Removal Lemmas with Polynomial Bounds

Lior Gishboliner * Asaf Shapira †

Abstract

A common theme in many extremal problems in graph theory is the relation between local and global properties of graphs. One of the most celebrated results of this type is the Ruzsa-Szemerédi triangle removal lemma, which states that if a graph is ε-far from being triangle free, then most subsets of vertices of size $C(\varepsilon)$ are not triangle free. Unfortunately, the best known upper bound on $C(\varepsilon)$ is given by a tower-type function, and it is known that $C(\varepsilon)$ is not polynomial in ε^{-1}. The triangle removal lemma has been extended to many other graph properties, and for some of them the corresponding function $C(\varepsilon)$ is polynomial. This raised the natural question, posed by Goldreich in 2005 and more recently by Alon and Fox, of characterizing the properties for which one can prove removal lemmas with polynomial bounds.

Our main results in this paper are new sufficient and necessary criteria for guaranteeing that a graph property admits a removal lemma with a polynomial bound. Although both are simple combinatorial criteria, they imply almost all prior positive and negative results of this type. Moreover, our new sufficient conditions allow us to obtain polynomially bounded removal lemmas for many properties for which the previously known bounds were of tower-type. In particular, we show that every *semi-algebraic* graph property admits a polynomially bounded removal lemma. This confirms a conjecture of Alon.

1 Introduction

The relation between local and global properties of graphs lies at the core of many of the most well studied problems in extremal graph theory. Perhaps the most natural problem of this type is whether the fact that a graph is "far" from satisfying a property $\mathcal{P}$ implies that it does not satisfy it locally. All graph properties we will consider in this paper are hereditary (i.e. closed under removal of vertices). Note that for such properties, an induced subgraph of G that does not satisfy $\mathcal{P}$ is a "witness" to the fact that G itself does not satisfy $\mathcal{P}$. Thus, for such properties the problem can be phrased as follows: can we deduce from the fact that G is far from satisfying some hereditary property $\mathcal{P}$ that G contains a small subgraph which can witness this fact, and moreover, how many such small witnesses does G contain?

Let us turn the above abstract problem into the concrete one we will study in this paper. We say that a graph G on n vertices is ε-far from satisfying a property $\mathcal{P}$ if one needs to add/delete at least εn^2 edges in order to turn G into a graph satisfying $\mathcal{P}$. The following is the local-vs-global problem we will study in this paper.

Definition 1.1. *Let $\mathcal{P}$ be a hereditary graph property. We say that $\mathcal{P}$ is* testable *if there is a function $f_{\mathcal{P}}(\varepsilon) : (0,1) \to \mathbb{N}$ so that for every graph G which is ε-far from satisfying $\mathcal{P}$, a sequence of*

*School of Mathematics, Tel Aviv University, Tel Aviv 69978, Israel.

†School of Mathematics, Tel Aviv University, Tel Aviv 69978, Israel. Email: asafico@tau.ac.il. Supported in part by ISF Grant 1028/16 and ERC Starting Grant 633509.

$f_{\mathcal{P}}(\varepsilon)$ random vertices of G, sampled uniformly and independently, induces a graph which does not satisfy $\mathcal{P}$ with probability at least $2/3$. We say that $\mathcal{P}$ is easily testable *if (the optimal such) $f_{\mathcal{P}}(\varepsilon)$ is polynomial in ε^{-1}. If $\mathcal{P}$ is not easily testable then it is* hard to test.

Let us mention two famous results in extremal graph theory which fall into the above framework. The first is the celebrated triangle removal lemma of Ruzsa and Szemerédi [38], which is usually stated as saying that if a graph G is ε-far from being triangle free, then G contains at least $n^3/f(\varepsilon)$ triangles. It is easy to see that this statement is equivalent to asserting that the property of being triangle free is testable per Definition 1.1 with a similar bound. The original proof of the triangle removal lemma relied on Szemerédi's regularity lemma [40], which supplied tower-type upper bounds for $f(\varepsilon)$. Due to its intrinsic interest, as well as its relation to other fundamental combinatorial problems, a lot of effort was put into improving this tower-type bound. Unfortunately, the best known upper bound, due to Fox [19], is still a tower-type function. At the other direction, it is known that triangle freeness is not easily testable [38], but the corresponding super polynomial lower bound on $f(\varepsilon)$ is very far from the tower-type upper bound (see e.g. Theorem 4 for a similar bound).

A second classical theorem which falls into the framework of Definition 1.1 is a theorem of Rödl and Duke [34], which states that if G is ε-far from being 3-colorable then G contains a non 3-colorable subgraph on $f(\varepsilon)$ vertices. Actually, a close inspection of the proof in [34] reveals that it in fact shows that 3-colorability is testable. Just as in the case of the triangle removal lemma discussed above, the original proof in [34] relied on the regularity lemma and thus supplied only tower-type upper bounds for $f(\varepsilon)$. However, the situation of this problem changed dramatically when Goldreich, Goldwasser and Ron [24] obtained a new proof of the Rödl-Duke theorem, which avoided the use of the regularity lemma and supplied a polynomial upper bound for $f(\varepsilon)$, thus showing that 3-colorability is easily testable. Actually, the authors of [24] proved a more general result, showing that every so called "partition property" is easily testable.

We now pause for a moment and make two observations regarding Definition 1.1. We first observe that showing that a hereditary property $\mathcal{P}$ is testable per Definition 1.1 is equivalent to proving a removal lemma for $\mathcal{P}$, that is, to proving that if G is ε-far from satisfying $\mathcal{P}$ then G contains at least $n^h/g_{\mathcal{P}}(\varepsilon)$ induced copies of some graph $H \notin \mathcal{P}$ on $h \leq h_{\mathcal{P}}(\varepsilon)$ vertices. As it turns out, it will be more convenient to work with Definition 1.1, especially when dealing with hereditary properties that cannot be characterized by a finite number of forbidden induced subgraphs. The second observation, is that the notion of testability from Definition 1.1 has an interesting algorithmic implication. Suppose we want to design an algorithm that will distinguish with some constant probability, say $2/3$, between graphs satisfying $\mathcal{P}$ and graphs that are ε-far from satisfying it. An immediate corollary of the fact that a property is testable, is that one can solve the above relaxed decision problem in time that depends only on ε and not on the size of the input. Indeed, all the algorithm has to do is sample $f_{\mathcal{P}}(\varepsilon)$ vertices and check if the induced subgraph spanned by these vertices satisfies $\mathcal{P}$. Such an algorithm is called a *property tester*, hence the name we used in Definition 1.1. This notion of testing graph properties was introduced by Goldreich, Goldwasser and Ron [24]. Following [24], numerous other property testing algorithms were designed in various other combinatorial settings.

Given the fact that some hereditary properties are testable, and in light of the algorithmic applications mentioned in the previous paragraph, it is natural to ask which hereditary properties are testable. This question was answered by Alon and Shapira [8] who proved that in fact every hereditary property is testable. This result was later reproved by Lovász and Szegedy [32], and generalized to the setting of hypergraphs by Rödl and Schacht [35] and by Austin and Tao [12]. Unfortunately, since all these proofs relied on some form of Szemerédi's regularity lemma [40], the bounds involved are of tower-type. It was also shown in [10] that there are cases where bounds of this type are unavoidable. However, the examples involved rely on ad-hoc constructions of families of forbidden subgraphs.

It is thus natural to ask which hereditary graph properties are easily testable, or at least which "natural" hereditary properties are easily testable. In other words, for which properties can we prove a removal lemma while avoiding the use of the regularity lemma. This problem was raised in 2005 by Goldreich [23] and recently also by Alon and Fox [6]. Our main results in this paper address this problem by giving very simple yet general combinatorial sufficient and necessary conditions for a hereditary property to be easily testable. In particular, we obtain polynomially bounded removal lemmas for many natural graph properties for which it was not previously known how to obtain a removal lemma without using the regularity lemma.

1.1 The case of finitely many forbidden subgraphs

From this point on, it will be more natural to think of a hereditary property in terms of its forbidden subgraphs. Given a family of graphs $\mathcal{F}$, let $\mathcal{P}^*_{\mathcal{F}}$ be the property of being induced $\mathcal{F}$-free, i.e. not containing an induced copy of each of the graphs of $\mathcal{F}$. When $\mathcal{F}$ consists of a single graph F we will use the notation $\mathcal{P}^*_F$. Note that the family of properties $\mathcal{P}^*_{\mathcal{F}}$ is precisely the family of hereditary properties.

In this subsection we describe our new results concerning hereditary properties which can be characterized by forbidding a finite number of induced subgraphs, that is, the properties $\mathcal{P}^*_{\mathcal{F}}$ with $\mathcal{F}$ being a finite set. We will describe both a sufficient and a necessary condition that a finite family of graphs needs to satisfy in order to guarantee that $\mathcal{P}^*_{\mathcal{F}}$ is easily testable, starting with the former.

We say that a graph F is *co-bipartite* if $V(F)$ can be partitioned into two cliques, and say that F is a *split* graph if $V(F)$ can be partitioned into two sets, one spanning a clique and the other spanning an independent set. Our main positive result regarding finite families is the following simple combinatorial condition, guaranteeing that $\mathcal{P}^*_{\mathcal{F}}$ is easily testable.

Theorem 1. *If $\mathcal{F}$ is a finite family of graphs that contains a bipartite graph, a co-bipartite graph and a split graph then $\mathcal{P}^*_{\mathcal{F}}$ is easily testable.*

We now mention some immediate applications of Theorem 1, starting with known results that follow as special cases of Theorem 1. Let P_k denote the path on k vertices. Alon and Shapira [7] proved that $\mathcal{P}^*_{P_3}$ is easily testable by relying on the fact that a graph satisfies $\mathcal{P}^*_{P_3}$ if and only if it is a disjoint union of cliques. Observing that P_3 is bipartite, co-bipartite and split, Theorem 1 gives the same result. In the same paper [7], it was shown that for any F other than P_2, P_3, P_4, C_4 and their complements, the property $\mathcal{P}^*_F$ is not easily testable. The two cases that were left open were $\mathcal{P}^*_{P_4}$ and $\mathcal{P}^*_{C_4}$. The case of $\mathcal{P}^*_{P_4}$ was settled only very recently by Alon and Fox [6] who used the structural characterization of induced P_4-free graphs in order to show that $\mathcal{P}^*_{P_4}$ is easily testable. As in the case of P_3, since P_4 is bipartite, co-bipartite and split, Theorem 1 gives the result of Alon and Fox [6] as a special case. Finally, a famous theorem of Alon [2] states that the property of being (not necessarily induced) F-free is easily testable if and only if F is bipartite. It is easy to see that the 'if part' of this theorem follows immediately from Theorem 1. Indeed, this follows from the simple observation that being F-free is equivalent to satisfying $\mathcal{P}^*_{\mathcal{F}}$, where $\mathcal{F}$ consists of all supergraphs of F on $|V(F)|$ vertices.

Let us turn to derive some new testability results from Theorem 1. It is well known that the property of being a *line graph* is equivalent to $\mathcal{P}^*_{\mathcal{F}}$, where $\mathcal{F}$ is a family of 9 graphs, each having at most 6 vertices (see [29]). One of these graphs is $K_{1,3}$, which is both bipartite and split, and another one is a complete graph on 5 vertices minus a single edge, which is co-bipartite. Hence, Theorem 1 implies that the property of being a line graph is easily testable. Two other graph properties which can be shown to be easily testable via Theorem 1 are being a *threshold graph* and a *trivially perfect graph*. Since both properties are equivalent to $\mathcal{P}^*_{\mathcal{F}}$ for an appropriate finite $\mathcal{F}$, where in both cases $P_4 \in \mathcal{F}$ (see [26, 27]), we immediately deduce from Theorem 1 that both are easily testable.

We now turn to describe our necessary condition for being easily testable. Recall that our sufficient condition from Theorem 1 asks $\mathcal{F}$ to contain a bipartite graph, a co-bipartite graph and a split graph. The next theorem shows that having at least one bipartite graph and at least one co-bipartite graph is a necessary condition.

Theorem 2. *Let $\mathcal{F}$ be a finite family for which $\mathcal{P}^*_{\mathcal{F}}$ is easily testable. Then $\mathcal{F}$ contains a bipartite graph and a co-bipartite graph.*

As we mentioned above, Alon [2] proved that being F-free is easily testable if and only if F is bipartite. It is now easy to see that the 'only if' part of Alon's result follows from Theorem 2. As we mentioned above, Alon and Shapira [7] proved that $\mathcal{P}^*_F$ is not easily testable for every F other than P_2, P_3, P_4, C_4 and their complements. Again, this result follows as a special case of Theorem 2.

Having given both a necessary and a sufficient condition, it is natural to ask if one of them in fact characterizes the finite families $\mathcal{F}$ for which $\mathcal{P}^*_{\mathcal{F}}$ is easily testable. Unfortunately, none do. It is known that being a split graph is equivalent to $\mathcal{P}^*_{\mathcal{F}}$ where $\mathcal{F} = \{C_5, C_4, \overline{C_4}\}$ (see [26]). While $\mathcal{F}$ does not satisfy the condition of Theorem 1 (it does not contain a split graph), the property of being a split graph is easily testable since it is one of the partition properties that were shown to be easily testable in [24]. Therefore, the sufficient condition in Theorem 1 is not necessary. Showing that the necessary condition of Theorem 2 is not sufficient is a bit harder, and is stated in the following theorem.

Theorem 3. *There is a bipartite F_1 and a co-bipartite F_2 such that $\mathcal{P}^*_{\{F_1,F_2\}}$ is not easily testable.*

Thus the above theorem also implies that in Theorem 1 we cannot drop the requirement that $\mathcal{F}$ should contain a split graph. The fact that we cannot drop the requirement that $\mathcal{F}$ should contain a bipartite graph follows from [38] where it was (implicitly) proved that triangle-freeness is not easily testable. By symmetry, the same holds for the co-bipartite graph.

We conclude our discussion on the case of finite forbidden families with the following theorem, which turns out to be the key step in the proof of Theorem 2. We will comment on the importance of this theorem in Subsection 1.3.

Theorem 4. *For every $h \geq 3$ there is $\varepsilon_0 = \varepsilon_0(h)$ such that the following holds for every $\varepsilon < \varepsilon_0$ and for every non-bipartite graph H on h vertices. For every $n \geq n_0(\varepsilon)$ there is a graph on n vertices which is ε-far from being induced H-free and yet contains at most $\varepsilon^{\Omega(\log(1/\varepsilon))} n^h$ (not necessarily induced) copies of H.*

1.2 The case of infinitely many forbidden subgraphs

We now turn to consider properties $\mathcal{P}^*_{\mathcal{F}}$ when $\mathcal{F}$ is a (possibly) infinite family. We start by introducing an important feature of a hereditary graph property.

Definition 1.2. *Let F be a graph with vertex set $V(F) = \{1, \ldots, p\}$ and let $g : V(F) \to \{0, 1\}$. We say that a graph G is a g-*blowup *of F if G admits a vertex partition $V(G) = P_1 \cup \cdots \cup P_p$ with the following properties.*

1. *For every $1 \leq i < j \leq p$, if $(i, j) \in E(F)$ then (P_i, P_j) is a complete bipartite graph, and if $(i, j) \notin E(F)$ then (P_i, P_j) is an empty bipartite graph.*

2. *For every $1 \leq i \leq p$, if $g(i) = 1$ then P_i is a clique and if $g(i) = 0$ then P_i is an independent set.*

Definition 1.3. *We say that a graph property $\mathcal{P}$ is* closed under blowups *if for every graph F which satisfies $\mathcal{P}$ there is a function $g : V(F) \to \{0, 1\}$ such that every g-blowup of F satisfies $\mathcal{P}$.*

Our main result regarding hereditary properties characterized by an infinite family of forbidden subgraphs $\mathcal{F}$ is the following.

Theorem 5. *Let $\mathcal{F}$ be a graph family such that*

1. *$\mathcal{F}$ contains a bipartite graph, a co-bipartite graph and a split graph.*
2. *$\mathcal{P}^*_{\mathcal{F}}$ is closed under blowups.*

Then $\mathcal{P}^*_{\mathcal{F}}$ is easily testable.

We now describe what we consider the most important result of this paper. Let us recall the definition of a *semi-algebraic graph property*. A semi-algebraic graph property $\mathcal{P}$ is given by an integer $k \geq 1$, a set of real $2k$-variate polynomials $f_1, \dots, f_t \in \mathbb{R}[x_1, \dots, x_{2k}]$ and a Boolean function $\Phi : \{\text{true}, \text{false}\}^t \to \{\text{true}, \text{false}\}$. A graph G satisfies the property $\mathcal{P}$ if one can assign a point $p_v \in \mathbb{R}^k$ to each vertex $v \in V(G)$ in such a way that a pair of distinct vertices u, v are adjacent if and only if

$$\Phi\Big(f_1(p_u, p_v) \geq 0, \dots, f_t(p_u, p_v) \geq 0\Big) = \text{true}.$$

In the expression $f_i(p_u, p_v)$, we substitute p_u into the first k variables of f_i and p_v into the last k variables of f_i. In what follows, we call the points p_v *witnesses*[1] to the fact that G satisfies $\mathcal{P}$.

Some examples of semi-algebraic graph properties are those that correspond to being an intersection graph of certain semi-algebraic sets in $\mathbb{R}^k$. For example, a graph is an *interval graph* if one can assign an interval in $\mathbb{R}$ to each vertex so that u, v are adjacent iff their intervals intersect. Similarly, a graph is a *unit disc graph* if it is the intersection graph of unit discs in $\mathbb{R}^2$.

The family of semi-algebraic graph properties has been extensively studied by many researchers, see e.g. [22] and its references. Alon [3] conjectured that every semi-algebraic graph property is easily testable. As we now show, this conjecture can be easily derived from Theorem 5.

Theorem 6. *Every semi-algebraic graph property is easily testable.*

Proof. (sketch) Fix a semi-algebraic graph property $\mathcal{P}$. Let $\mathcal{F}$ be the family of all graphs which do not satisfy $\mathcal{P}$. As $\mathcal{P}$ is a hereditary property, we have $\mathcal{P} = \mathcal{P}^*_{\mathcal{F}}$. To prove the theorem, it is enough to show that $\mathcal{P} = \mathcal{P}^*_{\mathcal{F}}$ satisfies Conditions 1 and 2 in Theorem 5. The fact that $\mathcal{F}$ satisfies Condition 1 of Theorem 5 follows directly from the well known fact that every graph satisfying $\mathcal{P}$ has a bounded VC-dimension (we will give the definition of the VC-dimension of a graph in the detailed proof of Theorem 6, see Subsection 2.2). As for Condition 2, assume F satisfies $\mathcal{P}$, and $\{p_v : v \in V(F)\}$ are points witnessing this fact. Then setting $g(v) = 1$ if and only if $\Phi\Big(f_1(p_v, p_v) \geq 0; \dots; f_t(p_v, p_v) \geq 0\Big) = \text{true}$, it is easy to see that every g-blowup of F satisfies $\mathcal{P}$. Indeed, the points witnessing the fact that a g-blowup of F satisfies $\mathcal{P}$ are obtained by taking each of the points p_v an appropriate number of times. ■

[1]Note that a graph G might have many assignments of points witnessing the fact that it satisfies $\mathcal{P}$.

The reader can find a more detailed proof of Theorem 6 in Subsection 2.2. Returning to the discussion at the beginning of the paper, observe that an immediate corollary of Theorem 6 is that for every semi-algebraic graph property $\mathcal{P}$ there is $c = c(\mathcal{P})$, so that if G is ε-far from satisfying $\mathcal{P}$ (and ε is small enough), then G contains a subgraph on ε^{-c} vertices which does not satisfy $\mathcal{P}$.

The concept of VC-dimension (implicitly) plays a key role in our proofs of Theorems 1, 5 and 6 (see [11, Chapter 14] for an overview of this concept). In fact, as we (implicitly) show later in the paper, a hereditary property $\mathcal{P}$ satisfies Condition 1 of Theorem 5 (i.e., it forbids a bipartite graph, a co-bipartite graph and a split graph), if and only if it has bounded VC dimension[2], in the sense that the VC-dimension of any graph satisfying $\mathcal{P}$ is bounded from above by some constant depending only on $\mathcal{P}$. Another aspect of the role played by VC-dimension in our results is the fact that the main tool we use, i.e. the "conditional" regularity lemma of [5] (stated here as Lemma 2.7), can be roughly stated as saying that graphs with bounded VC-dimension have small and highly-structured regular partitions (see [31] for a similar result). The proof of this lemma in [5] uses properties of VC-dimension.

It is worth mentioning that by now there are several works concerning efficient (i.e. polynomial) regularity lemmas for special classes of graphs, such as graphs with bounded VC-dimension [5, 31] (which, as mentioned above, play a key role in the present paper); semi-algebraic graphs and hypergraphs [21, 22, 41] and more generally distal graphs [15, 16, 39]; and graphs excluding an induced bipartite half-graph [33].

Given Theorem 1, it is natural to ask if Condition 1 in Theorem 5 already guarantees that a property is easily testable. In light of the above discussion, this is equivalent to the (aesthetically pleasing) statement that every hereditary property of bounded VC dimension is easily testable. As our final theorem shows, this is regretfully not the case.

Theorem 7. *There is a family of graphs $\mathcal{F}$ that contains a bipartite graph, a co-bipartite graph and a split graph, for which $\mathcal{P}^*_{\mathcal{F}}$ is not easily testable.*

1.3 Some nuggets about the proofs

We start with some comments regarding the proofs of Theorems 1 and 5. One key observation needed for these proofs is that given a bipartite graph A_1, a co-bipartite graph A_2, and a split graph A_3, there is a bipartite graph B on vertex sets X, Y, so that no matter which graphs one puts on X and on Y, one always gets a graph containing an *induced* copy of either A_1, A_2 or A_3 (see Lemma 2.2). This means that if $\mathcal{F}$ satisfies the assumption of Theorem 1 and G satisfies $\mathcal{P}^*_{\mathcal{F}}$ then G has no induced copy of any graph obtained by adding edges to the two partition classes of B. If this is the case, then one can apply a "conditional regularity lemma" of Alon, Fischer and Newman [5] in order to find a highly structured partition of G (even more structured than the one produced by the regularity lemma [40]) which is of size only poly$(1/\varepsilon)$. This is in sharp contrast to the general argument of [8] that relied on Szemerédi's regularity lemma [40] which can only produce partitions of size Tower$(1/\varepsilon)$, see [28]. The proof of Theorem 5 is more involved, mainly due to having to handle an infinite number of forbidden subgraphs. What usually considerably complicates proofs of this type is the need to embed multiple vertices into the same cluster of the partition mentioned above. The difficulty arises from the fact that clusters of the partition are not highly structured (as opposed to the bipartite graphs between them). However, when dealing with properties satisfying Condition

[2]What we show (see Lemma 2.2) is that Condition 1 of Theorem 5 implies that every graph G satisfying $\mathcal{P}$ has no induced bipartite copy of some $k \times k$ bipartite graph. It is easy to see that this in turn implies that such a G has VC dimension at most $2k$.

2 of Theorem 5, it is enough to embed at most one vertex into each cluster. This feature is what makes it possible to prove Theorem 5.

As we mentioned above, the construction described in Theorem 4 is the key step in the proof of Theorem 2. Let us explain why in Theorem 4 we managed to overcome a difficulty that was not resolved in previous works. Alon's result [2] that being F-free is not easily testable for non-bipartite F relied on a construction of a graph that is ε-far from being F-free yet contains only $\varepsilon^{c \log(1/\varepsilon)} n^{v(F)}$ copies of F. He further asked for which F the property $\mathcal{P}^*_F$ is easily testable. The reason why the construction in [2] did not imply that $\mathcal{P}^*_F$ is hard for every non-bipartite F (or a complement of one) was that it did not produce a graph that is ε-far from being *induced* F-free. In fact, in most cases the graph was induced F-free. So what we do in Theorem 4 is reprove the result of [2] in a way that simultaneously resolves the open problem raised in that paper. To prove Theorem 4 we too use a construction based on Behrend's [13] example of a large set of integers S without 3-term arithmetic progressions, but with the following twist. First, we take a set S that does not contain a (non-trivial) solution to *any* convex[3] linear equation with small coefficients. Second, we carefully label the vertices/clusters in this construction in such a way that any copy of H in the construction will necessarily contain a *monotone* cycle, i.e. a cycle whose labels increase in value. This property guarantees that such a cycle corresponds to a solution of a convex linear equation with integers from S, but we know that S has no such solution.

1.4 Organization

The rest of the paper is organized as follows. In Section 2 we prove Theorems 1 and 5. We also give a more detailed proof of Theorem 6 in Subsection 2.2. In Section 3 we prove Theorems 2, 3, 4 and 7.

2 Easily Testable Properties

In this section we prove Theorems 1 and 5. Throughout the section, we assume that n, the number of vertices of the host graph G, is large enough (as a function of the other parameters, i.e. the property $\mathcal{P}$ and the approximation parameter ε). We note that the minimal n for which our arguments work is (only) polynomial in $1/\varepsilon$ (where the polynomial depends on $\mathcal{P}$). To keep the presentation clean, we will often implicitly assume[4] that n is divisible by various integers which are bounded from above by a function of $\mathcal{P}$ and ε (which is polynomial in $1/\varepsilon$). We start with some preliminary definitions. Let G be a graph on n vertices. For a set $X \subseteq V(G)$, we denote by $G[X]$ the subgraph of G induced by X. We say that X is *homogeneous* if it is either a clique or an independent set.

For a pair of disjoint sets $X, Y \subseteq V(G)$, let $e(X, Y)$ denote the number of edges with one endpoint in X and one endpoint in Y, and set $d(X,Y) = \frac{e(X,Y)}{|X||Y|}$. The number $d(X, Y)$ is called the *density* of the pair (X, Y). Note that $d(X, Y) = 1$ (resp. $d(X, Y) = 0$) if and only if the bipartite graph between X and Y is complete (resp. empty). We say that the pair (X, Y) is *homogeneous* if either $d(X, Y) = 1$ or $d(X, Y) = 0$. For $\delta \in (0, 1)$, we say that (X, Y) is δ-*homogeneous* if either $d(X, Y) \geq 1 - \delta$ or $d(X, Y) \leq \delta$. In cases where we consider several graphs at the same time, we write $d_G(X, Y)$ to refer to the density in G. The *weight* of (X, Y) is defined as $\frac{|X||Y|}{n^2}$.

Let $\mathcal{U} = \{U_1, \dots, U_r\}$ be a vertex-partition of G, i.e. $V(G) = U_1 \uplus \cdots \uplus U_r$. We say that $\mathcal{U}$ is an *equipartition* if $||U_i| - |U_j|| \leq 1$ for every $1 \leq i, j \leq r$. Evidently, if r divides n (which we will assume

[3] A linear equation is convex if it is of the form $a_1x_1 + \ldots + a_kx_k = (a_1 + \ldots + a_k)x_{k+1}$ with all $a_i > 0$.

[4] if one wishes to discard this assumption, then it may be necessary to slightly change some of the constants chosen in the course of the proofs appearing in this section.

to be the case, as mentioned above) then all parts $U_1, \dots, U_r$ have the same size. We say that $\mathcal{U}$ is δ-*homogeneous* if the sum of weights of non-δ-homogeneous pairs (U_i, U_j), $1 \le i \ne j \le r$, is at most δ. Note that if all parts in $\mathcal{U}$ have the same size then $\mathcal{U}$ is δ-homogeneous if and only if the number of (ordered) non-δ-homogeneous pairs (U_i, U_j) is at most δr^2.

We will need the following well-known quantitative version of Ramsey's theorem.

Claim 2.1 (see e.g. [14])**.** *Every graph on 4^k vertices contains a homogeneous set of size k.*

Let $H = (S \cup T, E)$ be a bipartite graph. A *completion* of H is any graph on $V(H)$ that agrees with H on the edges between S and T. In other words, a completion of H is any graph obtained by putting two arbitrary graphs on the sets S and T. We say that H is a *bipartite obstruction* for a graph property $\mathcal{P}$ if *no* completion of H satisfies $\mathcal{P}$. The first ingredient in the proofs of Theorems 1 and 5 is the following lemma.

Lemma 2.2. *Let $\mathcal{F}$ be a graph family. Then $\mathcal{P}^*_{\mathcal{F}}$ admits a bipartite obstruction if and only if $\mathcal{F}$ contains a bipartite graph, a co-bipartite graph and a split graph.*

Proof. We start with the "only if"-direction of the lemma. Let H be a bipartite obstruction for $\mathcal{P}^*_{\mathcal{F}}$ with sides S and T. By putting empty graphs on S and T we get a bipartite graph that does not satisfy $\mathcal{P}^*_{\mathcal{F}}$. This bipartite graph must then contain as an induced subgraph some element of $\mathcal{F}$, which is evidently also bipartite. This shows that $\mathcal{F}$ contains a bipartite graph. Similarly, by putting complete graphs on S and T (resp. a complete graph on S and an empty graph on T) we infer that $\mathcal{F}$ contains a co-bipartite (resp. split) graph, as required.

We now prove the "if"-direction of the lemma. Let $F_1, F_2, F_3 \in \mathcal{F}$ be such that F_1 is bipartite, F_2 is co-bipartite and F_3 is split, and write $V(F_1) = P_1 \cup Q_1$, $V(F_2) = P_2 \cup Q_2$, $V(F_3) = P_3 \cup Q_3$, where P_1, Q_1, P_3 are independent sets and P_2, Q_2, Q_3 are cliques. Put $f := v(F_1) + v(F_2) + 2v(F_3)$, and let h be some large integer, to be chosen later. Let $H = (S \cup T, E)$ be a random bipartite graph with $|S| = |T| = h$; that is, for each $s \in S, t \in T$, the edge (s,t) is included in H with probability $\frac{1}{2}$, independently. We will show that with positive probability, H is a bipartite obstruction for $\mathcal{F}$, thus proving the lemma. Let us set

$$r := \left\lfloor \frac{h - 4^f}{f} \right\rfloor .$$

An (f,r)-*family* is a $2r$-tuple $(S_1, ..., S_r; T_1, ..., T_r)$ such that $S_1, ..., S_r$ (resp. $T_1, ..., T_r$) are pairwise-disjoint subsets of S (resp. T) of size f each. The number of ways to choose an (f,r)-family is exactly

$$\left(\frac{h!}{(f!)^r (h - fr)!} \right)^2 \le h^{2h}.$$

We need the following definition. Let F and H be graphs and let $V(F) = P \cup Q$ and $V(H) = S \cup T$ be vertex-partitions. An *induced bipartite copy* of $F[P,Q]$ in $H[S,T]$ is an injection $\varphi : V(F) \to V(H)$ such that $\varphi(P) \subseteq S$, $\varphi(Q) \subseteq T$ and for every $p \in P$ and $q \in Q$ we have $(p,q) \in E(F)$ if and only if $(\varphi(p), \varphi(q)) \in E(H)$.

For an (f,r)-family $\mathcal{Q} = (S_1, ..., S_r; T_1, ..., T_r)$ and for $(i,j) \in [r]^2$, let $A_{\mathcal{Q}}(i,j)$ be the event that $H[S_i, T_j]$ contains induced bipartite copies of $F_1[P_1, Q_1]$, $F_2[P_2, Q_2]$, $F_3[P_3, Q_3]$ and $F_3[Q_3, P_3]$. We claim that for every completion H' of H, if S_i and T_j are homogeneous sets in H' and $A_{\mathcal{Q}}(i,j)$ happened, then H' is not induced $\mathcal{F}$-free (and hence does not satisfy $\mathcal{P}^*_{\mathcal{F}}$). Indeed, if S_i, T_j are independent sets (in H') then $H'[S_i \cup T_j]$ contains an induced copy of F_1; if S_i, T_j are cliques (in H') then $H'[S_i \cup T_j]$ contains an induced copy of F_2; and if S_i is a clique and T_j is an independent set or vice versa, then $H'[S_i \cup T_j]$ contains an induced copy of F_3.

Now let $\mathcal{A}$ be the event that for every (f,r)-family $\mathcal{Q}$, there is a pair $(i,j) \in [r]^2$ for which $A_{\mathcal{Q}}(i,j)$ happened. We now show that if $\mathcal{A}$ happened then H is a bipartite obstruction for $\mathcal{F}$. We will then show that $\mathcal{A}$ happens with positive probability. Let H' be a completion of H. By repeatedly applying Claim 2.1, we extract from S pairwise-disjoint homogeneous sets $S_1, S_2, \ldots, S_r$ of size f each. This is possible due to our choice of r. Similarly, we extract from T pairwise-disjoint homogeneous sets $T_1, T_2, \ldots, T_r$ of size f each. Consider the (f,r)-family $\mathcal{Q} = (S_1, ..., S_r; T_1, ..., T_r)$. Since $\mathcal{A}$ happened, there is $(i,j) \in [r]^2$ for which $A_{\mathcal{Q}}(i,j)$ happened. Since S_i and T_j are homogeneous in H', we get that H' does not satisfy $\mathcal{P}^*_{\mathcal{F}}$, as required.

So it remains to show that $\mathbb{P}[\mathcal{A}] > 0$. Let $\mathcal{Q} = (S_1, ..., S_r; T_1, ..., T_r)$ be an (f,r)-family. Since $|S_i| = |T_j| = f = v(F_1) + v(F_2) + 2v(F_3)$, it is possible to put a bipartite graph on (S_i, T_j) that will contain induced bipartite copies of $F_1[P_1, Q_1]$, $F_2[P_2, Q_2]$, $F_3[P_3, Q_3]$ and $F_3[Q_3, P_3]$. This implies that $\mathbb{P}\left[A_{\mathcal{Q}}(i,j)\right] \geq 2^{-f^2}$. Since the events $\{A_{\mathcal{Q}}(i,j) : i,j \in [r]\}$ are independent, the probability that $A_{\mathcal{Q}}(i,j)$ did not happen for any $(i,j) \in [r]^2$ is at most $\left(1 - 2^{-f^2}\right)^{r^2} \leq e^{-2^{-f^2} r^2} < h^{-2h}$, with the rightmost inequality holding provided that we choose h to be large enough (see our choice of r). Recall that there are at most h^{2h} ways to choose an (f,r)-family $\mathcal{Q}$. By the union bound over all (f,r)-families, we get $\mathbb{P}[\mathcal{A}^c] < 1$, as required. This completes the proof. ■

An *induced bipartite copy* of a bipartite graph $H = (S \cup T, E)$ in a graph G is an injection $\varphi : V(H) \to V(G)$ such that for every $s \in S$ and $t \in T$ we have $(s,t) \in E(H)$ if and only if $(\varphi(s), \varphi(t)) \in E(G)$. Notice that there is no restriction on the subgraphs of G induced by $\varphi(S)$ or by $\varphi(T)$ (in other words, the definition only "cares" about the edges between $\varphi(S)$ and $\varphi(T)$).

The following lemma is the main tool used in the proofs of Theorems 1 and 5. It is worth noting that the idea of taking a regular partition and a refinement thereof (with a better measure of regularity) was first introduced in [4]. This approach, tailored to regularity lemmas with polynomial bounds, was also applied in [20, 22].

Lemma 2.3. *There are functions $\rho_{2.3} : \mathbb{N} \times (0,1) \to (0,1)$ and $\zeta_{2.3} : \mathbb{N}^2 \times (0,1)^2 \to (0,1)$ such that*[5] *$\rho_{2.3}(h,\delta) = \mathrm{poly}(\delta)$, $\zeta_{2.3}(h,m,\delta,\gamma) = \mathrm{poly}(\delta,\gamma)$, and the following holds for every pair of integers $h, m \geq 1$, for every $\gamma, \delta \in (0,1)$ and for every $h \times h$ bipartite graph H. Every graph G on $n \geq n_0(h,m,\delta,\gamma) = \mathrm{poly}(1/\delta, 1/\gamma)$ vertices either contains at least $\zeta_{2.3}(h,m,\delta,\gamma) n^{2h}$ induced bipartite copies of H or satisfies the following. There is an equipartition $\mathcal{U} = \{U_1, ..., U_r\}$ of G with $\delta^{-1} \leq r \leq \rho_{2.3}(h,\delta)^{-1}$ parts, and for each $1 \leq i \leq r$ there is a set $W_i \subseteq U_i$ and pairwise-disjoint sets $W_{i,1}, \ldots, W_{i,m} \subseteq W_i$ satisfying*

1. *For all but at most δr^2 of the pairs $1 \leq i < j \leq r$, it holds that (U_i, U_j) is δ-homogeneous and $|d(W_i, W_j) - d(U_i, U_j)| \leq \frac{1}{4}$.*
2. *For every $1 \leq i < j \leq r$, (W_i, W_j) is γ-homogeneous and $|d(W_{i,s}, W_{j,t}) - d(W_i, W_j)| \leq \gamma$ for every $1 \leq s, t \leq m$.*
3. *For every $1 \leq i \leq r$, either $d(W_{i,s}, W_{i,t}) \geq 1 - \gamma$ for every $1 \leq s < t \leq m$ or $d(W_{i,s}, W_{i,t}) \leq \gamma$ for every $1 \leq s < t \leq m$.*
4. *$|W_{i,s}| \geq n \cdot \zeta_{2.3}(h,m,\delta,\gamma)$ for every $1 \leq i \leq r$ and $1 \leq s \leq m$.*

[5] By $\rho_{2.3}(h,\delta) = \mathrm{poly}(\delta)$ we mean that $\rho_{2.3}(h,\delta)$ is at least polynomial in δ. The particular polynomial may (and usually will) depend on h, but we omit this from the notation because in what follows, h will depend only on the property $\mathcal{P}$ (and not on ε). Similarly, $\zeta_{2.3}(h,m,\delta,\gamma) = \mathrm{poly}(\delta,\gamma)$ means that $\zeta_{2.3}(h,m,\delta,\gamma)$ is (at least) polynomial in δ, γ, where the polynomial may depend on h, m.

The last tool we need in the proofs of Theorems 1 and 5 is the following counting lemma.

Lemma 2.4. *Let F be a graph, say with vertex-set $V(F) = \{1, \dots, \ell\}$, and let $\lambda \in (0, 1)$. Let $W_1, ..., W_\ell$ be pairwise-disjoint vertex sets in an n-vertex graph G such that*

1. *For every $1 \le i < j \le \ell$, if $(i, j) \in E(F)$ then $d(W_i, W_j) \ge 1 - \frac{1}{2\ell^2}$ and if $(i, j) \notin E(F)$ then $d(W_i, W_j) \le \frac{1}{2\ell^2}$.*
2. *$|W_i| \ge \lambda n$ for every $1 \le i \le \ell$.*

Then with probability at least $\frac{2}{3}$, a random sequence of $12\ell/\lambda$ vertices of G, sampled uniformly and independently, contains an induced copy of F.

Proof. For each $1 \le i \le \ell$, sample a vertex $w_i \in W_i$ uniformly at random. For every $1 \le i < j \le \ell$, the assumption of the lemma gives that with probability at least $1 - \frac{1}{2\ell^2}$, if $(i, j) \in E(F)$ then $(w_i, w_j) \in E(G)$ and if $(i, j) \notin E(F)$ then $(w_i, w_j) \notin E(G)$. By the union bound over all pairs $1 \le i < j \le \ell$ we get that with probability at least $1 - \binom{\ell}{2}/2\ell^2 \ge \frac{3}{4}$, the set $\{w_1, \dots, w_\ell\}$ spans an induced copy of F in which w_i plays the role of i.

Now let $u_1, \dots, u_s \in V(G)$ be a random sequence of vertices, sampled uniformly and independently, where $s = 12\ell/\lambda$. Let $\mathcal{A}$ be the event that $U := \{u_1, \dots, u_s\}$ contains a vertex of W_i for every $1 \le i \le \ell$. What we proved in the previous paragraph implies that conditioned on $\mathcal{A}$ happening, $G[U]$ contains an induced copy of F with probability at least $\frac{3}{4}$. Hence, to finish the proof it is enough to show that $\mathbb{P}[\mathcal{A}^c] \le \frac{1}{12}$. For $1 \le i \le \ell$, the probability that $U \cap W_i = \emptyset$ is $\left(1 - \frac{|W_i|}{n}\right)^s \le (1 - \lambda)^s \le e^{-\lambda s} \le \frac{1}{12\ell}$. Here we used the assumption $|W_i| \ge \lambda n$ and our choice of s. By the union bound over all $1 \le i \le \ell$ we get that $\mathbb{P}[\mathcal{A}^c] \le \frac{1}{12}$, as required. ■

We are now ready to prove Theorems 1 and 5.

Proof of Theorem 1. Our goal is to prove that $\mathcal{P}^*_{\mathcal{F}}$ is testable per Definition 1.1 with $f_{\mathcal{P}}(\varepsilon) = \text{poly}(1/\varepsilon)$. By Lemma 2.2, $\mathcal{P}^*_{\mathcal{F}}$ has a bipartite obstruction H. We can assume (by adding additional vertices if needed) that the two sides of H are of the same size, which we denote by h. We set $m := \max_{F \in \mathcal{F}} v(F)$. Given $\varepsilon < \frac{1}{2}$, set

$$\zeta := \zeta_{2.3}\left(h, m, \frac{\varepsilon}{3}, \frac{1}{4m^2}\right),$$

noting that $\zeta = \text{poly}(\varepsilon)$ (as h and m depend only on $\mathcal{P}$). Let G be an n-vertex graph which is ε-far from being induced $\mathcal{F}$-free. If G contains at least ζn^{2h} induced bipartite copies of H, then a random sequence of $2h = |V(H)|$ vertices of G (sampled uniformly and independently) spans an induced bipartite copy of H with probability at least ζ. Hence, a random sequence of $4h \cdot \zeta^{-1} = \text{poly}(1/\varepsilon)$ vertices of G contains an induced bipartite copy of H with probability at least

$$1 - (1 - \zeta)^{1/\zeta} \ge 1 - e^{-2} \ge \frac{2}{3}.$$

Since H is a bipartite obstruction for $\mathcal{P}^*_{\mathcal{F}}$, every graph which contains an induced bipartite copy of H does not satisfy $\mathcal{P}^*_{\mathcal{F}}$. So we see that the assertion of the theorem holds in the case that G contains at least ζn^{2h} induced bipartite copies of H.

Suppose from now on that G contains less than ζn^{2h} induced bipartite copies of H. We apply Lemma 2.3 to G with parameters $\delta = \frac{\varepsilon}{3}$, $\gamma = \frac{1}{4m^2}$ and m as defined above, to get an equipartition

$\mathcal{U} = \{U_1, ..., U_r\}$, sets $W_i \subseteq U_i$ and pairwise-disjoint sets $W_{i,1}, \ldots, W_{i,m} \subseteq W_i$ with the properties stated in the lemma.

Let G' be the graph obtained from G by making the following changes.

(a) For every $1 \leq i < j \leq r$, if $d(W_i, W_j) \geq 1 - \frac{1}{4m^2}$ then turn (U_i, U_j) into a complete bipartite graph, and if $d(W_i, W_j) \leq \frac{1}{4m^2}$ then turn (U_i, U_j) into an empty bipartite graph. By Item 2 in Lemma 2.3, one of these options holds.

(b) For every $1 \leq i \leq r$, if $d(W_{i,s}, W_{i,t}) \geq 1 - \frac{1}{4m^2}$ (resp. $d(W_{i,s}, W_{i,t}) \leq \frac{1}{4m^2}$) for every $1 \leq s < t \leq m$, then turn U_i into a clique (resp. an independent set). By Item 3 in Lemma 2.3, one of these options holds.

We claim that the number of edge-changes made in items (a)-(b) is less than εn^2. To prove this, we define $\mathcal{H}$ to be the set of pairs $1 \leq i < j \leq r$ for which (U_i, U_j) is $\frac{\varepsilon}{3}$-homogeneous and $|d(W_i, W_j) - d(U_i, U_j)| \leq \frac{1}{4}$. Observe that if $(i,j) \in \mathcal{H}$ then at most $\frac{\varepsilon}{3}|U_i||U_j|$ edge-changes were made in the bipartite graph (U_i, U_j) in Item (a) above; indeed, in the case that $d(W_i, W_j) \geq 1 - \frac{1}{4m^2} \geq \frac{3}{4}$ we have $d(U_i, U_j) \geq \frac{1}{2}$ and hence actually $d(U_i, U_j) \geq 1 - \frac{\varepsilon}{3}$; and the case that $d(W_i, W_j) \leq \frac{1}{4m^2}$ is symmetrical. By Item 1 in Lemma 2.3, the number of pairs $1 \leq i < j \leq r$ *not* belonging to $\mathcal{H}$ is at most $\frac{\varepsilon}{3}r^2$. It follows that the overall number of changes made in Item (a) is at most $|\mathcal{H}| \cdot \frac{\varepsilon}{3} \cdot \left(\frac{n}{r}\right)^2 + \frac{\varepsilon}{3}r^2 \cdot \left(\frac{n}{r}\right)^2 \leq \frac{2\varepsilon}{3}n^2$. As for item (b), the number of edge-changes made there is at most $r\binom{n/r}{2} < \frac{n^2}{r} \leq \frac{\varepsilon}{3}n^2$, where in the last inequality we used the fact that $r \geq \frac{3}{\varepsilon}$, which is guaranteed by Lemma 2.3. In conclusion, the number of edge-changes made when turning G into G' is less than εn^2.

Since G is ε-far from being induced $\mathcal{F}$-free, G' must contain an induced copy of some $F \in \mathcal{F}$. Suppose without loss of generality that $U_1, \ldots, U_p$ are the parts of $\mathcal{U}$ which contain vertices of this copy, and let X_i be the set of vertices of this copy which lie in U_i (for $1 \leq i \leq p$). From the definition of G' it follows that the sets $X_1, \ldots, X_p$ and the bipartite graphs (X_i, X_j), $1 \leq i < j \leq p$, are homogeneous. Note that by our choice of m we clearly have $\ell := v(F) \leq m$, and in particular $|X_i| \leq m$ for every $1 \leq i \leq p$.

We now show that the sets $W_{i,s}$, where $1 \leq i \leq p$ and $1 \leq s \leq |X_i|$, satisfy Condition 1 of Lemma 2.4 (with respect to F) *in the graph* G. First, for every $1 \leq i \leq p$, if X_i is a clique (resp. an independent set) then $G'[U_i]$ is a clique (resp. an independent set), which implies that $d_G(W_{i,s}, W_{i,t}) \geq 1 - \frac{1}{4m^2} \geq 1 - \frac{1}{2\ell^2}$ (resp. $d_G(W_{i,s}, W_{i,t}) \leq \frac{1}{4m^2} \leq \frac{1}{2\ell^2}$) for every $1 \leq s < t \leq m$ (see Item (b) above). Second, let $1 \leq i < j \leq p$. If (X_i, X_j) is a complete bipartite graph then $d_{G'}(U_i, U_j) = 1$ and hence $d_G(W_i, W_j) \geq 1 - \frac{1}{4m^2}$ (by Item (a) above). Now Item 2 in Lemma 2.3 implies that $d_G(W_{i,s}, W_{j,t}) \geq d_G(W_i, W_j) - \frac{1}{4m^2} \geq 1 - \frac{1}{2m^2} \geq 1 - \frac{1}{2\ell^2}$ for every $1 \leq s,t \leq m$. Similarly, if (X_i, X_j) is an empty bipartite graph then $d_{G'}(U_i, U_j) = 0$ and hence $d_G(W_i, W_j) \leq \frac{1}{4m^2}$. This implies that $d_G(W_{i,s}, W_{j,t}) \leq d_G(W_i, W_j) + \frac{1}{4m^2} \leq \frac{1}{2m^2} \leq \frac{1}{2\ell^2}$ for every $1 \leq s,t \leq m$

We now apply Lemma 2.4 to the graph F, the sets $W_{i,s}$ (where $1 \leq i \leq p$ and $1 \leq s \leq |X_i|$) and $\lambda = \zeta$ (noting that $|W_{i,s}| \geq \zeta n$ for every i, s, as guaranteed by Item 4 of Lemma 2.3). By Lemma 2.4, a sample of $\frac{12\ell}{\zeta} \leq \frac{12m}{\zeta} = \text{poly}(1/\varepsilon)$ vertices from G contains an induced copy of F (and hence does not satisfy $\mathcal{P}^*_{\mathcal{F}}$) with probability at least $\frac{2}{3}$. This completes the proof of the theorem. ■

Proof of Theorem 5. By Lemma 2.2, $\mathcal{P}^*_{\mathcal{F}}$ has a bipartite obstruction H. We may and will assume that both sides of H have the same size, h (as otherwise we can just add vertices to one of the sides). Let $\varepsilon < \frac{1}{2}$, and set

$$\gamma := \frac{1}{2} \cdot \rho_{2.3}\left(h, \frac{\varepsilon}{3}\right)^2,$$

and

$$\zeta := \zeta_{2.3}\left(h, 1, \frac{\varepsilon}{3}, \gamma\right).$$

Note that $\gamma = \mathrm{poly}(\varepsilon)$ and hence also $\zeta = \mathrm{poly}(\varepsilon)$.

Let G be an n-vertex graph which is ε-far from satisfying $\mathcal{P}^*_{\mathcal{F}}$. If G contains at least ζn^{2h} induced bipartite copies of H, then, just as in the proof of Theorem 1, a random sequence of $4h \cdot \zeta^{-1} = \mathrm{poly}(1/\varepsilon)$ vertices of G (sampled uniformly and independently) contains an induced bipartite copy of H, and hence does not satisfy $\mathcal{P}^*_{\mathcal{F}}$, with probability at least $\frac{2}{3}$. Thus, in this case the required result holds.

Suppose, then, that G contains less than ζn^{2h} induced bipartite copies of H. We apply Lemma 2.3 to G with parameters $\delta = \frac{\varepsilon}{3}$, γ defined as above and $m = 1$, to get an equipartition $\mathcal{U} = \{U_1, ..., U_r\}$ and sets $W_i \subseteq U_i$ with the properties stated in the lemma.

Define a graph F on $[r]$ as follows. For $1 \leq i < j \leq r$, if $d(W_i, W_j) \geq 1 - \gamma$ then $(i, j) \in E(F)$ and if $d(W_i, W_j) \leq \gamma$ then $(i, j) \notin E(F)$ (by Item 2 of Lemma 2.3, one of these options must hold). We will show that F does not satisfy $\mathcal{P}^*_{\mathcal{F}}$. Let us first complete the proof based on this fact. By Lemma 2.3 we have $v(F) = r \leq \rho_{2.3}(h, \frac{\varepsilon}{3})^{-1} = \mathrm{poly}(1/\varepsilon)$ and hence $\gamma \leq \frac{1}{2r^2}$. So by the definition of F, the sets $W_1, \ldots, W_r$ satisfy condition 1 of Lemma 2.4. By Item 4 of Lemma 2.3 we have $|W_i| \geq \zeta n$ for every $1 \leq i \leq r$. So Lemma 2.4 with $\lambda = \zeta$ implies that a sample of $12r/\zeta = \mathrm{poly}(1/\varepsilon)$ vertices of G, sampled uniformly at random and independently, contains an induced copy of F, and hence does not satisfy $\mathcal{P}^*_{\mathcal{F}}$, with probability at least $\frac{2}{3}$.

It thus remains to show that F does not satisfy $\mathcal{P}^*_{\mathcal{F}}$. Assume, by contradiction, that F satisfies $\mathcal{P}^*_{\mathcal{F}}$. Since $\mathcal{P}^*_{\mathcal{F}}$ is closed under blowups (recall Definition 1.3), there is a function $g : V(F) \to \{0, 1\}$ such that every g-blowup of F satisfies $\mathcal{P}^*_{\mathcal{F}}$. Now let G' be the graph obtained from G by making the following changes.

(a) For every $1 \leq i < j \leq r$, if $(i, j) \in E(F)$ then turn (U_i, U_j) into a complete bipartite graph and if $(i, j) \notin E(F)$ then turn (U_i, U_j) into an empty bipartite graph.

(b) For every $1 \leq i \leq r$, if $g(i) = 1$ then turn U_i into a clique and if $g(i) = 0$ then turn U_i into an independent set.

Since G' is a g-blowup of F (see Definition 1.2), G' satisfies $\mathcal{P}^*_{\mathcal{F}}$. We now show that the number of edge-changes made in Items (a)-(b) is less than εn^2, which will stand in contradiction to the fact that G is ε-far from satisfying $\mathcal{P}^*_{\mathcal{F}}$.

The definitions of F and G' imply the following: for every $1 \leq i < j \leq r$, if the bipartite graph (U_i, U_j) is complete (resp. empty) in G' then $d_G(W_i, W_j) \geq 1 - \gamma$ (resp. $d_G(W_i, W_j) \leq \gamma$). As in the proof of Theorem 1, let $\mathcal{H}$ be the set of pairs $1 \leq i < j \leq r$ such that (U_i, U_j) is $\frac{\varepsilon}{3}$-homogeneous (in G) and such that $|d_G(W_i, W_j) - d_G(U_i, U_j)| \leq \frac{1}{4}$. Observe that if $(i, j) \in \mathcal{H}$ then the number of edge-changes made in the bipartite graph (U_i, U_j) is at most $\frac{\varepsilon}{3}|U_i||U_j|$. Indeed, let $(i, j) \in \mathcal{H}$ and suppose first that (U_i, U_j) is a complete bipartite graph in G'. Then $d_G(W_i, W_j) \geq 1 - \gamma \geq \frac{3}{4}$, implying that $d_G(U_i, U_j) \geq d_G(W_i, W_j) - \frac{1}{4} \geq \frac{1}{2}$. Hence actually $d_G(U_i, U_j) \geq 1 - \frac{\varepsilon}{3}$. The case that (U_i, U_j) is an empty bipartite graph in G' is symmetrical.

By Item 1 in Lemma 2.3, there are at most $\frac{\varepsilon}{3}r^2$ pairs $1 \leq i < j \leq r$ which are not in $\mathcal{H}$. It follows that the overall number of edge-changes made in Item (a) is at most $|\mathcal{H}| \cdot \frac{\varepsilon}{3} \cdot \left(\frac{n}{r}\right)^2 + \frac{\varepsilon}{3}r^2 \cdot \left(\frac{n}{r}\right)^2 \leq \frac{2\varepsilon}{3}n^2$. As for item (b), the number of edge-changes made there is at most $r\binom{n/r}{2} < \frac{n^2}{r} \leq \frac{\varepsilon}{3}n^2$, where in the last inequality we used the fact that $r \geq \frac{3}{\varepsilon}$ (as guaranteed by Lemma 2.3). Thus, the overall number of edge-changes made in Items (a)-(b) is less than εn^2, as required. ∎

2.1 Proof of Lemma 2.3

In this subsection we prove Lemma 2.3. This lemma is proved using a "conditional regularity lemma" due to Alon, Fischer and Newman [5]. In order to state this result we first need some additional definitions. Let A be an $n \times n$ matrix with 0/1 entries whose rows and columns are indexed by $1, ..., n$. For two sets $R, C \subseteq [n]$, the *block* $R \times C$ is the submatrix of A whose rows are the elements of R and whose columns are the elements of C. The *density* of the block $R \times C$, denoted by $d(R \times C)$, is the fraction of 1's in the block. For $\delta \in (0,1)$, we say that $R \times C$ is δ*-homogeneous* if either $d(R \times C) \geq 1-\delta$ or $d(R \times C) \leq \delta$. The *weight* of $R \times C$ is $\frac{|R||C|}{n^2}$. A partition of A is a pair $(\mathcal{R}, \mathcal{C})$, where $\mathcal{R}$ and $\mathcal{C}$ are partitions of $[n]$. We think of $\mathcal{R}$ as a partition of the rows of A, and of $\mathcal{C}$ as a partition of the columns of A. We say that $(\mathcal{R}, \mathcal{C})$ is δ*-homogeneous* if the sum of weights of non-δ-homogeneous blocks $R \times C$, where $R \in \mathcal{R}$ and $C \in \mathcal{C}$, is at most δ. In the case that A is the adjacency matrix of a graph G, these definitions are analogous to the definitions given in the beginning of Section 2. Indeed, every pair of disjoint sets $X, Y \subseteq V(G) = [n]$ satisfies $d(X,Y) = d(X \times Y)$ (where the quantity on the left-hand side is the edge density in the bipartite graph with sides X, Y, and the quantity on the right-hand side is the density of the block $X \times Y$ in A). Moreover, if $\mathcal{P}$ is a partition of $[n]$ such that $(\mathcal{P}, \mathcal{P})$ is a δ-homogeneous partition of A, then[6] $\mathcal{P}$ is a δ-homogeneous partition of G.

A partition $(\mathcal{R}', \mathcal{C}')$ is a *refinement* of a partition $(\mathcal{R}, \mathcal{C})$ if every block of $\mathcal{R}' \times \mathcal{C}'$ is contained in some block of $\mathcal{R} \times \mathcal{C}$. We will need the following two lemmas.

Lemma 2.5. *Let $\delta \in (0,1)$. If $(\mathcal{R}, \mathcal{C})$ is a $\frac{\delta^2}{2}$-homogeneous partition of an $n \times n$ matrix A, then every refinement of $(\mathcal{R}, \mathcal{C})$ is a δ-homogeneous partition of A.*

Proof. Let $(\mathcal{R}', \mathcal{C}')$ be a refinement of $(\mathcal{R}, \mathcal{C})$. Let $\mathcal{N}$ be the set of non-δ-homogeneous blocks of $(\mathcal{R}', \mathcal{C}')$. Our goal is to show that the sum of weights of blocks $R' \times C' \in \mathcal{N}$ is at most δ. Let $\mathcal{N}_1$ (resp. $\mathcal{N}_2$) be the set of blocks $R' \times C' \in \mathcal{N}$ that are contained in a $\frac{\delta^2}{2}$-homogeneous (resp. non-$\frac{\delta^2}{2}$-homogeneous) block of $(\mathcal{R}, \mathcal{C})$. Since $(\mathcal{R}, \mathcal{C})$ is a $\frac{\delta^2}{2}$-homogeneous partition, the sum of weights of blocks $R' \times C' \in \mathcal{N}_2$ is at most $\frac{\delta^2}{2}$. Since $\mathcal{N} = \mathcal{N}_1 \cup \mathcal{N}_2$ and $\frac{\delta}{2} + \frac{\delta^2}{2} \leq \delta$, it is enough to show that the sum of weights of blocks $R' \times C' \in \mathcal{N}_1$ is at most $\frac{\delta}{2}$.

Let $R \times C$ be a $\frac{\delta^2}{2}$-homogeneous block of $(\mathcal{R}, \mathcal{C})$ and suppose without loss of generality that $d(R \times C) \leq \frac{\delta^2}{2}$ (the case that $d(R \times C) \geq 1 - \frac{\delta^2}{2}$ is symmetrical). Let $R'_1, \dots, R'_k$ (resp. $C'_1, \dots, C'_\ell$) be the parts of $\mathcal{R}'$ (resp. $\mathcal{C}'$) which are contained in R (resp. C). By averaging we have

$$d(R \times C) = \sum_{i=1}^{k} \sum_{j=1}^{\ell} \frac{|R'_i||C'_j|}{|R||C|} \cdot d(R'_i \times C'_j).$$

By Markov's inequality, the total weight of blocks $R'_i \times C'_j$ for which $d(R'_i, C'_j) > \delta$ is less than $\frac{\delta}{2} \cdot \frac{|R||C|}{n^2}$. In conclusion, for every $\frac{\delta^2}{2}$-homogeneous block $R \times C$ of $(\mathcal{R}, \mathcal{C})$ it holds that the total weight of blocks $R' \times C' \in \mathcal{N}_1$ contained in $R \times C$ is less than $\frac{\delta}{2} \cdot \frac{|R||C|}{n^2}$. By summing over all $\frac{\delta^2}{2}$-homogeneous blocks of $(\mathcal{R}, \mathcal{C})$ we get that the total weight of blocks $R' \times C' \in \mathcal{N}_1$ is less than $\frac{\delta}{2}$, as required. ■

Lemma 2.6. *Let A be an $n \times n$ matrix, let $\delta \in (0,1)$, let $\mathcal{P}_0$ be an equipartition of $[n]$, and let $(\mathcal{R}, \mathcal{C})$ be a $\frac{\delta^2}{8}$-homogeneous partition of A. Then there is an equipartition $\mathcal{U}$ of $[n]$ such that $(\mathcal{U}, \mathcal{U})$ is a δ-homogeneous partition of A, and such that $\mathcal{U}$ refines $\mathcal{P}_0$ and has $r := \lceil 4/\delta \rceil \cdot |\mathcal{P}_0| \cdot |\mathcal{R}| \cdot |\mathcal{C}|$ parts.*

[6] The other direction is not necessarily true, because the definition of a δ-homogeneous partition of a matrix takes into account the "diagonal" blocks $X \times X$, while the definition of a δ-homogeneous partition of a graph does not.

Proof. Let $\mathcal{S}$ be the common refinement of $\mathcal{R}$, $\mathcal{C}$ and $\mathcal{P}_0$, i.e. $\mathcal{S} = \{R \cap C \cap P : R \in \mathcal{R}, C \in \mathcal{C}, P \in \mathcal{P}_0\}$. Partition every $S \in \mathcal{S}$ into equal parts of size $\frac{n}{r}$ and an additional part of size less than $\frac{n}{r}$. Denote the resulting partition by $\mathcal{T}$. For each $P \in \mathcal{P}_0$, let Z_P be the union of all additional parts contained in P, and note that $|Z_P| < |\mathcal{R}| \cdot |\mathcal{C}| \cdot \frac{n}{r}$. Set $Z = \bigcup_{P \in \mathcal{P}_0} Z_P$, noting that $|Z| \leq |\mathcal{P}_0| \cdot |\mathcal{R}| \cdot |\mathcal{C}| \cdot \frac{n}{r} \leq \frac{\delta n}{4}$. As $(\mathcal{T}, \mathcal{T})$ is a refinement of $(\mathcal{R}, \mathcal{C})$ and $(\mathcal{R}, \mathcal{C})$ is a $\frac{\delta^2}{8}$-homogeneous partition of A, Lemma 2.5 (with $\frac{\delta}{2}$ in place of δ) implies that $(\mathcal{T}, \mathcal{T})$ is a $\frac{\delta}{2}$-homogeneous partition of A.

Let $\mathcal{U}$ be the equipartition obtained from $\mathcal{T}$ by partitioning each of the sets Z_P $(P \in \mathcal{P}_0)$ into parts of size $\frac{n}{r}$. It is clear that $\mathcal{U}$ refines $\mathcal{P}_0$ and has r parts. We claim that $(\mathcal{U}, \mathcal{U})$ is a δ-homogeneous partition of A. Observe that if $X \times Y$ is a non-δ-homogeneous block of $(\mathcal{U}, \mathcal{U})$, then either $X \times Y$ is a non-δ-homogeneous block of $(\mathcal{T}, \mathcal{T})$, or one of the sets X, Y is contained in Z. Since $|Z| \leq \frac{\delta n}{4}$, the sum of weights of blocks $X \times Y$ for which X or Y is contained in Z is at most $\frac{2|Z|n}{n^2} \leq \frac{\delta}{2}$. Combining this with the fact that $(\mathcal{T}, \mathcal{T})$ is $\frac{\delta}{2}$-homogeneous, we get that $(\mathcal{U}, \mathcal{U})$ is δ-homogeneous, as required. ■

Let B be a $0/1$-valued $h \times h$ matrix. A *copy* of B in a matrix A is a sequence of rows $r_1 < \cdots < r_h$ and a sequence of columns $c_1 < \cdots < c_h$ such that $A_{r_i, c_j} = B_{i,j}$ for every $1 \leq i, j \leq h$. We are now ready to state the Alon-Fischer-Newman Regularity Lemma.

Lemma 2.7 (Alon-Fischer-Newman [5])**.** *There is a constant c_0 such that the following holds for every integer $h \geq 1$ and $\delta \in (0,1)$. For every $0/1$-valued matrix A of size $n \times n$ with $n > (h/\delta)^{c_0 h}$, either A has a δ-homogeneous partition $(\mathcal{R}, \mathcal{C})$ with $|\mathcal{R}|, |\mathcal{C}| \leq (h/\delta)^{c_0 h}$, or for every $0/1$-valued $h \times h$ matrix B, A contains at least $(\delta/h)^{c_0 h^2} n^{2h}$ copies of B.*

The next lemma is an application of Lemma 2.7 to adjacency matrices of graphs. We assume that the vertex set of the graph G is $[n]$.

Lemma 2.8. *There is a function $\rho_{2.8} : \mathbb{N} \times (0,1) \to (0,1)$ such that $\rho_{2.8}(h, \delta) = \text{poly}(\delta)$, and such that for every integer $h \geq 1$, for every $h \times h$ bipartite graph $H = (S \cup T, E)$ and for every $\delta \in (0,1)$, the following holds: let G be a graph on $n \geq n_0(h, \delta) = \text{poly}(1/\delta)$ vertices and let $\mathcal{P}_0$ be an equipartition of $V(G) = [n]$. Then G either contains at least $\rho_{2.8}(h, \delta) n^{2h}$ induced bipartite copies of H or admits a δ-homogeneous equipartition $\mathcal{U}$ which refines $\mathcal{P}_0$ and has at most $|\mathcal{P}_0| \cdot \rho_{2.8}(h, \delta)^{-1}$ parts.*

Proof. We prove the lemma with $\rho = \rho_{2.8}(h, \delta) := (\frac{\delta^2}{8h})^{3c_0 h^2}$ (where c_0 is from Lemma 2.7). Let $A = A(G)$ be the adjacency matrix of G. Let B be the *bipartite adjacency matrix* of H; that is, B is an $h \times h$ matrix, indexed by $S \times T$, in which $B_{s,t} = 1$ if $(s,t) \in E(H)$ and $B_{s,t} = 0$ otherwise. Suppose first that A contains at least $(\frac{\delta^2}{8h})^{c_0 h^2} n^{2h}$ copies of B. Observe that a copy of B which does not intersect the main diagonal of A corresponds to an induced bipartite copy of H in G. The number of $h \times h$ submatrices of A which intersect its main diagonal is $O(n^{2h-1})$. Thus, G contains at least $(\frac{\delta^2}{8h})^{c_0 h^2} n^{2h} - O(n^{2h-1}) \geq (\frac{\delta^2}{8h})^{2c_0 h^2} n^{2h} \geq \rho n^{2h}$ induced bipartite copies of H, as required.

Now suppose that A contains less than $(\frac{\delta^2}{8h})^{c_0 h^2} n^{2h}$ copies of B. By Lemma 2.7, applied with approximation parameter $\frac{\delta^2}{8}$, A admits a $\frac{\delta^2}{8}$-homogeneous partition $(\mathcal{R}, \mathcal{C})$ with $|\mathcal{R}|, |\mathcal{C}| \leq (\frac{8h}{\delta^2})^{c_0 h}$. By Lemma 2.6, there is an equipartition $\mathcal{U}$ of $[n]$ which refines $\mathcal{P}_0$, has

$$\lceil 4/\delta \rceil \cdot |\mathcal{P}_0| \cdot |\mathcal{R}| \cdot |\mathcal{C}| \leq 8\delta^{-1} |\mathcal{P}_0| \cdot \left(\frac{8h}{\delta^2}\right)^{2c_0 h} \leq |\mathcal{P}_0| \cdot \left(\frac{8h}{\delta^2}\right)^{3c_0 h} \leq |\mathcal{P}_0| \cdot \rho^{-1}$$

parts, and satisfies that $(\mathcal{U}, \mathcal{U})$ is a δ-homogeneous partition of A. This implies that $\mathcal{U}$ is a δ-homogeneous partition of G. The lemma follows. ■

The following lemma is a "conditional" variant of a well-known corollary of Szemerédi's regularity lemma (see e.g. [4]).

Lemma 2.9. *There is a function $\zeta_{2.9} : \mathbb{N}^2 \times (0,1) \to (0,1)$ such that $\zeta_{2.9}(h,m,\gamma) = \mathrm{poly}(\gamma)$, and such that the following holds for every $h, m \geq 1$, for every $h \times h$ bipartite graph H, and for every $\gamma \in (0,1)$. Every graph G on $n \geq n_0(h,m,\gamma) = \mathrm{poly}(1/\gamma)$ vertices either contains at least $\zeta_{2.9}(h,m,\gamma)n^{2h}$ induced bipartite copies of H or there are pairwise-disjoint subsets $W_1, \dots, W_m \subseteq V(G)$ with the following properties:*

1. *Either $d(W_i, W_j) \geq 1 - \gamma$ for every $1 \leq i < j \leq m$, or $d(W_i, W_j) \leq \gamma$ for every $1 \leq i < j \leq m$.*
2. *$|W_i| \geq n \cdot \zeta_{2.9}(h,m,\gamma)$ for every $1 \leq i \leq m$.*

Proof. Set $\delta := \min\{4^{-m-1}, \gamma\}$. We prove the lemma with $\zeta = \zeta_{2.9}(h,m,\gamma) := 4^{-m-1} \cdot \rho_{2.8}(h,\delta)$, where $\rho_{2.8}$ is from Lemma 2.8. We assume that G contains less than ζn^{2h} (and hence also less than $\rho_{2.8}(h,\delta)n^{2h}$) induced bipartite copies of H and prove that the other alternative in the lemma holds. Let $\mathcal{P}_0$ be an arbitrary equipartition of $V(G)$ into 4^{m+1} parts. Apply Lemma 2.8 with δ as defined above, to obtain a δ-homogeneous equipartition $\mathcal{W}$ of G which refines $\mathcal{P}_0$ and has at most $|\mathcal{P}_0| \cdot \rho_{2.8}(h,\delta)^{-1} = 4^{m+1} \cdot \rho_{2.8}(h,\delta)^{-1} = \zeta^{-1}$ parts. Then every $W \in \mathcal{W}$ satisfies $|W| \geq \zeta n$.

Set $w := |\mathcal{W}|$, noting that $w \geq 4^{m+1}$ because $\mathcal{W}$ refines $\mathcal{P}_0$. Define an auxiliary graph J on the set $\mathcal{W}$ in which (W, W') is an edge if and only if the pair (W, W') is δ-homogeneous. Since $\mathcal{W}$ is a δ-homogeneous partition, we have

$$e(J) \geq \binom{w}{2} - \delta w^2 \geq \binom{w}{2} - 4^{-m-1}w^2 > \left(1 - \frac{1}{4^m - 1}\right)\frac{w^2}{2}.$$

By Turán's Theorem (see e.g. [14]), there is a subset $\mathcal{W}' \subseteq \mathcal{W}$ of size $|\mathcal{W}'| = 4^m$ which spans a clique in J. Then for every $W, W' \in \mathcal{W}'$, the pair (W, W') is δ-homogeneous and hence also γ-homogeneous. Define a new graph on $\mathcal{W}'$ as follows: for $W, W' \in \mathcal{W}'$, put an edge between W and W' if and only if $d(W, W') \geq 1 - \gamma$ (the other option being that $d(W, W') \leq \gamma$). By Ramsey's theorem (see Claim 2.1), this graph contains a homogeneous set of size m, which we denote by $\{W_1, \dots, W_m\}$. Then we have either $d(W_i, W_j) \geq 1 - \gamma$ for every $1 \leq i < j \leq m$, or $d(W_i, W_j) \leq \gamma$ for every $1 \leq i < j \leq m$, depending on whether $\{W_1, \dots, W_m\}$ is a clique or an independent set. This completes the proof. $\blacksquare$

We are now (almost) ready to prove Lemma 2.3. The last ingredient we will need is the following simple claim, whose proof is straightforward and thus omitted.

Claim 2.10. *Let $\gamma, \eta \in (0,1)$, let X, Y be disjoint vertex-sets and let $X' \subseteq X$, $Y' \subseteq Y$ be such that $|X'| \geq (\eta/\gamma)^{1/2}|X|$ and $|Y'| \geq (\eta/\gamma)^{1/2}|Y|$. If (X, Y) is η-homogeneous then $|d(X', Y') - d(X, Y)| \leq \gamma$.*

Proof of Lemma 2.3. Put

$$\rho := \frac{\delta}{2} \cdot \rho_{2.8}\left(h, \frac{\delta}{5}\right),$$

$$\eta := \min\left\{\rho^4,\ \gamma \cdot \zeta_{2.9}(h,m,\gamma)^2\right\},$$

$$\rho_1 := \rho \cdot \rho_{2.8}(h, \eta),$$

$$\zeta := \min\left\{\rho,\ \ \zeta_{2.9}(h,m,\gamma) \cdot \rho_1^{2h},\ \ (\eta/\gamma)^{1/2} \cdot \rho_1\right\},$$

where $\rho_{2.8}$ is from Lemma 2.8 and $\zeta_{2.9}$ is from Lemma 2.9. We prove the lemma with $\rho_{2.3}(h, \delta) := \rho$ and $\zeta_{2.3}(h, m, \delta, \gamma) := \zeta$. It is easy to check (using the guarantees of Lemmas 2.8 and 2.9) that ρ is polynomial in δ, and that ζ is polynomial in δ and γ, as required.

We assume that G contains less than ζn^{2h} induced bipartite copies of H and prove that the other alternative in the statement of the lemma holds. Since $\zeta \leq \rho \leq \rho_{2.8}(h, \frac{\delta}{5})$, G contains less than $\rho_{2.8}(h, \frac{\delta}{5})n^{2h}$ induced bipartite copies of H. Let $\mathcal{P}_0$ be an arbitrary equipartition of $V(G)$ into $\lceil 1/\delta \rceil$ parts. By Lemma 2.8 with approximation parameter $\frac{\delta}{5}$, there is a $\frac{\delta}{5}$-homogeneous equipartition $\mathcal{U} = \{U_1, \dots, U_r\}$ of G which refines $\mathcal{P}_0$, and satisfies

$$|\mathcal{U}| = r \leq |\mathcal{P}_0| \cdot \rho_{2.8}\left(h, \frac{\delta}{5}\right)^{-1} \leq \frac{2}{\delta} \cdot \rho_{2.8}\left(h, \frac{\delta}{5}\right)^{-1} = \rho^{-1} = \rho_{2.3}(h, \delta)^{-1}.$$

Note also that $r \geq \delta^{-1}$, as $\mathcal{U}$ is a refinement of $\mathcal{P}_0$.

Since $\zeta \leq \rho_1 \leq \rho_{2.8}(h, \eta)$, our assumption in the beginning of the proof implies that G contains less than $\rho_{2.8}(h, \eta)n^{2h}$ induced bipartite copies of H. Thus, by Lemma 2.8 with approximation parameter η and with $\mathcal{P}_0 = \mathcal{U}$, G admits an η-homogeneous equipartition $\mathcal{W}$ that refines $\mathcal{U}$ and has at most $|\mathcal{W}| \leq |\mathcal{U}| \cdot \rho_{2.8}(h, \eta)^{-1} \leq \rho^{-1} \cdot \rho_{2.8}(h, \eta)^{-1} = \rho_1^{-1}$ parts. Hence, for every $W \in \mathcal{W}$ we have $|W| \geq \rho_1 n$.

For each $1 \leq i \leq r$ define $\mathcal{W}_i = \{W \in \mathcal{W} : W \subseteq U_i\}$. Sample a part $W_i \in \mathcal{W}_i$ uniformly at random. Let $\mathcal{A}_1$ be the event that all pairs (W_i, W_j) are η-homogeneous. By using the fact that $\mathcal{W}$ is η-homogeneous, we get that for every $1 \leq i < j \leq r$, the probability that (W_i, W_j) is not η-homogeneous is at most

$$\frac{\eta|\mathcal{W}|^2}{(|\mathcal{W}|/|\mathcal{U}|)^2} = \eta|\mathcal{U}|^2 = \eta r^2 \leq \eta\rho^{-2} \leq \rho^2.$$

Thus, by the union bound over all $\binom{r}{2} < \frac{1}{2\rho^2}$ pairs $1 \leq i < j \leq r$, we get that $\mathbb{P}[\mathcal{A}_1] > \frac{1}{2}$.

A pair $1 \leq i < j \leq r$ is called *good* if (U_i, U_j) is $\frac{\delta}{5}$-homogeneous and $|d(W_i, W_j) - d(U_i, U_j)| \leq \frac{1}{4}$; otherwise (i, j) is called *bad*. Let $\mathcal{A}_2$ be the event that there are at most δr^2 bad pairs. Note that if $\mathcal{A}_2$ happened then Item 1 of the lemma is satisfied. We claim that $\mathbb{P}[\mathcal{A}_2] > \frac{1}{2}$. To this end, note that if (U_i, U_j) is $\frac{\delta}{5}$-homogeneous and $|d(W_i, W_j) - d(U_i, U_j)| > \frac{1}{4}$, then either $d(U_i, U_j) \geq 1 - \frac{\delta}{5}$ and $d(W_i, W_j) < \frac{3}{4}$, or $d(U_i, U_j) \leq \frac{\delta}{5}$ and $d(W_i, W_j) > \frac{1}{4}$; in either case, the probability that this happens is less than $\frac{\delta/5}{1/4} = \frac{4\delta}{5}$. It follows that the expected number of pairs $1 \leq i < j \leq r$ for which (U_i, U_j) is $\frac{\delta}{5}$-homogeneous but $|d(W_i, W_j) - d(U_i, U_j)| > \frac{1}{4}$, is less than $\frac{4\delta}{5}\binom{r}{2} < \frac{2\delta}{5}r^2$. By Markov's inequality, the probability that there are more than $\frac{4\delta}{5}r^2$ such pairs is smaller than $\frac{1}{2}$. Now, since all but at most $\frac{\delta}{5}r^2$ of the pairs (U_i, U_j) are $\frac{\delta}{5}$-homogeneous, our assertion that $\mathbb{P}[\mathcal{A}_2] > \frac{1}{2}$ follows. Thus, we proved that $\mathbb{P}[\mathcal{A}_i] > \frac{1}{2}$ for both $i = 1, 2$. This implies that $\mathbb{P}[\mathcal{A}_1 \cap \mathcal{A}_2] > 0$. From now on we fix a choice of $W_1, \dots, W_r$ for which both $\mathcal{A}_1$ and $\mathcal{A}_2$ happened.

Let $1 \leq i \leq r$, and observe that $G[W_i]$ contains less than $\zeta_{2.9}(h, m, \gamma) \cdot |W_i|^{2h}$ induced bipartite copies of H. Indeed, this follows from the fact that $|W_i| \geq \rho_1 n$, the fact that $\zeta \leq \zeta_{2.9}(h, m, \gamma) \cdot \rho_1^{2h}$, and our assumption that G contains less than ζn^{2h} induced bipartite copies of H. So by Lemma 2.9, applied to the graph $G[W_i]$, there are pairwise-disjoint sets $W_{i,1}, \dots, W_{i,m} \subseteq W_i$ such that

$$|W_{i,s}| \geq \zeta_{2.9}(h, m, \gamma) \cdot |W_i| \geq (\eta/\gamma)^{1/2} \cdot |W_i| \geq (\eta/\gamma)^{1/2} \cdot \rho_1 n \geq \zeta n$$

for every $1 \leq s \leq m$ (where in the second inequality we used our choice of η), and such that either $d(W_{i,s}, W_{i,t}) \geq 1 - \gamma$ for every $1 \leq s < t \leq m$ or $d(W_{i,s}, W_{i,t}) \leq \gamma$ for every $1 \leq s < t \leq m$. This establishes Item 3-4 of the lemma. Item 1 is guaranteed by our choice of $W_1, \dots, W_r$ (i.e. by the assumption that $\mathcal{A}_2$ happened). It thus remains to establish Item 2. The fact that all pairs (W_i, W_j) are γ-homogeneous follows from our assumption that $\mathcal{A}_1$ happened and the fact that $\eta \leq \gamma$. Now let $1 \leq i < j \leq r$ and $1 \leq s, t \leq m$. Note that (W_i, W_j) is η-homogeneous (as $\mathcal{A}_1$ happened), and that $|W_{i,s}| \geq (\eta/\gamma)^{1/2} \cdot |W_i|$ and $|W_{j,t}| \geq (\eta/\gamma)^{1/2} \cdot |W_j|$. So by Claim 2.10 with $X = W_i$, $Y = W_j$, $X' = W_{i,s}$ and $Y' = W_{j,t}$, we have $|d(W_{i,s}, W_{j,t}) - d(W_i, W_j)| \leq \gamma$, as required. ∎

2.2 Detailed Proof of Theorem 6

Let $\mathcal{P}$ be a semi-algebraic graph property defined by polynomials $f_1, \ldots, f_t \in \mathbb{R}[x_1, \ldots, x_{2k}]$ and a boolean function $\Phi : \{\text{true}, \text{false}\}^t \to \{\text{true}, \text{false}\}$. Let $\mathcal{F}$ be the family of all graphs which do not satisfy $\mathcal{P}$. As $\mathcal{P}$ is a hereditary property, we have $\mathcal{P} = \mathcal{P}^*_{\mathcal{F}}$. To prove the theorem, we only need to show that Conditions 1-2 of Theorem 5 are satisfied.

We start with Condition 1. The *VC-dimension* of a binary matrix A is the maximal integer $d \geq 0$ for which there is a $d \times 2^d$ submatrix B of A, such that the set of columns of B is the set of all 2^d binary vectors of length d. The VC-dimension of a graph is defined as the VC-dimension of its adjacency matrix. It is known[7] that for every semi-algebraic graph property $\mathcal{P}$ there is[8] $d = d(\mathcal{P})$ such that every graph which satisfies $\mathcal{P}$ has VC-dimension strictly less than d. Now let B be a $d \times 2^d$ binary matrix whose columns are all 2^d binary vectors of length d. Let H be the bipartite graph with sides $X = \{x_1, \ldots, x_d\}$ and $Y = \{y_1, \ldots, y_{2^d}\}$ such that $(x_i, y_j) \in E(H)$ if and only if $B_{i,j} = 1$. It is easy to see that no matter which graphs one puts on X and on Y (without changing the edges between X and Y), the resulting graph on $X \cup Y$ will not satisfy $\mathcal{P}$ since its VC-dimension will be at least $d = d(\mathcal{P})$. This means that H is a bipartite obstruction for $\mathcal{P}$, which implies (via Lemma 2.2) that $\mathcal{F}$ contains a bipartite graph, a co-bipartite graph and a split graph, as required.

As for Condition 2, let F be a graph on $V(F) = [p]$ which satisfies $\mathcal{P}$, and let $x_1, \ldots, x_p \in \mathbb{R}^k$ be witnesses to the fact that F satisfies $\mathcal{P}$. That is, for every $1 \leq i \neq j \leq p$ we have $(i, j) \in E(F)$ if and only if $\Phi\Big(f_1(x_i, x_j) \geq 0; \ldots; f_t(x_i, x_j) \geq 0\Big) = \text{true}$. We define a function $g : V(F) \to \{0, 1\}$ as follows: $g(i) = 1$ if

$$\Phi\Big(f_1(x_i, x_i) \geq 0; \ldots; f_t(x_i, x_i) \geq 0\Big) = \text{true}$$

and $g(i) = 0$ otherwise. We now show that every g-blowup of F satisfies $\mathcal{P}$. Let G be a g-blowup of F with a vertex partition $V(G) = P_1 \cup \cdots \cup P_p$ (as in Definition 1.2). Then for every $1 \leq i \leq p$, we simply assign the point x_i to every vertex of P_i. From the definition of a g-blowup and from our choice of g, it follows that for every $1 \leq i, j \leq p$ and for every pair of distinct vertices $v_i \in P_i$, $v_j \in P_j$ we have that $(v_i, v_j) \in E(G)$ if and only if $\Phi\Big(f_1(x_i, x_j) \geq 0; \ldots; f_t(x_i, x_j) \geq 0\Big) = \text{true}$. Thus we have shown that $\mathcal{P}$ is closed under blowups, completing the deduction of Theorem 6 from Theorem 5.

3 Hard to Test Properties

This section is organized as follows. In Subsection 3.1 we describe a variant of the well-known Ruzsa-Szemerédi construction which we use in the proofs of Theorems 2, 3 and 4. We then prove Theorem 3 in Subsection 3.2. In Subsection 3.3 we introduce some definitions that are needed in order to handle graph families (and not just individual graphs), leading to the proof of Theorem 2 in Subsection 3.4. The main step in the proof of Theorem 2 is Theorem 8 (see Subsection 3.4), which also implies Theorem 4. Finally, in Subsection 3.5 we prove Theorem 7.

[7] This follows from Warren's theorem on sign patterns of systems of polynomials, see for example [1].

[8] In fact, d can be bounded from above by a function of k,t, and the degrees of the polynomials $f_1, \ldots, f_t$.

In some of the proofs we use the following simple claim.

Claim 3.1. *For every pair of integers $m \geq 1$ and $h \geq 2$ there is a collection $\mathcal{S} \subseteq [m]^h$ of size at least m^2/h^2 such that every two h-tuples in $\mathcal{S}$ have at most one identical entry.*

Proof. We construct the collection $\mathcal{S}$ greedily: we start with an empty set, add an arbitrary element of $[m]^h$ to it, discard all h-tuples that coincide in more than one entry with the h-tuple we added and repeat. At the beginning, all m^h of the h-tuples in $[m]^h$ are potential elements of $\mathcal{S}$. At each step we discard at most $\binom{h}{2}m^{h-2}$ tuples. Therefore, at the end of the process we have a collection of size at least

$$\frac{m^h}{1+\binom{h}{2}m^{h-2}} \geq \frac{m^h}{h^2 m^{h-2}} = \frac{m^2}{h^2},$$

as required. ■

3.1 The Construction of the Graph R

We start with the following lemma, which plays a key role in our constructions.

Lemma 3.2. *For every $k \geq 2$ there is $\alpha = \alpha(k)$ such that for every integer m there is a set $S \subseteq [m]$, $|S| \geq \frac{m}{e^{\alpha\sqrt{\log m}}}$, with the following property: Let $2 \leq \ell \leq k$ and let $a_1, ..., a_\ell \geq 1$ be integers satisfying $a_1 + \cdots + a_\ell \leq k$. Then the only solutions to the equation*

$$a_1 s_1 + a_2 s_2 + \cdots + a_\ell s_\ell = (a_1 + \cdots + a_\ell)s_{\ell+1}$$

with $s_1, ..., s_{\ell+1} \in S$ are the trivial ones, i.e. $s_1 = s_2 = \cdots = s_\ell = s_{\ell+1}$.

Lemma 3.2 is a variant of Behrend's construction [13] of a large subset of $[m]$ without a 3-term arithmetic progression (note that the case $k = \ell = 2$ and $a_1 = a_2 = 1$ exactly corresponds to a 3-term arithmetic progression). It is easy to show (see e.g. [37] and [2]) that the same exact proof actually works for any fixed convex equation, and that moreover, it works "simultaneously" for all convex equations (for fixed k), thus establishing the above lemma. We therefore omit its proof.

The following lemma is our variant of the Ruzsa-Szemerédi construction (see [38]), and is the key ingredient in the proofs of Theorems 2, 3 and 4.

Lemma 3.3. *For every $h \geq 3$ there are $\delta_0 = \delta_0(h)$ and $\beta = \beta(h)$ such that for every $\delta < \delta_0$ there is a graph $R = R(h, \delta)$ with a vertex-partition $V(R) = V_1 \uplus \cdots \uplus V_h$, such that the following holds.*

1. *$|V(R)| \geq (1/\delta)^{\beta \log(1/\delta)}$.*
2. *$E(R)$ is the union of at least $\delta|V(R)|^2$ pairwise edge-disjoint h-cliques, each of the form $\{v_1, \ldots, v_h\}$ with $v_i \in V_i$ $(1 \leq i \leq h)$.*
3. *For every $3 \leq t \leq h$ and for every sequence $1 \leq i_1 < i_2 \cdots < i_t \leq h$, R contains at most $|V(R)|^2$ (not necessarily induced) cycles of the form $v_{i_1}v_{i_2}\ldots v_{i_t}v_{i_1}$ with $v_{i_j} \in V_{i_j}$ $(1 \leq j \leq t)$.*

Proof. Let $0 < \delta < \delta_0$ (for $\delta_0 = \delta_0(h)$ to be chosen later), and let m be the largest integer satisfying

$$\delta \leq \frac{1}{(h+1)^4 e^{\alpha\sqrt{\log m}}} \tag{1}$$

where $\alpha = \alpha(h-1)$ is from Lemma 3.2. It is easy to check that

$$m \geq e^{\alpha^{-2}\log^2\left(\frac{1}{\delta(h+1)^4}\right)} \geq (1/\delta)^{\beta\log(1/\delta)}, \tag{2}$$

where the second inequality holds provided that we choose $\beta = \beta(h)$ to be small enough, and provided that δ is sufficiently small with respect to h (we choose δ_0 accordingly).

Let $S \subseteq [m]$ be the set obtained by applying Lemma 3.2 with $k = h-1$. For each $j = 1,\dots,h$ set $V_j = \{1,2,...,jm\}$. With a slight abuse of notation, we think of $V_1,...,V_h$ as disjoint sets. The vertex-set of R is $V(R) = V_1 \uplus \cdots \uplus V_h$. By (2) we have $|V(R)| = \binom{h+1}{2}m \geq (1/\delta)^{\beta\log(1/\delta)}$, as required.

We now specify the edges of R. For every $x \in [m]$ and $s \in S$, set

$$A(x,s) := \{x, x+s, x+2s, \dots, x+(h-1)s\},$$

and put a clique on $A(x,s)$, in which $x+(j-1)s$ is taken from V_j for every $j = 1,\dots,h$. Note that for every $(x,s),(x',s') \in [m]\times S$, if $(x,s) \neq (x',s')$ then $|A(x,s)\cap A(x',s')| \leq 1$. Indeed, if $|A(x,s)\cap A(x',s')| \geq 2$ then there are $0 \leq i < j \leq h-1$ for which $x+is = x'+is'$ and $x+js = x'+js'$. Solving this system of equations yields $(x,s) = (x',s')$, as required. So the cliques that we defined are edge-disjoint, as required in Item 2 of the lemma. By the lower bound on $|S|$ in Lemma 3.2, and by (1), the number of these cliques is

$$m\cdot|S| \geq \frac{m^2}{e^{\alpha\sqrt{\log m}}} \geq \delta(h+1)^4 m^2 \geq \delta|V(R)|^2 .$$

To finish the proof, it remains to establish Item 3. Fix any $t \geq 3$ and any sequence of indices $1 \leq i_1 < i_2 \cdots < i_t \leq h$. We will show that for every cycle of the form $v_{i_1}v_{i_2}\dots v_{i_t}v_{i_1}$ with $v_{i_j} \in V_{i_j}$ there are $x \in [m]$ and $s \in S$ such that $v_{i_1}, v_{i_2}, \dots, v_{i_t} \in A(x,s)$. This will show that the cycles of this form are pair-disjoint[9], implying that there are at most $|V(R)|^2$ such cycles.

Let $v_{i_1}v_{i_2}\dots v_{i_t}v_{i_1}$ be a cycle in R with $v_{i_j} \in V_{i_j}$ for every $1 \leq j \leq t$. By the construction of R, for every $j = 1,\dots,t$ there is $(x_j,s_j) \in [m]\times S$ such that $\{v_{i_j}, v_{i_{j+1}}\} \subseteq A(x_j,s_j)$, with indices taken modulo t. This means that

$$v_{i_{j+1}} - v_{i_j} = (i_{j+1}-i_j)s_j \tag{3}$$

for every $1 \leq j \leq t-1$, and also that $v_{i_t} - v_{i_1} = (i_t - i_1)s_t$. Setting $a_j := i_{j+1} - i_j$ for $1 \leq j \leq t-1$, we see that

$$a_1s_1 + a_2s_2 + \cdots + a_{t-1}s_{t-1} = (a_1 + \cdots + a_{t-1})s_t.$$

Since S was chosen via Lemma 3.2, we must have $s_1 = s_2 = \cdots = s_t$. Hence, by (3) we have that $v_{i_j} = v_{i_1} + (i_j - i_1)s_1$ for every $1 \leq j \leq t$. By the definition of $A(x_1,s_1)$ and by the fact that $v_{i_1} \in A(x_1,s_1)$, we get that $v_{i_1}, v_{i_2}, \dots, v_{i_t} \in A(x_1,s_1) = \cdots = A(x_t,s_t)$, as required. ■

3.2 Proof of Theorem 3

Let M be the *complement* of the 7-vertex graph with vertex-set $\{1,2,3,4,5,6,7\}$ and edge-set $\{\{1,2\},\{3,4\},\{5,6\}\}$. It is easy to see that M is co-bipartite. We will prove Theorem 3 with $F_1 = C_8$ (the cycle on 8 vertices) and $F_2 = M$. We need the following lemma, which we prove later.

Lemma 3.4. *Let G be a graph admitting a vertex partition $V(G) = X_1 \cup \cdots \cup X_8$ such that*

[9] Two subgraphs are *pair disjoint* if they share at most one vertex.

- X_1, X_3, X_5, X_7 *are cliques and* X_2, X_4, X_6, X_8 *are independent sets.*
- *The only edges between the parts* $X_1, \dots, X_8$ *are between consecutive parts; that is, for every* $1 \le i \ne j \le 8$, *we have* $E(X_i, X_j) = \emptyset$ *unless* $|i - j| \equiv \pm 1 \pmod 8$.

Then the following holds.

1. *Every induced copy of* C_8 *in* G *is of the form* $x_1x_2...x_8x_1$, *where* $x_i \in X_i$.
2. G *is induced* M*-free.*

Proof of Theorem 3. Set $F_1 = C_8$ and $F_2 = M$. We will show that for every sufficiently small $\varepsilon > 0$ and for every $n \ge n_0(\varepsilon)$ there is a graph G on n vertices which is ε-far from being induced $\{F_1, F_2\}$-free yet contains at most[10] $\varepsilon^{\Omega(\log 1/\varepsilon)} n^{v(F_i)}$ induced copies of F_i for both $i = 1, 2$. This will imply that $\mathcal{P}^*_{\{F_1,F_2\}}$ is not easily testable.

Let $\varepsilon \in \left(0, \frac{\delta_0(8)}{64}\right)$, where $\delta_0(8)$ is from Lemma 3.3. Let $R = R(8, 64\varepsilon)$ be the graph obtained by applying Lemma 3.3. Recall that $V(R) = V_1 \uplus \dots \uplus V_8$, and put $r = |V(R)|$. For simplicity of presentation, we assume that n is divisible by r. We define a graph G on an $\frac{n}{r}$-blowup of R; that is, we replace each vertex $v \in V(R)$ with a vertex-set $B(v)$ of size $\frac{n}{r}$, where the sets $(B(v) : v \in V(R))$ are pairwise-disjoint. Put $B(V_i) := \bigcup_{v \in V_i} B(v_i)$ for $1 \le i \le 8$. The edges of G are defined as follows: $B(V_1), B(V_3), B(V_5), B(V_7)$ are cliques and $B(V_2), B(V_4), B(V_6), B(V_8)$ are independent sets. To define the edges between the sets $B(V_1), \dots, B(V_8)$, recall that by Lemma 3.3, R contains a collection $\mathcal{H}$ of at least $64\varepsilon r^2$ pairwise edge-disjoint cliques, each of the form $\{v_1, \dots, v_8\}$ with $v_i \in V_i$. For each such clique $\{v_1, \dots, v_8\} \in \mathcal{H}$ we put a blowup of C_8 on the sets $B(v_1), \dots, B(v_8)$; namely, for each $(x_1, \dots, x_8) \in B(v_i) \times \dots \times B(v_8)$, $x_1x_2 \dots x_8x_1$ is an induced 8-cycle in G. Notice that G satisfies the assumptions of Lemma 3.4 with $X_i = B(V_i)$. Thus, G is induced M-free, and every induced copy of C_8 in G is of the form $x_1x_2 \dots x_8x_1$ with $x_i \in B(V_i)$. Let $x_1x_2 \dots x_8x_1$ be an induced copy of C_8 in G and let $v_i \in V_i$ be such that $x_i \in B(v_i)$. From the construction of G it follows that $v_1v_2 \dots v_8v_1$ is a (not necessarily induced) cycle in R. By Item 3 in Lemma 3.3 (with parameters $t = 8$ and $i_j = j$ for $1 \le j \le 8$), the number of such cycles is at most r^2. We conclude that G contains at most $r^2 (n/r)^8 \le n^8/r$ induced copies of C_8. By Item 1 in Lemma 3.3 we have $r \ge (\frac{1}{64\varepsilon})^{\beta \log(1/64\varepsilon)} \ge (\frac{1}{\varepsilon})^{\Omega(\log 1/\varepsilon)}$ (where $\beta = \beta(8)$ is from Lemma 3.3). Therefore, the number of induced copies of C_8 in G is at most $\varepsilon^{\Omega(\log 1/\varepsilon)} n^8$, as required.

To finish the proof, we show that G contains εn^2 pair-disjoint induced copies of C_8, which will imply that G is ε-far from being induced $\{C_8, M\}$-free. By Claim 3.1 and the construction of G, for every clique $\{v_1, \dots, v_8\} \in \mathcal{H}$ there is a collection $\mathcal{S}_{v_1,\dots,v_8}$ of at least $(n/8r)^2$ pair-disjoint induced copies of C_8 of the form $(x_1, \dots, x_8) \in B(v_i) \times \dots \times B(v_8)$. Since the cliques in $\mathcal{H}$ are pair-disjoint, copies of C_8 that come from different cliques are pair-disjoint. In other words, for every pair of distinct $\{v_1^{(1)}, \dots, v_8^{(1)}\}, \{v_1^{(2)}, \dots, v_8^{(2)}\} \in \mathcal{H}$ and for every $(x_1^{(i)}, \dots, x_8^{(i)}) \in \mathcal{S}_{v_1^{(i)},\dots,v_8^{(i)}}$ (for $i = 1, 2$), it holds that $|\{x_1^{(1)}, \dots, x_8^{(1)}\} \cap \{x_1^{(2)}, \dots, x_8^{(2)}\}| \le 1$. We thus conclude that $\mathcal{S} := \bigcup_{\{v_1,\dots,v_8\} \in \mathcal{H}} \mathcal{S}_{v_1,\dots,v_8}$ is a collection of at least $|\mathcal{H}| \cdot (n/8r)^2 \ge 64\varepsilon r^2 (n/8r)^2 = \varepsilon n^2$ pair-disjoint induced copies of C_8 in G (where in the first inequality we used the fact that $|\mathcal{H}| \ge 64\varepsilon r^2$). This completes the proof. ■

Proof of Lemma 3.4. We start by proving Item 1. Let $C = x_1x_2...x_8x_1$ be an induced copy of C_8 in G. Our goal is to show that $|C \cap X_i| = 1$ for every $1 \le i \le 8$. First, assume by contradiction, that $|C \cap X_i| \ge 2$ for some $i \in \{1, 3, 5, 7\}$, say $i = 1$ (without loss of generality). Since X_1 is a clique,

[10] In fact, G will be induced F_2-free.

there must be some $j \in \{1, ..., 8\}$ for which $x_j, x_{j+1} \in X_1$ (with indices taken modulo 8). We assume, without loss of generality, that $j = 1$, i.e. that $x_1, x_2 \in X_1$. Note that $|C \cap X_1| < 3$, as otherwise C would contain a triangle. So we see that $C \cap X_1 = \{x_1, x_2\}$. As $(x_2, x_3), (x_1, x_8) \in E(G)$ but $x_3, x_8 \notin X_1$, we must have $x_3, x_8 \in X_2 \cup X_8$. First we consider the case that x_3 and x_8 are in the same part, say $x_3, x_8 \in X_2$. Then $x_4, x_7 \in X_3$ because $(x_4, x_3), (x_7, x_8) \in E(G)$, X_2 is an independent set and $x_4, x_7 \notin X_1$. Since X_3 is a clique, we get that $(x_4, x_7) \in E(G)$, in contradiction to the fact that C is an induced cycle. Now we consider the case that x_3 and x_8 are in different parts, say $x_3 \in X_2$, $x_8 \in X_8$. The path $P = x_3 x_4 ... x_8$ cannot go through X_1, and hence it must contain at least one vertex from each of the seven parts $X_2, ..., X_8$. But this is impossible as P consists of 6 vertices.

In the previous paragraph we showed that $|C \cap X_i| \leq 1$ for every $i \in \{1, 3, 5, 7\}$. Define the sets $X_{odd} := X_1 \cup X_3 \cup X_5 \cup X_7$ and $X_{even} := X_2 \cup X_4 \cup X_6 \cup X_8$. Since X_{even} is an independent set and $\alpha(C_8) = 4$, we have $|C \cap X_{even}| \leq 4$. Thus $|C \cap X_{odd}| \geq 4$, implying that $|C \cap X_i| = 1$ for every $i \in \{1, 3, 5, 7\}$. In order to finish the proof (of Item 1) it is enough to show that $|C \cap X_i| \geq 1$ for each $i \in \{2, 4, 6, 8\}$. Suppose, by contradiction, that $C \cap X_i = \emptyset$ for some $i \in \{2, 4, 6, 8\}$, say $i = 2$. Let $j, k \in \{1, \dots, 8\}$ be such that $x_j \in X_1$ and $x_k \in X_3$. In the cycle C there is a path between x_j and x_k with at most 5 vertices (including x_j and x_k). This path cannot intersect X_2, so it must contain at least one vertex from each of the seven parts $X_1, X_3, X_4, ..., X_8$, which is impossible.

We now prove Item 2. Suppose by contradiction that $Y \subseteq V(G)$ spans an induced copy of M. As before, define $X_{odd} = X_1 \cup X_3 \cup X_5 \cup X_7$ and $X_{even} = X_2 \cup X_4 \cup X_6 \cup X_8$, and notice that X_{even} is an independent set and that X_{odd} is a disjoint union of cliques and hence induced P_3-free (where P_3 is the path with 3 vertices). It is easy to check that every set of 5 vertices of M contains an induced copy of P_3. We conclude that $|Y \cap X_{odd}| \leq 4$. Moreover, $|Y \cap X_{even}| \leq 2$ because $\alpha(M) = 2$. All in all we get that $|Y| \leq 6 < 7 = |V(M)|$, a contradiction. ■

3.3 Homomorphisms and Cores

Recall that a *homomorphism* from a graph G_1 to a graph G_2 is a map $f : V(G_1) \to V(G_2)$ such that for every $u, v \in V(G_1)$, if $(u, v) \in E(G_1)$ then $(f(u), f(v)) \in E(G_2)$. We write $G_1 \leq_{\text{hom}} G_2$, and say that G_1 is homomorphic to G_2, if there is a homomorphism from G_1 to G_2. Notice that the relation $\leq_{\text{hom}}$ is transitive. For a graph G, the *core* of G, denoted $C(G)$, is an induced subgraph of G to which there is[11] a homomorphism from G, and which has the smallest number of vertices among all such induced subgraphs of G. We say that a graph G is *a core* if $C(G) = G$. Observe that the core of any graph is a core, and that every homomorphism from a core to itself is an isomorphism. It is now easy to check that for every pair of cores C_1, C_2, if $C_1 \leq_{\text{hom}} C_2$ and $C_2 \leq_{\text{hom}} C_1$ then C_1 and C_2 are isomorphic. This in turn implies that the core of a graph is defined uniquely, up to isomorphism. We refer the reader to [30] for detailed proofs of these claims, as well as an overview of the topic of graph homomorphisms and cores.

Let $\mathcal{F}$ be a finite family of graphs and consider the set $\mathcal{C} = \mathcal{C}(\mathcal{F}) = \{C(F) : F \in \mathcal{F}\}$. As we explained above, $(\mathcal{C}, \leq_{\text{hom}})$ is a poset in the following sense: for every $C_1, C_2 \in \mathcal{C}$, if $C_1 \leq_{\text{hom}} C_2$ and $C_2 \leq_{\text{hom}} C_1$, then C_1 and C_2 are isomorphic. Namely, $\leq_{\text{hom}}$ is a partial order on the set of equivalence classes of $\mathcal{C}$ under the equivalence relation of graph isomorphism. Let $K(\mathcal{F})$ be a minimal element of the poset $(\mathcal{C}, \leq_{\text{hom}})$; i.e., $K(\mathcal{F})$ is an (arbitrary) element of an (arbitrary) minimal equivalence class. The minimality of $K(\mathcal{F})$ implies that for every $C \in \mathcal{C}$, if there is a homomorphism from C to $K(\mathcal{F})$ (namely, if $C \leq_{\text{hom}} K(\mathcal{F})$), then C is isomorphic to $K(\mathcal{F})$. The key property of the graph $K(\mathcal{F})$ is described in the following proposition.

[11] We note that our definition of a core is a bit different (but equivalent) to the usual definition of a core, see e.g. [30].

Proposition 3.5. *For every $F \in \mathcal{F}$ and for every homomorphism $f : F \to K(\mathcal{F})$, there is a set $X \subseteq V(F)$ such that $f|_X$ is an isomorphism onto $K(\mathcal{F})$.*

Proof. Let $C = C(F)$ be the core of F. Since $f|_{V(C)}$ is a homomorphism from C to $K(\mathcal{F})$, and since $K(\mathcal{F})$ is minimal (in the sense described above), we have that C is isomorphic to $K(\mathcal{F})$. Fix an isomorphism $g : K(\mathcal{F}) \to C$. Then $f|_{V(C)} \circ g$ is a homomorphism from $K(\mathcal{F})$ to itself, and is hence an isomorphism (since $K(\mathcal{F})$ is a core). As g and $f|_{V(C)} \circ g$ are both isomorphisms, $f|_{V(C)}$ must also be an isomorphism. So the assertion of the proposition holds with $X = V(C)$. ■

3.4 Proof of Theorems 2 and 4

Theorems 2 and 4 follow easily from the following theorem.

Theorem 8. *For every $h \geq 3$ there is $\varepsilon_0 = \varepsilon_0(h)$ such that the following holds for every $\varepsilon < \varepsilon_0$ and for every non-bipartite graph H on h vertices. Let K be the core of H. For every $n \geq n_0(\varepsilon)$ there is a graph on n vertices with the following properties.*

1. *G is homomorphic to K.*
2. *G is ε-far from being induced-H-free.*
3. *G contains at most $\varepsilon^{\Omega(\log(1/\varepsilon))} n^k$ (not necessarily induced) copies of K, where $k = |V(K)|$.*

Proof. Fix a homomorphism $\varphi : H \to K$. Since H is not bipartite, and since the homomorphic image of a non-bipartite graph is itself non-bipartite, we get that K is not bipartite, and hence contains an odd cycle. Label the vertices of K by $a_1, \dots, a_k$ so that $a_1 a_2 \dots a_t a_1$ is an odd cycle. Define $H_i = \varphi^{-1}(a_i)$ for $i = 1, \dots, k$. Label the vertices of H by $1, \dots, h$ so that for each $1 \leq i < j \leq k$, the labels of the vertices in H_i are smaller than the labels of the vertices in H_j.

Let $\varepsilon > 0$. We will assume that ε is small enough where needed (in other words, we will choose $\varepsilon_0(h)$ implicitly). Assuming that $\varepsilon < \delta_0(h)/h^2$ (where $\delta_0(h)$ is from Lemma 3.3), let $R = R(h, h^2\varepsilon)$ be the graph from Lemma 3.3. Recall that $V(R) = V_1 \uplus \dots \uplus V_h$, and put $r := |V(R)|$.

We now define a graph S on $V(R)$ as follows. By Item 2 of Lemma 3.3, R contains a collection $\mathcal{H}$ of at least $\varepsilon h^2 r^2$ pair-disjoint h-cliques, each of the form $\{v_1, \dots, v_h\}$ with $v_i \in V_i$ ($1 \leq i \leq h$). For every $\{v_1, \dots, v_h\} \in \mathcal{H}$, we let $S[\{v_1, \dots, v_h\}]$ span an induced copy of H in which v_i plays the role of i for every $i \in [h] = V(H)$. The resulting graph is S. It is clear from the definition that $\mathcal{H}$ is a collection of pair-disjoint induced copies of H in S.

Let n be a large integer which we assume, for simplicity of presentation, to be divisible by $r = |V(S)|$. Let G be the $\frac{n}{r}$-blowup of S; that is, G is the graph obtained by replacing each vertex $v \in V(S)$ with an independent set $B(v)$ of size $\frac{n}{r}$ (where distinct vertices are replaced by disjoint sets), replacing edges with complete bipartite graphs and replacing non-edges with empty bipartite graphs. Clearly $|V(G)| = n$. For $1 \leq i \leq h$ put $B(V_i) := \bigcup_{v \in V_i} B(v)$. Observe that the map which sends $\bigcup_{i \in H_j} B(V_i)$ to a_j for every $1 \leq j \leq k$, is a homomorphism from G to K. This establishes Item 1 in the statement of the theorem.

As we already showed, $\mathcal{H}$ is a collection of at least $\varepsilon h^2 r^2$ pair-disjoint induced copies of H in S. We call these copies the *base copies* of H. For every base copy $\{v_1, \dots, v_h\} \in \mathcal{H}$, Claim 3.1 gives a collection of at least $(n/rh)^2$ pair-disjoint induced copies of H in G, each of the form $\{x_1, \dots, x_h\}$ with $x_i \in B(v_i)$. We say that these copies are *derived* from $\{v_1, \dots, v_h\}$. Since the base copies are pair-disjoint, two copies which are derived from different base copies are also pair-disjoint. Thus, G

contains a collection of at least $|\mathcal{H}| \cdot (n/rh)^2 \geq \varepsilon h^2 r^2 \cdot (n/rh)^2 = \varepsilon n^2$ pair-disjoint induced copies of H. This shows that G is ε-far from being induced H-free.

To finish the proof it remains to show that G contains at most $\varepsilon^{\Omega(\log(1/\varepsilon))} n^k$ copies of K. To this end, we now show that copies of K in G must be of a special form, which will allow us to bound their number. The details follow. Consider a copy of K in G. For each $j = 1, \dots, k$, let $U_j \subseteq V(G)$ be the set of vertices of this copy that are contained in $\bigcup_{i \in H_j} B(V_i)$. Notice that the map that sends U_j to a_j (for each $j = 1, \dots, k$) is a homomorphism from K to itself. By the property of a core (see Subsection 3.3), this map is an isomorphism. Thus, $|U_j| = 1$ for every $1 \leq j \leq k$. Write $U_j = \{u_j\}$, and note that for each $1 \leq i < j \leq k$ we have $(u_i, u_j) \in E(G)$ if and only if $(a_i, a_j) \in E(K)$. Now the fact that $a_1 a_2 \dots a_t a_1$ is a cycle in K implies that $u_1, \dots, u_t, u_1$ is a cycle in G. For each $1 \leq j \leq k$, let $i_j \in H_j$ be such that $u_j \in B(V_{i_j})$ and let $v_{i_j} \in V_{i_j}$ be such that $u_j \in B(v_{i_j})$. Then $i_1 < i_2 < \cdots < i_t$ due to the way we labeled the vertices of H. Moreover $v_{i_1} v_{i_2} \dots v_{i_t} v_{i_1}$ is a cycle in S because G is a blowup of S. Finally, it follows from the definition of S that $v_{i_1} v_{i_2} \dots v_{i_t} v_{i_1}$ must be a cycle in R.

We thus proved that every copy of K in G contains vertices $u_1, \dots, u_t$ with the following property: there is an increasing sequence $1 \leq i_1 < i_2 < \cdots < i_t \leq h$ and vertices $v_{i_j} \in V_{i_j}$ (for $1 \leq j \leq t$) such that $u_j \in B(v_{i_j})$ and such that $v_{i_1} v_{i_2} \dots v_{i_t} v_{i_1}$ is a cycle in R. For every increasing sequence $(i_1, i_2, \dots, i_t)$, Lemma 3.3 states that R contains at most r^2 cycles of the form $v_{i_1} v_{i_2} \dots v_{i_t} v_{i_1}$ with $v_{i_j} \in V_{i_j}$. Therefore, the number of copies of K in G that correspond to a specific increasing sequence is at most $r^2 (n/r)^t n^{k-t} \leq n^k / r$ (here we used the obvious fact that $t \geq 3$). By taking the union bound over all $\binom{h}{t}$ increasing sequences $(i_1, i_2, \dots, i_t)$ and using the inequality $r \geq (1/h^2\varepsilon)^{\beta(h) \log(1/h^2\varepsilon)} \geq (1/\varepsilon)^{\Omega(\log 1/\varepsilon)}$ (which is guaranteed by Item 1 of Lemma 3.3), we get that the number of copies of K in G is at most $\binom{h}{t} n^k / r \leq \varepsilon^{\Omega(\log 1/\varepsilon)} n^k$. This completes the proof. ■

Proof of Theorem 4. By Theorem 8, for every sufficiently small $\varepsilon > 0$ and for every $n \geq n_0(\varepsilon)$ there is a graph G on n vertices which is ε-far from being induced H-free yet contains at most $\varepsilon^{\Omega(\log 1/\varepsilon)} n^k$ (not necessarily induced) copies of K, the core of H. As K is a subgraph of H, G contains at most $\varepsilon^{\Omega(\log 1/\varepsilon)} n^k \cdot n^{h-k} = \varepsilon^{\Omega(\log 1/\varepsilon)} n^h$ (not necessarily induced) copies of H. ■

Proof of Theorem 2. Write $\mathcal{F} = \{F_1, \dots, F_\ell\}$. By symmetry (with respect to graph complementation), it is enough to prove that there is $1 \leq i \leq \ell$ for which F_i is bipartite. Assume, by contradiction, that F_i is not bipartite for every $1 \leq i \leq \ell$. We will show that for every sufficiently small $\varepsilon > 0$ and for every $n \geq n_0(\varepsilon)$, there is a graph G which is ε-far from being induced $\mathcal{F}$-free and yet contains at most $\varepsilon^{\Omega(\log 1/\varepsilon)} n^{v(F_i)}$ copies of F_i for every $1 \leq i \leq \ell$ (where the implicit constant in the exponent depends only on $\mathcal{F}$). This will imply that $\mathcal{P}^*_{\mathcal{F}}$ is not easily testable, a contradiction.

Let $K = K(\mathcal{F})$ be the graph defined in Subsection 3.3. Let us assume, without loss of generality, that K is the core of F_1. We claim that the graph G, obtained by applying Theorem 8 to $H = F_1$ (and K), satisfies our requirements. Clearly, G is ε-far from being induced $\mathcal{F}$-free because it is ε-far from being induced F_1-free.

By Theorem 8, there is a homomorphism $g : G \to K$. Let $1 \leq i \leq \ell$ and consider an embedding $f : F_i \to G$ of F_i into G. Then $g \circ f$ is a homomorphism from F_i to K. By Proposition 3.5, there is a set $X \subseteq V(F_i)$ such that $(g \circ f)|_{V(X)}$ is an isomorphism onto K. This means that $f(V(F_i)) \subseteq V(G)$ contains a copy of K. We conclude that every copy of F_i in G contains a copy of K. By Theorem 8, G contains at most $\varepsilon^{\Omega(\log 1/\varepsilon)} n^k$ copies of K. It follows that G contains at most $\varepsilon^{\Omega(\log 1/\varepsilon)} n^k \cdot n^{v(F_i)-k} = \varepsilon^{\Omega(\log 1/\varepsilon)} n^{v(F_i)}$ copies of F_i, as required. ■

3.5 Proof of Theorem 7

Let K be a graph with vertex set $[k]$. We say that a graph F is a *blowup* of K if F admits a vertex-partition $V(F) = X_1 \cup \cdots \cup X_k$ such that $X_1, \dots, X_k$ are independent sets and for every $1 \le i < j \le k$, if $(i,j) \in E(K)$ then (X_i, X_j) is a complete bipartite graph and if $(i,j) \notin E(K)$ then (X_i, X_j) is an empty bipartite graph. We say that F is the *s-blowup* of K if $|X_1| = \cdots = |X_k| = s$.

Throughout this subsection, C_m denotes the cycle of length m. In the proof of Theorem 7 we use the following simple proposition, whose proof appears at the end of this subsection.

Proposition 3.6. *Let k be an odd integer and let G be a blowup of C_k. Then G is induced C_6-free and (not necessarily induced) C_ℓ-free for every odd $3 \le \ell < k$.*

Recall the definition of a graph homomorphism from Subsection 3.3. We will use the simple fact that $C_{2\ell+1}$ has a homomorphism into C_{2k+1} if and only if $\ell \ge k$ (this fact accounts for the second part of Proposition 3.6). For the proof of Theorem 7 we need the following lemma from [9].

Lemma 3.7. *[9] Let K be a graph on k vertices, let F be a graph on f vertices which has a homomorphism into K and let G be the $\frac{n}{k}$-blowup of K where $n \ge n_0(f)$. Then G is $\frac{1}{2k^2}$-far from being (not necessarily induced) F-free.*

For a graph F, denote by $SG(F)$ the set of supergraphs of F (namely, the set of all graphs on $V(F)$ obtained from F by adding edges). Note that being (not necessarily induced) F-free is equivalent to being induced $SG(F)$-free. We are now ready to prove Theorem 7.

Proof of Theorem 7. Define a sequence $\{a_i\}_{i\ge1}$ as follows: set $a_1 = 3$ and $a_{i+1} = 2^{2(a_i+2)^2} + 1$. Note that a_i is odd for every $i \ge 1$. We prove the theorem with the graph family

$$\mathcal{F} = \{C_6\} \cup \bigcup_{i\ge1} SG(C_{a_i}).$$

Since $a_1 = 3$ we have $C_3 \in \mathcal{F}$. Note that C_6 is a bipartite graph and that C_3 is both a co-bipartite graph and a split graph. For $i \ge 1$ put $\varepsilon_i = \frac{1}{2(a_i+2)^2}$. We will show that $f_{\mathcal{P}^*_{\mathcal{F}}}(\varepsilon_i) \ge 2^{1/\varepsilon_i}$ for every $i \ge 1$ (recall Definition 1.1), which implies that $\mathcal{P}^*_{\mathcal{F}}$ is not easily testable.

Let $i \ge 1$ and put $k = a_i + 2$ and $f = a_{i+1}$. Since a_i is odd and $a_i \ge 3$, we have that k is odd and $k \ge 5$. Fix $n \ge n_0(f)$ which is divisible by k (where $n_0(f)$ is from Lemma 3.7), and let G be the $\frac{n}{k}$-blowup of C_k. By our choice of ε_i and k we have $\varepsilon_i = \frac{1}{2k^2}$. Since C_f has a homomorphism into C_k, Lemma 3.7 implies that G is ε_i-far from being C_f-free and hence is ε_i-far from being induced $SG(C_f)$-free. As $SG(C_f) \subseteq \mathcal{F}$, we conclude that G is ε_i-far from being induced $\mathcal{F}$-free.

Proposition 3.6 implies that G is induced C_6-free and that for every odd $3 \le \ell < k$, G is C_ℓ-free and hence induced $SG(C_\ell)$-free. By the definition of $\mathcal{F}$, if $F \in \mathcal{F}$ is an induced subgraph of G then $|V(F)| \ge a_{i+1} > 2^{2(a_i+2)^2} = 2^{1/\varepsilon_i}$. Here we used the definition of the sequence $\{a_i\}_{i\ge1}$ and our choice of ε_i. We conclude that every set $Q \subseteq V(G)$ of size less than $2^{1/\varepsilon_i}$ is induced $\mathcal{F}$-free, implying that $f_{\mathcal{P}^*_{\mathcal{F}}}(\varepsilon_i) \ge 2^{1/\varepsilon_i}$, as required. ■

We remark that using essentialy the same proof as above, we could have proven the following strengthening of Theorem 7. For every function $g : (0, 1/2) \to \mathbb{N}$ there is a graph family $\mathcal{F}$ that contains a bipartite graph, a co-bipartite graph and a split graph, and there is a decreasing sequence $\{\varepsilon_i\}_{i\ge1}$ with $\varepsilon_i \to 0$ such that $f_{\mathcal{P}^*_{\mathcal{F}}}(\varepsilon_i) > g(\varepsilon_i)$ for every $i \ge 1$.

Proof of Proposition 3.6. As G is a blow-up of C_k, it has a partition $V(G) = X_1 \cup \cdots \cup X_k$ into independent sets such that (X_i, X_j) is a complete bipartite graph if $|i - j| \equiv \pm 1 \pmod{k}$ and an empty bipartite graph otherwise. For the first part of the proposition, assume, by contradiction, that there is $Z \subseteq V(G)$ such that $G[Z]$ is isomorphic to C_6. Since C_6 is not a subgraph of C_k, there must be $1 \leq i \leq k$ such that $|Z \cap X_i| \geq 2$. Assume without loss of generality that there are distinct $u, v \in Z \cap X_1$. By the structure of C_6, there are distinct $x, y \in Z$ such that $(u, x), (u, y) \in E(G)$. Then $x, y \in X_2 \cup X_k$, implying that $(v, x), (v, y) \in E(G)$. Thus, $uxvy$ is a 4-cycle, in contradiction to the fact that $G[Z]$ is isomorphic to C_6.

For the second part of the proposition, simply observe that every subgraph of G with less than k vertices is bipartite. ■

Added note: After posting this paper online, we learned that a statement similar to Lemma 2.2 was proved in [31].

References

[1] N. Alon, Ramsey graphs cannot be defined by real polynomials, Journal of Graph Theory 14 (1990), 651-661.

[2] N. Alon, Testing subgraphs in large graphs, Random Structures and Algorithms 21 (2002), 359-370.

[3] N. Alon, private communication, 2013.

[4] N. Alon, E. Fischer, M. Krivelevich and M. Szegedy, Efficient testing of large graphs, Combinatorica 20 (2000), 451-476.

[5] N. Alon, E. Fischer, and I. Newman. Testing of bipartite graph properties. SIAM Journal on Computing 37 (2007), 959-976.

[6] N. Alon and J. Fox, Easily testable graph properties, Combin. Probab. Comput. 24 (2015), 646-657.

[7] N. Alon and A. Shapira, A characterization of easily testable induced subgraphs, Combin. Probab.Comput. 15 (2006), 791-805.

[8] N. Alon and A. Shapira, A characterization of the (natural) graph properties testable with one-sided error, SIAM Journal on Computing 37 (2008), 1703-1727.

[9] N. Alon and A. Shapira, A separation theorem in property testing, Combinatorica 28 (2008), 261-281.

[10] N. Alon and A. Shapira, Every monotone graph property is testable, SIAM Journal on Computing 38 (2008), 505-522.

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

[12] T. Austin and T. Tao, On the testability and repair of hereditary hypergraph properties, Random Structures and Algorithms 36 (2010), 373-463.

[13] F. A. Behrend, On sets of integers which contain no three terms in arithmetic progression, Proc. National Academy of Sciences USA 32 (1946), 331-332.

[14] B. Bollobás, **Extremal graph theory**, Courier Corporation, 2004.

[15] A. Chernikov and S. Starchenko, Regularity lemma for distal structures. arXiv preprint arXiv:1507.01482, 2015.

[16] A. Chernikov and S. Starchenko, Definable regularity lemmas for NIP hypergraphs. arXiv preprint arXiv:1607.07701, 2016.

[17] D. Conlon and J. Fox, Bounds for graph regularity and removal lemmas, GAFA **22** (2012), 1191-1256.

[18] P. Erdős, On circuits and subgraphs of chromatic graphs, Mathematika 9 (1962), 170-175.

[19] J. Fox, A new proof of the graph removal lemma, Ann. of Math. **174** (2011), 561–579.

[20] J. Fox, L. Gishboliner, A. Shapira and R. Yuster, The removal lemma for tournaments. Journal of Combinatorial Theory, Series B, to appear. 2018.

[21] J. Fox, M. Gromov, V. Lafforgue, A. Naor, and J. Pach, Overlap properties of geometric expanders. Journal für die reine und angewandte Mathematik (Crelles Journal), 2012(671), 49-83, 2012.

[22] J. Fox, J. Pach and A. Suk, A polynomial regularity lemma for semi-algebraic hypergraphs and its applications in geometry and property testing. arXiv preprint arXiv:1502.01730, 2015.

[23] O. Goldreich, Contemplations on testing graph properties, ECCC Technical Report 2005. Also, Studies in Complexity and Cryptography 6650 (2011), 547-554.

[24] O. Goldreich, S. Goldwasser, and D. Ron, Property testing and its connection to learning and approximation, J. ACM 45 (1998), 653-750.

[25] O. Goldreich and L. Trevisan, Three theorems regarding testing graph properties, Random Structures & Algorithms 23 (2003), 23-57.

[26] M. C. Golumbic, **Algorithmic graph theory and perfect graphs** (Vol. 57), 2004, Elsevier.

[27] M. C. Golumbic, Trivially perfect graphs, Discrete Mathematics 24 (1978), 105-107.

[28] T. Gowers, Lower bounds of tower type for Szemerédi's uniformity lemma, Geom. Funct. Anal. **7** (1997), 322–337.

[29] F. Harary, **Graph Theory**, Massachusetts: Addison-Wesley (1972), 71-83.

[30] P. Hell and J. Nesetril, **Graphs and homomorphisms**. Oxford University Press, 2004.

[31] L. Lov´asz and B. Szegedy, Regularity partitions and the topology of graphons, in An irregular mind, 415-445, Bolyai Soc. Math. Stud. 21, János Bolyai Math. Soc., Budapest, 2010.

[32] L. Lovász and B. Szegedy, Testing properties of graphs and functions, Israel J. Math. 178 (2010), 113-156.

[33] M. Malliaris and S. Shelah, Regularity lemmas for stable graphs. Trans. Amer. Math. Soc. 366, 1551-1585, 2014.

[34] V. Rödl and R. Duke, On graphs with small subgraphs of large chromatic number, Graphs and Combinatorics 1 (1985), 91-96.

[35] V. Rödl and M. Schacht, Generalizations of the removal lemma, Combinatorica 29, 2009, 467-501.

[36] R. Rubinfeld and M. Sudan, Robust characterizations of polynomials with applications to program testing, SIAM Journal on Computing 25 (1996), 252-271.

[37] I. Ruzsa, Solving a linear equation in a set of integers I, Acta Arithmetica 65 (1993), 259-282.

[38] I. Z. Ruzsa and E. Szemerédi, Triple systems with no six points carrying three triangles, in Combinatorics (Keszthely, 1976), Coll. Math. Soc. J. Bolyai 18, Volume II, 939-945.

[39] P. Simon, A note on "Regularity lemma for distal structures", Proceedings of the American Mathematical Society, 144(8), 3573-3578, 2016.

[40] E. Szemerédi, Regular partitions of graphs, In: *Proc. Colloque Inter. CNRS*, 1978, 399–401.

[41] T. Tao, Szemerédi's regularity lemma revisited. Contributions to Discrete Mathematics, 1(1), 2006.