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). We show that half-graphs are the *only* essential difficulty, by giving a much stronger version of Szemerédi's regularity lemma for stable graphs, 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 graphs with the non-k_*-order property. The first theorem gives a partition of any stable graph into indiscernible components (i.e., each component is either a complete or an empty graph) whose interaction is strongly uniform. To prove this, we first give a finitary version of the classic model-theoretic fact that stable theories admit large sets of indiscernibles, by showing that in finite stable 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 ϵ 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 [18] 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 A, B are finite graphs with disjoint vertex sets, the density $d(A, B) = \frac{|R \cap (A \times B)|}{|A||B|}$ and we say that (A, B) is ϵ-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 ϵ, 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 k with $m \leq k \leq m'$ and a partition $X = X_1 \cup \cdots \cup X_k$ satisfying:*

(1) $||X_i| - |X_j|| \leq 1$ *for all $i, j \leq k$*
(2) *all but at most ϵk^2 of the pairs (X_i, X_j) are ϵ-regular.*

As explained in §1.8 of [8], "Are there exceptional pairs?" it was not known for some time whether the ϵk^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 $a_i R b_j$ iff $i < j$, shows that exceptional pairs are necessary.

It is therefore natural to ask whether "half-graphs" are the main difficulty, i.e.:

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 theorem of this paper, Theorem 5.18, shows that indeed, half-graphs are the only essential difficulty: the result 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].

Theorem 5.18. *Let k_* and therefore k_{**} (Definition 2.10) be given. Let G be a finite graph with the non-k_*-order property. Then for any $\epsilon > 0$ there exists $m = m(\epsilon)$ such that for all sufficiently large $A \subseteq G$, there is an equitable partition of A into at most m pieces, in which each of the pieces is ϵ-excellent (Definition 5.2), all of the pairs are (ϵ, ϵ)-uniform (Claim 5.6), and if $\epsilon < \frac{1}{2^{k_{**}}}$, then $m \leq (3 + \epsilon) \left(\frac{8}{\epsilon}\right)^{k_{**}}$.*

The proof, 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 ϵ, due to the definition of "indivisibility," ϵ-excellence. This settles Question 1.1 and thus characterizes the class of graphs admitting ϵ-regular partitions with no irregular pairs. The bulk of the paper is devoted to a more general model-theoretic analysis, which we now describe.

1.2. The phenomenon of regularity in stable 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 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) is 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 ϵ-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* stable 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 $\leq 2k_*$ $b \in G_j$, $aRb \equiv \mathbf{t}(G_i, G_j)$.*

Moreover, $N = n_1 n_2$ 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 $f(x) = x^\epsilon$) 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 4.22. 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 k_{**} be given, and suppose G is a finite graph with the non-k_*-order property. Let $A \subset G$, $|A| = n$ with $n^{\epsilon^{k_{**}}} > k_{**}$. Then there is $\zeta < \epsilon^{k_{**}}$ and a partition $\langle A_i : i < i(*) \rangle$ of A such that:*

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

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, we show how to solve this instance of irregular pairs at the cost of a larger remainder:

Theorem 4.22. *Let k_* and therefore k_{**} be given. Let G be a graph with the non-k_*-order property, and let $k = k_*$.*

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

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

Finally, §5 gives the theorem quoted above, obtaining a partition whose size depends only on ϵ 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).

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 number of vertices.*

Remark 2.2. *In model theory, a formula $\varphi(\overline{x}; \overline{y})$ is said to have the order property with respect to a theory T if there exist a model $M \models T$ and sequence $\langle \overline{a}_i, \overline{b}_i : i < \omega \rangle$ of tuples from M such that $M \models \varphi(\overline{a}_i; \overline{b}_j)$ iff $i < j$. T is called* stable *if no formula has the order property with respect 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 theory of G. The exception is §3 which occasionally asks that R-rank be defined for certain sets of formulas Δ.*

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

G has the order property *if it has the k-order property for all k. Cf. Remark 2.2.*

Remark 2.4. *By the symmetry of R, it is enough to rule out the order in one direction (i.e. the non-k-order propery also implies that for no such sequence does $i < j \implies \neg(a_i R_G b_j) \wedge (a_j R_G b_i)$.*

Claim 2.5. *Suppose G is a graph with the non-k-order property. Then for any finite $A \subseteq G$, $|\{\{a \in A : aR_G b\} : b \in G\}| \leq |A|^k$, more precisely $\leq \Sigma_{i \leq k} \binom{|A|}{i}$.*

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

Definition 2.6. (Indiscernibility) *Let M be a model, let Γ be a set of formulas in the language of M and α an ordinal. Recall that a sequence $\langle a_i : i < \alpha \rangle$ of elements of M is said to be a Γ-indiscernible sequence if for any $n < \omega$, 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 α, 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 2.7. *Let φ 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 φ^X, where X is an expression which evaluates to either true or false, is simply φ or $\neg\varphi$, as appropriate. Likewise, the intended interpretation of expressions like "$xRa \equiv \mathbf{t}$", where $\mathbf{t} \in \{0,1\}$ or is an expression which evaluates to true or false, is "xRa if and only if $\mathbf{t} = 1$," or equivalently, iff $\mathbf{t}$ is true.*

Definition 2.8. (The set Δ_k) *Let Δ_k be the set of formulas $\{x_0 R x_1\} \cup \{\varphi^i_{k,m} : m \leq k, i \in \{1,2\}\}$ where*

$$\varphi^i_{k,m} = \varphi^i_{k,m}(x_0, \ldots, x_{k-1}) = (\exists y) \left(\bigwedge_{\ell < m} (x_\ell R y)^{if\ (i=1)} \wedge \bigwedge_{m \leq \ell < k} (x_\ell R y)^{if\ (i=2)} \right)$$

Observation 2.9. *Let H be a finite graph, and let $A = \langle a_i : i < \alpha \rangle$ be a Δ_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 ℓ such that $0 \leq \ell \leq k-1$, $bR_H a_{i_\ell}$ and*
- *for all ℓ 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 $\varphi^1_{k,m}(c_0, \ldots c_{k-1})$ is true in H. As any two increasing subsequences of A of length k satisfy the same Δ_k-formulas, this easily gives the k-order property.

Note that if we had assumed the inverse, i.e. for all ℓ such that $0 \leq \ell \leq k-1$, $\neg b R_H a_{i_\ell}$ and for all ℓ such that $k \leq \ell < 2k$, $b R_H a_{i_\ell}$, we again get the k-order property (the same proof works with φ^2 replacing φ^1, in the notation of Definition 2.8). $\square$

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_{**} on tree height:

Definition 2.10. (The tree bound k_{**}) *Suppose G does not have the k_*-order property. Let $k_{**} < \omega$ be minimal so that there do not exist sequences $\overline{a} = \langle a_\eta : \eta \in {}^{k_{**}}2 \rangle$ and $\overline{b} = \langle b_\rho : \rho \in {}^{k_{**}>}2 \rangle$ of elements of G such that if $\rho^\frown\langle \ell \rangle \trianglelefteq \eta \in {}^{k_{**}}2$ then $(a_\eta R b_\rho) \equiv (\ell = 1)$.*

Remark 2.11. *In general, given a formula $\varphi(x; y)$, (i) if φ has the non-k_*-order property, then k_{**} exists and $k_{**} < 2^{k_*+2} - 2$. Conversely (ii) if k_{**} is as in Definition 2.10, then φ has the non-$2^{k_{**}+1}$-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 2.12. (Truth values $\mathbf{t}$) *By a truth value $\mathbf{t} = \mathbf{t}(X, Y)$ for $X, Y \subset G$, we mean an element of $\{0, 1\}$, where these are identified with "false" and "true" respectively. When $X = \{x\}$, write $\mathbf{t} = \mathbf{t}(x, Y)$. The criteria for assigning this value will be given below.*

Definition 2.13. (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 2.14. (Distinguished symbols) *Throughout this article, ϵ, ζ, ξ are real numbers in $(0, 1)$. We use ρ, η for zero-one valued sequences ${}^n2$, 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, ℓ, m, n denote natural numbers, with the occasional exception of the standard logical notation $\ell(\overline{x})$, i.e. the length of the tuple $\overline{x}$. 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_*, k_{**}, m_*, m_{**} are distinguished. When relevant (the conventions are given at the beginning of each section), k_* is such that the graph G under consideration has the non-k_*-order property, Definition 2.3, and k_{**} is the associated tree bound, Definition 2.10. (The one exception is §3.2, in which k_* is such that the graph under consideration is k_*-dependent.) The relevant sections all compute bounds based on k_*, so it is useful, but not necessary, to assume k_* is minimal for this property. Likewise, various arguments in the paper involve construction of a rapidly decreasing sequence $\langle m_\ell : \ell < k_{**} \rangle$ of natural numbers. In context, we use m_* and m_{**} to refer to the first and last elements of the relevant sequence, i.e., m_0 and $m_{k_{**}-1}$ respectively.

G is a large graph, usually finite; $A, B, X, Y...$ are finite subgraphs of the ambient G. Alternately, one could let G be infinite, while restricting consideration to its finite subgraphs.

3. A PARTITION INTO INDISCERNIBLE PIECES

Classically, the hypothesis of stability implies that in infinite models, one can extract large indiscernible sequences (Definition 2.6 above). More precisely, given λ an infinite cardinal, M a model whose theory is stable in λ and $A, I \subseteq M$ 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 $\mathbf{t}_{A,B}$ such that there are at most a constant number of exceptional edges $(aRb \not\equiv \mathbf{t}_{A,B})$. 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 Δ_{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 $2k_*$ elements. Let i_1 be the k_*th element of X and let i_2 be the k_*th element of Y. Clearly $i_1 \neq i_2$.

Case 1: $i_1 < i_2$. By assumption, we can find a subsequence $a_{j_1} < \cdots < a_{j_{k_*}} < a_{j_{k_*+1}} < \cdots < a_{j_{2k_*}} \leq a_m$ such that $\{j_1 < \cdots < j_{k_*} = i_1\} \subseteq X$ and $\{i_2 = j_{k+1} < \cdots < j_{2k_*}\} \subset Y$. Observation 2.9 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 Γ be a set of formulas, n_1 a cardinal and n_2 an ordinal (for our purposes these will both be finite). Then $n_1 \to (n_2)_{T,\Gamma,1}$ means: for every sequence $\langle a_i : i < n_1 \rangle$ of elements of G, there is a non-constant sub-sequence $\langle a_{i_j} : j < n_2 \rangle$ which is a Γ-indiscernible sequence, Definition 2.6. Replacing 1 by ℓ means that the tuples a_i in the sequence have length ℓ. Usually we suppress mention of $T = Th(G)$ and assume $\ell(a_i) = 1$, and therefore simply write $n_1 \to (n_2)_\Gamma$.*

Claim 3.4. *If $n_1 \implies (n_2)_{2^{|\Delta_{k_*}|}}^{k_*}$ in the usual arrow notation then*

$$n_1 \to (n_2)_{\Delta_{k_*}}$$

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 Δ_{k_*} hold on the sequence, and so extract a homogeneous subsequence of order type n_2. $\square$

As explained in this section's introduction, the advantage of the next theorem is not in showing the existence of indiscernible subsequences, which could be obtained by Ramsey's theorem since Δ 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 $\beth_{k_*}(n_2 + 2^{|\Delta|})$, 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, Δ are such that:*
  (a) *Δ 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 φ 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 $A \subset G$, $|A| \geq 2$, 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 $f(x) = \left|\frac{x}{t}\right|^{\frac{1}{tr+t+1}} - k$ and r, t, k are constants depending only on Δ: r is from (1) of the theorem, t is a stability constant (the R-rank of Δ), and k is the number of free variables.*
   - *$\langle a_i : i \in u \rangle$ is Δ-indiscernible.*

(3) *In particular, for Δ_{k_*} from Definition 2.8, we have that $n_1 \to (n_2)_{T, \Delta_{k_*}, 1}$ for any $n_1 > (cn_2)^{(2tr)^{k_*}}$, for a constant c depending on k_*, i.e. $c = c(\Delta_{k_*}) = c(r, t, k_*)$ where r, t are computed for Δ_{k_*}.*

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 $f(x) = \left|\frac{x}{t}\right|^{\frac{1}{tr+t+1}} - k$ 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 ν is an initial segment of ρ. 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 $W_\ell \subseteq u_m \setminus \bigcup_{j < \ell} W_j$ and a tree order $<_\ell$ on $W_{\leq \ell} = \bigcup_{j \leq \ell} W_j$ such that

   - if $i <_\ell j$ then a_j, a_i realize the same Δ^m-type over $\{a_{\mathbf{i}} : \mathbf{i} <_\ell i\}$
   - if $\neg(i <_\ell j)$ and $\neg(j <_\ell i)$ then a_j, a_i realize different Δ^m-types over $\{a_s : s <_\ell i, s <_\ell j\}$

Call any such tree a *type tree*. That is, a type tree W_* is given by a tree ordering $<_* := \bigcup_\ell <_\ell$ 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 $<_*$, 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_* by induction, $<_*$ implies $<$ 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 Δ^m,

$$\varphi(a_{i_0}, \ldots a_{i_{k-1}}) \iff \varphi(a_{i_0}, \ldots a_{i_{m-1}}, a_{\ell_m - m}, \ldots a_{\ell_m - 1})$$

by inductive hypothesis, since the first m indices agree.

$$\varphi(a_{i_0}, \ldots a_{i_{m-1}}, a_{\ell_m - m}, \ldots a_{\ell_m - 1}) \iff \varphi(a_{i_0}, \ldots a_{i_{m-2}}, a_{j_{m-1}}, a_{\ell_m - m}, \ldots a_{\ell_m - 1})$$

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

$$\varphi(a_{i_0}, \ldots a_{j_{m-1}}, a_{\ell_m - m}, \ldots a_{\ell_m - 1}) \iff \varphi(a_{j_0}, \ldots a_{j_{k-1}})$$

by inductive hypothesis. We have verified that

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

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_* 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_*, i.e. (I) *of the induction.* As we have established that any branch through a type tree W_* 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_*. For $s \leq t$ let $S^s = \{i \in W_* : \text{above } i \text{ in the tree we can embed } ^s 2 \text{ but no more}\}$. W_* will be the disjoint union of $\{S^s : s \leq t\}$, and if $i_1 \in S^1 \wedge i_2 \in S^2 \wedge i_i \leq_* i_2$ then $s_1 \geq s_2$.

Thus every node in the type 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 $j \leq_* i$ as well as the m constants named as part of Δ^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 $\leq_*$.

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_ℓ^s denote the number of nodes in the tree of rank s and height ℓ, i.e. $N_\ell^s = |\{i : i \in S^s, \text{ht}(i) = \ell\}|$. We can further write $N_\ell^s = X_\ell^s + Y_\ell^s$, where X_ℓ^s counts the nodes in N_ℓ^s whose immediate predecessor also has rank s, and Y_ℓ^s counts the nodes in N_ℓ^s 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 ℓ, $X_{\ell+1}^s \leq N_\ell^s$

   (ii) For all $s < t$ and all ℓ, $Y_{\ell+1}^s \leq N_\ell^{s+1} \cdot (\ell + m)^r$

   (iii) Thus for all $s < t$ and all ℓ, $N_{\ell+1}^s \leq N_\ell^s + N_\ell^{s+1} \cdot (\ell + m)^r$

   (iv) For all $1 \leq s \leq t$, $N_0^{t-s} = 0$.

   (v) For all ℓ, $N_\ell^t \leq 1$.

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

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

$$N_\ell^{t-(s+1)} + N_\ell^{t-s} \cdot (\ell + m)^r \leq \sum_{j \leq \ell} (\ell + m)^{s(r+1)} \cdot (j + m)^r \leq (\ell + m)^{(s+1)(r+1)}$$

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

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

$$N_{\ell+1} \leq \sum_{s \leq t} N_{\ell+1}^{t-s} \leq \sum_{s \leq t} (\ell + m)^{s(r+1)} \leq t(\ell + m)^{t(r+1)}$$

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

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

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

$$t(h + m)^{tr+t+1} < |u_m|$$

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

$$f(x) = \left| \frac{x}{t} \right|^{\frac{1}{tr+t+1}} - k$$

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 $f(x) = \left| \frac{x}{t} \right|^{\frac{1}{tr+t+1}} - k$. Recall that r, t, k depend only on the set of formulas Δ: r is from (1) of the theorem, t is the R-rank of Δ and k is the number of free variables.

(3) The "in particular" clause gives a simpler, though less accurate, bound, assuming 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 $g(x) = x^{\frac{1}{2tr}}$. Specialized to the set of formulas Δ_{k_*} from Definition 2.8, this implies there is a constant $c = c(\Delta_{k_*}) = c(r, t, k_*) \leq \alpha(r, t, k_*)$ such that we can extract from any set of size $n_1 > (cn_2)^{(2tr)^{k_*}}$ a Δ_{k_*}-indiscernible subsequence of size at least n_2.

This completes the proof of the theorem. □

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* $A = \langle a_i : i < s_1 \rangle$, $B = \langle b_j : j < s_2 \rangle$ *are* Δ_{k_*}-*indiscernible sequences. Suppose that* $s_1 \geq 2k_*$ *and* $s_2 > (2k_*)^2$.
Let $\mathcal{U} = \{i < s_1 : \exists^{\geq 2k_*} j < s_2)(a_j R b_i)\}$. *Then either* $|\mathcal{U}| \leq 2k_*$ *or* $|\mathcal{U}| \geq s_1 - 2k_*$.

Proof. Suppose the conclusion fails. Let i_1 be the k_*th member of $\mathcal{U}$, and let i_2 be the k_*th member of $\{0, \ldots s_1 - 1\} \setminus \mathcal{U}$. Clearly $i_1 \neq i_2$.
Case 1: $i_1 < i_2$. Choose elements $j_0 < \cdots < j_{k_*-1}$ from $\mathcal{U}$ and elements $j_{k_*} < \cdots < j_{k_*+k_*-1} < s_1$ from $\{0, \ldots s_1 - 1\} \setminus \mathcal{U}$ satisfying $j_{k_*-1} \leq i_1 < i_2 \leq j_{k_*}$. Recall by Claim 3.2 that each a_{j_ℓ} partitions B into a small and large set; for each $\ell < 2k_*$, let the "small set" be

$$W_\ell = \{i < s_2 : a_{j_\ell} R b_i \leftrightarrow \left((\exists^{\geq 2k_*} i < s_2) \neg (a_{j_\ell} R b_i) \right)\}$$

9

By Claim 3.2 and the definition of $\mathcal{U}$, each $|W_\ell| < 2k_*$. Thus

$$\left| \bigcup_{\ell < 2k_*} W_\ell \right| \leq (2k_*)^2 < |B|$$

Choose $n \in \{0, \ldots s_2 - 1\} \setminus \bigcup_{\ell < 2k_*} W_\ell$. Then for all ℓ such that $0 \leq \ell \leq k_* - 1$, $b_n R a_\ell$ and for all ℓ such that $k_* \leq \ell < 2k_*$, $\neg b_n R a_\ell$. By Observation 2.9, G has the k_*-order property, contradiction.
Case 2: $i_2 < i_1$. Similar, interchanging R and $\neg R$. $\qquad \square$

Conclusion 3.7. *Recall the hypotheses of this section: G is a finite graph with the non-k_*-order property.*
If (A) then (B).

(A) (1) $n_1 \to (n_2)_{T, \Delta_{k_*}, 1}$
(2) $n > n_1 n_2$ and $n_2 \geq (2k_*)^2$

(B) *if $A \subseteq G$, $|A| = n$, 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) $n = n_2 m_1 + m_2$, $m_2 < n_1 \leq m_1$
(d) *For each i, $|A_i| \in \{n_2, n_2 + 1\}$*
(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 $\mathbf{t}(A_i A_j) \in \{0, 1\}$, we have that for all but $\leq 2k_*$ $a \in A_i$, for all but $\leq 2k_*$ $b \in A_j$, $aRb \equiv \mathbf{t}(A_i, A_j)$.*

Proof. First, choose m_1 satisfying $n_2 m_1 \leq n < n_2 m_1 + n_1$ (so $m_1 \geq n_1$ by (A)(2)), and let $<_*$ be a linear order on A. Second, we choose A'_ℓ by induction on $\ell < m_1$ to satisfy:

- $A'_\ell \subseteq A \setminus \{A'_j : j < \ell\}$
- $|A'_\ell| = n_2$
- if we list the elements of A in $<_*$-increasing order as $\langle a_{\ell, i} : i < n_2 \rangle$, this is a Δ_{k_*}-indiscernible sequence.

The existence of such A'_ℓ is guaranteed by the hypothesis (A)(1). Since Δ_{k_*} includes $\{xRy\}$, we will have (e) by the symmetry of R.

Third, let $\langle a_i^* : i < m_2 \rangle$ list the remaining elements, i.e. those of $A \setminus \bigcup \{A'_\ell : \ell < m_1\}$. Let $A_\ell := A'_\ell \cup \{a_\ell^*\}$ if a_ℓ^* 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. $\qquad \square$

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 $\overline{G} = \langle G_i \rangle$ 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 $\leq 2k_*$ $b \in G_j$, $aRb \equiv \mathbf{t}(G_i, G_j)$.*

Moreover, $N = n_1 n_2$ suffices for any $n_1 > (cn_2)^{(2tr)^{k_*}}$, as computed in Theorem 3.5 in the case where $\Delta = \Delta_{k_*}$ (for a constant c depending only on Δ_{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 ϵ-regularity in the following senses.*
- *It is clearly hereditary for $C_i \subseteq A_i, |C_i| \geq |2k_*|^2$.*
- *The density of exception is small:*

$$\frac{|\{(a,b) \in A_i \times A_j : (aRb) \equiv \neg\mathbf{t}_{i,j}\}|}{|A_i||A_j|} \leq \frac{2k_*}{|A_i|} + \frac{2k_*}{|A_j|}$$

- *If $|A_i|, |A_j|$ are not too small, $\mathbf{t}_{i,j} = \mathbf{t}_{j,i}$.*
- *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 Φ of binary relations satisfying the non-k_*-order property. Third, we can replace 2-place by $\mathbf{n}(*)$-place where $\mathbf{n}(*) \leq \omega$, so Φ is a set of formulas of the form $\varphi(x_0, \ldots x_{n-1})$, $n < \mathbf{n}(*)$. In this case, the assumption (A)(1) of Conclusion 3.7 becomes

$$\lambda_1^+ \to_{T,\Delta} (\lambda_2)_1$$

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

(1) by Ramsey: $\lambda_1 \to (\lambda_2)^{<\mathbf{n}(*)}_{2^{|\Delta|}}$
(2) by Erdös-Rado, similarly
(3) using Erdös cardinals
    or
(4) use stability: [15] Chapter II, or better (in one model) [?] §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 dependent graphs, Definition 3.11. In subsequent sections, we return to stable 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 2.5 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.

Hypothesis 3.10. *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 $<^G$, along with a symmetric binary edge relation R. We assume that G is k_*-dependent.*

Definition 3.11. *Let $k_*<\omega$ be given. We say that G is k_*-dependent* when there are no $a_\ell \in G$ *(for $\ell < k_*$) and $b_u \in G$ (for $u \subseteq k_*$) such that $a_\ell R b_u$ iff $\ell \in u$.*

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 $i \leq n$,*

$$\models (\exists x)\left(\bigwedge_{j\leq i} \neg(xRa_j) \wedge \bigwedge_{j>i} xRa_j\right)$$

(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 $u \subseteq n$,*

$$\models (\exists x)\left(\bigwedge_{j\in u} \neg(xRa_j) \wedge \bigwedge_{j\notin u} xRa_j\right)$$

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 $\mathrm{mem}_A(\ell)$ be the ℓth member of A under $<^G$, for $\ell < |A|$ (for infinite A such that $<^G |_A$ is well ordered, $\ell < (\mathrm{otp}(A), <^G |_A)$, an ordinal)*
(2) *If $A \subseteq G$, $<^G$ is a well-ordering of A then $\mathrm{mem}_A(\ell)$ is defined similarly.*

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

Definition 3.15. *Let $\Delta_{k_*} = \{\varphi_\eta(x_0, \ldots x_{k_*-1}) : \eta \in {}^{k_*}2\}$ where*

$$\varphi_\eta(x_0, \ldots x_{k_*-1}) = (\exists y) \bigwedge_{\ell < k_*} (x_\ell Ry)^{if\ \eta(\ell)=1}$$

Claim 3.16. *Suppose $A, B \subset G$ are disjoint, both A, B are Δ_{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 half-$(1, k_*)$-nice. So there is $a \in A$ which "alternates" k_* times on B. That is, we may choose a $<^G$-increasing sequence of elements $b_{i_0}, \ldots b_{i_{2k_*=1}} \subseteq B$ 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 $J = \langle j_0, \ldots j_{k_*-1}\rangle$ be any set of indices of elements of B, of size k_*. Then for any $\sigma \subseteq J$, by indiscernibility, we have that

$$\models \exists x\left(\bigwedge_{\ell\in\sigma} xRa_{j_\ell} \wedge \bigwedge_{\ell\in J\setminus\sigma} \neg xRa_{j_\ell}\right)$$

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

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

    (1) *We say that $\overline{A}$ is half-**f**-nice when $i < j < i(*)$ implies (A_i, A_j) is half-**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) *$m_1 \to (m_2)^{\leq k_*}_{2^{|\Delta_{k_*}|}}$ 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_is 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 Δ_{k_*}-indiscernible*

    (6) *$\overline{A}$ is $(1, k_*)$-nice*

Proof. Straightforward. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

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 $\frac{n}{m_3}$ pieces in which, moreover, for any $i_1 \neq i_2 < i(*)$,*

    (1) *$\{(j_i, j_2) : (A_{i_1 j_i}, A_{i_2 j_2}) \text{ is neither full nor empty}\}$ 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) *$|\{(j_i, j_2) : j_1, j_2 < \frac{m_2}{m_3}\}| = \left(\frac{m_2}{m_3}\right)^2$*

    (3) *so the density of bad pairs is $\leq \frac{k_* m_3^2}{m_2}$*

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 ϵ-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 4.22, 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 4.22, 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 §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 2.10. (3) By convention f, g are nondecreasing functions from $\mathbb{N}$ to $\mathbb{N} \setminus \{0\}$.*

Definition 4.2. (ϵ- and f-indivisible)

(1) *Let $\epsilon \in (0,1)_{\mathbb{R}}$. We say that $A \subseteq G$ is ϵ-indivisible if for every $b \in G$, for some truth value $\mathbf{t}$, the set $\{a \in A : aRb \equiv \mathbf{t}\}$ 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 $\mathbf{t}$ such that $|\{a \in A : aRb \not\equiv \mathbf{t}\}| < f(|A|)$. 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 $\langle b_\eta : \eta \in {}^{k>}2 \rangle$ and $\langle A_\eta : \eta \in {}^{k\leq}2 \rangle$ such that:

(1) $A_{\langle\rangle} = A$
(2) $A_{\eta^\frown\langle i\rangle} \subset A_\eta$
(3) $A_{\eta^\frown\langle 0\rangle} \cap A_{\eta^\frown\langle 1\rangle} = \emptyset$
(4) $|A_\eta| = m_{\lg(\eta)}$
(5) $b_\eta \in G$
(6) $A_{\eta^\frown\langle i\rangle} = \{a \in A_\eta : aRb_\eta \equiv (i=1)\}$

There is no problem at $k = 0$, but let us verify that the induction cannot continue past k_{**}. For each $\eta \in {}^k2$ $A_\eta \neq \emptyset$ by (4), so choose $a_\eta \in A_\eta$. If for all $\eta \in {}^{k_{**}>}2$ there exists b_η such that $A_{\eta^\frown\langle 1\rangle}, A_{\eta^\frown\langle 2\rangle}$ 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 2.10. So for at least one η, it must be that no such b_η can be found in G, i.e that for any $b \in G$, either $|\{a \in A_\eta : aRb\}|$ or $|\{a \in A_\eta : \neg aRb\}|$ 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. $\square$

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, $|A_j| \in \{m_\ell : \ell \leq k_{**}\}$*
(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. $\square$

Claim 4.5. *Assume $\epsilon \in (0, \frac{1}{2})_{\mathbb{R}}$, $n^{\epsilon^{k_{**}}} > k_{**}$. Let $\langle m_\ell : 0 \leq \ell \leq k_{**}\rangle$ be a sequence of integers satisfying $n \geq m_0$, $m_{k_{**}} > k_{**}$ and for all ℓ s.t. $0 \leq \ell \leq k_{**}$, $m_{\ell+1} = \lfloor (m_\ell)^\epsilon \rfloor$.*
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 $\leq$-increasing*
(3) *for each $i < i(*)$ for some $\ell = \ell(i) < k_{**}$, $|A_\ell| = m_\ell$ and A_ℓ is ϵ-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. $\square$

14

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 $f(|A|) \cdot g(|B|) < \frac{1}{2}|B|$. Then for some truth value $\mathbf{t} = \mathbf{t}(A, B)$ for all but $< f(|A|)$ of the $a \in A$ for all but $< g(|B|)$ of the $b \in B$, we have that $aRb \equiv \mathbf{t}$.*

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 $\mathbf{t}_a = \mathbf{t}_a(a, B)$ such that $|\{b \in B : aRb \equiv \mathbf{t}_a\}| < g(|B|)$. For $i \in \{0, 1\}$, let $\mathcal{U}_i = \{a \in A : \mathbf{t}_a = i\}$. If $|U_i| < f(|A|)$ for either i, we are done, so assume this fails. Choose $W_i \subset \mathcal{U}_i$ so that $|W_i| = f(|A|)$ for $i \in \{0, 1\}$. Again we gather the exceptions: let $V = \{b \in B : (\exists a \in W_1)(\neg aRb) \vee (\exists a \in W_0)(aRb)\}$. 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. $\qquad\square$

Remark 4.7. *When $f(n) = n^\epsilon$, $g(n) = n^\zeta$ the translated condition is: if $|A|^\epsilon |B|^\zeta < \frac{1}{2}|B|$.*

Claim 4.8. *Let A be ζ-indivisible and B be ϵ-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 $A' \subset A, B' \subset B$, $|A'| \geq |A|^{\zeta+\zeta_1}$, $|B'| \geq |B|^{\epsilon+\epsilon_1}$, then:*

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

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 $|B|^\zeta$ corresponding exceptional points $b \in B'$. Thus we compute:

$$\frac{|A|^\zeta \cdot |B'| + (|A'| - |A|^\zeta)|B|^\epsilon}{|A'||B'|} = \frac{|A|^\zeta}{|A'|} + \left(\frac{|A'| - |A|^\zeta}{|A'|} \right) \frac{|B|^\epsilon}{|B'|}$$

$$\leq \frac{|A|^\zeta}{|A'|} + \frac{|B|^\epsilon}{|B'|} \leq \frac{|A|^\zeta}{|A|^{\zeta+\zeta_1}} + \frac{|B|^\epsilon}{|B|^{\epsilon+\epsilon_1}} = \frac{1}{|A|^{\zeta_1}} + \frac{1}{|B|^{\epsilon_1}}$$

A similar result holds for f-indivisible replacing ϵ-indivisible. $\qquad\square$

We single out the following special case for Theorem 4.22 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 $\zeta_1 \in (0, 1-\frac{c}{|A|})$, $\epsilon_1 \in (0, 1-\frac{c}{|B|})$, we have: if $A' \subset A, B' \subset B$, $|A'| \geq c|A|^{\zeta_1}$, $|B'| \geq c|B|^{\epsilon_1}$, then:*

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

Returning to the general argument, choosing ϵ_1 (here called ζ) 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 $\zeta \in (0, \epsilon^{k_{**}})$, we have in addition that for every $i < j < i(*)$, if $A \subset A_i, |A| \geq |A_i|^{\epsilon+\zeta}$, $B \subset A_j, |B| \geq |A_j|^{\epsilon+\zeta}$ and $\mathbf{t}_{i,j} = \mathbf{t}(A_i, A_j)$ is the associated truth value, then*

$$\left| \frac{\{(a,b) \in (A, B) \; : \; aRb \equiv \neg\mathbf{t}_{i,j}\}}{|A||B||} \right| \leq \frac{1}{|A_i|^\zeta} + \frac{1}{|A_j|^\zeta} \leq \frac{1}{|A|^\zeta} + \frac{1}{|B|^\zeta}$$

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 $\mathbf{t}_{i,j}$ is defined. $\qquad\square$

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 $A \setminus \bigcup_i A_i$, 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 4.22, we obtain much stronger indivisibility conditions on the components and no irregular pairs, at the price of a somewhat larger remainder, Theorem 4.22.

4.1. Towards a proof of Theorem 4.16.

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

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

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

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

Then $\mathrm{Prob}(\mathcal{E}^+_{A_i^s,A_j^t,m}) \ge 1 - \frac{2}{n^{\epsilon^{k_{**}}}}$.

Proof. By choice of A_i, A_j we have that $\mathbf{t} = \mathbf{t}(A_i, A_j)$ is well defined. Let $\mathcal{U}_1 = \{a \in A_i : |\{b \in A_j : aRb \equiv \neg\mathbf{t}\}| \ge |A_j|^\epsilon\}$, and for each $a \in A_i \setminus \mathcal{U}_1$, let $\mathcal{U}_{2,a} = \{b \in A_j : aRb \equiv \neg\mathbf{t}\}$. By definition of $\mathbf{t}$, $|\mathcal{U}_1| \le |A_i|^\epsilon$ and for each relevant a, $|\mathcal{U}_{2,a}| \le |A_j|^\epsilon$.

We first consider A_i^s. The probability P_1 that $A_i^s \cap \mathcal{U}_1 \ne \emptyset$ is bounded by the following:

$$P_1 < \frac{m|\mathcal{U}_1|}{|A_i| - m} \le \frac{n^\zeta |A_i|^\epsilon}{|A_i| - m} < \frac{n^{2\zeta}\left(n^{\epsilon^{\ell_a+2}}\right)}{n^{\epsilon^{\ell_a+1}}}$$

$$\le \frac{1}{n^{\epsilon^{\ell_a+1} - \epsilon^{\ell_a+2} - 2\zeta}} = \frac{1}{n^{\epsilon^{\ell_a+1}(1-\epsilon) - 2\zeta}} < \frac{1}{n^{\epsilon^{\ell_a+1}}} \le \frac{1}{n^{\epsilon^{k_{**}}}}$$

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

$$P_2 < \frac{m|\bigcup_{a \in A_i^s} \mathcal{U}_{2,a}|}{|A_j| - m} \le \frac{m \cdot m \cdot |A_j|^\epsilon}{|A_j| - m} \le \frac{n^{2\zeta}|A_j|^\epsilon}{|A_j| - n^\zeta} \le \frac{1}{n^{\epsilon^{k_{**}}}}$$

by the analogous calculation. So $\mathrm{Prob}(\mathcal{E}^+_{A_i^s,A_j^t,m}) \ge (1 - \frac{1}{n^{\epsilon^{k_{**}}}})^2 \ge 1 - \frac{2}{n^{\epsilon^{k_{**}}}}$. $\square$

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

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

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

    (1) *$\langle C_i : i < r\rangle$ refines the partition $\langle A_i : i < i(*)\rangle$*
    (2) *$|C_i| = m_{**}$ for each $i < r$*
    (3) *For all but at most $\frac{2}{n^{\epsilon^{k_{**}}}} r^2$ 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 $\{(a, b) \in C_i \times C_j : aRb \not\equiv \mathbf{t}(A_i, A_j)\} = \emptyset$.*

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 $(1 - \frac{1}{n^{\epsilon^{k_{**}}}})^2$ 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 $m_0 = m_*$; in this case the number of ways of choosing a pair C_i, C_j from the same original component is at most

$$\binom{\frac{m_*}{m_{**}}}{2} \frac{n}{m_*} \quad \text{out of a possible} \quad \binom{\frac{n}{m_{**}}}{2}$$

so the ratio is approximately

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

Recall that by hypothesis, $\frac{m_*}{n} < \frac{1}{n^{2\epsilon^{k_{**}}}}$.

Combining (a) and (b), the total fraction of irregular pairs does not exceed $\frac{2}{n^{\epsilon^{k_{**}}}}$. $\qquad\square$

Remark 4.15. *In Claim 4.14, the hypothesis on m_* could obviously be weakened, or dropped at the expense of increasing the fraction of irregular pairs by $\frac{m_*}{n}$, as the calculation in part (b) of proof shows.*

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

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

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

Proof. Recall that $\epsilon = \frac{1}{r}$. (This hypothesis is just to ensure divisibility, and could be modified or dropped in favor of allowing for slight rounding errors.) Choose m_{**} maximal so that $(m_{**})^{r^{k_{**}}} \leq n$, and subject to the constraint that $\frac{m_*}{n} < \frac{1}{n^{2\epsilon^{k_{**}}}}$. (One can drop this constraint, by Remark 4.15, at

the cost of increasing the fraction in item (3) by $\frac{m_*}{n}$.) By hypothesis $m_{**} > k_{**}$. Then the sequence $\langle m_\ell : \ell < k_{**} \rangle$ satisfies the hypotheses of Claims 4.3 and 4.5, and furthermore m_{**} divides m_ℓ for $\ell < k_{**}$. Apply Claim 4.5 to obtain a decomposition into ϵ-indivisible pieces A_i' such that for each i and some ℓ, $|A_i'| = m_\ell$. Claim 4.14 gives a further partition into pieces $\langle A_i : i < i(*) \rangle$ each of size m_{**}; additionally, we partition the remainder from Claim 4.5 into pieces of size 1. Let ζ be such that $m_{**} = n^\zeta$. This gives clause (1), and clause (2) holds by Claim 4.5(d) since $m_0 = (m_{**})^{r^{k_{**}-1}}$. Condition (3) holds by Claim 4.13. Finally,

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

$\square$

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

4.2. **Towards a proof of Theorem 4.22.** 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 ϵ-indivisible for $\epsilon \in (0,1)_{\mathbb{R}}$ was shorthand 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
   If (1) $|A| = n$
     (2) $|\mathcal{P}| \leq n^{\frac{1}{\zeta}}, \mathcal{P} \subseteq \mathcal{P}(A)$
     (3) $(\forall B \in \mathcal{P})(|B| \leq n^\epsilon)$
  then there exists $\mathcal{U} \subseteq A$, $|\mathcal{U}| = \lfloor n^\xi \rfloor$ such that $(\forall B \in \mathcal{P}) \, (|\mathcal{U} \cap B| \leq c)$.

Lemma 4.19. *If the reals ϵ, ζ, ξ and the natural numbers n, c satisfy:*
  (a) $\epsilon \in (0, 1)$, $\zeta > 0$
  (b) $0 < \xi < \min(1 - \epsilon, \frac{1}{2})$
  (c) n *sufficiently large, i.e.* $n > \mathbf{n}(\epsilon, \zeta, \xi, c)$ *from Remark 4.20*
  (d) $c > \frac{1}{\zeta(1-\xi-\epsilon)}$
  then $\bigoplus[n, \epsilon, \zeta, \xi, c]$ 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*

$$ \frac{1}{n^{1-2\xi}} + \frac{1}{n^{(1-\xi-\epsilon)c - 1/\zeta}} < 1 $$

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 $\mathbf{n}(\epsilon, \zeta, \xi, c)$ be the minimal $n \in \mathbb{N}$ for which the displayed equation is true.

Claim 4.21. *Suppose that we are given constants $k, c \in \mathbb{N}$ and $\epsilon, \xi, \zeta \in \mathbb{R}$ such that:*
  (1) $A \subseteq G$ *implies* $|\{\{a \in A : aRb\} : b \in G\}| \leq |A|^k$
  (2) $\epsilon \in (0, \frac{1}{2})$

(3) $\xi \in (0, \frac{1}{2})$ is such that $\xi < \epsilon^{k_{**}}$ and for all natural numbers $\ell \leq k_{**}$, $\frac{\xi}{\epsilon^\ell} < \frac{1}{2} < 1 - \epsilon$.

(4) the constant c satisfies

$$c > \frac{1}{\zeta(1 - \frac{\xi}{\epsilon^{k_{**}}} - \epsilon)}$$

Then for every sufficiently large $n \in \mathbb{N}$ (meaning $n > \mathbf{n}(\epsilon, \zeta, \xi, c)$ 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 $\mathbf{t} \in \{0,1\}$ such that for all but k elements $a \in Z$, $aRb \equiv t$

Proof. Since G has the non-k_*-order property (see the second paragraph of 2.14), $k = k_*$ will satisfy condition (1) by Claim 2.5, and recall that k_{**} is the associated tree bound from 2.10. For lower bounds on the size of n, see Remark 4.20.

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_ℓ by induction on $\ell < k_{**}$ so that $m_{\ell+1} = \lfloor (m_\ell)^\epsilon \rfloor$. By Claim 4.3 there is $\ell < k_{**}$ and $A_1 \subseteq A$ such that $|A_1| = m_\ell$ and A_1 is ϵ-indivisible. Let $\mathcal{P}_1 = \{\{a \in A_1 : aRb\} : b \in G\}$. So $|\mathcal{P}_1| \leq |A|^k = (m_\ell)^k$, by choice of k.

We would like to apply Lemma 4.19 to conclude that $\bigoplus[\epsilon, \frac{1}{k}, \frac{\xi}{\epsilon^\ell}, c]$ holds for $A = A_1$, $\mathcal{P} = \mathcal{P}_1$. Let us verify that the hypotheses of that Lemma hold:

- (1),(2) hold as $|A_1| = m_\ell$, and $|\mathcal{P}_1| = (m_\ell)^k = (m_\ell)^{\frac{1}{\frac{1}{k}}}$
- (3) holds by definition of $\mathcal{P}_1$, as A_1 is ϵ-indivisible
- (a) clear
- (b), (d) by choice of ξ and c in this Claim
- (c) by choice of n "sufficiently large"

We conclude that there is $Z \subseteq A_1$ (i.e. the $\mathcal{U}$ guaranteed by Lemma 4.19) which is c-indivisible and satisfies

$$|Z| = \left\lfloor (m_\ell)^{\frac{\xi}{\epsilon^\ell}} \right\rfloor = \left\lfloor (n^{\epsilon^\ell})^{\frac{\xi}{\epsilon^\ell}} \right\rfloor = \left\lfloor n^\xi \right\rfloor$$

which completes the proof. $\qquad \square$

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

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

We will show that for $n > \mathbf{n}(\epsilon, \zeta, \xi, c)$ in the sense of Remark 4.20, there is nonzero probability that a sequence $\eta \in \mathcal{F}_*$ satisfies (1) all the elements of η 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) $(1 - \xi - \epsilon)c - \frac{1}{\zeta} > 0$ and c is a natural number. We proceed to compute several probabilities.

$\circledast_2$ *The probability that η 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.)

19

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

$\circledast_3$ *The probability that η 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

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

$\circledast_4$ *The probability that η intersects some $B \in \mathcal{P}$ in more than c places.* By $\circledast_3$,

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

$$\le n^{\frac{1}{\zeta}} \cdot \frac{1}{n^{(1-\xi-\epsilon)c}} = \frac{1}{n^{(1-\xi-\epsilon)c - \frac{1}{\zeta}}}$$

As remarked above, it suffices for the Lemma to show that the sum of the probabilities $\circledast_2 + \circledast_4 < 1$, i.e. that

$$\frac{1}{n^{1-2\xi}} + \frac{1}{n^{(1-\xi-\epsilon)c - \frac{1}{\zeta}}} < 1$$

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

Theorem 4.22. *Let k_* and therefore 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 $\epsilon, \zeta \in \mathbb{R}$ which, along with k, satisfy the hypotheses of Claim 4.21, and any $\theta \in \mathbb{R}$, $0 < \theta < 1$,

there exists $N = N(k_*, k, c, \epsilon, \zeta, \theta)$ such that

for any $A \subseteq G$, $|A| = n > N$, there is $i(*) \in \mathbb{N}$ and a partition $\langle A_i : i < i(*)\rangle$ of A into disjoint pieces (plus a remainder) satisfying:

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

Remark 4.23. *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 4.22) Assume that n is large enough so that $n^\theta > \mathbf{n}(\epsilon, \zeta, \xi, c) + 1$, where $\mathbf{n}(...)$ is the lower bound from Lemma 4.19 and Remark 4.20. Note that by choice of k, k satisfies Claim 2.5.

We are aiming for pieces of uniform size $n^{\theta\zeta}$. First, given θ, define by induction a decreasing sequence $\langle m_\ell : \ell \le k_{**}\rangle$ by $m_{k_{**}-1} = \lfloor n^\theta\rfloor$, $m_{k_{**}} = \lfloor (m_{k_{**}-1})^\epsilon\rfloor$ and for each $1 < j \le k_{**}$, $m_{k_{**}-j} = \left\lceil (m_{k_{**}-j+1})^{\frac{1}{\epsilon}}\right\rceil$. 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 B_i, using the decreasing sequence $\langle m_\ell : \ell \le k_{**}\rangle$ just defined, to obtain an ϵ-indivisible $B_i \subseteq R_i$. By construction, for some

$1 \leq \ell \leq k_{**}$ this set B_i will have cardinality $m_{k_{**}-\ell} = \left\lfloor (n^\theta)^{\frac{1}{\ell-1}} \right\rfloor$. (Note that ϵ-indivisibility need not be preserved under taking subsets.)

By the first line of the proof (recall $m_{k_{**}-1} = \lfloor n^\theta \rfloor$), we have $|B_i| > \mathbf{n}(\epsilon, \zeta, \xi, c)$. 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 $|B_i|^\zeta$. That is, $|Z_i| = |B_i|^\zeta = \left\lfloor n^{\theta \cdot \frac{1}{\ell-1} \cdot \zeta} \right\rfloor$ for some $1 \leq \ell \leq k_{**}$. 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 $\lfloor n^{\theta\zeta} \rfloor$. This completes the construction at stage i.

This construction can continue as long as the remainder B_i has size at least $m_0 = \left\lceil n^{\frac{\theta}{\epsilon^{k_{**}-1}}} \right\rceil$, as required by Claim 4.3; the remainder must have strictly smaller size, which completes the proof. $\square$

5. REGULARITY FOR STABLE GRAPHS

Thus far, we have given several regularity lemmas for the class of stable (or, in Section 3.2, dependent) 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_*-order property which unilaterally improves the usual result, Theorem 5.18: for each ϵ, there is $m = m(\epsilon)$ such that all sufficiently large G with the non-k_*-order property admit an equitable distribution such that (1) there are *no* irregular pairs, (2) each component satisfies a strong indivisibility condition, called ϵ-excellence, and (3) the bounds are much improved. For most of the construction, "regularity" of pairs means ϵ-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.

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_* fixed throughout this section, G has the non-k_*-order property, Definition 2.3 and* (c) k_{**} *is the corresponding bound on the height of a 2-branching tree, Definition 2.10. Throughout this section ϵ, ζ, ξ are reals $\in (0, \frac{1}{2})$.*

Definition 5.2. (Good, excellent)
   (1) *We say that $A \subseteq G$ is ϵ-good when for every $b \in G$ for some truth value $\mathbf{t} = \mathbf{t}(b, A) \in \{0, 1\}$ we have $|\{a \in A : (aRb) \not\equiv \mathbf{t}\}| < \epsilon|A|$. As $\epsilon < \frac{1}{2}$, this is meaningful.*
   (2) *We say that $A \subseteq G$ is (ϵ, ζ)-excellent when*
      (a) *A is ϵ-good and moreover*
      (b) *if $B \subseteq G$ is ζ-good then for some truth value $\mathbf{t} = \mathbf{t}(B, A)$,*
          $|\{a \in A : \mathbf{t}(a, B) \neq \mathbf{t}(B, A)\}| < \epsilon|A|$.
      Again, as $\epsilon < \frac{1}{2}$ the average is meaningful. When $\epsilon = \zeta$, we will just write ϵ-excellent.

Remark 5.3. *Any set $A \subset G$ satisfying condition* (b) *for ϵ-excellence must also be ϵ-good, since any singleton set $\{b\}$ is clearly ϵ-good (in fact, ϵ-excellent). Any B which satisfies $(\forall a \in G) \bigvee_{\mathbf{t} \in \{0,1\}} (\forall b \in B)(aRb \equiv \mathbf{t})$ will also be excellent.*

The next claim, which will be used repeatedly, gives a way to extract ϵ-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_ℓ to the levels of the tree.

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

(I) *For every $A \subseteq G$, $|A| \geq \frac{1}{\epsilon^{k_{**}}}$, there is A' such that:*
  (a) $A' \subseteq A$
  (b) $|A'| \geq \epsilon^{k_{**}-1}|A|$
  (c) A' *is ϵ-excellent*
(II) *Alternately, suppose we are given a decreasing sequence of natural numbers $\langle m_\ell : \ell < k_{**}\rangle$ such that $\epsilon m_\ell \geq m_{\ell+1}$ for $\ell < k_{**}-1$, and $m_{k_{**}-1} > k_{**}$. Then for every $A \subseteq G$, $|A| \geq \frac{1}{\epsilon^{k_{**}}}$, there is A' such that (a),(b)',(c)' hold, where:*
  (b)' $|A'| = m_\ell$ *for some $\ell < k_{**}$*
  (c)' A' *is $\frac{m_{\ell+1}}{m_\ell}$-excellent (so in particular, ϵ-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 $k \leq k_{**}$ such that:

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

Note that $\mathbf{t}(a, B_\eta)$ is well defined in (8) as B_η is good. When $k = 0$, define $A_{\langle\rangle} = A$. Now suppose $k = m + 1$. In Case (I), suppose that for all $\eta \in {}^m 2$, A_η fails to be ϵ-excellent. By definition, for each such η, there is some set $B_\eta \subset G$ which is ϵ-good and such that

$$|\{a \in A_\eta : \mathbf{t}(a, B) \neq 1\}| \geq \epsilon|A_\eta| \text{ and } |\{a \in A_\eta : \mathbf{t}(a, B) \neq 0\}| \geq \epsilon|A_\eta|$$

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

Meanwhile, in case (II), we are interested in whether A_η is $\frac{m_{k+1}}{m_k}$-excellent rather than ϵ-excellent; if not, there an $\frac{m_{k+1}}{m_k}$-good set B_η such that the displayed equation holds with "$\geq \frac{m_{k+1}}{m_k}|A_\eta|$" in place of "$\geq \epsilon|A_\eta|$". In this case, choose $A_{\eta^\frown\langle i\rangle}$ to be a subset of $\{a \in A_\eta : \mathbf{t}(a, B_\eta) \neq i\}$ 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. Suppose we have defined A_η for $\eta \in {}^{k_{**}}2$ and B_ν for $\nu \in {}^{k_{**}>}2$ satisfying (1)-(8). For each η, since we assumed either (I) $\epsilon^{k_{**}}|A| > 0$ or (II) that $|A_\eta| = m_\ell \geq m_{k_{**}-1} > k_{**}$, we have that $A_\eta \neq \emptyset$ so we may choose $a_\eta \in A_\eta$. Furthermore, for each $\nu \in {}^{k_{**}>}2$ and $\eta \in {}^{k_{**}}2$ such that $\nu \triangleleft \eta$, we may define

$$\mathcal{U}_{\nu,\eta} = \{b \in B_\nu : (a_\eta R b) \not\equiv t(a_\eta, B_\nu)\}$$

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

$$\left|\bigcup\{\mathcal{U}_{\nu,\eta} : \nu \triangleleft \eta \in {}^{k_{**}}2\}\right| < 2^{k_{**}}\epsilon|B_\nu| < |B_\nu|$$

by the hypothesis on the size of ϵ. In particular, for each $\nu \in {}^{k_{**}>}2$ we may choose an element $b_\nu \in B_\nu \setminus \bigcup\{\mathcal{U}_{\nu,\eta} : \nu \triangleleft \eta \in {}^{k_{**}}2\}$. Now the sequences $\langle a_\eta : \eta \in {}^{k_{**}}2\rangle$ and $\langle b_\nu : \nu \in {}^{k_{**}>}2\rangle$ contradict Definition 2.10, i.e. the choice of k_{**}.

We have shown that for some $k < k_{**}$ the induction must stop. Hence for some $\nu \in {}^{k}2$, A_ν is ϵ-excellent [if in case (II), A_ν is $\frac{m_{k+1}}{m_k}$-excellent, so in particular ϵ-excellent] and satisfies condition (5), which completes the proof. $\qquad\square$

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

Claim 5.6. *If A is ϵ-excellent and B is ζ-good then the pair (A, B) is (ϵ, ζ)-uniform, where we say that (A, B) is (ϵ, ζ)-uniform if for some truth value $\mathbf{t} = \mathbf{t}(A, B) \in \{0, 1\}$ we have: for all but $< \epsilon|A|$ of the elements of A, $\mathbf{t}(A, B) = \mathbf{t}(a, B)$.*
  In other words, for all but $< \epsilon|A|$ of the elements of $|A|$, for all but $< \zeta|B|$ of the elements of B, $(aRb) \equiv (\mathbf{t}(A, B) = 1)$. When $\epsilon = \zeta$, we will just write ϵ-uniform.

Proof. By the definition of *excellent*. $\qquad\square$

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

Discussion 5.8. *At this point, we have a way to obtain ϵ-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 ϵ-excellent sets. In preparation, then, we now apply several facts from probability to prove that sufficiently large ϵ-excellent sets can be equitably partitioned into a small number of pieces all of which are ϵ'-excellent for ϵ' close to ϵ.*

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*

$$\mathrm{Prob}\left(\frac{|A' \cap B|}{|A'|} \in \left(\frac{|B|}{|A|} - \zeta, \frac{|B|}{|A|} + \zeta\right)\right)$$

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 Erdös 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. $\qquad\square$

Fact 5.10. *(Vapnik and Chervonenkis, [19]) 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 $X_r = x_1, \ldots x_r$ of elements of X, A is said to induce the subset of $\{x_1, \ldots x_r\}$ consisting of those elements x_i which belong to A. The number of different subsamples of X_r induced by sets of S is denoted $\Delta^S(x_1, \ldots x_r)$. Define $m^S(r) = \max\{\Delta^S(x_1, \ldots x_r)\}$, where the maximum is taken over all samples of size r.
  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 ϵ, $\lim_{\ell \to \infty} \mathrm{Prob}(\pi^{(\ell)} > \epsilon) = 0$, is that there exist a finite k such that $m^S(\ell) \leq \ell^k + 1$ for all ℓ.

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) ([19] p. 272) *Given k from the last paragraph of Fact 5.10, if ℓ satisfies*

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

   then in any sample of size at least ℓ, with probability at least $(1-\eta)$, the relative frequencies differ from their corresponding probabilities by an amount less than ζ, 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 ϵ, ζ there is N_1 such that for all $n > N_1 = N_1(\epsilon, \zeta)$, if $A \subset G$, $|A| = n$, A is ϵ-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 $b \in G \implies \mathbf{t}(b, A') = \mathbf{t}(b, A)$.*
   (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 $\leq \xi$.*
   (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 ϵ-excellent, there exists a partition of A into r disjoint pieces of equal size each of which is ϵ'-excellent. Note that $N(\epsilon, \epsilon', r)$ increases with r.*

Proof. (1) Call $B \subset A$ an *exceptional set* if there is $b \in G$ such that $B = \{a \in A : aRb \not\equiv \mathbf{t}(b, A)\}$ 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_*-order property, Claim 2.5 in the case where $G = A$, $A = A'$ 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 ϵ be as given, $\zeta = \epsilon' - \epsilon$, and $\xi = \frac{1}{r+1}$. 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 $h^{-1}(s) \in [A]^{\frac{n}{r}}$. Since h was random, for each given s, each $B \in [A]^{\frac{n}{r}}$ 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 $1 - r\xi = 1 - \frac{r}{r+1} > 0$, 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. $\qquad\square$

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 $m_* := m_0$, $m_{**} := m_{k_{**}-1} \geq k_{**}$. 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) *$i \neq j \implies A_i \cap A_j = \emptyset$*
   (d) *each A_i is ϵ-excellent*
   (e) *<u>hence</u> if $i \neq j < j(*)$ then the pair (A_i, A_j) is (ϵ, ϵ)-uniform*
   (f) *$B := A \setminus \bigcup\{A_i : i < i(*)\}$ has $< m_*$ members*

(1A) *Suppose further that:*
   (i) *$m_{**}|m_k$ 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 $\langle A_i : i < i(*) \rangle$ to refer to this new partition) such that for each $i < i(*)$, $|A_i| = m_{**}$. Furthermore, each of these new pieces A_i is ϵ'-excellent.

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

$$|B_i| \in \left\{ \left\lfloor \frac{|B|}{i(*)} \right\rfloor, \left\lfloor \frac{|B|}{i(*)} \right\rfloor + 1 \right\}$$

Let $A'_i = A_i \cup B_i$ 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*

$$\zeta = \max\left\{ \frac{\epsilon'|A_i| + |B_i|}{|A_i| + |B_i|} : i < i(*) \right\} \leq \frac{\epsilon' m_{**} + \left\lceil \frac{m_*}{i(*)} \right\rceil}{m_{**} + \left\lceil \frac{m_*}{i(*)} \right\rceil}$$

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

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

Proof. (1) Applying Claim 5.4 we try to choose a sequence of ϵ-excellent sets A_i, each of size m_ℓ 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 ms to satisfy $m_{**} \approx \epsilon^{k_{**}} m_*$ so that $N = N(\epsilon, \epsilon', \epsilon^{-k_{**}})$ 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 ζ 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, $|B_i| \le 2\epsilon'|A_i|$, as then we would have

$$\frac{\epsilon'|A_i| + |B_i|}{|A_i| + |B_i|} \le \frac{\epsilon'|A_i| + 2\epsilon'|A_i|}{|A_i|} = 3\epsilon'$$

We verify that the assumption on m_* is enough to give this bound. By definition, as the B_is arise from an equitable partition of the remainder B, $|B_i| \le \frac{m_* - 1}{i(*)} + 1$, 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 $|B| \le m_* - 1$, we can bound $i(*)$ by $\frac{n}{m_{**}} \ge i(*) \ge \frac{n - m_* + 1}{m_{**}} > \frac{n - m_*}{m_{**}}$. Thus

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

We had assumed that $m_* \le \frac{\epsilon'n + 1}{1 + \epsilon'}$, and so:

$$m_*(1 + \epsilon') \le \epsilon'n + 1$$
$$m_* - 1 \le (n - m_*)\epsilon'$$
$$\frac{m_* - 1}{n - m_*} \le \epsilon'$$

We had also assumed that $\frac{1}{\epsilon'} < m_{**}$, 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. □

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

(1) *$\epsilon_1, \epsilon_3 \in \mathbb{R}$ such that $0 < \epsilon_3 < \epsilon_2 := \frac{\epsilon_1}{3} < \epsilon_1 < 2^{-k_{**}}$*
(2) *A sequence of positive integers $\langle m_\ell : \ell < k_{**} \rangle$ such that:*
    (a) *$m_{\ell+1} < \epsilon_3 m_\ell$ for each $\ell < k_{**}$*
    (b) *$m_{**} | m_\ell$ for each $\ell < k_{**}$*
    (c) *$\log\log m_0 \le m_{**}$*
    (d) *$m_{**} := m_{k_{**}-1} \ge \max(k_{**}, \frac{1}{\epsilon_2})$*
    (e) *$m_{k_{**}-2} > N(\epsilon_3, \epsilon_2, \frac{m_0}{m_{**}})$, from Claim 5.13(3)*
(3) *$A \subseteq G$, $|A| = n$ where n satisfies $m_0 \le \frac{\epsilon_2 n + 1}{1 + \epsilon_2}$*

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

- *for all $i < j < i(*)$, $||A_i| - |A_j|| \le 1$*
- *each A_i is ϵ_1-excellent*
- *Each pair (A_i, A_j) is ϵ_1-uniform*

Proof. By Claim 5.14, using $\epsilon = \epsilon_3$, $\epsilon' = \epsilon_2$ and $3\epsilon' = \epsilon_1$; 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 $\epsilon = \epsilon_1$, and then choose ϵ_3 to run the proof of Corollary 5.15. The role of the respective ϵs appears in conditions (2)(a) and (2)(e) of this Corollary. On one hand, ϵ_3 determines the rate of decrease of the sequence of ms, thus the size of m_{**}, and ultimately the number of components in the partition: so one would usually want to choose ϵ_3 close to $\epsilon_1 = \epsilon$.*

On the other hand as ϵ_3 approaches ϵ_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 ϵ-uniformity to ϵ-regularity. As the following calculation shows, η-uniform pairs will be ρ-regular when ρ (the parameter for a lower bound on the size of a subset chosen) is sufficiently large relative to η (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 ϵ-regular pairs will be near 0 or 1.

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

    (1) $\frac{|Z|}{|A||B|} < \epsilon + \zeta$

    (2) $\frac{|Z'|}{|A'||B'|} < \frac{\epsilon+\zeta}{\xi}$

In particular, if the pair (A, B) is ϵ_0-uniform for $\epsilon_0 \leq \frac{\epsilon^2}{2}$ then (A, B) is also ϵ-regular.

Proof. Let A', B' be given. For $a \in A$, let $\mathcal{W}_a = \{b \in B : aRb \not\equiv \mathbf{t}(A, B)\}$, and let $\mathcal{U} = \{a \in A : |\mathcal{W}_a| > \epsilon|A|\}$. So $|\mathcal{U}| < \epsilon|A|$, and $a \in A \setminus \mathcal{U} \implies |\mathcal{W}_a| < \zeta|B|$. Since

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

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

we can bound the cardinalities as follows:

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

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

and likewise

$$\frac{|Z'|}{|A' \times B'|} = \frac{|\mathcal{U}||B'| + |A'| \cdot \max\{|\mathcal{W}_a| : a \in \mathcal{U}\}}{|A'||B'|}$$

$$< \frac{\epsilon|A|\xi|B| + \xi|A|\zeta|B|}{|A'||B'|} \cdot \frac{|A||B|}{|A||B|} = (\epsilon\xi + \xi\zeta) \cdot \frac{|A||B|}{|A'||B'|} = \frac{\xi(\epsilon+\zeta)}{\xi^2} = \frac{\epsilon+\zeta}{\xi}$$

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

For the "in particular" clause, let $d(X, Y) = \frac{e(X,Y)}{|X||Y|}$ be the usual edge density. We have shown that if $\mathbf{t}(A, B) = 1$, $d(A, B) > 1 - (\epsilon + \zeta)$ 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 $\frac{\epsilon+\zeta}{\xi}$. If (A, B) is (ϵ_0, ϵ_0)-uniform and ϵ is such that $|A'| \geq \epsilon|A|$, $|B'| \geq \epsilon|B|$ where $\epsilon_0 \leq \frac{\epsilon^2}{2}$ then the difference in densities is bounded by $\frac{\epsilon^2}{\epsilon} = \epsilon$, which completes the proof. $\square$

We now give the main result of this section. Recall the definitions of non-k_*-order property (Definition 2.3), k_{**} (Definition 2.10), ϵ-excellent (Definition 5.2), and ϵ-uniform (Claim 5.6).

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

(1) *for all $i, j < i(*)$, $||A_i| - |A_j|| \leq 1$*
(2) *each of the pieces A_i is ϵ-excellent*
(3) *all of the pairs (A_i, A_j) are (ϵ, ϵ)-uniform*
(4) *if $\epsilon < \frac{1}{2^{k_{**}}}$, then $m \leq (3 + \epsilon)\left(\frac{8}{\epsilon}\right)^{k_{**}}$*

Proof. Without loss of generality, assume $\epsilon < \frac{1}{2^{k_{**}}}$. (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: $\epsilon_3 = \frac{\epsilon}{4}$, $\epsilon_2 = \frac{\epsilon}{3}$, and $\epsilon_1 = \epsilon$.

Step 1: Fixing q. Given ϵ_3, let $q = \left\lceil \frac{1}{\epsilon_3} \right\rceil \in \mathbb{N}$. It follows that $\frac{2}{\epsilon_3} \geq q \geq \frac{1}{\epsilon_3}$ and thus $\frac{\epsilon_3}{2} \leq \frac{1}{q} \leq \epsilon_3$. In particular, any sequence $\langle m_\ell : \ell < k_{**} \rangle$ such that $m_{**} := m_{k_{**}-1} \in \mathbb{N}$ and $m_\ell = q m_{\ell+1}$ for all $\ell < k_{**}$ will satisfy $m_{\ell+1} = \frac{1}{q} m_\ell \leq \epsilon_3 m_\ell$, $m_\ell \in \mathbb{N}$ 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 $\langle m_\ell : \ell < k_{**} \rangle$ 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

$$m_* \leq \frac{\frac{\epsilon}{3} n + 1}{1 + \frac{\epsilon}{3}} \text{ so it suffices to choose } m_* \leq \frac{\frac{\epsilon}{3} n}{1 + \frac{\epsilon}{3}} = \frac{\epsilon n}{3 + \epsilon}$$

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

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

Thus setting $m_{**} := \max\{c \in \mathbb{N} : c > k_{**}, \ c > \frac{1}{\epsilon_2}, \ q^{k_{**}-1} c \leq \frac{\epsilon n}{3+\epsilon}\}$ 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_ℓ 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 $\langle m_\ell : \ell < k_{**} \rangle$ 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, $\frac{\epsilon n}{3+\epsilon} - q^{k_{**}-1} < q^{k_{**}-1} m_{**}$, so assuming n is not too small [again (H1) in Step 5],

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

Step 4: Bounding $\frac{n}{m_{**}}$. Applying Step 3, an inequality from Step 1, and the definition of ϵ_3,

$$\frac{n}{m_{**}} \leq \frac{n}{\frac{1}{2}\left(\frac{\epsilon n}{3+\epsilon}\right)\left(\frac{1}{q^{k_{**}-1}}\right)} = \frac{2(3+\epsilon)q^{k_{**}-1}}{\epsilon} \leq \frac{2(3+\epsilon)}{\epsilon}\left(\frac{2}{\epsilon_3}\right)^{k_{**}-1} = (3 + \epsilon)\left(\frac{2}{\epsilon}\right)\left(\frac{2}{\frac{\epsilon}{4}}\right)^{k_{**}-1}$$

$$\leq (3 + \epsilon)\left(\frac{8}{\epsilon}\right)^{k_{**}}$$

Note that a choice of ϵ_3 closer to ϵ_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 $\frac{\epsilon n}{3+\epsilon} - q^{k_{**}-1} > k_{**}q^{k_{**}-1}$
   and also ensures that $n > 2q^{k_{**}-1}\left(\frac{3+\epsilon}{\epsilon}\right)$, for the calculation in Step 3

(H2) n is large enough for the sequence $\langle m_\ell : \ell < k_{**} \rangle$ 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 $n \geq N(\frac{\epsilon}{3}, \frac{\epsilon}{2}, q^{k_{**}-1}) \cdot \left(\frac{3+\epsilon}{\epsilon}\right) \cdot q^{k_{**}-2}$ where $N(\cdot,\cdot,\cdot)$ 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 $\langle m_\ell : \ell < k_{**} \rangle$ 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. $\qquad\square$

Corollary 5.19. *Let G be a graph with the non-k_*-order property. For every $\epsilon \in (0, \frac{1}{2})$ there are N, m such that if $A \subseteq G$, $|A| \geq N$, then for some $\ell \leq m$, there is a partition $A = \langle A_i : i < \ell \rangle$ such that each A_i is ϵ-excellent, and for every $0 \leq i < j < \ell$,*

- $||A_i| - |A_j|| \leq 1$
- (A_i, A_j) *is ϵ-regular and*
- *if $B_i \in [A_i]^{\geq \epsilon|A_i|}$ and $B_j \in [A_j]^{\geq \epsilon|A_j|}$ then*

$$\left(d(B_i, B_j) < \epsilon\right) \vee \left(d(B_i, B_j) \geq 1 - \epsilon\right)$$

Proof. This is a slight weakening of Theorem 5.18, which also replaces "ϵ-uniform," as defined in Claim 5.6, by the more familiar ϵ-regular via Claim 5.17. For ϵ-excellent, see Definition 5.2. $\qquad\square$

References

[1] N. Alon, R. Duke, H. Leffman, V. Rödl, R. Yuster, "The algorithmic aspects of the regularity lemma." FOCS 33(1992), 479-481; Journal of Algorithms 16 (1994), 80-109.

[2] P. Erdös and A. Rényi, "On the central limit theorem for samples from a finite population." Publications of the Mathematical Institute of the H.A.S. 4 (1959), 49–61, available at www.renyi.hu/∼p_erdos.

[3] W. Feller, *An introduction to probability theory and its applications* vol I, Wiley, 3rd edition (1968).

[4] W. T. Gowers, "Lower bounds of tower type for Szemerédi's uniformity lemma." GAFA, vol. 7 (1997) 322-337.

[5] E. Hrushovski, Y. Peterzil, and A. Pillay, "Groups, measures, and the NIP." J. Amer. Math. Soc. 21 (2008), no. 2, 563-596.

[6] W. Hodges, *Model Theory*, Encyclopedia of Mathematics, Cambridge University Press.

[7] I. Kaplan and S. Shelah, "Examples in dependent theories." arXiv:1009.5420v1 (2010)

[8] J. Komlós and M. Simonovits, "Szemerédi's Regularity Lemma and its applications in graph theory." (1996) In *Combinatorics: Paul Erdös is Eighty*, Vol. 2 (D. Miklós, V. T. Sós and T. Szönyi, eds), Bolyai Society Math. Studies, Keszthely, Hungary, pp. 295-352.

[9] L. Lovász and B. Szegedy, "Regularity partitions and the topology of graphons." *An Irregular Mind (Szemerédi is 70)*, Bolyai Society Mathematical Studies, 21 (2010) pp. 415-446.

[10] M. C. Laskowski, "Vapnik-Chervonenkis classes of definable sets." J. London Math Soc. (2) 45 (1992) 377-384.

[11] M. Malliaris, "The characteristic sequence of a first-order formula." JSL, 75, 4 (2010) pp. 1415-1440.

[12] M. Malliaris, "Edge distribution and density in the characteristic sequence." Annals of Pure and Applied Logic 162, 1, (2010) 1–19.

[13] M. Malliaris, "Hypergraph sequences as a tool for saturation of ultrapowers." Submitted (2010).

[14] W. L. Nicholson, "On the Normal Approximation to the Hypergeometric Distribution." Ann. Math. Statist. Volume 27, Number 2 (1956), 471-483.

[15] S. Shelah, *Classification Theory and the number of non-isomorphic models*, rev. ed. North-Holland, 1990.

[16] S. Shelah, "A dependent dream and recounting types." (2009) http://shelah.logic.at

[17] S. Shelah, Classification theory for elementary classes with the dependence property - a modest beginning." Scientiae Mathematicae Japonicae 59, No. 2; (special issue: e9, 503–544) (2004), 265–316, math.LO/0009056

[18] E. Szemerédi, "On sets of integers containing no k elements in arithmetic progression," Acta Arith. 27 (1975), 199–245.

[19] V. Vapnik and A. Chervonenkis, "On the uniform convergence of relative frequencies of events to their probabilities." Theory of Probability and its Applications, vol. XVI, no. 2 (1971) 264-280.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF CHICAGO, 5734 S. UNIVERSITY AVENUE, CHICAGO, IL 60637, USA

E-mail address: `mem@math.uchicago.edu`

EINSTEIN INSTITUTE OF MATHEMATICS, EDMOND J. SAFRA CAMPUS, GIVAT RAM, THE HEBREW UNIVERSITY OF JERUSALEM, JERUSALEM, 91904, ISRAEL, AND DEPARTMENT OF MATHEMATICS, HILL CENTER - BUSCH CAMPUS, RUTGERS, THE STATE UNIVERSITY OF NEW JERSEY, 110 FRELINGHUYSEN ROAD, PISCATAWAY, NJ 08854-8019 USA

E-mail address: `shelah@math.huji.ac.il`

URL: `http://shelah.logic.at`