilon)$ such that the number of n-vertex graphs with VC-dimension d is $2^{\Omega(n^{2-\varepsilon})}$, while the Milnor-Thom theorem can be used to deduce that the number of n-vertex semi-algebraic graphs coming from a relation with bounded "description complexity" is only $2^{O(n \log n)}$. Furthermore, it is known [3] that semi-algebraic graphs have the *strong Erdős-Hajnal property*, that is, there exists a constant $\delta > 0$ such that every n-vertex semi-algebraic graph of bounded complexity contains a complete or an empty *bipartite* graph whose parts are of size at least δn. This is not true, in general, for graphs with bounded VC-dimension. In particular, the probabilistic construction in Section 4 shows the following.

Theorem 1.6. *For fixed $d \geq 5$ and for every sufficiently large n, there is an n-vertex graph $G = (V, E)$ with VC-dimension at most d with the property that there are no two disjoint subsets $A, B \subset V(G)$ such that $|A|, |B| \geq 4n^{4/d} \log n$ and (A, B) is homogeneous, that is, either $A \times B \subset E(G)$ or $(A \times B) \cap E(G) = \emptyset$.*

It follows from a result of Alon *et al.* [3] that a stronger regularity lemma holds for semi-algebraic graphs of bounded description complexity, where all but an ε-fraction of the pairs of parts in the equitable partition are complete or empty, instead of just ε-homogeneous as in the bounded VC-dimension case (see [37]). This result was further extended to k-uniform hypergraphs by Fox et al. [23], and the authors [25] recently showed that it holds with a polynomial number of parts.

Organization. In the next section, we prove Theorem 1.3. In Section 3, we prove Theorem 1.1, which nearly settles the Erdős-Hajnal conjecture for graphs with bounded VC-dimension. In Section 4, we prove Theorems 1.2 and 1.6. In Section 5, we prove Theorem 1.5. We conclude by

[1] A binary semi-algebraic relation E on a point set $P \subset \mathbb{R}^d$ is the set of pairs of points (p, q) from P whose $2d$ coordinates satisfy a boolean combination of a fixed number of polynomial inequalities.

discussing a number of other results for graphs and hypergraphs with bounded VC-dimension. We systemically omit floors and ceilings whenever they are not crucial for sake of clarity in our presentation. All logarithms are natural logarithms.

2 Regularity partition for hypergraphs with bounded VC-dimension

In this section, we prove Theorem 1.3. We start by recalling several classic results on set systems with bounded VC-dimension. Let $\mathcal{F}$ be a set system on a ground set V. The *primal shatter function* of $\mathcal{F}$ is defined as

$$\pi_{\mathcal{F}}(z) = \max_{V' \subset V, |V'|=z} |\{A \cap V' : A \in \mathcal{F}\}|.$$

In other words, $\pi_{\mathcal{F}}(z)$ is a function whose value at z is the maximum possible number of distinct intersections of the sets of $\mathcal{F}$ with a z-element subset of V. The *dual shatter function* of $(V, \mathcal{F})$, denoted by $\pi^*_{\mathcal{F}}$, whose value at z is defined as the maximum number of equivalence classes on V defined by a z-element subfamily $\mathcal{Y} \subset \mathcal{F}$, where two points $x, y \in V$ are *equivalent* with respect to $\mathcal{Y}$ if x belongs to the same sets of $\mathcal{Y}$ as y does. In other words, the dual shatter function of $\mathcal{F}$ is the primal shatter function of the dual set system $\mathcal{F}^*$.

The VC-dimension of $\mathcal{F}$ is closely related to its shatter functions. A famous result of Sauer [38], Shelah [41], Perles, and Vapnik-Chervonenkis [49] states the following.

Lemma 2.1. *If $\mathcal{F}$ is a set system with VC-dimension d, then*

$$\pi_{\mathcal{F}}(z) \leq \sum_{i=0}^{d} \binom{z}{i}.$$

On the other hand, suppose that the primal shatter function of $\mathcal{F}$ satisfies $\pi_{\mathcal{F}}(z) \leq cz^d$ for all z. Then, if the VC-dimension of $\mathcal{F}$ is d_0, we have $2^{d_0} \leq c(d_0)^d$, which implies $d_0 \leq 4d\log(cd)$. It is known that if $\mathcal{F}$ has VC-dimension d, then $\mathcal{F}^*$ has VC-dimension at most $2^d + 1$.

Given two sets $A_1, A_2 \in \mathcal{F}$, the *symmetric difference* of A_1 and A_2, denoted by $A_1 \triangle A_2$, is the set $(A_1 \cup A_2) \setminus (A_1 \cap A_2)$. We say that the set system $\mathcal{F}$ is *δ-separated* if for any two sets $A_1, A_2 \in \mathcal{F}$ we have $|A_1 \triangle A_2| \geq \delta$. The following *packing lemma* was proved by Haussler in [29].

Lemma 2.2. *Let $\mathcal{F}$ be a set system on a ground set V such that $|V| = n$ and $\pi_{\mathcal{F}}(z) \leq cz^d$ for all z. If $\mathcal{F}$ is δ-separated, then $|\mathcal{F}| \leq c_1(n/\delta)^d$ where $c_1 = c_1(c, d)$.*

We will use Lemma 2.2 and the following lemma to prove Theorem 1.3.

Lemma 2.3. *Let $0 < \varepsilon < 1/2$ and $H = (W_1 \cup \cdots \cup W_k, E)$ be a k-partite k-uniform hypergraph such that $|W_i| = m$ for all i. If $(W_1, \ldots, W_k)$ is not ε-homogeneous, then there are at least $\varepsilon(1-\varepsilon)m^{k+1}$ pairs of k-tuples (e, e'), where $|e \cap e'| = k - 1$, $e \in E(H)$, $e' \notin E(H)$, and $|e \cap W_i| = |e' \cap W_i| = 1$ for all i.*

Proof. Let ε_j be the fraction of pairs of k-tuples (e, e'), each containing one vertex in each W_i and agree on all vertices except in W_j, and e is an edge and e' is not an edge. It suffices to show that $\varepsilon_1 + \varepsilon_2 + \cdots + \varepsilon_k \geq \varepsilon(1-\varepsilon)$.

Pick vertices $a_i, b_i \in W_i$ uniformly at random with repetition for $i = 1, 2, \ldots, k$. For $0 \leq i \leq k$, let $e_i = \{a_j : j \leq i\} \cup \{b_j : j > i\}$. In particular, $e_k = (a_1, \ldots, a_k)$ and $e_0 = (b_1, \ldots, b_k)$. Then let

X be the event that e_0 and e_k have different adjacency, that is, e_0 is an edge and e_k is not an edge, or e_0 is not an edge and e_k is an edge. Then we have

$$Pr[X] \geq 2\varepsilon(1-\varepsilon),$$

since $(W_1, \dots, W_k)$ is not ε-homogeneous. Let X_i be the event that e_i and e_{i+1} have different adjacency, and let Y be the event that at least one event X_i occurs. Then by the union bound, we have

$$Pr[Y] \leq Pr[X_0] + Pr[X_1] + \cdots + Pr[X_{k-1}] = 2\varepsilon_1 + 2\varepsilon_2 + \cdots + 2\varepsilon_k.$$

On the other hand, if X occurs, then Y occurs. Therefore $2\varepsilon_1 + 2\varepsilon_2 + \cdots + 2\varepsilon_k \geq Pr[Y] \geq Pr[X] \geq 2(1-\varepsilon)\varepsilon$, which completes the proof. □

Proof of Theorem 1.3. Let $0 < \varepsilon < 1/4$, $k \geq 2$, and $H = (V, E)$ be an n-vertex k-uniform hypergraph with dual VC-dimension d. For every vertex $v \in V$, let $N(v)$ denote the set of $(k-1)$-tuples $S \in \binom{V}{k-1}$ such that $v \cup S \in E(H)$. Let $\mathcal{F}$ be the set-system whose ground set is $\binom{V}{k-1}$, and $A \in \mathcal{F}$ if and only if $A = N(v)$ for some vertex $v \in V$. Hence $\mathcal{F} = \{N(v) : v \in V\}$ has VC-dimension d. Set $\delta = \frac{\varepsilon^2}{4k^2}\binom{n}{k-1}$. By examining each vertex and its neighborhood one by one, we greedily construct a maximal set $S \subset V(H)$ such that $\mathcal{F}' = \{N(s) : s \in S\}$ is δ-separated. By Lemma 2.2, we have $|S| \leq c_1(4k^2/\varepsilon^2)^d$. Let $S = \{s_1, s_2, \dots, s_{|S|}\}$.

We define a partition $\mathcal{Q} : V = U_1 \cup \cdots \cup U_{|S|}$ of the vertex set such that $v \in U_i$ if i is the smallest index such that $|N(v) \triangle N(s_i)| < \delta$. Such an i always exists, since S is maximal. By the triangle inequality, for $u, v \in U_i$, we have $|N(u) \triangle N(v)| < 2\delta$. Set $K = 8k|S|/\varepsilon$. Partition each part U_i into parts of size $|V|/K = n/K$ and possibly one additional part of size less than n/K. Collect these additional parts and divide them into parts of size $|V|/K$ to obtain an equitable partition $\mathcal{P} : V = V_1 \cup \cdots \cup V_K$ into K parts. The number of vertices of V belonging to parts V_i that are not fully contained in one part of $\mathcal{Q}$ is at most $|S||V|/K$. Hence, the fraction of (unordered) k-tuples $(V_{i_1}, \dots, V_{i_k})$ such that at least one of the parts is not fully contained in some part of $\mathcal{Q}$ is at most $k|S|/K = \varepsilon/8$. Let X denote the set of unordered k-tuples of parts $(V_{i_1}, \dots, V_{i_k})$ such that each part is fully contained in a part of $\mathcal{Q}$ (though, in not necessarily the same part) and $(V_{i_1}, \dots, V_{i_k})$ is not ε-homogeneous.

Let T be the set of pairs of k-tuples (e, e'), such that $|e \cap e'| = k-1$, $e \in E(H)$, $e' \notin E(H)$, $|e \cap V_{i_j}| = |e' \cap V_{i_j}| = 1$ for $j = 1, 2, \dots, k$, and $(V_{i_1}, \dots, V_{i_k}) \in X$. Notice that for $(e, e') \in T$, such that $e \cap V_{i_j} = b$, $e' \cap V_{i_j} = b'$, $b \neq b'$, and V_{i_j} lies completely inside a part in $\mathcal{Q}$, we have $|N(b) \triangle N(b')| \leq 2\delta$. Therefore

$$|T| \leq K\left(\frac{n}{K}\right)^2 2\delta \leq \frac{\varepsilon^2}{2Kk^2}n^2\binom{n}{k-1}.$$

On the other hand, by Lemma 2.3, every k-tuple of parts $(V_{i_1}, \dots, V_{i_k})$ that is not ε-homogeneous gives rise to at least $\varepsilon(1-\varepsilon)(n/K)^{k+1}$ pairs (e, e') in T. Hence $|T| \geq |X|\varepsilon(1-\varepsilon)(n/K)^{k+1}$. Since $\varepsilon < 1/4$ and $k \geq 2$, the inequalities above imply that

$$|X| \leq (2\varepsilon/3)\binom{K}{k}.$$

Thus, the fraction of k-tuples of parts in $\mathcal{P}$ that are not ε-homogeneous is at most $\varepsilon/8 + 2\varepsilon/3 < \varepsilon$, and $K \leq c(1/\varepsilon)^{2d+1}$ where $c = c(k, d)$.

Finally, it remains to show that the partition $\mathcal{P}$ can be computed in $O(n^k)$ time. Given two vertices $s, v, \in V$, we have $|N(s)\triangle N(v)| = |N(s)| + |N(v)| - 2|N(s) \cap N(v)|$. Therefore we can determine if $|N(s)\triangle N(u)| < \delta$ in $O(n^{k-1})$ time. Hence the maximal set $S \subset V$ described above (and therefore the partition $\mathcal{Q}$) can be computed in $O(n^k)$ time since $|S| \leq n$. The final equitable partition $\mathcal{P}$ requires an additional $O(n)$ time, which gives a total running time of $O(n^k)$. $\square$

We now establish Theorem 1.4 which shows that the partition size in Theorem 1.3 is tight up to an absolute constant factor in the exponent.

Proof of Theorem 1.4. Given two vertex subsets X, Y of a graph G, we write $e_G(X, Y)$ for the number of edges between X and Y in G, and write $d_G(X, Y)$ for the density of edges between X and Y, that is, $d_G(X, Y) = \frac{e_G(X,Y)}{|X||Y|}$. The pair (X, Y) is said to be (ε, δ)-*regular* if for all $X' \subset X$ and $Y' \subset Y$ with $|X'| \geq \delta|X|$ and $|Y'| \geq \delta|Y|$, we have $|d_G(X, Y) - d_G(X', Y')| \leq \varepsilon$. In the case that $\varepsilon = \delta$, we just say ε-*regular*. We will make use of the following construction due to Conlon and Fox.

Lemma 2.4 ([11])**.** *For $d \geq 16$ and $\varepsilon \in (0, 1/100)$, there is a graph H on $n = \lceil(5\varepsilon)^{-d/2}\rceil$ vertices such that for every equitable vertex partition of H with at most $\sqrt{n}$ parts, there are at least an ε-fraction of the pairs of parts which are not $(4/5)$-regular.*

Let $H = (V, E)$ be the graph obtained from Lemma 2.4 on $n = \lceil(5\varepsilon)^{-d/2}\rceil$ vertices, where $\varepsilon \in (0, 1/100)$ and $d \geq 16$, and consider a random subgraph $G \subset H$ by picking each edge in E independently with probability $p = n^{-2/d} = 5\varepsilon$. Then we have the following.

Lemma 2.5. *In the random subgraph G, with probability at least $1 - n^{-2}$, every pair of disjoint subsets $X, Y \subset V$, with $|X| \leq |Y|$, satisfy*

$$|e_G(X, Y) - p \cdot e_H(X, Y)| < \sqrt{g}, \tag{1}$$

where $g = 2|X||Y|^2 \ln(ne/|Y|)$.

Proof. For fixed sets $X, Y \subset V(G)$, where $|X| = u_1$ and $|Y| = u_2$, let $E_H(X, Y) = \{e_1, \dots, e_m\}$. We define $S_i = 1$ if edge e_i is picked and $S_i = 0$ otherwise, and set $S = S_1 + \cdots + S_m$. A Chernoff-type estimate (see Theorem A.1.4 in [5]) implies that for $a > 0$, $Pr[|S - pm| > a] < 2e^{-2a^2/m}$. Since $m \leq u_1u_2$, the probability that (1) does not hold is less than $2e^{-2g/(u_1u_2)}$. By the union bound, the probability that there are disjoint sets $X, Y \subset V(G)$ for which (1) does not hold is at most

$$\begin{aligned}
\sum_{u_2=1}^{n}\sum_{u_1=1}^{u_2}\binom{n}{u_2}\binom{n-u_2}{u_1}2e^{-2g/(u_1u_2)} &\leq \sum_{u_2=1}^{n}\sum_{u_1=1}^{u_2}\left(\frac{ne}{u_2}\right)^{u_2}\left(\frac{ne}{u_1}\right)^{u_1}2e^{-2g/(u_1u_2)} \\
&\leq \sum_{u_2=1}^{n}\sum_{u_1=1}^{u_2}2\left(\frac{ne}{u_2}\right)^{-2u_2} \leq n^{-2}.
\end{aligned}$$

$\square$

By the analysis in Section 4, the probability that G has VC-dimension at least $d+1$ is at most

$$\binom{n}{d+1} n^{2^{d+1}} p^{(d+1)2^d} \leq n^{d+1} n^{-2^{d+1}/d} < \frac{1}{10},$$

since $d \geq 16$. Therefore, the union bound implies that there is a subgraph $G \subset H$ such that G has VC-dimension at most d, and every pair of disjoint subsets $X, Y \subset V$, with $|X| \leq |Y|$, satisfy

$$|e_G(X,Y) - p \cdot e_H(X,Y)| < \sqrt{2|X||Y|^2 \ln(ne/|Y|)}. \tag{2}$$

We will now show that for every equitable vertex partition of G into fewer than $\sqrt{n} = (5\varepsilon)^{-d/4}$ parts, there are at least an ε-fraction of the pairs of parts which are not ε-homogenous.

Let $\mathcal{P}$ be a equitable partition on V into t parts, where $t < \sqrt{n} = (5\varepsilon)^{-d/4}$. By Lemma 2.4, there are at least $\varepsilon\binom{t}{2}$ pairs of parts in $\mathcal{P}$ which are not (4/5)-regular in H. Let (X, Y) be such a pair. Then there are subsets $X' \subset X$ and $Y' \subset Y$ such that $|X'| \geq 4|X|/5$, $|Y'| \geq 4|Y|/5$, and

$$|d_H(X,Y) - d_H(X',Y')| \geq 4/5.$$

Moreover, by (2), we have

$$|e_G(X,Y) - p \cdot e_H(X,Y)| \leq \sqrt{2}\left(\frac{n}{t}\right)^{3/2} \ln(te) \leq \frac{\sqrt{2}\ln(te)}{n^{1/4}}(n/t)^2.$$

Since $d \geq 16$ and $\varepsilon \in (0, 1/100)$, this implies

$$|e_G(X,Y) - p \cdot e_H(X,Y)| \leq (5\varepsilon)^2\sqrt{2}\ln(te)(n/t)^2 \leq \frac{\varepsilon}{4}(n/t)^2.$$

Hence $|d_G(X,Y) - p \cdot d_H(X,Y)| \leq \varepsilon/4$. Therefore we have

$$|d_G(X',Y') - d_G(X,Y)| \geq p \cdot |d_H(X',Y') - d_H(X,Y)| - 2\frac{\varepsilon}{4} \geq 4\varepsilon - \frac{\varepsilon}{2} > 3\varepsilon.$$

Finally, it is easy to see that (X, Y) is not ε-homogeneous in G. Indeed if (X, Y) were ε-homogeneous, then we have either $d_G(X, Y) < \varepsilon$ or $d_G(X, Y) > 1 - \varepsilon$. In the former case we have $d_G(X', Y') > 3\varepsilon$, which implies

$$e_G(X,Y) \geq e_G(X',Y') > 3\varepsilon\frac{4|X|}{5}\frac{4|Y|}{5} > \varepsilon|X||Y|,$$

contradiction. In the latter case, we have $d(X', Y') < 1 - 3\varepsilon$, and a similar analysis shows that $e_G(X, Y) < (1 - \varepsilon)|X||Y|$, contradiction.

Thus, any equitable vertex partition on G such that all but an ε-fraction of the pairs of parts are ε-homogeneous, requires at least $(5\varepsilon)^{-d/4}$ parts. □

3 Proof of Theorem 1.1

The family $\mathcal{G}$ of all *complement reducible graphs*, or *cographs*, is defined as follows: The graph with one vertex is in $\mathcal{G}$, and if two graphs $G, H \in \mathcal{G}$, then so does their disjoint union, and the graph obtained by taking their disjoint union and adding all edges between G and H. Clearly, every induced subgraph of a cograph is a cograph, and it is well known that every cograph on n vertices contains a clique or independent set of size $\sqrt{n}$.

Let $f_d(n)$ be the largest integer f such that every graph G with n vertices and VC-dimension at most d has an induced subgraph on f vertices which is a cograph. Cographs are perfect graphs, so that Theorem 1.1 is an immediate consequence of the following result.

Theorem 3.1. *For any $\delta \in (0, 1/2)$ and for every integer $d \geq 1$, there is a $c = c(d, \delta)$ such that $f_d(n) \geq e^{c(\log n)^{1-\delta}}$ for every n.*

Proof. For simplicity, let $f(n) = f_d(n)$. The proof is by induction on n. The base case $n = 1$ is trivial. For the inductive step, assume that the statement holds for all $n' < n$. Let $\delta > 0$ and let $G = (V, E)$ be an n-vertex graph with VC-dimension d. We will determine $c \in (0, 1)$ later.

Set $\varepsilon = (1/32)e^{-3c(\log n)^{1-\delta}}$. We apply Theorem 1.3 to obtain an equitable partition $\mathcal{P} : V = V_1 \cup \cdots \cup V_K$ into at most $K \leq \varepsilon^{-c_4}$ parts, where $c_4 = O(d)$, such that all but an ε-fraction of the pairs of parts are ε-homogeneous. We call an unordered pair of distinct vertices (u, v) *bad* if at least one of the following holds:

1. (u, v) lie in the same part, or
2. $u \in V_i$ and $v \in V_j$, $i \neq j$, where (V_i, V_j) is not ε-homogeneous, or
3. $u \in V_i$ and $v \in V_j$, $i \neq j$, $uv \in E(G)$ and $|E(V_i, V_j)| < \varepsilon|V_i||V_j|$, or
4. $u \in V_i$ and $v \in V_j$, $i \neq j$, $uv \notin E(G)$ and $|E(V_i, V_j)| > (1-\varepsilon)|V_i||V_j|$.

By Theorem 1.3, the number of bad pairs of vertices in G is at most

$$K\binom{n/K}{2} + \left(\frac{n}{K}\right)^2 \varepsilon \binom{K}{2} + \varepsilon \left(\frac{n}{K}\right)^2 (1-\varepsilon)\binom{K}{2} \leq 2\varepsilon\binom{n}{2}.$$

By Turán's Theorem, there is a subset $R \subset S$ of at least $\frac{1}{4\varepsilon}$ vertices such that R does not contain any bad pairs. This implies that all vertices of R are in distinct parts of $\mathcal{P}$. Furthermore, if uv are adjacent in R, then the corresponding parts V_i, V_j satisfy $|E(V_i, V_j)| \geq (1-\varepsilon)|V_i||V_j|$, and if uv are not adjacent, then we have $|E(V_i, V_j)| < \varepsilon|V_i||V_j|$. Since the induced graph $G[R]$ has VC-dimension at most d, $G[R]$ contains a cograph U_0 of size $t = f(1/(4\varepsilon))$, which, by the induction hypothesis, is a set of size at least $e^{c(\log(1/4\varepsilon))^{1-\delta}}$. Without loss of generality, we denote the corresponding parts of U_0 as $V_1, \ldots, V_t$. Each part contains n/K vertices.

For each vertex $u \in V_1$, let $d_b(u)$ denote the number of bad pairs uv, where $v \in V_i$ for $i = 2, \ldots, t$. Then there is a subset $V_1' \subset V_1$ of size $\frac{n}{2K}$, such that each vertex $u \in V_1'$ satisfies $d_b(u) < 8t\varepsilon(n/K)$. Indeed, otherwise at least $n/(2K)$ vertices in V_1 satisfies $d_b(u) \geq 8t\varepsilon(n/K)$, which implies

$$\frac{n}{2K}\frac{8t\varepsilon n}{K} \leq \sum_{u \in V_1'} d_b(u) \leq \sum_{u \in V_1} d_b(u) \leq \varepsilon(t-1)\left(\frac{n}{K}\right)^2,$$

and hence a contradiction. By the induction hypothesis, we can find a subset $U_1 \subset V_1'$ such that the induced subgraph $G[U_1]$ is a cograph of size $f(n/(2K))$. If the inequality

$$f\left(\frac{n}{2K}\right) 8t\varepsilon\frac{n}{K} > \frac{n}{4tK}$$

is satisfied, then we have

$$f^3(n) \geq f\left(\frac{n}{2K}\right) t^2 > \frac{1}{32\varepsilon}.$$

By setting ε such that $\frac{1}{\varepsilon} = 32e^{3c(\log n)^{1-\delta}}$, we have $f(n) \geq e^{c(\log n)^{1-\delta}}$ and we are done.

Therefore, we can assume that

$$f\left(\frac{n}{2K}\right) 8t\varepsilon \frac{n}{K} \leq \frac{n}{4tK}.$$

Hence, by deleting any vertex $v \in V_2 \cup \cdots \cup V_t$ that is in a bad pair with a vertex in U_1, we have deleted at most $\frac{n}{4tK}$ vertices in each V_i for $i = 2, \ldots, t$.

We repeat this entire process on the remaining vertices in $V_2, \ldots, V_t$. At step i, we will find a subset $U_i \subset V_i$ that induces a cograph of size

$$f\left(\frac{n}{2K} - i\frac{n}{4Kt}\right) \geq f\left(\frac{n}{4K}\right),$$

and again, if the inequality

$$f\left(\frac{n}{4K}\right) 8t\varepsilon \frac{n}{K} > \frac{n}{4tK}$$

is satisfied, then we are done by the same argument as above. Therefore we can assume that our cograph $G[U_i]$ has the property that there are at most $n/(4tK)$ bad pairs between U_i and V_j for $j > i$. At the end of this process, we obtain subsets $U_1, \ldots, U_t$ such that the union $U_1 \cup \cdots \cup U_t$ induces a cograph of size at least $tf\left(\frac{n}{4K}\right)$. Therefore we have

$$\begin{aligned} f(n) &\geq f\left(\tfrac{1}{4\varepsilon}\right) f\left(\tfrac{n}{4K}\right) \\ &\geq f\left(e^{3c(\log n)^{1-\delta}}\right) f\left(e^{\log n - c\cdot c_5(\log n)^{1-\delta}}\right) \\ &\geq e^{c\left(3c(\log n)^{1-\delta}\right)^{1-\delta}} e^{c\left(\log n - c\cdot c_5(\log n)^{1-\delta}\right)^{1-\delta}}, \end{aligned} \tag{3}$$

where $c_5 = c_5(d)$. Notice we have the following estimate:

$$\begin{aligned} \left(\log n - c\cdot c_5(\log n)^{1-\delta}\right)^{1-\delta} &= (\log n)^{1-\delta}\left(1 - \tfrac{c\cdot c_5}{\log^{\delta} n}\right)^{1-\delta} \\ &\geq (\log n)^{1-\delta}\left(1 - \tfrac{c\cdot c_5}{(\log n)^{\delta}}\right) \\ &\geq (\log n)^{1-\delta} - c\cdot c_5(\log n)^{1-2\delta}. \end{aligned} \tag{4}$$

Plugging (4) into (3) gives

$$\begin{aligned} f(n) &\geq e^{c\left(3c(\log n)^{1-\delta}\right)^{1-\delta}} \cdot e^{c(\log n)^{1-\delta} - c^2\cdot c_5(\log n)^{1-2\delta}} \\ &= e^{c(\log n)^{1-\delta}} \cdot e^{\left(3^{1-\delta}c^2(\log n)^{1-2\delta+\delta^2} - c^2 c_5(\log n)^{1-2\delta}\right)}. \end{aligned} \tag{5}$$

The last inequality follows from the fact that $c < 1$. Let $n_0 = n_0(d,\delta)$ be the minimum integer such that for all $n \geq n_0$ we have

$$3^{1-\delta}(\log n)^{1-2\delta+\delta^2} - c_5(\log n)^{1-2\delta} \geq 0.$$

We now set $c = c(d,t)$ to be sufficiently small such that the statement is trivial for all $n < n_0$. Hence we have $f(n) \geq e^{c(\log n)^{1-\delta}}$ for all n. $\square$

4 Random constructions

Here we prove Theorems 1.2 and 1.6. The proof of Theorem 1.2 uses the Lovász Local Lemma [19] in a similar manner as Spencer [44] to give a lower bound on Ramsey numbers.

Lemma 4.1 (Lovász Local Lemma). *Let $\mathcal{A}$ be a finite set of events in a probability space. For $A \in \mathcal{A}$ let $\Gamma(A)$ be a subset of $\mathcal{A}$ such that A is independent of all events in $\mathcal{A} \setminus (\{A\} \cup \Gamma(A))$. If there is a function $x : \mathcal{A} \to (0,1)$ such that for all $A \in \mathcal{A}$,*

$$Pr[A] \leq x(A) \prod_{B \in \Gamma(A)} (1 - x(B)),$$

then $Pr\left[\bigcap_{A \in \mathcal{A}} \overline{A}\right] \geq \prod_{A \in \mathcal{A}} (1 - x(A))$. In particular, with positive probability no event in $\mathcal{A}$ holds.

Proof of Theorem 1.2. Let s and d be positive integers such that $d > s + 2$. Let $G(n,p)$ denote the random graph on n vertices in which each edge appears with probability p independently of all the other edges, where $p = n^{-2/(s+1)}$ and n is a sufficiently large number. For each set S of s vertices, let A_S be the event that S induces a complete graph. For each set T of t vertices, let B_T be the event that T induces an empty graph. Clearly, we have $Pr[A_S] = p^{\binom{s}{2}}$ and $Pr[B_T] = (1-p)^{\binom{t}{2}}$.

For each set D of d vertices, let C_D be the event that D is shattered. Then

$$\begin{aligned}
Pr[C_D] &\leq \prod_{W \subset D} Pr[\exists v \in V(G) : N(v) \cap D = W] \\
&= \prod_{W \subset D} \left(1 - \left(1 - p^{|W|}(1-p)^{d-|W|}\right)^n\right) \\
&= \prod_{j=0}^{d} \left(1 - \left(1 - p^j (1-p)^{d-j}\right)^n\right)^{\binom{d}{j}} \\
&\leq \prod_{j=1}^{d} \left(n \cdot p^j (1-p)^{d-j}\right)^{\binom{d}{j}} \\
&\leq \prod_{j=1}^{d} n^{\binom{d}{j}} \cdot p^{j\binom{d}{j}} \\
&\leq n^{2^d} \cdot p^{d2^{d-1}}.
\end{aligned}$$

Next we estimate the number of events dependent on each A_S, B_T and C_D. Let $S \subset V$ such that $|S| = s$. Then the event A_S is dependent on at most $\binom{s}{2}\binom{n}{s-2} \leq s^2 n^{s-2}$ events $A_{S'}$, where $|S'| = s$. Likewise, A_S is dependent on at most $\binom{n}{t}$ events B_T where $|T| = t$. Finally A_S is dependent on at most $\binom{s}{2}\binom{n}{d-2} \leq s^2 n^{d-2}$ events C_D where $|D| = d$.

Let $T \subset V$ be a set of vertices such that $|T| = t$. Then the event B_T is dependent on at most $\binom{t}{2}\binom{n}{s-2} \leq t^2 n^{s-2}$ events A_S where $|S| = s$. Likewise, B_T is dependent on at most $\binom{n}{t}$ events $B_{T'}$ where $|T'| = t$. Finally B_T is dependent on at most $\binom{t}{2}\binom{n}{d-2} \leq t^2 n^{d-2}$ events C_D where $|D| = d$.

Let $D \subset V$ be a set of vertices such that $|D| = d$. Then the event C_D is dependent on at most $\binom{d}{2}\binom{n}{s-2} \leq d^2 n^{s-2}$ events A_S where $|S| = s$. Likewise, C_D is dependent on at most $\binom{n}{t}$ events B_T where $|T| = t$. Finally C_D is dependent on at most $\binom{d}{2}\binom{n}{d-2} \leq d^2 n^{d-2}$ events $C_{D'}$ where $|D'| = d$.

By Lemma 4.1, it suffices to find three real numbers $x, y, z \in (0, 1)$ such that

$$p^{\binom{s}{2}} \leq x(1-x)^{s^2 n^{s-2}} (1-y)^{\binom{n}{t}} (1-z)^{s^2 n^{d-2}}, \tag{6}$$

$$(1-p)^{\binom{t}{2}} \leq y(1-x)^{t^2 n^{s-2}} (1-y)^{\binom{n}{t}} (1-z)^{t^2 n^{d-2}}, \tag{7}$$

and

$$n^{2^d} \cdot p^{d2^{d-1}} \leq z(1-x)^{d^2 n^{s-2}} (1-y)^{\binom{n}{t}} (1-z)^{d^2 n^{d-2}}. \tag{8}$$

Recall $p = n^{\frac{-2}{s+1}}$, $s \geq 3$, and $d > s + 2$. We now set $t = c_1 n^{\frac{2}{s+1}}(\log n)$, $x = c_2 n^{\frac{-2\binom{s}{2}}{s+1}}$, $y = e^{-c_3 n^{\frac{2}{s+1}}(\log n)^2}$, and $z = c_4 n^{2^d - \frac{2}{s+1} d 2^{d-1}}$, where c_1, c_2, c_3, c_4 only depend on s and d. By letting $c_1 > 10c_3$, setting c_1, c_2, c_3, c_4 sufficiently large, an easy (but tedious) calculation shows that (6), (7), (8) are satisfied when n is sufficiently large. By Lemma 4.1, there is an n-vertex K_s-free graph G with VC-dimension at most d and independence number at most $c_1 n^{\frac{2}{s+1}} \log n$. $\square$

Proof of Theorem 1.6. Let $d \geq 5$ and n be a sufficiently large integer that will be determined later. Consider the random n-vertex graph $G = G(n, p)$, where each edge is chosen independently with probability $p = n^{-4/d}$. By choosing n sufficiently large, the union bound and the analysis above implies that the probability that G has VC-dimension at least d is at most $1/3$.

Let $A, B \subset V(G)$ be vertex subsets, each of size k. The probability that (A, B) is homogenous is at most

$$p^{k^2} + (1-p)^{k^2} \leq n^{-4k^2/d} + e^{-n^{-4/d}k^2}.$$

The probability that G contains a homogeneous pair (A, B), where $|A|, |B| = k$, is at most

$$\binom{n}{k}\binom{n-k}{k}\left(n^{-4k^2/d} + e^{-n^{-4/d}k^2}\right) < 1/3,$$

for $k = 4n^{4/d}\log n$ and n sufficiently large. Thus, again by the union bound, there is a graph with VC-dimension less than d, with no two disjoint subsets $A, B \subset V(G)$ such that (A, B) is homogeneous and $|A|, |B| = 4n^{4/d}\log n$. $\square$

5 Ramsey-Turán numbers for graphs with bounded VC-dimension

In this section we prove Theorem 1.5. First let us recall a classical theorem in graph theory.

Theorem 5.1 (Turán)**.** *Let $G = (V, E)$ be a K_p-free graph with n vertices. Then the number of edges in G is at most $\frac{1}{2}\left(1 - \frac{1}{p-1} + o(1)\right)n^2$.*

Together with a sampling argument of Varnavides [50], we have the following lemma (see also Lemma 2.1 in [31]).

Lemma 5.2. *For $\varepsilon > 0$, every n-vertex graph $G = (V, E)$ with $|E| \geq \frac{1}{2}\left(1 - \frac{1}{p-1} + \varepsilon\right) n^2$ has at least δn^p copies of K_p, where $\delta = \delta(p, \varepsilon)$.*

In order to establish the upper bound in Theorem 1.5, it suffices to show

$$\mathbf{RT}_d(n, K_{2p}, o(n)) \leq \frac{1}{2}\left(1 - \frac{1}{p-1}\right) n^2 + o(n^2),$$

since we have $\mathbf{RT}_d(n, K_{2p-1}, o(n)) \leq \mathbf{RT}(n, K_{2p-1}, o(n))$. The following theorem implies the inequality above.

Theorem 5.3. *Let $\varepsilon > 0$ and let $G = (V, E)$ be an n-vertex graph with VC-dimension d. If G is K_{2p}-free and $|E| > \frac{1}{2}\left(1 - \frac{1}{p-1} + \varepsilon\right) n^2$, then G contains an independent set of size γn, where $\gamma = \gamma(d, p, \varepsilon)$.*

Proof. By Lemma 5.2, G contains at least δn^p copies of K_p, where $\delta = \delta(\varepsilon, p)$. Without loss of generality, we can assume that δ is sufficiently small and will be determined later. We apply the regularity lemma (Lemma 1.3) with approximation parameter $\delta/4$ to obtain a (near) equipartition $\mathcal{P} : V = V_1 \cup \cdots \cup V_K$ such that $4/\delta \leq K \leq c\,(4/\delta)^{2d+1}$, where $c = c(d)$, and all but a $\frac{\delta}{4}$-fraction of the pairs of parts in $\mathcal{P}$ are $(\delta/4)$-homogeneous.

By deleting all edges inside each part, we have deleted at most

$$K\binom{n/K}{2} \leq \frac{n^2}{2K} \leq \frac{n^2}{8}\delta$$

edges. By deleting all edges between pairs of parts that are not $(\delta/4)$-homogeneous, we have deleted an additional

$$\left(\frac{n}{K}\right)^2 \frac{\delta}{4}\binom{K}{2} \leq \frac{n^2}{8}\delta$$

edges. Finally, by deleting all edges between pairs (V_i, V_j) with density less than $\delta/4$, we have deleted at most

$$\frac{\delta}{4}\left(\frac{n}{K}\right)^2\binom{K}{2} \leq \frac{n^2}{8}\delta$$

edges, which implies we have deleted in total less than $n^2\delta/2$ edges in G. The only edges remaining in G are edges between pairs of parts (V_i, V_j) with density greater than $1 - \frac{\delta}{4}$. Since each edge lies in at most n^{p-2} copies of K_p, we have deleted at most $\delta n^p/2$ K_p-s in G. Therefore there is at least one copy of K_p remaining, which implies that there are p parts $V_{i_1}, \ldots, V_{i_p} \in \mathcal{P}$ that pairwise have density at least $1 - \frac{\delta}{4}$, with $|V_{i_j}| = n/K$. Set $\delta_1 = \delta/4$.

For fixed $j \in \{2, \ldots, p\}$, notice that there are at least $(1 - 1/(2p))(n/K)$ vertices $v \in V_{i_1}$ such that $|N(v) \cap V_{i_j}| \geq (1 - 4\delta_1 p)n/K$. Indeed, otherwise we would have

$$|E(V_{i_1}, V_{i_j})| \leq (n/K)\left(1 - \frac{1}{2p}\right)\left(\frac{n}{K}\right) + \frac{n/K}{2p}\,(1 - 4\delta_1 p)\,\frac{n}{K} = \left(\frac{n}{K}\right)^2 - 2\delta_1(n/K)^2.$$

On the other hand, $|E(V_{i_1}, V_{i_j})| \geq (1 - \delta_1)\,(n/K)^2$. This implies $2\delta_1 < \delta_1$ which is a contradiction.

Therefore there is a subset $V'_{i_1} \subset V_{i_1}$ with $|V'_{i_1}| \geq |V_{i_1}|/2$ such that each vertex $v \in V'_{i_1}$ satisfies $|N(v) \cap V_{i_j}| \geq (1 - 4\delta_1 p)|V_{i_j}|$ for all $j = 2, \ldots, p$. If V'_{i_1} is an independent set, then we are done since $|V'_{i_1}| \geq n/(2K)$. Otherwise we have an edge uv in V'_{i_1}. For $j = 2, \ldots, p$, the pigeonhole principle implies that $|V_{i_j} \cap N(u) \cap N(v)| \geq \frac{n}{K}(1 - 8\delta_1 p)$. We define $V_{i_j}^{(2)}$ to be a set of exactly $\frac{n}{K}(1 - 8\delta_1 p)$ elements in $V_{i_j} \cap N(u) \cap N(v)$. Notice that the graph induced on the vertex set $V_{i_2}^{(2)} \cup \cdots \cup V_{i_p}^{(2)}$ is K_{2p-2}-free. Moreover, the density between each pair of parts $(V_{i_j}^{(2)}, V_{i_\ell}^{(2)})$ is at least $(1 - \delta_2)$ where $\delta_2 = \delta_1 + 16\delta_1 p$. We repeat this process on the remaining $p - 1$ parts $V_{i_2}^{(2)}, \ldots, V_{i_p}^{(2)}$.

After j steps, we have either found an independent set of size at least

$$\frac{n}{2K}(1 - 8\delta_1 p)(1 - 8\delta_2(p-1)) \cdots (1 - 8\delta_{j-1}(p - j + 2)),$$

where δ_k is defined recursively as $\delta_1 = \delta/4$ and $\delta_k = \delta_{k-1} + 16\delta_{k-1}p$, or we have obtained subsets $V_{i_j}^{(j)}, \ldots, V_{i_p}^{(j)}$ such that

$$|V_{i_\ell}^{(j)}| = \frac{n}{K}(1 - 8\delta_1 p)(1 - 8\delta_2(p-1)) \cdots (1 - 8\delta_{j-1}(p - j)),$$

for $\ell = j, \ldots, p$, $V_{i_j}^{(j)} \cup \cdots \cup V_{i_p}^{(j)}$ is K_{2p-2j}-free, and the density between each pair of parts $(V_{i_k}^{(j)}, V_{i_\ell}^{(j)})$ is at least $1 - \delta_j$.

By letting $\delta = \delta(\varepsilon, p)$ be sufficiently small such that $\delta_k < \frac{1}{100p}$ for all $k \leq p$, we obtain an independent set of size γn, where $\gamma = \gamma(d, p, \varepsilon)$.

$\square$

The lower bound on $\mathbf{RT}_d(n, K_{2p-1}, o(n))$ and $\mathbf{RT}_d(n, K_{2p}, o(n))$ in Theorem 1.5 follows from a geometric construction of Fox et al. in [24] (see page 15), which is a graph with VC-dimension at most four.

6 Concluding remarks

Many interesting results arose in our study of graphs and hypergraphs with bounded VC-dimension. In particular, we strengthen several classical results from extremal hypergraph theory for hypergraphs with bounded VC-dimension. Below, we briefly mention two of them.

Hypergraphs with bounded VC-dimension. Erdős, Hajnal, and Rado [17] showed that every 3-uniform hypergraph on n vertices contains a clique or independent set of size $c \log \log n$. A famous open question of Erdős asks if $\log \log n$ is the correct order of magnitude for Ramsey's theorem for 3-uniform hypergraphs. According to the best known constructions, there are 3-uniform hypergraphs on n vertices with no clique or independent set of size $c'\sqrt{\log n}$. For $k \geq 4$, the best known lower and upper bounds on the size of the largest clique or independent set in every n-vertex k-uniform hypergraph is of the form $c \log^{(k-1)} n$ (the $(k-1)$-times iterated logarithm) and $c'\sqrt{\log^{(k-2)} n}$, respectively (see [13] for more details). By combining Theorem 1.1 with an argument of Erdős and Rado [20], one can significantly improve these bounds for hypergraphs of bounded (neighborhood) VC-dimension.

Theorem 6.1. *Let $k \geq 3$ and $d \geq 1$. Every k-uniform hypergraph on n vertices with VC-dimension d contains a clique or independent set of size $e^{\left(\log^{(k-1)} n\right)^{1-o(1)}}$.*

Geometric constructions given by Conlon et al. [12] show that Theorem 6.1 is tight apart from the $o(1)$ term in the second exponent. That is, for fixed $k \geq 3$, there are k-uniform hypergraphs on n vertices with VC-dimension $d = d(k)$ such that the largest clique or independent set is of size $O(\log^{(k-2)} n)$.

The Erdős-Hajnal conjecture for tournaments. A *tournament* $T = (V, E)$ on a set V is an orientation of the edges of the complete graph on the vertex set V, that is, for $u, v \in V$ we have either $(u, v) \in E$ or $(v, u) \in E$, but not both. A tournament with no directed cycle is called *transitive*. If a tournament has no subtournament isomorphic to T, then it is called T-free.

An old result due to Entringer, Erdős, and Harner [14] and Spencer [45] states that every tournament on n vertices contains a transitive subtournament of size $c \log n$, which is tight apart from the value of the constant factor. Alon, Pach, and Solymosi [4] showed that the Erdős-Hajnal conjecture is equivalent to the following conjecture.

Conjecture 6.2. *For every tournament T, there is a positive $\delta = \delta(T)$ such that every T-free tournament on n vertices has a transitive subtournament of size n^{δ}.*

In particular, it is known that every T-free tournament on n vertices contains a transitive subtournament of size $e^{c\sqrt{\log n}}$, where $c = c(T)$. Here we note that this bound can be improved in the special case that the forbidden tournament $T = (V, E)$ is *2-colorable*, that is, there is a 2-coloring on $V(T)$ such that the each color class induces a transitive subtournament.

Theorem 6.3. *For fixed integer $k > 0$, let T be a 2-colorable tournament on k vertices. Then every T-free tournament on n vertices contains a transitive subtournament of size $e^{(\log n)^{1-o(1)}}$.*

The idea of the proof of Theorem 6.3 is to use the fact that a tournament T is 2-colorable if and only if the outneighborhood set system of every T-free tournament has VC-dimension at most $c(T)$. There is a straightforward analogue of Theorem 1.3 for tournaments whose outneighborhood set system has bounded VC-dimension, and with this analogous tool, the proof of Theorem 6.3 is essentially the same as the proof of Theorem 1.1.

7 Acknowledgements

We would like to thank Lisa Sauermann for pointing out a small error in an earlier version of the proof of Lemma 2.3.

References

[1] M. Ajtai, J. Komlós, and E. Szemerédi, A note on Ramsey numbers, *J. Combin. Theory Ser. A* **29** (1980), 354–360.

[2] N. Alon, E. Fischer, and I. Newman, Efficient testing of bipartite graphs for forbidden induced subgraphs, *SIAM J. Comput.* **37** (2007), 959–976.

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

[4] N. Alon, J. Pach, J. Solymosi, Ramsey-type theorems with forbidden subgraphs, Combinatorica 21 (2001), 155–170.

[5] N. Alon and J. H. Spencer, *The probabilistic method*, 3rd ed., Wiley, 2008.

[6] M. Anthony, G. Brightwell, and C. Cooper, The Vapnik-Chervonenkis dimension of a random graph, 14th British Combinatorial Conference (Keele, 1993). *Discrete Math.* **138** (1995), 43–56.

[7] T. Bohman, The triangle-free process, *Adv. Math.* **221** (2009), 1653–1677.

[8] T. Bohman and P. Keevash, The early evolution of the H-free process, *Invent. Math.* **181** (2010), 291–336.

[9] B. Bollobás and P. Erdős, On a Ramsey-Turán type problem, *J. Combin Theory Ser. B* **21** (1976), 166–168.

[10] B. Chazelle, *The Discrepancy Method: Randomness and Complexity,* Cambridge University Press, New York, 2000.

[11] D. Conlon and J. Fox, Bounds for graph regularity and removal lemmas, *Geom. Funct. Anal.* **22** (2012), 1191–1256.

[12] D. Conlon, J. Fox, J. Pach, B. Sudakov, and A. Suk, Ramsey-type results for semi-algebraic relations, *Trans. Amer. Math. Soc.* **366** (2014), 5043–5065.

[13] D. Conlon, J. Fox, and B. Sudakov, Hypergraph Ramsey numbers, *J. Amer. Math. Soc.* **23** (2010), 247–266.

[14] R. C. Entringer, P. Erdős, and C. C. Harner, Some extremal properties concerning transitivity in graphs, *Periodica Math. Hungar.* **3** (1973), 275–279.

[15] P. Erdős, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* **53** (1947), 292–294.

[16] P. Erdős and A. Hajnal, Ramsey-type theorems, *Discrete Appl. Math.* **25** (1989), 37–52.

[17] P. Erdős, A. Hajnal, and R. Rado, Partition relations for cardinal numbers, *Acta Math. Acad. Sci. Hungar.* **16** (1965), 93–196.

[18] P. Erdős, A. Hajnal, V.T. Sós, and E. Szemerédi, More results on Ramsey-Turán type problem, *Combinatorica* **3** (1983) 69–82.

[19] P. Erdős and L. Lovász, Problems and results on 3-chromatic hypergraphs and some related questions. *Infinite and finite sets* (Colloq., Keszthely, 1973; dedicated to P. Erds on his 60th birthday), Vol. II, pp. 609–627. Colloq. Math. Soc. Janos Bolyai, Vol. 10, North-Holland, Amsterdam, 1975.

[20] P. Erdős and R. Rado, Combinatorial theorems on classifications of subsets of a given set, *Proc. London Math. Soc.* **3** (1952), 417–439.

[21] P. Erdős and V.T. Sós, Some remarks on Ramsey's and Turán's theorem, in: Combinatorial theory and its applications, II (Proc. Colloq., Balatonfüred, 1969), North-Holland, Amsterdam, 1970, 395–404.

[22] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compos. Math.* **2** (1935), 463–470.

[23] J. Fox, M Gromov, V. Lafforgue, A. Naor, and J. Pach, Overlap properties of geometric expanders, *J. Reine Angew. Math.* **671** (2012), 49–83.

[24] J. Fox, J. Pach, and A. Suk, Semi-algebraic colorings of complete graphs, arXiv:1505.07429.

[25] J. Fox, J. Pach, and A. Suk, A polynomial regularity lemma for semi-algebraic hypergraphs and its applications in geometry and property testing, *SIAM J. Comput.* **45** (2016), 2199–2223.

[26] J. Fox and B. Sudakov, Density theorems for bipartite graphs and related Ramsey-type results, *Combinatorica* **29** (2009), 153–196.

[27] J. E. Goodman, J. O'Rourke, and C. D. Tóth (eds.), *Handbook of Discrete and Computational Geometry,* Chapman Hill/CRC Press, Boca Raton, 2017.

[28] S. Har-Peled, *Geometric Approximation Algorithms,* Mathematical Surveys and Monographs, Vol. 173, Amer. Math. Soc., Providence, 2011.

[29] D. Haussler, Sphere packing numbers for subsets of the Boolean n-cube with bounded Vapnik-Chervonenkis dimension, *J. Combin. Theory Ser. A*, **69** (1995), 217–232.

[30] D. Haussler and E. Welzl, ε-nets and simplex range queries, *Discrete Comput. Geom.* **2** (1987), 127–151.

[31] P. Keevash, Hypergraph Turán problems, *Surveys in Combinatorics*, Cambridge University Press, 2011, 83–140.

[32] J. Komlós, J. Pach, and G. Woeginger, Almost tight bounds for ε-nets, *Discrete Comput. Geom.* **7** (1992), 163–173.

[33] E. Kranakis, D. Krizanc, B. Ruf, J. Urrutia, and G. Woeginger, The VC-dimension of set systems defined by graphs, *Discrete Appl. Math.* **77** (1997), 237–257.

[34] D. Larman, J. Matoušek, J. Pach, and J. Törőcsik, A Ramsey-type result for convex sets, *Bull. London Math. Soc.* **26**(2) (1994), 132–136.

[35] L. Lovász and B. Szegedy, Regularity partitions and the topology of graphons, An Irregular Mind, Imre Bárány, József Solymosi, and Gábor Sági editors, *Bolyai Society Mathematical Studies* **21** (2010), 415–446.

[36] J. Matoušek, *Lectures on Discrete Geometry*, Springer-Verlag, New York, 2002.

[37] J. Pach and J. Solymosi, Structure theorems for systems of segments, *Proceeding of the Japanese Conference on Discrete and Computational Geometry* (2000), 308–317.

[38] N. Sauer, On the density of families of sets, *J. Combin. Theory Ser. A* **13** (1972), 145–147.

[39] M. Schaefer, Deciding the Vapnik-Cervonenkis dimension is Σ_p^3-complete, *J. Comput. System Sci.* **58** (1999), 177–182.

[40] M. Sharir and P. K. Agarwal, *Davenport-Schinzel Sequences and Their Geometric Applications*, Cambridge University Press, Cambridge, 1995.

[41] S. Shelah, A combinatorial problem; stability and order for models and theories in infinitary languages, *Pacific J. Math.* (1972) **41**, 247–261.

[42] M. Simonovits and V. Sós, Ramsey-Turán theory, *Discrete Math.* **229** (2001), 293–340.

[43] V. T. Sós, On extremal problems in graph theory, in *Proc. Calgary Internat. Conf. on Combinatorial Structures and their Application*, 1969, 407–410.

[44] J. Spencer, Asymptotic lower bounds for Ramsey functions, *Discrete Math.* **20** (1977), 69–76.

[45] J. Spencer, Random regular tournaments, *Periodica Mathematica Hungarica* **5** (1974), 105–120.

[46] R. Stanley, *Combinatorics and Commutative Algebra,* Birkhäuser, Boston, 1996.

[47] E. Szemerédi, On graphs containing no complete subgraphs with 4 vertices, *Mat. Lapok* **23** (1972) 111–116 (in Hungarian).

[48] R. Tamassia, ed., *Handbook of Graph Drawing and Visualization*, Chapman and Hall/CRC Press, Boca Raton, 2013.

[49] V. Vapnik and A. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Theory Probab. Appl.* **16** (1971), 264–280.

[50] P. Varnavides, On certain sets of positive density, *J. London Math. Soc.* **34** (1959), 358–360.