REGULARITY LEMMAS FOR STABLE GRAPHS

M. MALLIARIS AND S. SHELAH

Abstract. We develop a framework in which Szemerédi’s celebrated Regularity Lemma for graphs interacts with core model-theoretic ideas and techniques. Our work relies on a coincidence of ideas from model theory and graph theory: arbitrarily large half-graphs coincide with model-theoretic instability, so in their absence, structure theorems and technology from stability theory apply. In one direction, we address a problem from the classical Szemerédi theory. It was known that the “irregular pairs” in the statement of Szemerédi’s regularity lemma cannot be eliminated, due to the counterexample of half-graphs (i.e., the order property, corresponding to model-theoretic instability). We show that half-graphs are the only essential difficulty, by giving a much stronger version of Szemerédi’s regularity lemma for models of stable theories of graphs (i.e. graphs with the non-$k_\ast$-order property), in which there are no irregular pairs, the bounds are significantly improved, and each component satisfies an indivisibility condition. In the other direction, we take a more model-theoretic approach, and give several new Szemerédi-type partition theorems for models of stable theories of graphs. The first theorem gives a partition of any such graph into indiscernible components, meaning here that each component is either a complete or an empty graph, whose interaction is strongly uniform. This relies on a finitary version of the classic model-theoretic fact that stable theories admit large sets of indiscernibles, by showing that in models of stable theories of graphs one can extract much larger indiscernible sets than expected by Ramsey’s theorem. The second and third theorems allow for a much smaller number of components at the cost of weakening the “indivisibility” condition on the components. We also discuss some extensions to graphs without the independence property. All graphs are finite and all partitions are equitable, i.e. the sizes of the components differ by at most 1. In the last three theorems, the number of components depends on the size of the graph; in the first theorem quoted, this number is a function of $\epsilon$ only as in the usual Szemerédi regularity lemma.

1. Introduction

1.1. Half-graphs and the order property. The starting point for this paper was the question of irregular pairs in Szemerédi’s celebrated Regularity Lemma. The Regularity Lemma has been a major tool in graph theory and related fields since its invention in [19] to prove that sets of positive upper density in $\mathbb{N}$ contain arbitrarily long arithmetic progressions. For more on the history and wide applicability of this lemma, see the excellent survey [8].

We first state Szemerédi’s lemma. Recall that if $G$ is a graph with edge relation $R$, and $A, B \subseteq G$ are finite subgraphs with disjoint vertex sets, the density $d(A, B) = |R \cap (A \times B)| / |A||B|$ and we say that $(A, B)$ is $\epsilon$-regular if for all $A' \subseteq A, B' \subseteq B$ with $|A'| \geq \epsilon |A|$, $|B'| \geq \epsilon |B|$, we have that $|d(A, B) - d(A', B')| < \epsilon$.

Theorem A. (Szemerédi’s regularity lemma) For every $\epsilon, m$ there exist $N = N(\epsilon, m)$, $m' = m(\epsilon, m)$ such that given any finite graph $X$, of size at least $N$, there is $\ell$ with $m \leq \ell \leq m'$ and a partition $X = X_1 \cup \cdots \cup X_\ell$ satisfying:

$$\sum_{i,j} ||X_i| - |X_j|| \leq 1 \text{ for all } i, j \leq \ell$$
(2) all but at most $\ell^2$ of the pairs $(X_i, X_j)$ are $\epsilon$-regular.

As explained in §1.8 of [8], “Are there exceptional pairs?” it was not known for some time whether the $\ell^2$ irregular pairs allowed in clause (2) were necessary. Several researchers (Lovasz, Seymour, Trotter, as well as Alon, Duke, Leffman, Rödl, and Yuster in [1]) then independently observed that the half-graph, i.e., the bipartite graph with vertex sets $\{a_i : i < n\} \cup \{b_i : i < n\}$ (for arbitrarily large $n$) such that $a_iRb_j$ if $i < j$, shows that exceptional pairs are necessary.

We may therefore ask whether “half-graphs” are the main difficulty. [This appears to be a new question, but it is a very natural one.]

**Question 1.1.** Consider the class of graphs which admit a uniform finite bound on the size of an induced sub-half-graph. It is possible to give a stronger regularity lemma for such graphs in which there are no irregular pairs?

This is a particularly suggestive question for a model theorist, who will recognize the half-graph as an instance of the order property, see Definition 2.3. One result of this paper, Conclusion 5.19, shows that half-graphs are indeed the only essential difficulty: for graphs with the non-$k_s$-order property, it both eliminates irregular pairs and also significantly improves the tower-of-exponential bounds of the usual Szemerédi lemma, which are necessary by work of Gowers [4].

**Conclusion 5.19.** For every $k_s \in \mathbb{N}$ and $\epsilon \in (0, \frac{1}{2})$ there are $N = N(\epsilon, k_s), m = m(\epsilon, k_s)$ such that for every finite graph $G$ with the non-$k_s$-order property (Definition 2.3) and every $A \subseteq G$ with $|A| \geq N$, there is $\ell \leq m$ and a partition $A = (A_i : i < \ell)$ such that each $A_i$ is $\epsilon$-excellent, and for every $0 \leq i < j < \ell$,

- $||A_i| - |A_j|| \leq 1$
- $(A_i, A_j)$ is $\epsilon$-regular, and moreover if $B_i \in [A_i]_{2\epsilon|A_i|}$ and $B_j \in [A_j]_{2\epsilon|A_j|}$ then
  $$
  (d(B_i, B_j) < \epsilon) \lor (d(B_i, B_j) \geq 1 - \epsilon)
  $$
- if $\epsilon < \frac{1}{2^{k_s}}$, then $m \leq (3 + \frac{2}{\epsilon})(\frac{1}{\epsilon})^{2k_s}$

Conclusion 5.19 follows directly from a stronger structural result, Theorem 5.18, which we now quote and briefly explain. The notion of uniformity is generally more informative than regularity in our context. Note that the proof of Theorem 5.18 does not use the Szemerédi lemma.

**Theorem 5.18.** Let $k_s$ and therefore a bound for $k_{ss}$ (Definition 2.11) be given. Then for all $\epsilon > 0$ there exists $m = m(\epsilon, k_s)$ and $N = N(\epsilon, k_s)$ such that for every finite graph $G$ with the non-$k_s$-order property and every $A \subseteq G$ with $|A| \geq N$, there is a partition $(A_i : i < i(\epsilon) \leq m)$ of $A$ into at most $m$ pieces, where each of the pieces is $\epsilon$-excellent (Definition 5.2), all of the pairs are $\epsilon$-uniform (Claim 5.6), and if $\epsilon < \frac{1}{2^{k_s}}$, then $m \leq (3 + \epsilon)(\frac{2}{\epsilon})^{k_{ss}}$.

The proof of Theorem 5.18, contained in §5, has two main features. First, bounds on the size of certain induced subgraphs (orders and binary trees), familiar from stability, allow for the partition of the graph into relatively large “indivisible” pieces. However, the size of these pieces may vary along a fixed rapidly decreasing sequence. We then use the absence of the independence property to show that one can refine the given partition by randomly partitioning each piece to obtain equally sized components while preserving “indivisibility.” This is an instance of another classic connection between model theory and graph theory, the link of NIP to the Vapnik-Chervonenkis theory first noticed by Laskowski [10]. Finally, we verify that the number of pieces can be bounded by a function of $\epsilon$, due to the definition of “indivisibility,” $\epsilon$-excellence. This settles Question 1.1 and thus characterizes the class of graphs admitting $\epsilon$-regular partitions with no irregular pairs.

The bulk of the paper is devoted to a more general model-theoretic analysis of stable theories of graphs, which we now describe.
1.2. The phenomenon of regularity in stable theories of graphs. Recall that theories in which no formula has the order property are called stable. Such theories have been fundamental to model theory since the second author’s work in [15], see e.g. the “Unstable Formula Theorem” II.2.2, p. 30. Generally speaking, one contribution of such model-theoretic analysis is to characterize global structural properties, such as number of models, existence of indiscernible sets, number of types, and so on, in terms of local combinatorial properties, such as the order property in some formula. What does the phenomenon of regularity mean for stable theories of graphs, and what are its essential parameters?

The usual regularity lemma and its associated tools (the Key Lemma on extracting configurations which appear in the reduced graph) are able to give uniform approximations to any sufficiently large graph thanks to two degrees of freedom: in the Szemerédi partition, (1) some pairs of components can fail to be $\epsilon$-regular, and (2) the structure within each component can be quite chaotic. Having characterized the class of graphs admitting a regular partition with no irregular components, we show more generally throughout the paper that not only (1) but also (2) fails in a strong sense for this class: our regularity lemmas work from the inside out by extracting large “indivisible” pieces from the given graph, and then obtaining the uniform interaction essentially for free.

The order of the results was chosen to illustrate a spectrum of tradeoffs between structure of the components of the partition on one hand and size of the partition on the other.

The first partition theorem, Theorem 3.8, gives a partition in which structure of the components is optimized: all of the pieces are either complete graphs or empty graphs. This uses a Ramsey-type result, Theorem 3.5, of independent interest which applies $R(x = x, \Delta, 2)$ (stability) rank to show that in finite models of a stable theory of graphs one can extract much larger indiscernible sets than predicted by Ramsey’s theorem.

**Theorem 3.8.** Let $k_*, n_2$ be given with $n_2 > (2k_*)^2$. Then there is $N = N(n_2, k_*)$ such that any finite graph $G$, $|G| > N$ with the non-$k_*$-order property admits an equitable partition $\overline{G} = \langle G_i \rangle$ into disjoint pieces $G_i$ which, after possibly omitting one element from each $G_i$, satisfies: each $G_i$ is either a complete graph or an empty graph, and for all pairs $G_i, G_j \in \overline{G}$ there exists a truth value $t(G_i, G_j) \in \{0,1\}$ such that for all but $\leq 2k_*$ $a \in G_i$, for all but $\leq 2k_*$ $b \in G_j$, $aRb \equiv t(G_i, G_j)$.

Moreover, $N = n_1n_2$ suffices if $n_1 > (cn_2)^{(2tr)k_*}$ for constants $c, t, r$ depending on $k_*$, see 3.5 (3).

Due to the strength of the indiscernibility requirement, the number of components is forced to grow with the size of the graph (this is relaxed only in §5). §4 then builds a more general theory: in this section “indivisibility” of the components is defined in terms of an arbitrary nondecreasing function $f$ (at times specialized to $f(x) = x^c$) which specifies the allowed number of exceptions. With no further assumptions on $f$ the number of components must still grow with $|G|$. The main results of §4 are Theorem 4.16 and Theorem 4.23. The first relies on a probabilistic argument to extract an equitable partition, at the cost of irregular pairs:

**Theorem 4.16.** Let $\epsilon = \frac{1}{r} \in (0, \frac{1}{2})$, $k_*$ and therefore a bound for $k_{**}$ be given, and suppose $G$ is a finite graph with the non-$k_*$-order property. Let $A \subset G$, $|A| = n$ with $n^{k_{**}} > k_{**}$. Then there is $\zeta < \epsilon^{k_{**}}$ and a partition $\langle A_i : i < i(*) \rangle$ of $A$ such that:

1. for all $i$, either $|A_i| = n^\epsilon$ or $|A_i| = 1$, with $\{|i : |A_i| = 1\}| \leq n^\epsilon$
2. $\frac{2}{n^{k_{**}}} \geq \frac{1}{(\frac{1}{2})^c} \{|\{(i,j) : |A_i| = 1, |A_j| = 1, \text{ or } (a,b) \in A_i \times A_j : aRb \notin \{A_i \times A_j, \emptyset\}\}|$

Moreover, the total number of pieces is at most $n^c$ where $c = c(\epsilon) = 1 - \epsilon^{k_{**}+1} - 2\epsilon^{2k_{**}+1}$.

Note, however, the entirely uniform interaction of the regular pairs. After proving a combinatorial lemma, in Theorem 4.23 we show how to solve this instance of irregular pairs at the cost of a larger remainder.

3
Finally, §5 gives the theorem quoted above, obtaining a partition whose size depends only on \( \epsilon \) at the cost of weakening the “indivisibility” condition to allow for a linear number of exceptions.

Though the results of the paper form a natural sequence, each section is self-contained and can be read independently. The authors are working on improving the bounds in 3.2 and 3.8 and in 4.11 and on the parallel to §5 for \( k_* \)-dependent graphs (necessarily with exceptional pairs).

**Acknowledgements.** Thanks to David Galvin, Chris Laskowski, Thomas Scanlon, and Pierre Simon for comments on an earlier draft of this paper. We are also grateful to the referee for many useful suggestions.

## 2. Preliminaries

**Notation 2.1.** (Graphs) We consider graphs model-theoretically, that is, as structures \( G \) in a language with equality and a symmetric irreflexive binary relation \( R \), whose domain consists of a set of vertices, and where the interpretation \( R_G \) consists of all pairs of vertices \((a, b)\) connected by an edge. We will often write \( aRb \) to indicate that \((a, b) \in R_G \), and write \( G \) for the domain of \( G \). In particular, \(|G|\), the cardinality of \( G \), is the cardinality of the set of vertices.

**Remark 2.2.** In model theory, a formula \( \varphi(\overline{x}, \overline{y}) \) is said to have the order property relative to a theory \( T \) if there exist a model \( M \models T \) and sequence \( (\overline{a}_i, \overline{b}_i) \) of tuples from \( M \) such that \( M \models \varphi(\overline{a}_i, \overline{b}_j) \) if \( i < j \). A theory \( T \) is called stable if no formula has the order property relative to \( T \). In what follows we will focus on the order property for the graph edge relation \( R \). That is, we only require that the formula \( xRy \) be stable, not the full theory of \( G \). The exception is §3 which occasionally asks that \( R \)-rank be defined for certain sets of formulas \( \Delta \).

**Definition 2.3.** (The non-\( k \)-order property) A graph \( G \) has the non-\( k \)-order property when there are no \( a_i, b_j \in G \) for \( i < k \) such that \( i < j < k \implies (a_i R_G b_j) \land \neg (a_j R_G b_i) \). If such a configuration does exist, \( G \) has the k-\( \text{order property} \).

**Definition 2.4.** (Stable theory of graphs) We say that \( T \) is a stable theory of graphs if and only if \( T \) contains the axioms of graph theory and the formula \( xRy \) is stable for \( T \). (For the model-theoretic reader, we emphasize that we are only requiring stability of the formula \( xRy \), see Remark 2.2 above.) Note that \( T \) is a stable theory of graphs if and only if for some \( k_* \in \mathbb{N} \), \( T \) implies the non-\( k_* \)-order property.

**Remark 2.5.** By the symmetry of \( R \), it is enough to rule out the order in one direction (i.e. the non-\( k \)-order property also implies that for no such sequence does \( i < j \implies \neg(a_i R_G b_j) \land a_j R_G b_i \).

**Claim 2.6.** Suppose \( G \) is a graph with the non-\( k \)-order property. Then for any finite \( A \subseteq G \), \(|\{a \in A : a R_G b \} : b \in G\}| \leq |A|^k \), more precisely \( \leq \sum_{i \leq k} (|A|) \).

**Proof.** See [15] Theorem II.4.10(4) p. 72 and Theorem 1.7(2) p. 657. \( \square \)

**Definition 2.7.** (Indiscernibility) Let \( M \) be a model, let \( \Gamma \) be a set of formulas in the language of \( M \) and \( \alpha \) an ordinal. Recall that a sequence \( \langle a_i : i < \alpha \rangle \) of elements of \( M \) is said to be a \( \Gamma \)-indiscernible sequence if for any \( n < \min(\alpha, \omega) \), any formula \( \gamma = \gamma(x_0, \ldots, x_{n-1}) \in \Gamma \) and any two increasing sequences \( i_0 < \cdots < i_{n-1}, j_0 < \cdots < j_{n-1} \) from \( \alpha \), we have that \( M \models \gamma(a_{i_0}, \ldots, a_{i_{n-1}}) \iff M \models \gamma(a_{j_0}, \ldots, a_{j_{n-1}}) \).

**Notation 2.8.** Let \( \varphi \) be a formula. Then we identify \( \varphi^0 = \neg \varphi \), \( \varphi^1 = \varphi \). We also identify “true” with \( 1 \) and “false” with \( 0 \), so that in particular the intended interpretation of \( \varphi^X \), where \( X \) is an expression which evaluates to either true or false, is simply \( \varphi \) or \( \neg \varphi \), as appropriate. Likewise, the intended interpretation of expressions like “\( x R_a t \)” where \( t \in \{0, 1\} \) or is an expression which evaluates to true or false, is “\( x R_a t \) if and only if \( t = 1 \)” or equivalently, \( \varphi^t \) is true.
Definition 2.9. (The set $\Delta_k$) Let $\Delta_k$ be the set of formulas \( \{x_0Rx_1\} \cup \{\varphi_{k,m} : m \leq k, i \in \{1,2\}\} \) where
\[
\varphi_{k,m} = \varphi_{k,m}(x_0, \ldots, x_{k-1}) = (\exists y) \left( \bigwedge_{\ell < m} (x_\ell Ry)^{i\ell} \land \bigwedge_{m \leq \ell < k} (x_\ell Ry)^{i\ell} \right)
\]

Observation 2.10. Let $H$ be a finite graph, and let $A = \langle a_i : i < \alpha \rangle$ be a $\Delta_k$-indiscernible sequence of elements of $H$ where $\alpha \geq 2k$. Suppose that for some increasing sequence of indices $i_0 < \cdots < i_{2k-1} < \alpha$ and for some element $b \in H$ the following holds:

- for all $\ell$ such that $0 \leq \ell \leq k-1$, $bR_H a_{i_\ell}$ and
- for all $\ell$ such that $k \leq \ell < 2k$, $\neg bR_H a_{i_\ell}$.

Then $H$ has the $k$-order property.

Proof. For each $m$ with $0 \leq m \leq k-1$, consider the sequence $\langle c_j : 0 \leq j \leq k-1 \rangle$ given by $c_j := a_{m+i}$. Then $b$ witnesses that $\varphi_{k,m}(c_0, \ldots, c_{k-1})$ is true in $H$. As any two increasing subsequences of $A$ of length $k$ satisfy the same $\Delta_k$-formulas, this easily gives the $k$-order property.

Note that if we had assumed the inverse, i.e. for all $\ell$ such that $0 \leq \ell \leq k-1$, $\neg bR_H a_{i_\ell}$ and for all $\ell$ such that $k \leq \ell < 2k$, $bR_H a_{i_\ell}$, we again get the $k$-order property (the same proof works with $\varphi^2$ replacing $\varphi^1$, in the notation of Definition 2.9).

Below, we will consider graphs with the non-$k_*$-order property (and reserve the symbol $k_*$ for this bound). We define an associated bound $k_{\ast \ast}$ on tree height:

Definition 2.11. (The tree bound $k_\ast \ast$) Suppose $G$ does not have the $k_\ast$-order property. Let $k_\ast \ast = k_\ast \ast(G) < \omega$ be minimal so that there do not exist sequences $\pi = \langle a_\eta : \eta \in k_\ast \ast\rangle$ and $b = \langle b_\rho : \rho \in k_\ast \ast \rangle$ of elements of $G$ such that if $\rho \subseteq \langle \ell \rangle \leq \eta \in k_\ast \ast$ then $(a_\eta R b_\rho) \equiv (\ell = 1)$.

Remark 2.12. In general, given a formula $\varphi(x,y)$, we may define $k_\ast \ast = k_\ast \ast(\varphi)$ as in Definition 2.11 with $\varphi$ in place of $R$. Then (i) if $\varphi$ has the non-$k_\ast$-order property, $k_\ast \ast$ exists and $k_\ast \ast < 2^{k_\ast \ast + 1}$.

Conversely (ii) if $k_\ast \ast = k_\ast \ast(\varphi)$ exists, then $\varphi$ has the non-(2$^{k_\ast \ast + 1}$)-order property. See Hodges [6] Lemma 6.7.9 p. 313.

It will also be useful to speak about the average interaction of sets.

Definition 2.13. (Truth values t) By a truth value $t = t(X,Y)$ for $X,Y \subseteq G$, we mean an element of $\{0,1\}$, where these are identified with “false” and “true” respectively. When $X = \{x\}$, write $t = t(x,Y)$. The criteria for assigning this value will be given below.

Definition 2.14. (Equitable partitions) We will call a partition of $A \subseteq G$ into disjoint pieces $\langle A_i : i < m \rangle$ equitable if for all $i < j < m$, $|A_i| - |A_j| \leq 1$.

Notation 2.15. (Distinguished symbols) Throughout this article, $\epsilon, \zeta, \xi$ are real numbers in $(0,1)$.

We use $\rho, \eta$ for zero-one valued sequences $^\omega \mathbb{2}$, usually in the context of trees. (Following logical convention, a given natural number $n$ is often identified with $\{0, \ldots, n-1\}$.) The letters $x,y,z$ are variables, and $i,j,k,\ell,m,n$ denote natural numbers, with the occasional exception of the standard logical notation $\ell(\pi)$, i.e. the length of the tuple $\pi$. $T$ is a first-order theory, unless otherwise specified the theory of the graph $G$ under consideration in the language (=vocabulary) with equality and a binary relation symbol $R$.

The symbols $k_\ast, k_{\ast \ast}, m_\ast, m_{\ast \ast}$ are distinguished. When relevant (the conventions are given at the beginning of each section), $k_\ast$ is such that the graph $G$ under consideration has the non-$k_\ast$-order property, Definition 2.3, and $k_{\ast \ast}$ is the associated tree bound, Definition 2.11. (The one exception is $\S3.2$, in which $k_\ast$ is such that the graph under consideration is $k_\ast$-dependent.) The relevant sections all compute bounds based on $k_\ast$, so it is useful, but not necessary, to assume $k_\ast$ is minimal for this
3. A partition into indiscernible pieces

Classically, the hypothesis of stability implies that in infinite models, one can extract large indiscernible sequences (Definition 2.7 above). More precisely, given an infinite cardinal, a model whose theory is stable in $\lambda$ and $A, J \subseteq M$ with $|A| \leq \lambda < |I|$, there is $J \subseteq I$, $|J| > \lambda$ such that $J$ is an $A$-indiscernible sequence (in fact, an $A$-indiscernible set); see [15] Theorem 2.8 p. 14.

In this section, we begin by proving a finite analogue of this result, Theorem 3.5, which shows that a finite stable graph will have relatively large indiscernible subsequences (in fact, subsets) compared to what one could expect from Ramsey’s theorem. We apply this to give an equitable partition of any stable graph in which the number of pieces is much larger than the size of those pieces; the gain, however, is that the pieces in the partition are themselves indiscernible sets, there are no irregular pairs, and the condition of “regularity” is very strong. Namely, to each pair of pieces $(A, B)$ we may associate a truth value $t_{A,B}$ such that there are at most a constant number of exceptional edges $(aRb \not= t_{A,B})$. This is not superceded in later sections. Moreover we can extend some results to unstable dependent theories $T$, see §3.2.

**Hypothesis 3.1.** Throughout §3 $G$ is a finite graph with edge relation $R$ which has the non-$k_*$-order property.

The next claim will be applied to prove Crucial Observation 3.6 below.

**Claim 3.2.** If $m \geq 4k_*$ and $\langle a_i : i < m \rangle$ is a $\Delta_k$-indiscernible sequence in $G$, and $b \in G$, then either $|\{i : a_i R b\}| < 2k_*$ or $|\{i : \neg (a_i R b)\}| < 2k_*$.

**Proof.** Suppose for a contradiction that both $Y = \{i : a_i R b\}$ and $X = \{i : \neg (a_i R b)\}$ have at least 2$k_*$ elements. Let $i_1$ be the $k_*$th element of $X$ and let $i_2$ be the $k_*$th element of $Y$. Clearly $i_1 \not= i_2$.

- **Case 1:** $i_1 < i_2$. By assumption, we can find a subsequence $a_{j_1} < \cdots a_{j_{k_*}} < a_{j_{k_*+1}} < \cdots < a_{j_{k_*+2}} \leq a_m$ such that $\{j_1 < \cdots < j_{k_*} = i_1\} \subseteq X$ and $\{i_2 = j_{k_*+1} < \cdots < j_{k_*+2}\} \subseteq Y$. Observation 2.10 gives the $k_*$-order property, contradiction.

- **Case 2:** $i_2 < i_1$. Similar argument, replacing $R$ by $\neg R$ (since $R$ is symmetric, it is equivalent). $\square$

**Definition 3.3.** (the notation is from [7]) Let $\Gamma$ be a set of formulas, $n_1$ a cardinal and $n_2$ an ordinal (for our purposes these will both be finite). Then $n_1 \rightarrow (n_2)^\Gamma_{\text{fin}}$ means: for every sequence $\langle a_i : i < n_1 \rangle$ of elements of $G$, there is a non-constant sub-sequence $\langle a_i : j < n_2 \rangle$ which is a $\Gamma$-indiscernible sequence, Definition 2.7. Replacing 1 by $\ell$ means that the tuples $a_i$ in the sequence have length $\ell$. Usually we suppress mention of $T = Th(G)$ and assume $\ell(a_i) = 1$, and therefore simply write $n_1 \rightarrow (n_2)^\Gamma$.

**Claim 3.4.** If $n_1 \rightarrow (n_2)^{\Delta_{k_*}}_{\text{fin}}$ in the usual arrow notation then

$$n_1 \rightarrow (n_2)^{\Delta_{k_*}}$$

**Proof.** Given an increasing sequence of elements of $n_1$ of length $k_*$ we may color it according to which subset of the formulas of $\Delta_{k_*}$ hold on the sequence, and so extract a homogeneous subsequence of order type $n_2$. $\square$
is finite, but rather in showing that in our context they are much larger than expected: a priori, in Claim 3.4 the minimal $n_1$ is essentially $\sum_{k}(n_2 + 2^{2k})$, compared to (2)-(3) in the theorem below. It is possible that versions of this result exist (for infinitary versions see [15]). The statement and proof make use of some model-theoretic notions not needed elsewhere in the paper, i.e. types (consistent sets of formulas in the given free variables with parameters from a specified set) and R-rank (used in the “by definition” clause in Step 2A of the proof; see [15] p. 21, p. 31).

**Theorem 3.5.** Assume that $k, k_2, \Delta$ are such that:

(a) $\Delta$ is a finite set of formulas, each with $\leq k$ free variables, and closed under cycling the variables

(b) For each formula $\varphi(x_0, \ldots x_{k-1}) \in \Delta$ and any partition $\{x_0, \ldots x_{\ell}\}, \{x_{\ell+1}, \ldots x_{k-1}\}$ of the free variables of $\varphi$ into object and parameter variables, the formula $\varphi(x_0, \ldots x_{\ell}; x_{\ell+1}, \ldots x_{k-1})$ has the non-$k_2$-order property.

Then:

(1) There exists a natural number $r$ such that for any $A \in G$, $|A| \geq 2$, we have that $|S_\Delta(A)| \leq |A|^r$

(2) For each $A = \langle a_i : i < n \rangle$ there exists $u \subseteq n$ such that:

- $|u| \geq f^k(n)$ where $f(x) = \left| \frac{x}{2} \right| + 1 - k$ and $r, t, k$ are constants depending only on $\Delta$:
  - $r$ is from (1) of the theorem, $t$ is a stability constant (the R-rank of $\Delta$), and $k$ is the number of free variables.
- $\langle a_i : i \in u \rangle$ is $\Delta$-indiscernible.

(3) In particular, for $\Delta_{k_s}$ from Definition 2.9, we have that $n_1 \to (n_2)_{T, \Delta_{k_s}, 1}$ for any $n_1 > (cn_2)^{(2n)^{k_s}}$, for a constant $c$ depending on $k_s$, i.e. $c = c(\Delta_{k_s}) = c(r, t, k_s)$ where $r, t$ are computed for $\Delta_{k_s}$.

**Proof.** (1) See [15] Theorem II.4.10(4) and II.4.11(4) p. 74.

(2) Adding dummy variables if necessary, we may suppose that each $\varphi \in \Delta$ has the free variables $x_0, \ldots x_{k-1}$. (We may then have to omit $k$ elements at the end.)

We prove by induction on $m \leq k$ that there is $u_m \subseteq n$ such that:

(I) $|u_{m+1}| \geq f(n)$ for $f(x) = \left| \frac{x}{2} \right| + 1 - k$ where $r$ is from clause (1) of the theorem and $t$ is a constant defined below

(II) if $i_0 < \cdots < i_{k-1}, j_0 < \cdots < j_{k-1}$ are from $u_m$, $\bigwedge_\ell(\ell < k - m \implies i_\ell = j_\ell)$, and $\varphi \in \Delta$, then

$$\models \varphi(a_{i_0}, \ldots a_{i_{k-1}}) = \varphi(a_{j_0}, \ldots a_{j_{k-1}})$$

The case $m = 0$. Trivial: $u = n$.

The case $m+1$. Let $u_m$ be given, and suppose $|u_m| = \ell_m$. Let $\Delta^m = \{ \varphi(x_0, \ldots x_{m-1}, a_{\ell_m-m}, \ldots a_{\ell_m-1}) : \varphi \in \Delta \}$. This case will be broken up into several steps.

**Step 0:** Arranging the elements of $u_m$ into a type tree. The natural partial order on $\omega > \omega$ is given by $\nu \leq \rho$ if $\nu$ is an initial segment of $\rho$. A tree order on a finite set of vertices is a partial order which is order-isomorphic to some downward closed subset of $\omega > \omega$ under the natural partial order.

By induction on $\ell < \ell_m$ choose sets $W_\ell \subseteq u_m \setminus \bigcup_{\ell < j} W_j$ and a tree order $<_\ell$ on $W_{\leq \ell} = \bigcup_{\ell \leq \ell} W_j$ such that

- if $i <_\ell j$ then $a_i, a_j$ realize the same $\Delta^m$-type over $\{a_i : i <_\ell j\}$
- if $\neg(i <_\ell j)$ and $\neg(j <_\ell i)$ then $a_j, a_i$ realize different $\Delta^m$-types over $\{a_s : s <_\ell i, s <_\ell j\}$
Call any such tree a type tree. That is, a type tree $W_s$ is given by a tree ordering $<_s := \bigcup_\ell <_\ell$ on $u_m$ which satisfies the above conditions.

**Step 1: Choosing a branch through the tree suffices, i.e. (II) of the induction.**

In this step, we verify that any branch through a type tree, i.e. any maximal subset linearly ordered by $<_s$ will satisfy the inductive hypothesis on indiscernibility. (Step 2 will show it satisfies the inductive hypotheses on size.) The key is that in every branch, the type does not depend on the last element.

More precisely, suppose $i_0 < \cdots < i_{k-1}$, $j_0 < \cdots < j_{k-1}$ are from $u_{m+1}$. $\bigwedge_\ell (\ell < k - m - 1 \implies i_\ell = j_\ell)$, and $\varphi \in \Delta$. (As we had built the tree $W_s$ by induction, $<_s$ implies $< \text{ in the sense of the order of the original sequence.})$ Without loss of generality, suppose $i_{m-1} < j_{m-1}$. Then, recalling the parameters used in the definition of $\Delta^m$,

$$\varphi(a_{i_0}, \ldots, a_{i_{k-1}}) <\equiv \varphi(a_{i_0}, \ldots, a_{i_{m-1}}, a_{\ell_{m-m}}, \ldots, a_{\ell_{m-1}})$$

by inductive hypothesis, since the first $m$ indices agree.

$$\varphi(a_{i_0}, \ldots, a_{i_{m-1}}, a_{\ell_{m-m}}, \ldots, a_{\ell_{m-1}}) <\equiv \varphi(a_{i_0}, \ldots, a_{i_{m-2}}, a_{j_{m-1}}, a_{\ell_{m-m}}, \ldots, a_{\ell_{m-1}})$$

since by construction $a_{i_{m-1}}, a_{j_{m-1}}$ realize the same $\Delta^m$-type over $a_{i_0}, \ldots, a_{i_{m-2}}$ (again, recall the parameters used), and finally

$$\varphi(a_{i_0}, \ldots, a_{j_{m-1}}, a_{\ell_{m-m}}, \ldots, a_{\ell_{m-1}}) <\equiv \varphi(a_{j_0}, \ldots, a_{j_{k-1}})$$

by inductive hypothesis. We have verified that

$$\varphi(a_{i_0}, \ldots, a_{i_{k-1}}) <\equiv \varphi(a_{j_0}, \ldots, a_{j_{k-1}})$$

based only on the assumption that the first $m - 1$ indices coincide, which completes the inductive step. Having established that a branch through the tree $W_s$ will give condition (II) for the inductive step, we turn to computing a lower bound on the size of a branch.

**Step 2: Lower bounds on the length of a branch through $W_s$, i.e. (I) of the induction.** As we have established that any branch through a type tree $W_s$ would suffice for the inductive hypothesis (II), we now calculate how long a branch is expected to be. That is, we establish a lower bound on the length of a branch by showing that a maximally branching type tree with $|u_m|$ nodes must attain a certain height $h$.

This step will be split up into several parts. The key constraints are the decreasing rank of $2A$ and the bound on branching in $2B$. The remaining parts are computational.

**Step 2A: Partitioning the nodes of the tree using stability rank.** Let $t = R(x = x, \Delta^m, 2)$ where $R$ is the stability rank; then by definition of this rank, we cannot embed $t+1/2$ in $W_s$. For $s \leq t$ let

$$S^s = \{i \in W_s : \text{above } i \text{ in the tree we can embed } s \}$$

$W_s$ will be the disjoint union of $\{S^s : s \leq t\}$, and if $i_1 \in S^1 \land i_2 \in S^2 \land i_1 \leq_s i_2$ then $s_1 \geq s_2$.

Thus every node in the tree is assigned a “rank” i.e. a value in $\{0, \ldots, t\}$. Rank is non-increasing with order, i.e. if the rank of $i$ is $s$ then all nodes in the tree above $i$ have rank $\leq s$. Moreover, by definition, each node $i$ has at most one immediate successor of the same rank.

**Step 2B: A bound on branching.** By condition (1) of the theorem, if $i$ is a node in some type tree and $i$ has height $h$, then the number of immediate successors of $i$ is at most $(h + m)^r$. This is because by definition of type tree, any two distinct immediate successors must satisfy distinct types over their common initial segment, which includes all elements $j$ with $j \leq_s i$ as well as the $m$ constants named as part of $\Delta^m$.

Note that the shortest tree is attained when branching is maximal, i.e. when for each node $i$ of rank $s$, (a) $i$ has exactly one immediate successor of rank $s$ and (b) all of its other immediate
successors have rank \( s - 1 \). This is because any two distinct successors of \( i \) each of rank \( s \) must lie along the same branch, and the rank assignment is nonincreasing with \( \leq s \).

For the remainder of Step 2 we will assume maximal branching. This will simplify notation.

Step 2C: Counting nodes of a given rank and height assuming maximal branching. Let \( N^s_\ell \) denote the number of nodes in the tree of rank \( s \) and height \( \ell \), i.e. \( N^s_\ell = |\{ i : i \in S^s, \text{ht}(i) = \ell \}| \). We can further write \( N^s_\ell = X^s_\ell + Y^s_\ell \), where \( X^s_\ell \) counts the nodes in \( N^s_\ell \) whose immediate predecessor also has rank \( s \), and \( Y^s_\ell \) counts the nodes in \( N^s_\ell \) whose immediate predecessor has rank \( s + 1 \). By step 2B, assuming maximal branching, these values satisfy the following inequalities:

(i) For all \( s \leq t \) and all \( \ell \), \( X^s_{\ell+1} \leq N^s_\ell \)

(ii) For all \( s < t \) and all \( \ell \), \( Y^s_{\ell+1} \leq N^{s+1}_\ell \cdot (\ell + m)^r \)

(iii) Thus for all \( s < t \) and all \( \ell \), \( N^s_{\ell+1} \leq N^s_\ell + N^{s+1}_\ell \cdot (\ell + m)^r \)

(iv) For all \( 1 \leq s \leq t \), \( N^{t-s}_0 = 0 \).

(v) For all \( \ell \), \( N^1_\ell \leq 1 \).

Step 2D: An inductive bound. We now observe by induction on \( s \) that \( N^{t-s}_{\ell+1} \leq (\ell + m)^{s(r+1)} \).

For the base case \( s = 1 \), by a nested induction on \( \ell \), \( N^{t-1}_{\ell+1} \leq \sum_{j \leq \ell} (j + m)^r \leq (\ell + m)^{r+1} \) using (iii) and (v) of (2C). Then for \( s + 1 \), by a nested induction on \( \ell \) using (iii) and (iv) of (2C)

\[
N^{t-s}_{\ell+1} + N^{t-s}_{\ell} \cdot (\ell + m)^r \leq \sum_{j \leq \ell} (\ell + m)^{s(r+1)} \cdot (j + m)^r \leq (\ell + m)^{(s+1)(r+1)}
\]

and since the left hand side bounds \( N^{t-s}_{\ell+1} \) by (iii) this completes the inductive step \( s + 1 \).

We may now bound the total number \( N_{\ell+1} \) of nodes of height \( \ell + 1 \):

\[
N_{\ell+1} \leq \sum_{s \leq \ell} N^{t-s}_{\ell+1} \leq \sum_{s \leq \ell} (\ell + m)^{s(r+1)} \leq t(\ell + m)^{t(r+1)}
\]

Note \( N_0 = 1 \), so the total number of nodes in a tree of height \( h \) is thus at most:

\[
N = 1 + \sum_{\ell < h} N_{\ell+1} \leq 1 + \sum_{\ell < h} t(\ell + m)^{t(r+1)} < t(h + m)^{t(r+1)+1}
\]

Step 2D: Concluding that (II) holds. Since \( r, t \) are fixed as a function of \( \Delta \) and \( m \leq k \), the bound from Step 2C shows that whenever \( h \) is such that

\[
t(h + m)^{r+1} < |u_m|
\]

then no type tree of height \( h \) can exhaust the nodes of \( u_m \), and thus any such type tree must have a branch of length at least \( h + 1 \). In particular, this inequality will hold when \( h < f(|u_m|) \) where

\[
f(x) = \left| \frac{x}{t+1} - k \right|
\]

For uniformity, we subtract \( k \) instead of \( m \). This completes the inductive step.

Thus in \( k \) steps we extract a sequence of indices \( u \) for an indiscernible sequence. The size of \( u \) will be at least \( f^k(n) \) where \( f(x) = \left| \frac{x}{t+1} - k \right| \). Recall that \( r, t, k \) depend only on the set of formulas \( \Delta \): \( r \) is from (1) of the theorem, \( t \) is the \( R \)-rank of \( \Delta \) and \( k \) is the number of free variables.

(3) The “in particular” clause gives a simpler, though less accurate, bound, without loss of generality that \( r \geq 2 \). That is, when \( x \) is sufficiently large relative to \( t, r, k \), say \( x > \alpha = \alpha(r, t, k) \) the function \( f^k(x) \geq g^k(x) \) where \( g(x) = x^{\frac{1}{t+1}} \). Specialized to the set of formulas \( \Delta_k \), from Definition
2.9, this implies there is a constant $c = c(\Delta_{k_s}) = c(r, t, k_s) \leq \alpha(r, s, k_s)$ such that we can extract from any set of size $n_1 > (cn_2)^{(2r)^{k_s}}$ a $\Delta_{k_s}$-indiscernible subsequence of size at least $n_2$.

This completes the proof of the theorem. \hfill \Box

We now return to building a regularity lemma. From Claim 3.2 we know how individual elements interact with indiscernible sequences. The next observation shows a uniformity to the individual decisions made by elements in an indiscernible sequence.

Observation 3.6. (Crucial Observation) Suppose that $A = (a_i : i < s_1)$, $B = (b_j : j < s_2)$ are $\Delta_{k_s}$-indiscernible sequences. Suppose that $s_1 \geq 2k_s$ and $s_2 > (2k_s)^2$.

Let $U = \{i < s_1 : \exists_{2k_s} j < s_2|(a_j, Rb_i)\}$. Then either $|U| \leq 2k_s$ or $|U| \geq s_1 - 2k_s$.

Proof. Suppose the conclusion fails. Let $i_1$ be the $k_s$th member of $U$, and let $i_2$ be the $k_s$th member of $\{0, \ldots s_1 - 1\} \setminus U$. Clearly $i_1 \neq i_2$.

Case 1: $i_1 < i_2$. Choose elements $j_0 < \cdots < j_{k_s - 1}$ from $U$ and elements $j_{k_s} < \cdots < j_{k_s + k_s - 1} < s_1$ from $\{0, \ldots s_1 - 1\} \setminus U$ satisfying $j_{k_s - 1} \leq i_1 < i_2 \leq j_{k_s}$. Recall by Claim 3.2 that each $a_{j_{k_s}}$ partitions $B$ into a small and large set; for each $\ell < 2k_s$, let the “small set” be

$$W_\ell = \{i < s_2 : a_{j_{k_s}}Rb_i \leftrightarrow (\exists_{2k_s} j < s_2)\}$$

By Claim 3.2 and the definition of $U$, each $|W_\ell| < 2k_s$. Thus

$$\bigcup_{\ell < 2k_s} W_\ell \leq (2k_s)^2 < |B|$$

Choose $n \in \{0, \ldots s_2 - 1\} \setminus \bigcup_{\ell < 2k_s} W_\ell$. Then for all $\ell$ such that $0 \leq \ell \leq k_s - 1$, $b_nR\ell$ and for all $\ell$ such that $k_s \leq \ell < 2k_s$, $-b_nR\ell$. By Observation 2.10, $G$ has the $k_s$-order property, contradiction.

Case 2: $i_2 < i_1$. Similar, interchanging $R$ and $-R$. \hfill \Box

Conclusion 3.7. Recall the hypotheses of this section: $G$ is a finite graph with the non-$k_s$-order property.

If (A) then (B).

(A) (1) $n_1 \rightarrow (n_2)_T, \Delta_{k_s, 1}$

(2) $n > n_1n_2$ and $n_2 \geq (2k_s)^2$

(B) if $A \subseteq G$, $\vert A \vert = n$, then we can find $\overline{A}$, $m_1, m_2$ such that:

(a) $\overline{A} = \langle A_i : i < m_1 \rangle$

(b) $\overline{A}$ is a partition of $A$

(c) $n = n_2m_1 + m_2$, $m_2 \leq n_1 \leq m_1$

(d) For each $i$, $\vert A_i \vert \in \{n_2, n_2 + 1\}$

(e) Each $A_i$ is either a complete graph or an empty graph (after possibly omitting one element)

(f) If $i \neq j < m_1$, then (after possibly omitting one element of $A_i$ and/or $A_j$), for some truth value $t(A_i, A_j) \in \{0, 1\}$, we have that for all but $\leq 2k_s$ $a \in A_i$, for all but $\leq 2k_s$ $b \in A_j$, $aRb \equiv t(A_i, A_j)$.

Proof. First, choose $m_1$ satisfying $n_2m_1 \leq n < n_2m_1 + n_1$ (so $m_1 \geq n_1$ by (A)(2)), and let $<_s$ be a linear order on $A$. Second, we choose $A'_\ell$ by induction on $\ell < m_1$ to satisfy:

- $A'_\ell \subseteq A \setminus \{A'_j : j < \ell\}$
- $\vert A'_\ell \vert = n_2$
- if we list the elements of $A$ in $<_s$-increasing order as $\langle a_{\ell,i} : i < n_2 \rangle$, this is a $\Delta_{k_s}$-indiscernible sequence.
The existence of such \( A'_\ell \) is guaranteed by the hypothesis (A)(1). Since \( \Delta_{k_*} \) includes \( \{xRy\} \), we will have (e) by the symmetry of \( R \).

Third, let \( \{a'_i : i < m_2\} \) list the remaining elements, i.e. those of \( A \setminus \bigcup \{A'_\ell : \ell < m_1\} \). Let \( A_\ell := A'_\ell \cup \{a'_i\} \) if \( a'_i \) is well defined and \( A_\ell := A'_\ell \) otherwise. Condition (c) ensures there is enough room. This takes care of (a)-(e).

In Condition (f), we may want to delete the extra vertex added in the previous paragraph. Then the Crucial Observation 3.6 applied to any pair \((A_i, A_j)\) gives our condition, i.e. it shows that if we choose an element \( a \in A_i \) (provided we did not choose one of the at most \( 2k_* \) exceptional points) and then subsequently choose an element \( b \in A_j \) (all but at most \( 2k_* \) of them are good choices) we find that \( a, b \) will relate in the expected way. This completes the proof.

**Theorem 3.8.** Let \( k_*, n_2 \) be given with \( n_2 > (2k_*)^2 \). Then there is \( N = N(n_2, k_*) \) such that any finite graph \( G, |G| > N \) with the non-\( k_* \)-order property admits a partition \( G = \langle G_i \rangle \) into disjoint pieces \( G_i \) which satisfies:

1. for each \( G_i \in \overline{G}, |G_i| \in \{n_2, n_2 + 1\} \)
2. (after possibly omitting one element) each \( G_i \) is either a complete graph or an empty graph
3. for all pairs \( G_i, G_j \in \overline{G} \) (after possibly omitting one element from each) there exists a truth value \( t(G_i, G_j) \in \{0, 1\} \) such that for all but \( \leq 2k_* \) \( a \in G_i \), for all but \( \leq 2k_* \) \( b \in G_j \), \( aRb \equiv t(G_i, G_j) \).

Moreover, \( N = n_1n_2 \) suffices for any \( n_1 > (cn_2)^{(2\tau_*)^+} \), as computed in Theorem 3.5 in the case where \( \Delta = \Delta_{k_*} \) (for a constant \( c \) depending only on \( \Delta_{k_*} \)).

**Proof.** By Conclusion 3.7 and Theorem 3.5.

**Remark 3.9.**

1. Note that clause (f) of Conclusion 3.7 is stronger than the condition of \( \varepsilon \)-regularity in the following senses.
   - It is clearly hereditary for \( C_i \subseteq A_i, |G_i| \geq |2k_*|^2 \).
   - The density of exception is small:
     \[
     \left| \{(a, b) \in A_i \times A_j : (aRb) \equiv \neg t_{i,j}\} \right| \leq \frac{2k_*}{|A_i|} + \frac{2k_*}{|A_j|}
     \]
   - If \( |A_i|, |A_j| \) are not too small, \( t_{i,j} = t_{j,i} \).
   - If we weaken the condition that \( \bigcup_i A_i = A \) to the condition that \( |A \setminus \bigcup_i A_i| \leq m_2 \), we can omit the exceptional points. It may be better to have \( |A_i| \in \{n_2, 1\} \) with \( |i : |A_i| = 1| < n_1 \).

2. As for the hypotheses (A)(1)-(2) of the theorem: although Theorem 3.5 will not apply outside the stable case, some extensions to the wider class of dependent theories are discussed in Section 3.2, e.g. Claim 3.18.

### 3.1. Generalizations

Some natural directions for generalizing these results would be the following. First, as stated, the Szemerédi condition is not a priori meaningful for infinite sets, while the condition (f) from Conclusion 3.7 is meaningful, and we can generalize the results of this section replacing \( n, n_1, n_2, k \) by infinite \( \lambda_0, \lambda_1, \lambda_2, \kappa \). Second, we can allow \( G \) to be a directed graph and replace \( \{xRy\} \) with a set \( \Phi \) of binary relations satisfying the non-\( k_* \)-order property. Third, we can replace 2-place by \( m(*) \)-place where \( m(*) \leq \omega \), so \( \Phi \) is a set of formulas of the form \( \varphi(x_0, \ldots, x_{n-1}) \), \( n < m(*) \). In this case, the assumption (A)(1) of Conclusion 3.7 becomes

\[
\lambda_1^+ \rightarrow T, \Delta \ \lambda_2^+
\]

which can be justified by appeal to one of the following:

1. by Ramsey: \( \lambda_1 \rightarrow (\lambda_2)^{<m(*)}_{\kappa, \Delta} \)
2. by Erdős-Rado, similarly
(3) using Erdős cardinals
or
(4) use stability: [15] Chapter II, or better (in one model) [18] §5.
Finally, it would be natural to consider extending the results above to hypergraphs.

3.2. Remarks on dependent theories. In this brief interlude we discuss some extensions of §3 to
the more general class of theories of dependent graphs, Definition 3.10. In subsequent sections, we
return to stable theories of graphs. Although, as discussed in the introduction, the order property
is enough to cause irregularity, many of the properties considered in this paper are applicable to
dependent graphs, e.g. Claim 2.6 and Fact 5.10.

Dependent theories (theories without the independence property, see below) are a rich class
extending the stable theories (theories without the order property), and have been the subject of
recent research, see e.g. [17] and [5]. From the point of view of graph theory and combinatorics, the
Vapnik-Chervonenkis connection [10] makes this a particularly interesting class. Below, we indicate
how bounds on alternation can be used to easily deduce a weaker analogue of Theorem 3.8.

Definition 3.10. Let \( k_\ast < \omega \) be given. We say that \( G \) is \( k_\ast \)-dependent when there are no \( a_\ell \in G \)
(for \( \ell < k_\ast \)) and \( b_\ell \in G \) (for \( u \subseteq k_\ast \)) such that \( a_\ell R b_\ell \) iff \( \ell \in u \).

Hypothesis 3.11. \( G \) is an ordered graph (or a graph) meaning that it is given by an underlying
vertex set on which there is a linear order \( \prec^G \), along with a symmetric binary edge relation \( R \). We
assume that \( G \) is \( k_\ast \)-dependent.

Remark 3.12. Stable implies dependent, i.e. if \( xRy \) does not have the order property it will not
have the independence property; but the reverse is not true. More precisely, the formula \( xRy \) has
the order property if, for every \( n < \omega \), there exist elements \( a_0, \ldots, a_n \) such that for all \( i \leq n \),

\[
\models (\exists x) \left( \bigwedge_{j \leq i} \neg(xRa_j) \land \bigwedge_{j > i} xRa_j \right)
\]

(note this definition remains agnostic about the existence of an \( x \) connected to some partition out
of order) whereas the formula \( xRy \) has the independence property (is not dependent) if, for every
\( n < \omega \), there exist elements \( a_0, \ldots, a_n \) such that for all \( u \subseteq n \),

\[
\models (\exists x) \left( \bigwedge_{j \in u} \neg(xRa_j) \land \bigwedge_{j \notin u} xRa_j \right)
\]

We first discuss the results of §3. On the infinite partition theorem for dependent \( T \) (existence
of indiscernibles), see [7] for negative results, and [16] for positive results.

Notation 3.13. (1) If \( A \subseteq G \) then let \( \text{mem}_A(\ell) \) be the \( \ell \)th member of \( A \) under \( \prec^G \), for \( \ell < |A| \)
(for infinite \( A \) such that \( \prec^G \mid A \) is well ordered, \( \ell < (\text{otp}(A), \prec^G \mid A), \) an ordinal)
(2) If \( A \subseteq G \), \( \prec^G \) is a well-ordering of \( A \) then \( \text{mem}_A(\ell) \) is defined similarly.

Definition 3.14. (Compare Observation 3.6.)
(1) We say a pair \( (A, B) \) of subsets of \( G \) is semi-\( f \)-nice (for \( f = (f_1, f_2) \) when:
for all but \( f_1(|A|) \) members of \( A \),
for all but \( f_2(|B|) \) numbers \( \ell < |B| \) (or \( \text{otp}(B) \)), we have that
\( aR\text{mem}_B(\ell) \equiv aR\text{mem}_B(\ell + 1) \).
(2) (restated for clarity:) If \( f = (c_1, c_2) \) where \( c_1, c_2 \) are constants, then in (1) replace the
condition “all but \( f_i(|A|) \) members of \( A \)” with “all but \( c_i \) members of \( A \) for \( i = 1, 2 \).
(3) We say a pair \( (A, B) \) of subsets of \( G \) is \( f \)-nice when \( (A, B) \) and \( (B, A) \) are both semi-\( f \)-nice.
If $G$ is just a graph then the above $A, B$ should be replaced by $(A, <_A), (B, <_B)$.

**Definition 3.15.** Let $\Delta_k = \{\varphi_0(x_0, \ldots, x_{k-1}) : \eta \in k, 2\}$ where
\[\varphi_\eta(x_0, \ldots, x_{k-1}) = (\exists y) \bigwedge_{i < k} (x_i R y) \text{ if } \eta(i) = 1\]

**Claim 3.16.** Suppose $A, B \subset G$ are disjoint, both $A, B$ are $\Delta_k$-indiscernible sequences, and $|A| \geq 2k, |B| \geq 2k$. Then $(A, B)$ is $(1, k)$-nice.

**Proof.** Suppose not, so without loss of generality $(A, B)$ is not semi-$(1, k_*)$-nice. So there is $a \in A$ which “alters” $k$, times on $B$. That is, we may choose a $<^G$-increasing sequence of elements $b_0, \ldots, b_{2k} \subseteq B$ such that $aRb_j$ (for $j$ even) and $\neg aRb_j$ (for $j$ odd). Let us verify that this means $G$ is $k_*$-dependent. Let $J = \langle j_0, j_1, \ldots, j_{k-1} \rangle$ be any set of indices of elements of $B$, of size $k_*$. Then for any $\sigma \subseteq J$, by indiscernibility, we have that
\[|\sigma| \leq \exists x \left( \bigwedge_{i \in \sigma} x Ra_{j_i} \land \bigwedge_{i \in J \setminus \sigma} \neg x Ra_{j_i} \right)\]

since this is true when the appropriate increasing sequence of $k_*$-many indices (corresponding to the pattern of membership in $\sigma$) is chosen from among $i_0, \ldots, i_{2k-1}$.

**Definition 3.17.** Let $\overline{A} = \{A_i : i < i(*)\}$ be a set of subsets of $A \subset G$ (usually pairwise disjoint).

1. We say that $\overline{A}$ is semi-$f$-nice when $i < j < i(*)$ implies $(A_i, A_j)$ is semi-$f$-nice.
2. We say that $\overline{A}$ is $f$-nice when $i < j < i(*)$ implies $(A_i, A_j)$ is $f$-nice.

**Claim 3.18.** Assume that
1. $G$ is an ordered graph and is $k_*$-dependent
2. $m_1 \rightarrow (m_2)^{\leq k_2} \Delta_{k_2}$ in the sense of Ramsey’s theorem
3. $A \subset G$, $|A| = n$

Then we can find $\langle A_i : i < i(*)\rangle$ such that:
1. $|A_i| = m_2$ for all $i$
2. the $A_i$s are pairwise disjoint
3. each $A_i \subseteq A$ and is either complete or edge free
4. $B = A \setminus \bigcup\{A_i : i < i(*)\}$ has $m_1$ members
5. each $A_i$ is $\Delta_{k_*}$-indiscernible
6. $\overline{A}$ is $(1, k_*)$-nice

**Proof.** Straightforward.

**Conclusion 3.19.** Continuing with the notation of Claim 3.18, let $\langle A_i : i < i(*)\rangle$ be the partition of $A \subset G, |A| = n$ into pieces of size $m_2$ obtained there. Suppose that $m_3 | m_2$, and suppose that we divide each $A_i$ into convex intervals $A_{i,j}$ each of length $m_3$. If $m_2 > (m_3)^2$ then we obtain an equitable partition into $\frac{n}{m_2}$ pieces in which, moreover, for any $i_1 \neq i_2 < i(*)$,
1. $\{\langle j_1, j_2 \rangle : (A_{i_1,j_1}, A_{i_2,j_2}) \text{ is neither full nor empty}\}$ has cardinality $\leq \frac{m_2}{m_3} \times m_3 \times k_* = k_* m_2$
   (in each pair, each element’s alternations can be seen in at most $k_*$ other pairs)
2. $\left| \{\langle i_1, j_2 \rangle : j_1, j_2 < \frac{m_2}{m_3} \} \right| = \left( \frac{m_2}{m_3} \right)^2$
3. so the density of bad pairs is $\leq \frac{k_* m_2^2}{m_2}$

On the other hand, the density of pairs with $i_1 = i_2$ is $\leq \frac{m_2^2}{n^2}$.

**Remark 3.20.** Here we obtain quite small pieces (coming from Ramsey’s theorem) and there are exceptional pairs; but for the regular pairs there are no exceptions.
4. ON THE BOUNDS

In this section, we take a different approach, aimed at improving the bounds on the number of components. First, in a series of claims, we give conditions for partitioning a given graph with the non-\( k_* \)-order property into disjoint \( \epsilon \)-indivisible sets (Definition 4.2), and show when such sets interact uniformly. However, the procedure for extracting such sets does not ensure uniform size (Discussion 4.11). We solve this in two different ways. The first (probabilistic) approach, resulting in Theorem 4.16, gives a partition in which there are irregular pairs, but the “regular” pairs have no exceptional edges. The second, resulting in Theorem 4.23, proceeds by first proving a combinatorial lemma 4.19 which allows us to strengthen the “indivisibility” condition to one in which the number of exceptions is constant; thus in Theorem 4.23, there are no irregular pairs, at the cost of a somewhat larger remainder.

As mentioned in the introduction, one recurrent strategy in this paper is partitioning a given graph into “indivisible” components; compare Definition 4.2 with Definition 5.2. Reflecting the strength of Definition 4.2, the number of pieces in each of the two partition theorems of this section grows with the size of the graph, as in Theorem 3.8. In Section 5, under the weaker Definition 5.2, the number of pieces in the partition will be a constant \( c = c(\epsilon) \) as in the classical Szemerédi result.

**Hypothesis 4.1.** Throughout \( \S 4 \), we assume: (1) \( G \) is a finite graph. (2) \( G \) has the non-\( k_* \)-order property, and so \( k_* \) is the corresponding tree-height bound from Definition 2.11. (3) By convention \( f, g \) are nondecreasing functions from \( \mathbb{N} \) to \( \mathbb{N} \setminus \{0\} \).

**Definition 4.2.** (\( \epsilon \)- and \( f \)-indivisible)

1. Let \( \epsilon \in (0,1]_{\mathbb{R}} \). We say that \( A \subseteq G \) is \( \epsilon \)-indivisible if for every \( b \in G \), for some truth value \( t \), the set \( \{ a \in A : aRb \equiv t \} \) has \(< |A|^{\epsilon} \) members.

2. In general, we say that \( A \) is \( f \)-indivisible (where \( f : \omega \to \omega \)) if for any \( b \in G \), there exists a truth value \( t \) such that \(|\{ a \in A : aRb \neq t \}| < f(|A|) \). By convention in this section, we assume that \( f \) is nondecreasing.

**Claim 4.3.** Assume that \( m_0 > \cdots > m_{k_*} \) is a sequence of nonzero natural numbers and for all \( \ell < k_* \), \( f(m_\ell) \geq m_{\ell+1} \) (e.g. \( f(n) = n^\epsilon \)). If \( A \subseteq G \), \( |A| = m_0 \) then for some \( \ell < k_* \) there is an \( f \)-indivisible \( B \in |A|^{m_\ell} \).

**Proof.** Suppose not. So we will choose, by induction on \( k \leq k_* \), elements \( \langle b_\eta : \eta \in k \rangle \) and \( \langle A_\eta : \eta \in k \leq 2 \rangle \) such that:

1. \( A_0 = A \)
2. \( A_{\eta^{-1}(i)} \subseteq A_\eta \)
3. \( A_{\eta^{-1}(0)} \cap A_{\eta^{-1}(1)} = \emptyset \)
4. \( |A_\eta| = m_{\lg(\eta)} \)
5. \( b_\eta \in G \)
6. \( A_{\eta^{-1}(i)} = \{a \in A_\eta : aRb_\eta \equiv (i = 1)\} \)

There is no problem at \( k = 0 \), but let us verify that the induction cannot continue past \( k_* \).

For each \( \eta \in k_2 \), \( A_\eta \neq \emptyset \) by (4), so choose \( a_\eta \in A_\eta \). If for all \( \eta \in k_*+2 \) there exists \( b_\eta \) such that \( A_{\eta^{-1}(1)}, A_{\eta^{-1}(2)} \) are defined and satisfy the conditions, then the sequences \( \langle a_\eta : \eta \in k_*+2 \rangle \) and \( \langle b_\eta : \eta \in k_*+2 \rangle \) contradict the choice of \( k_* \), Definition 2.11. So for at least one \( \eta \), it must be that no such \( b_\eta \) can be found in \( G \), i.e. that for any \( b \in G \), either \(|\{ a \in A_\eta : aRb \}| \) or \(|\{ a \in A_\eta : aRb \}| \) is less than \( m_{\ell+1} \), for \( \ell = \lg(\eta) \). Let \( B = A_\eta \), so \( |B| = m_\ell \) and \( B \) is \( f(|B|) \)-indivisible, which completes the proof.

**Claim 4.4.** Assume \( f, \langle m_\ell : \ell \leq k_* \rangle \) are as in Claim 4.3. For any \( A \subseteq G \), we can find a sequence \( \langle A_j : j < m \rangle \) such that:


(a) For each j, A_j is f-indivisible
(b) For each j, |A_j| \in \{m_\ell : \ell \leq k_\ast\}
(c) A_j \subseteq A \setminus \bigcup\{A_i : i < j\}
(d) A \setminus \bigcup\{A_j : j < m\} has < m_0 members

Proof. We choose A_j by induction on j to satisfy (a)+(b)+(c). If |A| < m_0 we are trivially in case (d). By Claim 4.3, we can continue as long as there are at least m_0 elements remaining.

Claim 4.5. Assume \epsilon \in (0, 1/2)_\mathbb{R}, n^{k_\ast} > k_\ast. Let \langle m_\ell : 0 \leq \ell \leq k_\ast\rangle be a sequence of integers satisfying n \geq m_0, m_{k_\ast} > k_\ast and for all \ell s.t. 0 \leq \ell \leq k_\ast, m_{\ell+1} = \lceil(m_\ell)^\epsilon\rceil.

If A \subseteq G, |A| = n then we can find \overline{A} such that:
1. \overline{A} = \langle A_i : i < i(*)\rangle is a sequence of pairwise disjoint sets
2. \langle |A_i| : i < i(*)\rangle is \leq\text{-increasing
3. for each i < i(*) for some \ell = \ell(i) < k_\ast, |A_\ell| = m_\ell and A_\ell is \epsilon-indivisible
4. A \setminus \{A_i : i < i(*)\} has < m_0 elements

Proof. By Claim 4.4, using f(n) = n^\epsilon and renaming the sets A_i so that clause (3) holds.

The next claim says that for all sufficiently indivisible pairs of sets, averages exist (notice there is a potential asymmetry in the demand that B be large).

Claim 4.6. Suppose A is f-indivisible, B is g-indivisible and and f(|A|) \cdot g(|B|) < 1/2|B|. Then for some truth value t = t(A,B) for all but < f(|A|) of the a \in A for all but < g(|B|) of the b \in B, we have that aRb \equiv t.

Proof. Similar to the proof of Observation 3.6 above. For each a \in A there is, by g-indivisibility of B, a truth value \tau_a = t_a(a,B) such that |\{b \in B : aRb \equiv \tau_a\}| < g(|B|). For i \in \{0,1\}, let U_i = \{a \in A : \tau_a = i\}. If |U_i| < f(|A|) for either i, we are done, so assume this fails. Choose W_i \subseteq U_i so that |W_i| = f(|A|) for i \in \{0,1\}. Again we gather the exceptions: let V = \{b \in B : (\exists a \in W_1)(\neg aRb) \lor (\exists a \in W_0)(aRb)\}. Then |V| \leq |W_1| + |W_0| g(|B|) < |B| by hypothesis, so we may choose b_* \in B \setminus V. But then a \in W_1 \implies b_*Ra and a \in W_0 \implies \neg b_*Ra, contradicting the f-indivisibility of A.

Remark 4.7. When f(n) = n^\epsilon, g(n) = n^\xi the translated condition is: if |A'||B|^\xi < 1/2|B|.

Claim 4.8. Let A be \zeta-indivisible and B be \epsilon-indivisible. Suppose that the hypotheses of Claim 4.6 are satisfied, so averages exist. Then for all \zeta_1 \in (0, 1-\zeta), \epsilon_1 \in (0, 1-\epsilon), we have: if A' \subset A, B' \subset B, |A'| \geq |A|^\zeta_1, |B'| \geq |B|^\epsilon_1, then:

\[\left|\left\{(a,b) \in (A',B') : aRb \equiv \neg t(A,B)\right\}\right| \leq \frac{1}{|A|^\zeta_1} + \frac{1}{|B|^\epsilon_1}.

Proof. To bound the number of exceptional edges, recall that in A, hence also in A', there are at most |A|^\zeta elements which do not have the expected average behavior over B. Likewise, for each non-exceptional a \in A' there are no more than |B|^\xi corresponding exceptional points b \in B'. Thus we compute:

\[\frac{|A|^\zeta \cdot |B'| + (|A'| - |A|)|B|^\epsilon}{|A'||B'|} = \frac{|A|^\zeta}{|A'|} + \left(\frac{|A'| - |A|}{|A'|}\right)\frac{|B|^\epsilon}{|B'|} \leq \frac{|A|^\zeta}{|A'|} + \frac{|B|^\epsilon}{|B'|} \leq \frac{|A|^\zeta}{|A|^\zeta_1} + \frac{|B|^\epsilon}{|B|^\epsilon_1} = \frac{1}{|A|^\zeta_1} + \frac{1}{|B|^\epsilon_1}
\]

A similar result holds for f-indivisible replacing \epsilon-indivisible.

We single out the following special case for Theorem 4.23 below.
Corollary 4.9. Let $A, B$ be $f$-indivisible where $f(n) = c$ is a constant function. Suppose that the hypotheses of Claim 4.6 are satisfied, so averages exist. Then for all $\epsilon_1 \in (0, 1 - \frac{1}{|A|}), \epsilon_1 \in (0, 1 - \frac{1}{|B|})$, we have: if $A' \subset A, B' \subset B$, $|A'| \geq c|A|^{\epsilon_1}, |B'| \geq c|B|^{\epsilon_1}$, then:
\[
\left| \{ (a, b) \in (A', B') : aRb \equiv -t(A, B) \} \right| \leq \frac{1}{|A|^{|\epsilon_1|}} + \frac{1}{|B|^{|\epsilon_1|}}
\]

Returning to the general argument, choosing $\epsilon_1$ (here called $\zeta$) small enough means we can apply Claim 4.8 to any pair of elements from the partition in Claim 4.5:

Claim 4.10. In Claim 4.5, if $\zeta \in (0, e^{k_{***}})$, we have in addition that for every $i < j < i^*(*)$, if $A \subset A_i, |A| \geq |A_i|^{1+\zeta}, B \subset A_j, |B| \geq |A_j|^{1+\zeta}$ and $t_{i,j} = t(A_i, A_j)$ is the associated truth value, then
\[
\left| \{ (a, b) \in (A, B) : aRb \equiv -t_{i,j} \} \right| \leq \frac{1}{|A|^{|\zeta|}} + \frac{1}{|A|^{|\zeta|}} + \frac{1}{|B|^{|\zeta|}}
\]

Proof. By Claim 4.8. Note that the enumeration along with clause (2) of Claim 4.5 (i.e. $|A_i| \leq |A_j|$) ensures $t_{i,j}$ is defined. \qed

Discussion 4.11. In some respects Claim 4.10, applied to the partition of Claim 4.5, is quite strong: (a) There are no irregular pairs. (b) For each pair the number of exceptional edges is very low. On the other hand: (c) There is a remainder $A \backslash \bigcup A_i$, not serious as we can distribute the remaining elements among the existing $A_i$ without much loss, as was done in §3. (d) There is an inherent asymmetry: the result assumes $i < j < i^*(*)$, we have not discussed $j < i < i^*(*)$, but this is also not serious. (e) The cardinalities of the $A_i$ are not essentially constant: this seems more serious.

We give two different resolutions of (e) in the remainder of this section. In Theorem 4.16, we obtain an equitable partition at the price of allowing for irregular pairs. In Theorem 4.23, we obtain much stronger indivisibility conditions on the components and no irregular pairs, at the price of a somewhat larger remainder, Theorem 4.23.

4.1. Towards a proof of Theorem 4.16.

Definition 4.12. Assume that $A, B$ are $f$-indivisible (usually: $\epsilon$-indivisible), disjoint (for notational simplicity), and that $f(A) \times f(B) < \frac{1}{2}|B|$ (so $t(A, B)$ is well defined). Let $m$ divide $|A|$ and $|B|$.

We define a probability space: divide $A$ into $|A|/m$ pieces of size $m (A_i : i < i_A)$ and likewise divide $B$ into $|B|/m$ parts each of size $m, (B_j : j < j_B)$. Call this partition an equivalence relation $E$ on $A \cup B$.

For each $i < i_A, j < j_B$, let $E_{A_i, A_j, m}$ be the event: for all $a \in A_i$, for all $b \in B_j$, $aRb \equiv t(A, B)$.

Claim 4.13. Let $A_i, A_j$ be two sets from the conclusion of Claim 4.5. So $\epsilon \in (0, \frac{1}{2})$, $f(x) = [x^\epsilon]$. Ignoring a minor error due to rounding to natural numbers, suppose that $|A_i| = m \epsilon_n = n^\epsilon \epsilon_n+1, |A_j| = m \epsilon_n = n^{-\epsilon} \epsilon^{-1}$ and $|A_i| \leq |A_j|$. Let $m$ be an integer such that $m$ divides both $|A_i|$ and $|A_j|$, and $m = n^\zeta$ for some $\zeta < e^{k_{***}}$. Choose a random partition of $A_i$ and $A_j$ into pieces of size $m$. Let $A_{i}^{*}, A_{j}^{*}$ be pieces from $A_i$ and $A_j$, respectively, under this partition.

Then $\text{Prob}(E_{A_i, A_j, m}) \geq 1 - e^{-\frac{1}{n^{1+\epsilon}}}$.

Proof. By choice of $A_i, A_j$ we have that $t = t(A_i, A_j)$ is well defined. Let $U_1 = \{ a \in A_i : |\{ b \in A_j : aRb \equiv t \} | \geq |A_j|^\epsilon \}$, and for each $a \in A_i \setminus U_1$, let $U_{2, a} = \{ b \in A_j : aRb \equiv -t \}$. By definition of $t$, $|U_1| \leq |A_i|^\epsilon$ and for each relevant $a$, $|U_{2, a}| \leq |A_j|^\epsilon$.\[16\]
We first consider $A_i^a$. The probability $P_1$ that $A_i^a \cap \mathcal{U}_1 \neq \emptyset$ is bounded by the following:

$$P_1 < \frac{m|\mathcal{U}_1|}{|A_i| - m} \leq \frac{n^\epsilon|A_i|^\epsilon}{|A_i| - m} < \frac{n^{2\epsilon}(n^{\epsilon^a} + 2)}{n^{\epsilon^a + 1}} \leq \frac{1}{n^{\epsilon^a + 1 - \epsilon^a/2}} = \frac{1}{n^{\epsilon^a + 1 - 2\epsilon}} < \frac{1}{n^{\epsilon^a + 1}} \leq \frac{1}{n^{k^a}}.$$ 

Now if $A_i^a \cap \mathcal{U}_1 = \emptyset$ then $|\bigcup_{a \in A_i^a} \mathcal{U}_{2,a}| \leq m|A_j|^\epsilon$. So the probability $P_2$ that we have $A_j^a \cap \bigcup_{a \in A_i^a} \mathcal{U}_{2,a} \neq \emptyset$ is bounded by:

$$P_2 < \frac{m|\bigcup_{a \in A_i^a} \mathcal{U}_{2,a}|}{|A_j| - m} \leq \frac{m \cdot m \cdot |A_j|^\epsilon}{|A_j| - m} \leq \frac{n^{2\epsilon}|A_j|^\epsilon}{|A_j| - n^\epsilon} \leq \frac{1}{n^{k^a}}.$$ 

by the analogous calculation. So $\text{Prob}(\mathcal{E}_{A_i^a, A_j^a, m}) \geq (1 - \frac{1}{n^{k^a}})^2 \geq 1 - \frac{2}{n^{2k^a}}$. □

Claim 4.14. Let $(m_\ell : \ell < k^a)$ be a sequence which satisfies the hypotheses of Claim 4.5 and suppose that $m_{\ast}^a$ divides $m_\ell$ for $\ell < k^a$. Let $n$ be sufficiently large relative to $m_{\ast}$: it suffices that $m_\ast < \frac{n}{2\epsilon k^a}$ (see Remark 4.15).

Let $A \subset G$, $|A| = n$ and let $(A_i : i < i(\ast))$ be the partition of $A$ given by Claim 4.5 with respect to the sequence $(m_\ell : \ell < k^a)$ (we will temporarily ignore the remainder of size $\leq m_\ast$). Recall that $t(A_i, A_j)$ is well defined for $i < j$ by Claim 4.6.

Then there exists a partition $(C_i : i < r)$ of $\bigcup_{i < i(\ast)} A_i$ such that:

1. $(C_i : i < r)$ refines the partition $(A_i : i < i(\ast))$
2. $|C_i| = m_{\ast}$ for each $i < r$
3. For all but at most $\frac{2}{n^{k^a}} r^2$ of the pairs $(C_i, C_j)$, there are no exceptional edges: that is, if $i < j$, $C_i \subseteq A_i$, and $C_j \subseteq A_j$, then $\{(a, b) : a, b \in C_i \times C_j : a R b \neq t(A_i, A_j)\} = \emptyset$.

Proof. The potential irregularity in a pair $(C_i, C_j)$ comes from two sources.

(a) The case where $C_i \subset A_i$, $C_j \subset A_j$, $i \neq j$ and $(C_i, C_j)$ contains some exceptional edges. By Claim 4.13 and linearity of expectation, there exists a partition satisfying (1), (2), in which (3) holds when computed on pieces $C_i$, $C_j$ which came originally from distinct components $A_i, A_j$. In fact, this will be true for all but at most $(1 - \frac{1}{n^{k^a}})^2$ of such pairs by the calculation in the last line of Claim 4.13.

(b) The case where $C_i, C_j$ are both from the same original component $A_i$. Here we have no guarantee of uniformity. Let us compute a bound on the fraction of such pairs $C_i, C_j$. The maximum is attained when all of the original components were of maximal size $m_0 = m_{\ast}$; in this case the number of ways of choosing a pair $C_i, C_j$ from the same original component is at most

$$\left(\frac{m_{\ast}}{m_{\ast}} - \frac{n}{m_{\ast}}\right) n \text{ out of } \left(\frac{n}{m_{\ast}}\right)^2$$

so the ratio is approximately

$$\frac{\left(\frac{m_{\ast}}{m_{\ast}}\right)^2 n}{\left(\frac{n}{m_{\ast}}\right)^2} = \frac{m_{\ast}}{n}.$$ 

Recall that by hypothesis, $\frac{m_{\ast}}{n} < \frac{1}{2n^{k^a}}$.

Combining (a) and (b), the total fraction of irregular pairs does not exceed $\frac{2}{n^{k^a}}$. □
Remark 4.15. In Claim 4.14, the hypothesis on $m_s$ could obviously be weakened, or dropped at the expense of increasing the fraction of irregular pairs by $\frac{m_s}{n^c}$. As the calculation in part (b) of proof shows.

Theorem 4.16. Let $\epsilon = \frac{1}{\ell} \in (0, \frac{1}{2})$, $k_s$ and therefore a bound for $k_{ss}$ be given, and suppose $G$ is a finite graph with the non-$k_s$-order property. Let $A \subset G$, $|A| = n$ with $n^{k_{ss}} > k_{ss}$. Then there is $\zeta < e^{k_{ss}}$ and a partition $\langle A_i : i < (s) \rangle$ of $A$ such that:

1. For all $i$, either $|A_i| = \lfloor n^\zeta \rfloor$ or $|A_i| = 1$
2. $|\{ i : |A_i| = 1 \}| \leq n^\epsilon$
3. $\frac{2}{n^{k_{ss}+1}} \geq \frac{1}{(\ell^2)^{n^\epsilon}} |\{(i, j) : |A_i| = 1, |A_j| = 1, \text{ or } (a, b) \in A_i \times A_j : aRb \notin \{A_i \times A_j, \emptyset\}\}|$

Moreover, we may choose $\zeta \geq (1 - 2e^{k_{ss}})e^{k_{ss}+1}$, so the total number of pieces $n^{1-\zeta}$ is at most $n^c$ where $c = c(\epsilon) = 1 - e^{k_{ss}+1} - 2e^{2k_{ss}+1}$.

Proof. Recall that $\epsilon = \frac{1}{\ell}$. (This hypothesis is just to ensure divisibility, and could be modified or dropped in favor of allowing for slight rounding errors.) Choose $m_{ss}$ maximal so that $(m_{ss})^{k_{ss}} \leq n$, and subject to the constraint that $\frac{m_s}{n} < \frac{1}{n^{k_{ss}}}$. (One can drop this constraint, by Remark 4.15, at the cost of increasing the fraction in item (3) by $\frac{m_s}{n}$.) By hypothesis $m_{ss} > k_{ss}$. Then the sequence $\langle m_\ell : \ell < k_{ss} \rangle$ satisfies the hypotheses of Claims 4.3 and 4.5, and furthermore $m_{ss}$ divides $m_\ell$ for $\ell < k_{ss}$. Apply Claim 4.5 to obtain a decomposition into $\epsilon$-indivisible pieces $A_i'$ such that for each $i$ and some $\ell$, $|A_i'| = m_\ell$. Claim 4.14 gives a further partition into pieces $\langle A_j : i < (i) \rangle$ each of size $m_{ss}$; additionally, we partition the remainder from Claim 4.5 into pieces of size 1. Let $\zeta$ be such that $m_{ss} = n^{\zeta}$. This gives clause (1), and clause (2) holds by Claim 4.5(d) since $m_0 = (m_{ss})^{k_{ss}-1}$. Condition (3) holds by Claim 4.13. Finally,

$$n^\zeta = m_{ss} \approx (m_{ss})^{k_{ss}} \approx \left(\frac{n}{n^{2k_{ss}}}\right)^{k_{ss}}$$

Remark 4.17. Though the number of components grows (solved only in \S 5) and this regularity lemma admits irregular pairs, the regular pairs have no exceptional edges.

4.2. Towards a proof of Theorem 4.23. In this subsection we take a different approach, and obtain a regularity lemma in which there are no irregular pairs, at the price of a somewhat larger remainder. The strategy will be to base the partition on a sequence of $c$-indivisible sets, i.e. sets which are $f$-indivisible for a particular constant function $f(x) = c$: such sets will then interact in a strongly uniform way. [Recall from Definition 4.2 that $c$-indivisible for $\epsilon \in (0, 1)\mathbb{R}$ was short-hand for $f$-indivisible when $f(x) = x^\epsilon$: this was the only exception to standard notation, and in particular, $c$-indivisible for $c \in \mathbb{N}$ means $f(x) = c$.] The proof that such sets exist relies on a combinatorial lemma 4.19. To motivate the combinatorial lemma, the reader may wish to first look through the proof of the existence claim, Claim 4.21.

Definition 4.18. For $n, c \in \mathbb{N}$, $\epsilon, \zeta, \xi \in \mathbb{R}$ let $\bigoplus[n, \epsilon, \zeta, \xi, c]$ be the statement:

For any set $A$ and a family $\mathcal{P}$ of subsets of $A$, we have

1. $|A| = n$
2. $|\mathcal{P}| \leq n^{\zeta}, \mathcal{P} \subseteq \mathcal{P}(A)$
3. $\forall B \in \mathcal{P} \left(|B| \leq n^\xi \right)$

then there exists $\mathcal{U} \subseteq A$, $|\mathcal{U}| = \lfloor n^\xi \rfloor$ such that $\forall B \in \mathcal{P} \left(|\mathcal{U} \cap B| \leq c \right)$.

Lemma 4.19. If the reals $\epsilon, \zeta, \xi$ and the natural numbers $n, c$ satisfy:

1. $\epsilon \in (0, 1)$, $\zeta > 0$
Claim 4.21. Suppose that we are given constants $n$ of Lemma 4.19 imply that the two exponents are positive constants, so it is well defined to let $\sum_{[n, \epsilon, \xi, \zeta, c]}$ holds.

We delay the proof until after the next claim. Note that in clause (d) we have that $c > 0$ by (b).

Remark 4.20. In the statement of Lemma 4.19, for “$n$ sufficiently large” it suffices to choose $n$ such that
\[
\frac{1}{n^{1-2\xi}} + \frac{1}{n^{(1-\xi-\epsilon)(e-\xi)}} < 1
\]
See the last displayed equation in the proof of Lemma 4.19. As explained there, the hypotheses of Lemma 4.19 imply that the two exponents are positive constants, so it is well defined to let $n(\epsilon, \xi, \zeta, c)$ be the minimal $n \in \mathbb{N}$ for which the displayed equation is true.

Claim 4.21. Suppose that we are given constants $k, k, c \in \mathbb{N}$ and $\epsilon, \xi, \zeta \in \mathbb{R}$ such that:

1. $G$ is a graph with the non-$k$-order property, thus $k_*$ exists by 2.11
2. $A \subseteq G$ implies $|\{a \in A : aRb\} : b \in G| \leq |A|^k$
3. $\epsilon \in (0, \frac{1}{2})$
4. $\xi \in (0, e^{k_*}/2)$
5. The constant $c$ satisfies
\[
c > \frac{1}{\zeta(1 - \frac{2\xi}{\epsilon^{k_*}} - \epsilon)}
\]

Then for every sufficiently large $n \in \mathbb{N}$ (meaning $n > n(\epsilon, \xi, \zeta, c)$ in the sense of Remark 4.20) if $A \subseteq G$ with $|A| = n$ then there is $Z \subseteq A$ such that

(a) $|Z| = \lfloor n^\epsilon \rfloor$
(b) $Z$ is $c$-indivisible in $G$, i.e. for any $b \in G$ there is $t \in \{0, 1\}$ such that for all but $c$ elements $a \in Z$, $aRb \equiv t$

Remark 4.22. Since $G$ has the non-$k_*$-order property, $k = k_*$ will satisfy condition (2) by Claim 2.6, and recall that $k_*$ is the associated tree bound from 2.11. For lower bounds on the size of $n$, see Remark 4.20.

Proof. (of Claim 4.21) Let $A \subseteq G$, $|A| = n$ be given. For transparency of notation suppose that for each natural number $\ell \leq k_*$, $n^{\epsilon^\ell} \in \mathbb{N}$, and that $n^{\xi} \in \mathbb{N}$. We choose $m_\ell$ by induction on $\ell < k_*$ so that $m_{\ell+1} = \lfloor (m_\ell)^{\epsilon^\ell} \rfloor$. By Claim 4.3 there is $\ell < k_*$ and $A_1 \subseteq A$ such that $|A_1| = m_\ell$ and $A_1$ is $\epsilon$-indivisible. Let $P_1 = \{a \in A : aRb \in G\}$. So $|P_1| \leq |A|^k = (m_\ell)^k$, by choice of $k$.

We would like to apply Lemma 4.19 to conclude that $\sum_{[n, \epsilon, \xi, \zeta, c]}$ holds for $A = A_1$, $P = P_1$.

Let us verify that the hypotheses of Definition 4.18 and Lemma 4.19 hold:

- (1),(2) hold as $|A_1| = m_\ell$, and $|P_1| = (m_\ell)^k = (m_\ell)^{t/\xi}$
- (3) holds by definition of $P_1$, as $A_1$ is $\epsilon$-indivisible
- (a) clear
- (b), (d) by choice of $\xi$ and $c$ in this Claim
- (c) by choice of $n$ “sufficiently large”

We conclude that there is $Z \subseteq A_1$ (i.e. the $U$ guaranteed by Lemma 4.19) which is $c$-indivisible and satisfies
\[
|Z| = \lfloor (m_\ell)^{\xi^\ell} \rfloor = \lfloor (n^{\epsilon^\ell})^{\xi^\ell} \rfloor = \lfloor n^{\xi} \rfloor
\]
which completes the proof.

We now prove Lemma 4.19.
Proof. (of Lemma 4.19) Let \( m = \lfloor n^\xi \rfloor \); this is the size of the set \( \mathcal{U} \) we hope to build.

Let \( \mathcal{F}_* = m^\mathcal{A} \) be the set of sequences of length \( m \) from \( \mathcal{A} \), so \( |\mathcal{F}_*| = n^m \). We will use \( \eta \) for such a sequence and write \( \eta[\ell] \) for the value at the \( \ell \)th place.

Define a probability distribution \( \mu \) on \( \mathcal{F} \subseteq \mathcal{F}_* \) by: \( \mu(\mathcal{F}) = \frac{|\mathcal{F}|}{|\mathcal{F}_*|} \).

We will show that for \( n > n(\epsilon, \zeta, \xi, c) \) in the sense of Remark 4.20, there is nonzero probability that a sequence \( \eta \in \mathcal{F}_* \) satisfies (1) all the elements of \( \eta \) are distinct, i.e. as a set it has cardinality \( m \) and (2) for any \( B \in \mathcal{P} \) there are fewer than \( k \) integers \( t < m \) such that \( \eta[t] \in B \). This will prove the lemma.

We calculate the relevant probabilities in four steps.

\( \circledast_1 \) Verifying some inequalities. By assumption (b) of the Lemma, \( 1 - 2\xi > 0 \) and \( 1 - \xi - \epsilon > 0 \). So by assumption (d) \( (1 - \xi - \epsilon)c - \frac{1}{\xi} > 0 \) and \( c \) is a natural number. We proceed to compute several probabilities.

\( \circledast_2 \) The probability that \( \eta \) is not sequence of distinct elements. (This is bounded by the sum over \( s < t \) of the probability that \( \eta[s] = \eta[t] \): note we don’t mind if this happens for more than one pair.)

\[
\text{Prob} ( (\exists s < t < m)(\eta[s] = \eta[t]) ) \leq \binom{m}{2} \frac{n}{n^2} \leq \frac{m^2}{2n} \leq \frac{n^{\xi^2}}{2n} \leq \frac{1}{2n^{1-\twoxi}} < \frac{1}{n^{1-\twoxi}}
\]

\( \circledast_3 \) The probability that \( \eta \) intersects a given \( B \in \mathcal{P} \) in more than \( c \) places. (For the bound, choose \( c \) indices, then choose \( c \) values for those places from \( B \), over all possible choices of those values.)

Let \( B \in \mathcal{P} \) be given. Then

\[
\text{Prob} ( (\exists c \leq t < m)(\eta[t] \in B) ) \leq \binom{m}{c} \frac{|B|^c}{n^c} \leq \frac{m^c|B|^c}{n^c} \leq \frac{n^{\xi^c}n^c}{n^c} = \frac{1}{n^{(1-\xi-\epsilon)c}}
\]

\( \circledast_4 \) The probability that \( \eta \) intersects some \( B \in \mathcal{P} \) in more than \( c \) places. By \( \circledast_3 \),

\[
\text{Prob} ( (\exists B \in \mathcal{P})(\exists c \leq t < m)(\eta[t] \in B) ) \leq |\mathcal{P}| \cdot \left( \max\{\text{Prob} ( (\exists c \leq t < m)(\eta[t] \in B) ) : B \in \mathcal{P} \} \right)
\]

\[
\leq n^{\frac{\xi}{c} - \frac{1}{n^{(1-\xi-\epsilon)c}}} = \frac{1}{n^{(1-\xi-\epsilon)c - \xi}}
\]

As remarked above, it suffices for the Lemma to show that the sum of the probabilities \( \circledast_2 + \circledast_4 \) is smaller than 1, i.e. that

\[
\frac{1}{n^{1-\twoxi}} + \frac{1}{n^{(1-\xi-\epsilon)c - \xi}} < 1
\]

By \( \circledast_1 \), both exponents are nonzero, and moreover they are constant, so the sum will clearly eventually be smaller than 1.

\( \square \)

**Theorem 4.23.** There is a function \( N : \mathbb{N} \times \mathbb{R}^3 \times (0, 1) \to \mathbb{N} \) such that the following holds. Let \( k_* \) and therefore a bound for \( k_* \) be given. Let \( G \) be a graph with the non-\( k_* \)-order property, and let \( k = k_* \) as in the proof of Claim 4.21.

Then for any \( c \in \mathbb{N} \) and \( \epsilon, \zeta, \xi \in \mathbb{R} \) which, along with \( k \), satisfy the hypotheses of Claim 4.21, any \( \theta \in \mathbb{R} \), \( 0 < \theta < 1 \), and any \( A \subseteq G \), \( |A| = n > N(c, \epsilon, \zeta, \xi, \theta) \), there is \( i(*) \in \mathbb{N} \) and a partition \( \{A_i : i < i(*)\} \) of \( A \) into disjoint pieces (plus a remainder) satisfying:

1. \( |A_i| = \lfloor n^\xi \rfloor \) for each \( i < i(*) \)
2. each \( A_i \) is \( c \)-indivisible, i.e. indivisible with respect to the constant function \( f(x) = c \)
(3) \(|A \setminus \bigcup_{i \in I^*} A_i| \leq \left\lfloor \frac{n^\theta}{\ell^c} \right\rfloor |)

Remark 4.24. Recall that the interaction of any two distinct \(A_i, A_j\) given by this theorem will be highly uniform. Assuming \(n^\theta > 2c\), average types exist in the sense of Claim 4.6, and in particular the calculations of Corollary 4.9 apply.

Proof. (of Theorem 4.23) Assume that \(n\) is large enough so that \(n^\theta > n(\epsilon, \zeta, \xi, c) + 1\), where \(n(\ldots)\) is the lower bound from Lemma 4.19 and Remark 4.20. Note that by choice of \(k, k\) satisfies Claim 2.6.

We are aiming for pieces of uniform size \(n^\theta c\). First, given \(\theta\), define by induction a decreasing sequence \(\langle m_\ell : \ell \leq k_{ss} \rangle\) by \(m_{k_{ss} - 1} = [n^\theta]\), \(m_{k_{ss}} = \left\lfloor (m_{k_{ss} - 1})^c \right\rfloor\) and for each \(1 < j \leq k_{ss}\), \(m_{k_{ss} - j} = \left\lfloor (m_{k_{ss} - j + 1})^{\frac{1}{c}} \right\rfloor\). This sequence, fixed for the remainder of the proof, satisfies the hypotheses of Claim 4.3.

Second, choose a sequence of disjoint \(c\)-indivisible sets \(A_i\) by induction on \(i\), as follows. Let \(R_i\) denote the remainder \(A \setminus \bigcup_{j<i} A_j\) at stage \(i\). Apply Claim 4.3 to \(R_i\), using the decreasing sequence \(\langle m_\ell : \ell \leq k_{ss} \rangle\) just defined, to obtain an \(\epsilon\)-indivisible \(B_i \subseteq R_i\). By construction, for some \(1 \leq \ell \leq k_{ss}\) this set \(B_i\) will have cardinality \(m_{k_{ss} - \ell} = \left\lfloor (n^\theta)^{\frac{1}{c-1}} \right\rfloor\). (Note that \(\epsilon\)-indivisibility need not be preserved under taking subsets.)

By the first line of the proof (recall \(m_{k_{ss} - 1} = [n^\theta]\)), we have \(|B_i| > n(\epsilon, \zeta, \xi, c)\). Apply Claim 4.21 to \(B_i\), using \(c, k, \epsilon, \zeta, \xi\) as given, to extract a \(c\)-indivisible subset \(Z_i\) of size \(|B_i|^c\). That is, \(|Z_i| = |B_i|^c = \left\lfloor n^\theta \cdot \frac{1}{c-1} \right\rfloor\) for some \(1 \leq \ell \leq k_{ss}\). Since the property of being \(c\)-indivisible is preserved under taking subsets, choose \(A_i\) to be any subset of \(Z_i\) of cardinality exactly \([n^\theta \zeta]\). This completes the construction at stage \(i\).

This construction can continue as long as the remainder \(R_i\) has size at least \(m_0 = \left\lfloor \frac{n^\theta}{\ell^c} \right\rfloor\), as required by Claim 4.3; the final remainder will be strictly smaller, which completes the proof. \(\square\)

5. Regularity for stable theories of graphs

Thus far, we have given several regularity lemmas for stable (or, in Section 3.2, dependent) theories of graphs which in some senses improved the classic Szemerédi result, particularly in the “indivisibility” of the components; however, in each case the size of the partition given depended on \(|A|\). In this section, we obtain a partition theorem for any graph \(G\) with the non-\(k_s\)-order property which unilaterally improves the usual result, Theorem 5.18: for each \(\epsilon\), there is \(m = m(\epsilon, k_s)\) such that all sufficiently large \(G\) with the non-\(k_s\)-order property admit an equitable distribution such that (1) there are no irregular pairs, (2) each component satisfies a strong indivisibility condition, called \(\epsilon\)-excellence, and (3) the bounds are much improved. For most of the construction, “regularity” of pairs means \(\epsilon\)-uniformity, Claim 5.6 below; this is useful in our context as the density will be close to 0 or 1. A translation is given in Claim 5.17 and Conclusion 5.19.

This section relies on §2 (Preliminaries) for notation and definitions; nonetheless, definitions will be referenced the first time they are used. Although this section naturally extends the results and strategies of previous sections, it is self-contained and can be read independently.

Hypothesis 5.1. Throughout §5, we assume: (a) \(G\) is a finite graph, (b) for some \(k_s\) fixed throughout this section, \(G\) has the non-\(k_s\)-order property. Definition 2.3 and (c) \(k_s\) is the corresponding bound on the height of a 2-branching tree, Definition 2.11. Throughout this section \(\epsilon, \zeta, \xi\) are reals \(\in (0, \frac{1}{2})\).

Definition 5.2. (Good, excellent)
(1) We say that $A \subseteq G$ is $\epsilon$-good when for every $b \in G$ for some truth value $t = t(b, A) \in \{0, 1\}$ we have $|\{a \in A : (aRb) \neq t\}| < \epsilon|A|$. As $\epsilon < \frac{1}{2}$, this is meaningful.

(2) We say that $A \subseteq G$ is $(\epsilon, \zeta)$-excellent when

- (a) $A$ is $\epsilon$-good and moreover
- (b) if $B \subseteq G$ is $\zeta$-good then for some truth value $t = t(B, A)$, $|\{a \in A : t(a, B) \neq t(B, A)\}| < \epsilon|A|$.

Again, as $\epsilon < \frac{1}{2}$ the average is meaningful. When $\epsilon = \zeta$, we will just write $\epsilon$-excellent.

**Remark 5.3.** We think of $t(b, A)$, $t(B, A)$ as average truth values. Note that any set $A \subseteq G$ satisfying condition (b) for $\epsilon$-excellence must also be $\epsilon$-good, since any singleton set $\{b\}$ is clearly $\epsilon$-good (in fact, $\epsilon$-excellent). Any $B$ which satisfies $(\forall a \in G)\bigvee_{t \in \{0, 1\}} (\forall b \in B)(aRb \equiv t)$ will also be excellent.

The next claim, which will be used repeatedly, gives a way to extract $\epsilon$-excellent subsets of any given $A$ by inductively building a tree whose (full) branching must eventually stop. In the statement of the Claim, Case (II) abstracts from Case (I) by assigning cardinalities $m_\ell$ to the levels of the tree.

**Claim 5.4.** (Crucial claim) Assume $\epsilon < \frac{1}{2^{k_\ast}}$.

(I) For every $A \subseteq G$, $|A| \geq \frac{1}{\epsilon^{k_\ast}}$, there is $A'$ such that:

- (a) $A' \subseteq A$
- (b) $|A'| \geq \epsilon^{k_\ast-1}|A|$
- (c) $A'$ is $\epsilon$-excellent.

(II) Alternately, suppose we are given a decreasing sequence of natural numbers $(m_\ell : \ell < k_\ast)$ such that $\epsilon m_\ell \geq m_{\ell + 1}$ for $\ell < k_\ast - 1$, and $m_{k_\ast - 1} > k_\ast$. Then for every $A \subseteq G$, $|A| \geq \frac{1}{\epsilon^{k_\ast}}$, there is $A'$ such that (a),(b),(c) hold:

- (b) $|A'| = m_\ell$ for some $\ell < k_\ast$
- (c) $A'$ is $\frac{m_{\ell + 1}}{m_\ell}$-excellent (so in particular, $\epsilon$-excellent).

**Proof.** The strategy is as follows. Since the proof is essentially the same for Cases (I) and (II), we prove both simultaneously by giving the proof for Case (I), and pointing out when the cases differ.

We will try to choose $(\overline{A}_k, \overline{B}_k)$ by induction on $k \leq k_\ast$ such that:

1. $\overline{A}_k = \langle A_\eta : \eta \in \kappa^2 \rangle$
2. $\overline{A}_k$ is a partition of $A$, or of a subset of $A$
3. $A_\emptyset = A$
4. If $k = m + 1$, $\nu \in \kappa^2$ then $A_\nu$ is the disjoint union of $A_{\nu^-(0)}, A_{\nu^-(1)}$
5. $|A_\eta| \geq \epsilon^k|A|$ for $\eta \in \kappa^2$
   - or In case (II): $|A_\eta| \geq m_\ell$, with equality if desired
6. $\overline{B}_k = \langle B_\nu : \nu \in \kappa^2 \rangle$ (note that $B_\kappa$ is defined at stage $k + 1$)
7. Each $B_\nu \subseteq G$ is $\epsilon$-good
   - or In case (II): $B_\nu$ is $\frac{m_{k + 1}}{m_k}$-good
8. For all $\eta \in \kappa^{k - 1}$, $a \in A_{\eta^-(0)}$ implies $t(a, B_\eta) = 0$ and $a \in A_{\eta^-(1)}$ implies $t(a, B_\eta) = 1$.

Note that $t(a, B_\eta)$ is well defined in (8) as $B_\eta$ is good. When $k = 0$, define $A_\emptyset = A$. Now suppose $k = m + 1$. In Case (I), suppose that for all $\eta \in \kappa^2$, $A_\eta$ fails to be $\epsilon$-excellent. By definition, for each such $\eta$, there is some set $B_\eta \subseteq G$ which is $\epsilon$-good and such that

$$|\{a \in A_\eta : t(a, B) \neq 1\}| \geq \epsilon|A_\eta|$$

again noting that these two sets partition $A_\eta$ by goodness of $B_\eta$. So we can define $A_{\eta^-(i)} := \{a \in A_\eta : t(a, B_\eta) = i\}$ for $i \in \{0, 1\}$. 22

---

**modified: 2012-03-08**
Meanwhile, in case (II), we are interested in whether \( A_\eta \) is \( \frac{m_{k+1}}{m_k} \)-excellent rather than \( \epsilon \)-excellent; if not, there an \( \frac{m_{k+1}}{m_k} \)-good set \( B_\eta \) such that the displayed equation holds with \( \left| \frac{m_{k+1}}{m_k} |A_\eta| \right| \) in place of \( \left| \frac{m_{k+1}}{m_k} |A_\eta| \right| \). In this case, choose \( A_\eta^{(i)} \) to be a subset of \( \{ a \in A_\eta : t(a, B_\eta) \neq i \} \) of cardinality \( m_{k+1} \), for \( i = 0, 1 \).

This completes the inductive step, and satisfies conditions (1)-(8).

We now show that the induction cannot continue indefinitely for all \( k < k_{**} \). Suppose we have defined \( A_\eta \) for \( \eta \in k_* \), and \( B_\nu \) for \( \nu \in k_{**} \), satisfying (1)-(8). For each \( \eta \), since we assumed either (I) \( \epsilon^{k_*} |A| > 0 \) or (II) that \( |A_\eta| = m_\epsilon \geq m_{k_*-1} > k_{**} \), we have that \( A_\eta \neq \emptyset \) so we may choose \( a_\eta \in A_\eta \). Furthermore, for each \( \nu \in k_{**} \) and \( \eta \in k_{**} \) such that \( \nu \triangleleft \eta \), we may define

\[ U_{\nu, \eta} = \{ b \in B_\nu : (a_\eta Rb) \neq t(a_\eta, B_\nu) \} \]

i.e. the set of elements in \( B_\nu \) which do not relate to \( a_\eta \) in the expected way. By assumption \( \frac{m_{k+1}}{m_k} \leq \epsilon \), so in both Cases (I) and (II), \( |U_{\nu, \eta}| < \epsilon |B_\nu| \) by the goodness of \( B_\nu \). Hence for any such \( \nu \),

\[ \left| \bigcup \{U_{\nu, \eta} : \nu \triangleleft \eta \in k_{**} \} \right| < 2^{k_{**}} \epsilon |B_\nu| < |B_\nu| \]

by the hypothesis on the size of \( \epsilon \). In particular, for each \( \nu \in k_{**} \) we may choose an element \( b_\nu \in B_\nu \setminus \bigcup \{U_{\nu, \eta} : \nu \triangleleft \eta \in k_{**} \} \). Now the sequences \( \{a_\eta : \eta \in k_{**}\} \) and \( \{b_\nu : \nu \in k_{**}\} \) contradict Definition 2.11, i.e. the choice of \( k_{**} \).

We have shown that for some \( k < k_{**} \) the induction must stop. Hence for some \( \nu \in k \), \( A_\nu \) is \( \epsilon \)-excellent [if in case (II), \( A_\nu \) is \( \frac{m_{k+1}}{m_k} \)-excellent, so in particular \( \epsilon \)-excellent] and satisfies condition (5), which completes the proof. \( \square \)

**Remark 5.5.** Note that the tree construction just given naturally tends away from uniform size since we do not know when or where the induction will stop.

By definition, if \( A \) is \( \epsilon \)-excellent and \( B \) is \( \zeta \)-good, they will interact in a strongly uniform way, namely, most of the elements of \( A \) will have the same average \( t(a, B) \in \{0, 1\} \) over \( B \). Let us give this a name:

**Claim 5.6.** If \( A \) is \( \epsilon \)-excellent and \( B \) is \( \zeta \)-good then the pair \((A, B)\) is \( (\epsilon, \zeta) \)-uniform, where we say that \((A, B)\) is \( (\epsilon, \zeta) \)-uniform if for some truth value \( t = t(A, B) \in \{0, 1\} \) we have: for all \( \text{but} \) \( \epsilon |A| \) of the elements of \( A \), \( t(A, B) = t(a, B) \).

In other words, for all \( \text{but} \) \( \epsilon |A| \) of the elements of \( |A| \), for all \( \text{but} \) \( \zeta |B| \) of the elements of \( B \), \( (aRb) \equiv (t(A, B) = 1) \). When \( \epsilon = \zeta \), we will just write \( \epsilon \)-uniform.

**Proof.** By the definition of excellent. \( \square \)

**Remark 5.7.** So in some ways “\((A, B)\) is \( (\epsilon, \epsilon) \)-uniform” is stronger than being \( \epsilon \)-regular; see also Claim 5.17 below.

**Discussion 5.8.** At this point, we have a way to obtain \( \epsilon \)-excellent subsets of any given graph, whose sizes vary along a fixed sequence. Below, we will extract a collection of such sets as the first stage in obtaining a regularity lemma. However, the goal is a partition into pieces of approximately equal size, which will require an appropriate further division of the first-stage collection of \( \epsilon \)-excellent sets. In preparation, then, we now apply several facts from probability to prove that sufficiently large \( \epsilon \)-excellent sets can be equitably partitioned into a small number of pieces all of which are \( \epsilon' \)-excellent for \( \epsilon' \) close to \( \epsilon \).

**Fact 5.9.** Assume \( p, q > 0 \). If \( |A| = n \), \( B \subset A = p \), \( m \leq n \), \( \frac{m}{n} \geq q \), \( A' \) is a random subset of \( A \) with exactly \( m \) elements, then

\[
\text{Prob} \left( \frac{|A' \cap B|}{|A'|} \in \left( \frac{|B|}{|A|} - \zeta, \frac{|B|}{|A|} + \zeta \right) \right).
\]
can be modeled by a random variable which is asymptotically normally distributed.

Proof. That is, our hypergeometric distribution (sampling \( m \) elements from a set of size \( n \) without replacement) will be asymptotically approximated by the binomial distribution (sampling with replacement), and therefore by the normal distribution. See Erdős and Rényi [2] p. 52, Feller [3] p. 172, Nicholson [14]. Note that in our case \( m \) will remain relatively large as a fraction of \( n \).  

**Fact 5.10.** (Vapnik and Chervonenkis, [20]) Let \( X \) be a set of events on which a probability measure \( P_X \) is defined. Let \( S \) be a collection of random events, i.e. subsets of \( X \), measurable w.r.t. \( P_X \). Each sample \( x_1, \ldots, x_\ell \) and event \( A \in S \) determines a relative frequency \( v_A^{(\ell)} \) of \( A \) in this sample. Let \( P(A) \) be the probability of \( A \) and let \( \pi^{(\ell)} = \sup \{|v_A^{(\ell)} - P(A)| : A \in S\} \).

For each \( A \in S \) and finite sample \( X_\ell = x_1, \ldots, x_\ell \) of elements of \( X \), \( A \) is said to induce the subset of \( \{x_1, \ldots, x_\ell\} \) consisting of those elements \( x_i \) which belong to \( A \). The number of different subsamples of \( X_\ell \) induced by sets of \( S \) is denoted \( \Delta^S(x_1, \ldots, x_\ell) \). Define \( m^S(\ell) = \max(\Delta^S(x_1, \ldots, x_\ell)) \), where the maximum is taken over all samples of size \( \ell \).

Then a sufficient condition for the relative frequencies of events in \( S \) to converge uniformly over \( S \) (in probability) to their corresponding probabilities, i.e. for it to be true that for any \( \epsilon \), \( \lim_{\ell \to \infty} \text{Prob}(\pi^{(\ell)} > \epsilon) = 0 \), is that there exist a finite \( k \) such that \( m^S(\ell) \leq \ell^k + 1 \) for all \( \ell \).

**Remark 5.11.** The connection between the condition of Vapnik-Chervonenkis and the independence property, defined in Remark 3.12 above, was observed and developed by Laskowski [10].

**Fact 5.12.** (Rate of the almost sure convergence)

1. ([20] p. 272) Given \( k \) from the last paragraph of Fact 5.10, if \( \ell \) satisfies

\[
\ell \geq 16 \left( k \log \frac{16k}{\zeta^2} - \log \frac{\eta}{4} \right)
\]

then in any sample of size at least \( \ell \), with probability at least \( 1 - \eta \), the relative frequencies differ from their corresponding probabilities by an amount less than \( \zeta \), simultaneously over the entire class of events.

2. Bounds on the error of the normal approximation to the hypergeometric distribution may be found in Nicholson [14] p. 474 Theorem 2.

**Claim 5.13.** (Random partitions of excellent sets)

1. For every \( \epsilon, \zeta \) there is \( N_1 \) such that for all \( n > N_1 = N_1(\epsilon, \zeta) \), if \( A \subseteq G \), \( |A| = n \), \( A \) is \( \epsilon \)-good, \( n \geq m \geq \log \log(n) \), if we randomly choose an \( m \)-element subset \( A' \) from \( A \) then almost surely \( A' \) is \( (\epsilon + \zeta) \)-good. Moreover, we have that \( b \in G \Rightarrow t(b, A') = t(b, A) \).

1A) That is, in part (1), for each \( \xi \in (0, 1) \) there is \( N_2 = N_2(\epsilon, \zeta, \xi) \) such that the probability of failure is \( \xi \).

2. Similarly for “excellent” replacing “good”.

3. In particular, for all \( \epsilon' > \epsilon \) and \( r \geq 1 \) there exists \( N = N(\epsilon, \epsilon', r) \) such that if \( |A| = n > N \), \( r \) divides \( n \) and \( A \) is \( \epsilon' \)-excellent, there exists a partition of \( A \) into \( r \) disjoint pieces of equal size each of which is \( \epsilon' \)-excellent. Note that \( N(\epsilon, \epsilon', r) \) increases with \( r \).

Proof. (1) Call \( B \subseteq A \) an exceptional set if there is \( b \in G \) such that \( B = \{a \in A : aRb \neq t(b, A)\} \) and \( |B| \geq \epsilon m \). It suffices to show that almost surely \( A' \) satisfies: for all exceptional sets \( B \)

\[
\frac{|A' \cap B|}{|A'|} \in \left( \frac{|B|}{|A'|} - \zeta, \frac{|B|}{|A'|} + \zeta \right)
\]

By Fact 5.9, for \( n, m \) sufficiently large, we may approximate drawing a set of size \( m \) by the sum of \( m \) independent, identically and normally distributed random variables, where the probability of \( x \in B \) is just \( |B|/|A'| \). Since \( G \) has the non-\( k_\ast \)-order property, Claim 2.6 in the case where \( G = A \),
Claim 5.14. Assume that \( A = A' \) shows that the Vapnik-Chervonenkis sufficient conditions (Fact 5.10) are satisfied. (Recall the definition of exceptional set from the first line of the proof.)

(1A) By Fact 5.9 and Fact 5.12.
(2) Follows by the “moreover” in the previous clause.
(3) Let \( \epsilon \) be as given, \( \zeta = \epsilon' - \epsilon \), and \( \xi = \frac{2}{r}. \) Let us verify that \( N = N_2(\epsilon, \zeta, \xi) \) suffices. First, randomly choose a function \( h : A \to \{0, \ldots, r - 1\} \) such that for all \( s < r, \) \( |\{a \in A : h(a) = s\}| = \frac{n}{r}. \) Then each \( s < r \) induces a random choice of a subset of \( A, \) since for each \( s < r \) we have \( h^{-1}(s) \in [A]^{r}. \) Since \( h \) was random, for each given \( s, \) each \( B \in [A]^{r} \) is equally probable. By part (1), for each \( s < t \)

\[
1 - \xi \leq \text{Prob}\{h^{-1}(s) \text{ is } (\epsilon + \zeta)-\text{excellent}\}
\]

and therefore

\[
1 - r\xi \leq \text{Prob}\{\bigwedge_{s < r} h^{-1}(s) \text{ is } (\epsilon + \zeta)-\text{excellent}\}
\]

But since \( 1 - r\xi = 1 - \frac{2}{r} > 0, \) there exists an \( h \) which works, i.e. an \( h \) such that for each \( s < t, \) \( h^{-1}(s) \) is \( (\epsilon + \zeta) \)-excellent. Since \( \epsilon + \zeta = \epsilon', \) this finishes the proof.

The next claim forms the core of the proof of Theorem 5.18. The statement is laid out so as to make the strategy of construction clear (based on the claims established so far). A less transparent, but more compact, list of the requirements in this claim is summarized in Corollary 5.15. For the Theorem, it remains to construct an appropriate sequence \( \langle m_i : i < k_{**} \rangle \) which respects the various bounds collected here, and to show that this can be done while keeping \( m_{**} \) sufficiently large relative to \( |A| \).

Claim 5.14. Assume that \( \epsilon < \epsilon' < 2^{-k_{**}}. \) Suppose that \( A \subseteq G, |A| = n. \)

(1) Let \( \langle m_i : i < k_{**} \rangle \) be a sequence of natural numbers such that \( m_{i+1} \leq \epsilon m_i \) for \( i < k_{**}, \) and let \( m_* := m_0, m_{**} := k_{**} \geq k_{**}. \) Then there is \( \overline{A} \) such that:

(a) \( \overline{A} = \langle A_i : i < j(*) \rangle, \) for some \( j(*) \leq \frac{n}{m_{**}} \)
(b) For each \( i, A_i \subseteq A \) and \( |A_i| \in \{m_\ell : \ell < k_{**}\} \)
(c) \( i \neq j \Rightarrow A_i \cap A_j = \emptyset \)
(d) each \( A_i \) is \( \epsilon \)-excellent
(e) hence if \( i \neq j < j(*) \) then the pair \( (A_i, A_j) \) is \( (\epsilon, \epsilon) \)-uniform
(f) \( B := A \setminus \bigcup\{A_i : i < i(*)\} \) has \( < m_* \) members

(1A) Suppose further that:

(i) \( m_{**} \mid m_k \) for each \( k < k_{**} \)
(ii) \( m_{k_{**}-2} > N = N(\epsilon, \epsilon', \frac{m_{**}}{m_{**}}) \) (as in Claim 5.13)
(iii) \( \log \log m_* \leq m_{**} \)

Then for some \( i(*) \) with \( j(*) \leq i(*) \leq \frac{n}{m_{**}} \) there is a further refinement of the partition from (1) into \( i(*) \) disjoint pieces (in slight abuse of notation we will now use \( A_i : i < i(*) \)) to refer to this new partition) such that for each \( i < i(*)(A_i) = m_{**}. \) Furthermore, each of these new pieces \( A_i \) is \( \epsilon' \)-excellent.

(2) Let \( \langle A_i : i < i(*) \rangle \) be the partition into equally sized \( \epsilon' \)-excellent pieces from (1A). Then there exists a partition \( B_i : i < i(*) \rangle \) of the remainder \( B, \) allowing \( B_i = \emptyset \) for some \( i \) (i.e. \( \left\lceil \frac{|B|}{i(*)} \right\rceil \) may be 0) such that

\[
|B_i| \in \left\{ \left\lceil \frac{|B|}{i(*)} \right\rceil, \left\lceil \frac{|B|}{i(*)} \right\rceil + 1 \right\}
\]

Let \( A' = A_i \setminus B_i \) for \( i < i(*) \). Then:

(a) \( \langle A'_i : i < i(*) \rangle \) is a partition of \( A \)
(b) the sizes of the \( A'_i \) are almost equal, i.e. \( ||A'_i|| - ||A'_j|| \leq 1 \)

(c) if we let

\[
\zeta = \max \left\{ \frac{\alpha |A_i| + |B_i|}{|A_i| + |B_i|} : i < i(*) \right\} \leq \frac{\alpha m_{**} + \frac{m_1}{i(*)}}{m_{**} + \frac{m_{**}}{i(*)}}
\]

then \( i \neq j < i(*) \) implies \((A'_i, A'_j)\) is \((\zeta, \zeta)\)-uniform.

(3) If, moreover, \( m_{**} > \frac{1}{\ell} \) and \( m_\ell \leq \frac{\alpha n + 1}{\ell + \ell} \), then \( \zeta < 3\epsilon' \), where \( \zeta \) is as in (2)(c).

Proof. (1) Applying Claim 5.4 we try to choose a sequence of \( \varepsilon \)-excellent sets \( A_i \), each of size \( m_\ell \) for some \( \ell < k_{**} \), by induction on \( i \) from \( C_i := A \setminus \bigcup_{j<i} A_j \). We can continue as long as \( |C_i| \geq m_{**} \).

Note that condition (e) is immediate, for all pairs \((A_i, A_j)\) without exceptions, by Claim 5.6.

(1A) By Claim 5.13(3). Note that in the application below, we will build all relevant sequences of \( m_{**} \) to satisfy \( m_{**} \approx \epsilon_{k_{**}} m_{\ell} \) so that \( N = N(\varepsilon, \epsilon', \epsilon^{-k_{**}}) \) can be computed, if desired, before the sequence is chosen.

(2) Immediate: the partition remains equitable because the \( A_i \) all have size \( m_{**} \), and \( \zeta \) bounds the relative size of a “bad” subset of any given \( A_i \).

(3) Given the assumption of an equitable partition from (2)(b), it would suffice to show that for every \( i \), \( |B_i| \leq 2\epsilon'|A_i| \), as then we would have

\[
\frac{\alpha |A_i| + |B_i|}{|A_i| + |B_i|} \leq \frac{\alpha |A_i| + 2\epsilon'|A_i|}{|A_i|} = 3\epsilon'
\]

We verify that the assumption on \( m_{**} \) is enough to give this bound. By definition, as the \( B_i \)’s arise from an equitable partition of the remainder \( B \), \( |B_i| \leq \frac{m_{**} - 1}{n(i*)} + 1 \), where \( i(*) \) is the number of components from the partition (1A), by (2) above. Since the components \( A_i \) from (1A) all have size \( m_{**} \), and \( |B| \leq m_{**} - 1 \), we can bound \( i(*) \) by \( \frac{n}{m_{**}} \geq i(*) \geq \frac{n - m_{**} + 1}{m_{**}} \). Thus

\[
|B_i| - 1 \leq \frac{m_{**} - 1}{i(*)} < (m_{**} - 1) \left( \frac{n - m_{**}}{m_{**}} \right)^{-1} \quad \text{and so} \quad \frac{|B_i| - 1}{|A_i|} < \left( \frac{m_{**} - 1}{m_{**}} \right) \left( \frac{n - m_{**}}{m_{**}} \right)^{-1} = \frac{m_{**} - 1}{n - m_{**}}
\]

We had assumed that \( m_{**} \leq \frac{\alpha n + 1}{\ell + \ell} \), and so:

\[
m_{**}(1 + \epsilon') \leq \epsilon' n + 1 \\
m_{**} - 1 \leq (n - m_{**}) \epsilon' \\
\frac{m_{**} - 1}{n - m_{**}} \leq \epsilon'
\]

We had also assumed that \( \frac{1}{\ell} < m_{**} \), i.e. \( \frac{1}{m_{**}} < \epsilon' \). Since \( |A_i| = m_{**} \) (so \( \frac{|B_i| - 1}{|A_i|} = \frac{|B_i|}{|A_i|} - \frac{1}{m_{**}} \)), we conclude that

\[
\frac{|B_i|}{|A_i|} < \frac{m_{**} - 1}{n - m_{**}} + \frac{1}{m_{**}} < \epsilon' + \epsilon' = 2\epsilon'
\]

which completes the proof. \( \Box \)

**Corollary 5.15.** To summarize the requirements of Claim 5.14, suppose that \( k_* \) and therefore a bound for \( k_{**} \) are fixed in advance, \( G \) is a graph with the non-\( k_* \)-order property, and that we are given:

(1) \( \epsilon_1, \epsilon_2 \in \mathbb{R} \) such that \( 0 < \epsilon_3 < \epsilon_2 := \frac{4}{3} < \epsilon_1 < 2^{-k_*} \)

(2) A sequence of positive integers \( \langle m_\ell : \ell < k_{**} \rangle \) such that:

(a) \( m_{\ell+1} < \epsilon_3 m_\ell \) for each \( \ell < k_{**} \)

(26)
(b) \( m_* \mid m_\ell \) for each \( \ell < k_* \)
(c) \( \log \log m_0 \leq m_* \)
(d) \( m_* := m_{k_* - 1} \geq \max(k_* \cdot \frac{1}{1+\varepsilon_2}) \)
(e) \( m_{k_* - 2} > N(\varepsilon_3, \varepsilon_2, \frac{m_0}{m_*}) \), from Claim 5.13(3)

(3) \( A \subseteq G \), \( |A| = n \) where \( n \) satisfies \( m_0 \leq \frac{\varepsilon_0 n + 1}{1+\varepsilon_2} \).

Then there exists \( i(*) \leq \frac{n}{m_*} \) and a partition of \( A \) into disjoint pieces \( \langle A_i : i < i(*) \rangle \) such that:

- for all \( i < j < i(*) \), \( |A_i| - |A_j| \leq 1 \)
- each \( A_i \) is \( \varepsilon_1 \)-excellent
- Each pair \( (A_i, A_j) \) is \( \varepsilon_1 \)-uniform

Proof. By Claim 5.14, using \( \varepsilon = \varepsilon_3 \), \( \varepsilon' = \varepsilon_2 \) and \( 3\varepsilon' = \varepsilon_1 \); note that the partition we obtain was called \( \langle A'_i : i < i(*) \rangle \) in Claim 5.14.

Discussion 5.16. In practice, we are given \( \varepsilon = \varepsilon_1 \), and then choose \( \varepsilon_3 \) to run the proof of Corollary 5.15. The role of the respective \( \varepsilon \) appears in conditions (2)(a) and (2)(e) of this Corollary. On one hand, \( \varepsilon_3 \) determines the rate of decrease of the sequence of \( m_\ast \), thus the size of \( m_* \), and ultimately the number of components in the partition: so one would usually want to choose \( \varepsilon_3 \) close to \( \varepsilon_1 = \varepsilon \). On the other hand as \( \varepsilon_3 \) approaches \( \varepsilon_1 \), the lower bound on the size of the graph \( A \) may rise, via the \( N \) from (2)(e), which comes from Claim 5.13(3).

Before stating the main result of this section, Theorem 5.18, we consider more explicitly the relation of \( \varepsilon \)-uniformity to \( \varepsilon \)-regularity. As the following calculation shows, \( \eta \)-uniform pairs will be \( \rho \)-regular when \( \rho \) (the parameter for a lower bound on the size of a subset chosen) is sufficiently large relative to \( \eta \) (the parameter for an upper bound on the number of non-uniform edges). As mentioned above, uniformity is somewhat more precise in our context for large enough graphs, as the densities of sufficiently large \( \varepsilon \)-regular pairs will be near 0 or 1.

Claim 5.17. Suppose that \( \varepsilon, \zeta, \xi \in (0, \frac{1}{2}) \), and the pair \( (A, B) \) is \((\varepsilon, \zeta)\)-uniform. By uniformity, there is a truth value \( t(A, B) \in \{0, 1\} \). Let \( Z := \{(a, b) \in (A \times B) : aRb \neq t\} \) and likewise let \( Z' := \{(a, b) \in (A' \times B') : aRb \neq t\} \). Suppose also that \( A' \subseteq A \), \( |A'| \geq \xi |A| \), \( B' \subseteq B \), \( |B'| \geq \xi |B| \), and \( \frac{\varepsilon + \zeta}{\xi} < \frac{1}{2} \). Then:

\[(1) \frac{|Z|}{|A||B|} < \varepsilon + \zeta \]
\[(2) \frac{|Z'|}{|A'||B'|} < \frac{\varepsilon + \zeta}{\xi} \]

In particular, if the pair \( (A, B) \) is \( \varepsilon_0 \)-uniform for \( \varepsilon_0 < \frac{1}{2} \) then:

(a) \( (A, B) \) is \( \varepsilon \)-regular, and
(b) if \( A' \in [A]^{\geq \xi |A|} \) and \( B' \in [B]^{\geq \xi |B|} \) then \( d(A', B') < \varepsilon \) \( \lor \) \( d(A', B') \geq 1 - \varepsilon \)

Proof. Let \( A', B' \) be given. For \( a \in A \), let \( W_a = \{b \in B : aRb \neq t(A, B)\} \), and let \( U = \{a \in A : |W_a| > \varepsilon |A|\} \). So \( |U| < \varepsilon |A| \), and \( a \in A \setminus U \implies |W_a| < \zeta |B| \). Since

\[Z \subseteq U \times B \cup \bigcup \{(a, b) \in A \times B : b \in W_a, a \notin U\}\]
\[Z' \subseteq U \times B' \cup \bigcup \{(a, b) \in A' \times B' : b \in W_a, a \notin U\}\]

27
we can bound the cardinalities as follows:

$$|Z| \leq |U| \cdot |B| + |A| \cdot \max\{|W_a| : a \in U\}$$

$$\frac{|Z|}{|A \times B|} < \frac{\epsilon |A| + \zeta |B|}{|B|} = \epsilon + \zeta$$

and likewise

$$\frac{|Z'|}{|A' \times B'|} = \frac{|U||B'| + |A'| \cdot \max\{|W_a| : a \in U\}}{|A'||B'|} \leq \frac{\epsilon |A||B| + \zeta |A||B|}{|A'||B'|} = (\epsilon \xi + \zeta \xi) \cdot \frac{|A||B|}{|A'||B'|} = \frac{\epsilon \xi + \zeta \xi}{\xi^2} = \epsilon + \frac{\zeta}{\xi}$$

by the assumption on the size of $A', B'$. This completes the proof of (1) and (2).

For the “in particular” clause, let $d(X, Y) = \frac{|B||X \times Y|}{|X||Y|}$ be the usual edge density. We have shown that if $t(A, B) = 1$, $d(A, B) > 1 - (\epsilon + \zeta)$ while $d(A', B') > 1 - \frac{\epsilon + \zeta}{\xi}$, and likewise if $d(A, B) = 0$, $d(A, B) < (\epsilon + \zeta)$ while $d(A', B') < \frac{\epsilon + \zeta}{\xi}$. Thus the difference in density $|d(A, B) - d(A', B')|$ is bounded by $\frac{\epsilon + \zeta}{\xi} < \frac{1}{2}$, and moreover $d(A', B')$ differs from either 0 or 1 by at most $\frac{\epsilon + \zeta}{\xi}$.

Now if $\epsilon, \zeta$ are equal and $\leq \frac{1}{2}$, then the hypothesis $\xi \in (0, \frac{1}{2})$ guarantees that $\frac{\epsilon + \zeta}{\xi} < \frac{1}{2}$ is satisfied. Thus when $(A, B)$ is $(\epsilon_0, \epsilon_0)$-uniform and $\epsilon$ is such that $|A'| \geq \epsilon|A|$, $|B'| \geq \epsilon|B|$ where $\epsilon_0 \leq \frac{\epsilon}{2}$ and $\epsilon < \frac{1}{4}$, then $\frac{\epsilon + \zeta}{\xi}$ is bounded by $\frac{\epsilon}{2} = \epsilon$, which completes the proof. \qed

We now give the main result of this section. Recall the definitions of non-$k_*$-order property (Definition 2.3), $k_{**}$ (Definition 2.11), $\epsilon$-excellent (Definition 5.2), and $\epsilon$-uniform (Claim 5.6).

**Theorem 5.18.** Let $k_*$ and therefore a bound for $k_{**}$ be given. Then for all $\epsilon > 0$ there exists $m = m(\epsilon, k_*)$ and $N = N(\epsilon, k_*)$ such that for every finite graph $G$ with the non-$k_*$-order property and every $A \subseteq G$ with $|A| \geq N$, there is a partition $(A_i : i < i(\ast) \leq m)$ of $A$ into at most $m$ pieces, where:

1. for all $i, j < i(\ast)$, $|A_i| - |A_j| \leq 1$
2. each of the pieces $A_i$ is $\epsilon$-excellent
3. all of the pairs $(A_i, A_j)$ are $(\epsilon, \epsilon)$-uniform
4. if $\epsilon < \frac{1}{2m + 1}$, then $m \leq (3 + \epsilon) \left(\frac{2}{\xi}\right)^k_{**}$

**Proof.** Without loss of generality, assume $\epsilon < \frac{1}{2m + 1}$. (This is necessary for Claim 5.15, which uses Claim 5.4.)

We proceed in stages. Let $n = |A|$. When hypotheses are made about the minimum size of $n$, these will be labeled (Hx) and collected in Step 5.

**Step 0: Fixing epsilons.** When applying Corollary 5.15 we will use: $\epsilon_3 = \frac{\epsilon}{2}$, $\epsilon_2 = \frac{\epsilon}{4}$, and $\epsilon_1 = \epsilon$.

**Step 1: Fixing q.** Given $\epsilon_3$, let $q = \left\lceil \frac{1}{\epsilon_3}\right\rceil \in \mathbb{N}$. It follows that $\frac{q}{\epsilon_3} \geq q \geq \frac{1}{\epsilon_3}$ and thus $\frac{1}{\epsilon} \leq \frac{1}{q} \leq \epsilon_3$. In particular, any sequence $(m_\ell : \ell < k_{**})$ such that $m_{**} := m_{k_{**} - 1} \in \mathbb{N}$ and $m_\ell := qm_{\ell + 1}$ for all $\ell < k_{**}$ will satisfy $m_{\ell + 1} = \frac{1}{q} m_\ell \leq e_{3/2} m_\ell$, $m_\ell \in \mathbb{N}$ for each $\ell < k_{**}$, and $m_{**} | m_\ell$ for all $\ell < k_{**}$.

**Step 2: Choosing m_{**}.** In this step, the aim is to build a sequence $(m_\ell : \ell < k_{**})$ whose elements are as large as possible subject to the constraints (2)(a),(b),(d) and (3) of Corollary 5.15. In keeping
with prior notation, let $m_* := m_0$. Recalling $\epsilon_2 = \frac{\epsilon}{3}$ from Step 0, Condition 5.15(3) asks that

$$m_* \leq \frac{en + 1}{3 + \epsilon}$$

so it suffices to choose $m_* \leq \frac{en}{3 + \epsilon} = \frac{en}{3 + \epsilon}$

Let (H1) be the assumption that $n$ is not too small (see Step 5). Then there exists $c \in \mathbb{N}$, $c > k_*$, such that

$$q^{k_* - 1} c \in \left( \frac{en}{3 + \epsilon} - q^{k_* - 1}, \frac{en}{3 + \epsilon} \right]$$

Thus setting $m_* := \max\{c \in \mathbb{N} : c > k_*, \ c > \frac{1}{\epsilon_2}, \ q^{k_* - 1} c \leq \frac{en}{3 + \epsilon}\}$ is well defined, and $m_*$ will belong to the half-open interval just given. Having defined $m_*$, for each $\ell < k_*$ let $m_\ell := q^{k_* - \ell - 1} m_*$. By Step 1, the $m_\ell$ are integer valued and satisfy the required conditions on divisibility and size. By choice of $c$, $m_* = q^{k_* - 1} m_*$ satisfies the inequality 5.15(3).

We have defined a sequence $(m_\ell : \ell < k_*)$ of positive integers which satisfies conditions (2)(a),(b),(d) and (3) of Corollary 5.15. We fix this sequence for the remainder of the proof, and proceed to calculate various bounds in terms of it.

**Step 3: Bounding $m_*$.** By the definition of $m_*$ in Step 2, $\frac{en}{3 + \epsilon} - q^{k_* - 1} < q^{k_* - 1} m_*$, so assuming $n$ is not too small [again (H1) in Step 5],

$$\frac{en}{3 + \epsilon} (q^{k_* - 1} - 1) - 1 < m_* \implies \frac{1}{2} \frac{en}{3 + \epsilon} (q^{k_* - 1} - 1) \leq m_*$$

**Step 4: Bounding $\frac{n}{m_*}$.** Applying Step 3, an inequality from Step 1, and the definition of $\epsilon_3$,

$$\frac{n}{m_*} \leq \frac{n}{2} \left( \frac{en}{3 + \epsilon} \right) \left( \frac{1}{q^{k_* - 1}} \right) \leq \frac{2(3 + \epsilon) q^{k_* - 1}}{\epsilon} \leq \frac{2(3 + \epsilon)}{\epsilon} \left( \frac{2}{\epsilon_3} \right)^{k_* - 1} = \left( 3 + \epsilon \right) \left( \frac{8}{\epsilon} \right)^{k_* - 1}$$

Note that a choice of $\epsilon_3$ closer to $\epsilon_2$ would slightly improve this bound, at the cost of increasing the threshold size of $n$ in (H3) of Step 5.

**Step 5: Requirements for the lower bound on $n = |A|$.** We collect the necessary hypotheses on the size of the graph:

(H1) $n$ is large enough to allow for the choice of $m_*$ in the interval from Step 2 while preserving $m_* > k_*, \ m_* > \frac{1}{\epsilon_2}$:

- it suffices that $n > (k_* + 1) q^{k_* - 1} \left( \frac{3 + \epsilon}{\epsilon} \right)$, which ensures $\frac{en}{3 + \epsilon} - q^{k_* - 1} > k_* q^{k_* - 1}$
- and also ensures that $n > 2 q^{k_* - 1} \left( \frac{3 + \epsilon}{\epsilon} \right)$, for the calculation in Step 3

(H2) $n$ is large enough for the sequence $(m_\ell : \ell < k_*)$ to satisfy log-log $m_* \leq m_*$:

- it suffices that $n \geq (\log \log q^{k_* - 1}) (q^{k_* - 1}) \left( \frac{3 + \epsilon}{\epsilon} \right)$

(H3) $n$ is large enough for $m_{k_* - 2}$ to satisfy condition (2)(e) of Corollary 5.15:

- it suffices that $n \geq N \left( \frac{q}{3}, \frac{q - 1}{2} \right) \cdot \left( \frac{3 + \epsilon}{\epsilon} \right) \cdot q^{k_* - 2}$ where $N(\cdot, \cdot)$ is from Claim 5.13 and incorporates the bounds from Fact 5.12.

Under these assumptions the sequence constructed in Step 2 will also satisfy conditions (2)(c),(e) of Corollary 5.15. By Step 0 and Step 2, all the hypotheses of that Corollary are satisfied.

**Step 6: Obtaining the partition.** Assuming $n$ is sufficiently large, as described in Step 5, we have constructed a sequence $(m_\ell : \ell < k_*)$ so that the graph $A$ and the constructed sequence satisfy the
hypotheses of Corollary 5.15. Thus we obtain a partition of $A$ satisfying (1),(2),(3) of the Theorem. Condition (4) follows from Step 4, which completes the proof. □

**Conclusion 5.19.** For every $k_* \in \mathbb{N}$ and $\epsilon \in (0,\frac{1}{k_*})$ there are $N, m$ such that for every finite graph $G$ with the non-$k_*$-order property and every $A \subseteq G$ with $|A| \geq N$, there is $\ell \leq m$ and a partition $A = \langle A_i : i < \ell \rangle$ such that each $A_i$ is $\epsilon$-excellent, and for every $0 \leq i < j < \ell$,

- $||A_i| - |A_j|| \leq 1$
- $(A_i, A_j)$ is $\epsilon$-regular, and moreover if $B_i \in [A_i]_{\geq \epsilon|A_i|}$ and $B_j \in [A_j]_{\geq \epsilon|A_j|}$ then
  $$\left( d(B_i, B_j) < \epsilon \right) \lor \left( d(B_i, B_j) \geq 1 - \epsilon \right)$$
- if $\epsilon < \frac{1}{2^{k_*}}$, then $m \leq (3 + \frac{2}{\epsilon^2}) (\frac{2^m}{\epsilon})^{k_*}$

**Proof.** This is a slight weakening of Theorem 5.18: by applying that Theorem to hypotheses of Corollary 5.15. Thus we obtain a partition of $A$ satisfying (1),(2),(3) of the Theorem. Condition (4) follows from Step 4, which completes the proof. □

**REFERENCES**


