STABLE REGULARITY FOR RELATIONAL STRUCTURES

NATHANAEL ACKERMAN, CAMERON FREER, AND REHANA PATEL

Abstract. We generalize the stable graph regularity lemma of Malliaris and Shelah to the case of finite structures in finite relational languages, e.g., finite hypergraphs. We show that under the model-theoretic assumption of stability, such a structure has an equitable regularity partition of size polynomial in the reciprocal of the desired accuracy, and such that for each k-ary relation and k-tuple of parts of the partition, the density is close to either 0 or 1. In addition, we provide regularity results for finite and Borel structures that satisfy a weaker notion that we call *almost stability*.

1. Introduction

Szemerédi's regularity lemma for graphs is a fundamental tool in combinatorics. It can be viewed as saying that every finite graph can be approximated by one that has a small "structural skeleton" overlaid with randomness. Malliaris and Shelah [MS14] show that one can obtain more control over this approximation in the presence of a model-theoretic tameness condition known as *stability*, that is essentially combinatorial in nature. In this paper, we extend the result of Malliaris and Shelah to the case of arbitrary finite structures in a finite relational

December 26, 2017

language. In particular, our result yields better bounds on hypergraph regularity approximations in the presence of stability.

The Szemerédi regularity lemma can be expressed more formally as saying that for any finite graph there is a partition of the vertices, known as a *regularity partition*, such that the partition is *equitable* (i.e., the sizes of the parts differ by at most 1), and for all but a few pairs of (not necessarily distinct) parts of the partition, the induced subgraph on the vertices among that pair is close to a random bipartite graph (or random graph, if the parts are not distinct) having some edge density between 0 and 1. The pairs for which this does not hold are called *irregular*. The accuracy of the approximation yielded by a regularity partition is measured both in terms of having few irregular pairs, and by the closeness of each regular pair to a random (bipartite) graph. The regularity lemma provides an upper bound on the size of a regularity partition that depends only on the desired accuracy of the approximation, and not on the particular graph being approximated. For details, see, e.g., [RS10].

While this bound on the size of the regularity partition depends only on the desired accuracy, in general one cannot guarantee a bound better than a tower of exponentials (of height that is polynomial in the reciprocal of the accuracy) [Gow97]. Further, it has long been known that if a graph contains a large *half-graph* as an induced subgraph, then any regularity partition for the graph must have irregular pairs (independently observed by Lovász, Seymour, and Trotter and by Alon, Duke, Leffman, Rödl, and Yuster [ADL+94]).

Malliaris and Shelah [MS14] observed that the presence of a large induced half-graph corresponds to the absence of *stability*, a key property from model theory that provides a sense in which a combinatorial object is highly structured, or tame (for details, see [She90]). Malliaris and Shelah [MS14] show that when a graph is stable, it admits a regularity partition with no irregular pairs, with a number of parts that is merely polynomial in the reciprocal of the accuracy, and where for each pair of (not necessarily distinct) parts, the induced bipartite graph across the parts (or induced graph on the one part) is either complete or empty. In other words, this polynomial-size partition of the vertices is such that for every pair (V_1, V_2) of parts of the partition (possibly with $V_1 = V_2$), the induced subgraph on $V_1 \cup V_2$ can be modified by a small number of edges so that either between every pair of distinct elements, one from V_1 and the other from V_2, there is an edge, or between every pair of distinct elements, one from V_1 and the other from V_2, there is no edge. In this case, the graph is close in edit distance to an *equitable blow-up* of a small finite graph (possibly with self-loops).

The regularity lemma for graphs has been generalized to finite structures in a finite relational language (see, e.g., [AC14]), a key case of which are the k-uniform hypergraphs (see, e.g., [Tao06], [Gow07], [RS07], and [ES12]). The upper bounds on the partition size are even worse than for graphs, as Moshkovitz and Shapira have recently shown that the bounds are necessarily of Ackermann-type.

The model-theoretic notion of stability also makes sense in the context of finite relational languages. In this paper, we extend Malliaris and Shelah's results to show that every finite stable structure in a finite relational language admits an equitable partition with polynomially many parts such that for every relation R (of arity k, say) and every k-tuple $(V_1, \ldots, V_k)$ of parts (possibly with repetition), the induced substructure restricted to R on $V_1 \cup \cdots \cup V_k$ can be modified by a small number of "R-edges" so that either every k-tuple of elements in $V_1 \times \cdots \times V_k$ forms an R-edge, or every k-tuple of elements in $V_1 \times \cdots \times V_k$ does not form an R-edge. In particular, the relational structure is close in edit distance to an equitable blow-up of a small structure in the same language. This shows that in the stable case, not only is "randomness" in the R-edges eliminated in the approximation, but so are the "intermediate levels" that are a key complication of the general case of hypergraph regularity lemmas. Our proof closely follows the methods of [MS14].

In the case of finite relational structures that are *almost stable* (in a sense that we make precise), we again show that the structure is close in edit distance to an equitable blow-up of a small finite structure, albeit where the few edits may not be distributed as uniformly as we can require in the stable case. Finally, we provide a similar regularity lemma for almost stable relational structures that are Borel.

1.1. **Related work.** Expanding on Malliaris and Shelah's stable regularity lemma for graphs, Malliaris and Pillay [MP16] give a short proof of the stable regularity lemma for arbitrary Keisler measures. In this more general setting, they obtain most of the nice properties from the stable regularity lemma on graphs [MS14], but they do not get precise bounds on the size of the partition.

Independently from our work in the present paper, Chernikov and Starchenko [CS16] prove a stable regularity lemma for Keisler measures over finite and Borel structures in a language with a single relation. In the case of finite structures, their stable regularity lemma is closely related to our main result, Theorem 4.8, restricted to languages with a single relation. However, while the partitions they obtain are definable (unlike ours), they need not be equitable.

Chernikov and Starchenko also obtain two regularity lemmas for structures satisfying certain model-theoretic conditions other than stability, one for NIP structures that generalizes a result of Lovász and Szegedy [LS10], and one for distal structures, generalizing their earlier result [CS15].

Generalizing Green's group-theoretic regularity lemma [Gre05], Terry and Wolf obtain a stable version for vector spaces over finite fields [TW17], and Conant, Pillay, and Terry obtain a further generalization to arbitrary finite groups [CPT17].

1.2. **Road map of the proof of the main result.** Before beginning our technical construction, we here provide a road map of the proof of the main result, Theorem 4.8. We will first describe how to "augment" relations and give a quick proof outline in terms of such augmented relations. Then we will provide more detail on three key aspects: obtaining ε-excellent sets, making a partition equitable, and modifying the original structure so that it is a blow-up.

Let $\mathcal{L}$ be a finite relational language, and let $\widehat{\tau} \in \mathbb{N}$. Suppose that $\mathcal{G}$ is a finite $\mathcal{L}$-structure such that none of its relations has the so-called $\widehat{\tau}$-branching property. (In fact, a slightly weaker hypothesis will suffice.) In particular, $\mathcal{G}$ is stable.

We begin by augmenting every relation in $\mathcal{G}$. Each relation in $\mathcal{G}$ can be thought of as a $\{\top, \bot\}$-valued function of some arity. We replace each relation with a continuum-sized family of functions (indexed by $\varepsilon > 0$) each of which takes values in $\{\top, \bot, \uparrow\}$, and further allow each argument to be either an element or a subset of $\mathcal{G}$. In the case where exactly one argument is a subset of $\mathcal{G}$, this will be done by "polling" the elements in a subset and assigning a truth value ($\top$ or $\bot$) if and only if a sufficiently large majority (namely, a $(1-\varepsilon)$-fraction) of the elements agree on that truth value (when all other arguments are fixed), and $\uparrow$ otherwise. However, when more than one argument is a subset, the polling is more complicated. For a given order of arguments, we will define this notion of polling by induction on the number of arguments that are sets, in a way that depends on the order of arguments polled so far.

These augmented relations will be used to construct collections of so-called *ε-excellent sets,* that in particular are such that whenever all arguments of an augmented relation are ε-excellent then the (function indexed by ε of the) augmented relation has a truth value (i.e., is assigned $\top$ or $\bot$).

The proof outline is as follows. Assume that $\mathcal{G}$ is large enough (relative to $\widehat{\tau}$). We first find, using the augmented relations, an ε-excellent partition of a large subset of G. We then transform this into an equitable partition of G into $(\varepsilon + \zeta)$-excellent sets (where ζ depends only on ε). Finally, we show that it is possible to change some ε-fraction of the (original) relations so that an equitable partition now describes this modification of $\mathcal{G}$ as exactly the "blow-up" of a small finite structure, whose size (i.e., the number of parts of an equitable partition) is at most polynomial in ε, where the polynomial's exponent depends only on $\widehat{\tau}$ and the maximum arity of $\mathcal{L}$.

1.2.1. *ε-excellent sets.* Suppose $A \subseteq \mathcal{G}$. We now describe how to find an ε-excellent subset of A that is *big* in the sense that its size is among a particular collection of natural numbers determined by ε. We show that a witness to the non-ε-excellence of A can be taken to consist of a relation R, an order of its arguments, an index j among the arity(R)-many arguments, an (arity$(R) - 1$)-tuple of sets $\langle B_i \rangle_{i \neq j}$ (satisfying a certain additional property with respect to the order) and two big disjoint subsets A_0 and A_1, such that the truth value assigned by the augmentation of R (with polling based on the given ordering) to $\langle B_i \rangle_{i \neq j}$ along with A_0 in the jth coordinate is different from the truth value that it assigns to $\langle B_i \rangle_{i \neq j}$ along with A_1 in the jth coordinate. Having found such a witness to the non-ε-excellence of A, we then look for such witnesses to the non-ε-excellence of A_0 and of A_1. We repeat this process on big disjoint subsets of A_0 and of A_1, etc., and stop as soon as some branch can go no farther (because we have reached some big subset of A that itself has no such witness), after which the resulting binary tree of subsets of A is perfect. A *mesa* is an object of the following sort that arises from a perfect tree of such witnesses: a finite perfect binary tree, each node of which is labeled by a triple consisting of a relation symbol, an index for one of the arguments of the relation, an ordering for the arguments of the relation, and certain witnessing subsets. At least one node of a maximal mesa does not itself have witnesses; we call such a node a *cap*, and it turns out that the height of any maximal mesa can be bounded above in terms of $\widehat{\tau}$. The intuitive idea is that a mesa is not too "tall", by virtue of not being too "wide"; there can be many caps on it — by virtue of any of which it doesn't get too "tall".

Mesas have three important properties. First, as already mentioned, each chosen subset of A occurring in its tree is big (i.e., its size is in the special set of sizes). Second, also as already noted, if the mesa is maximal, then there must be at least one cap, whose corresponding subset must therefore be ε-excellent. Third, from any mesa such that every node has the same labels for the relation, argument index, and argument order, we can extract a witness to the branching property of $\mathcal{G}$ of the same height as the mesa.

Next, by a Ramsey-theoretic result, there is a function $f\colon \mathbb{N} \to \mathbb{N}$ such that $f = O(n \log n)$ with the following property: whenever $k \in \mathbb{N}$ and T is a perfect binary tree with height $f(k)$, each node of which is labeled by a triple consisting of a relation symbol, an index for one of the arguments of the relation, and an ordering for the arguments of the relation, there is a perfect subtree of T of height k such that every node of the subtree has the same label. In particular, this holds of a mesa. Hence from a bound on the branching property for $\mathcal{G}$ we may obtain a bound on the height of any mesa arising from $\mathcal{G}$.

Because we have bounds on how much the sets decrease in size as one proceeds down a mesa, the bound on the height of the mesa induces a bound on the size of the excellent sets. In aggregate, using the fact that no relation has the $\widehat{\tau}$-branching property, we can find a constant c_ε such that any set A has an ε-excellent set of size at least $c_\varepsilon \cdot |A|$.

1.2.2. *Equitable partitions.* We now describe in more detail how we find an equitable partition of "most" of $\mathcal{G}$ consisting of $(\varepsilon + \zeta)$-excellent sets. Using the method for extracting excellent subsets that have size at least a positive fraction, we repeat this procedure to get a partition of "most" of the structure where every element of the partition is excellent and the size of the partition is bounded in terms of ε. We then aim to modify this partition to an equitable one while only increasing the error slightly. The allowable sizes for a "big" set in fact were chosen so that their greatest common denominator is also in the set. Consider a random, equitable, refinement of the original partition where the size of each element is this greatest common divisor. Using the fact that all relations of $\mathcal{G}$ are appropriately stable, the limiting properties of certain hypergeometric distributions imply that with high probability a random such partition is $(\varepsilon+\zeta)$-excellent provided that the structure underlying the partition is “large”. In particular, this implies that there is some such equitable $(\varepsilon+\zeta)$-refinement.

1.2.3. *Modifying the original structure.* We now describe how to change the truth values of each relation on just an $(\varepsilon\cdot r)$-fraction of the elements (where r is the arity of the relation), so that the resulting structure is the blow-up of a finite structure of size bounded by a polynomial in ε^{-1}. This modification of the structure has two parts. First, we show that for any ε-excellent partition of “most” of $\mathcal{G}$, the relations may be modified on a small portion of the elements so as to obtain a partition of the same set which is “indiscernible” (i.e., a blow-up of a finite structure). Next we have to deal with the (small number of) elements of $\mathcal{G}$ not in any part of the original partition. We show that if we add such elements to parts of the partition arbitrarily (while keeping the partition equitable), we may then modify relations on these elements (with respect to the other elements) so that in the modified structure the relations agree with the other elements within the part to which they were assigned. In aggregate these actions only require us to change the relations on a small fraction of the elements, yielding a structure that is exactly a blow-up while being close to the original.

1.3. **Notation.** We now introduce some notation and conventions that we will use throughout the paper.

All logarithms are in base 2, and we will simply write log.

In this paper $\mathcal{L}$ will denote a fixed finite relational language. All $\mathcal{L}$-formulas will be first order. Equality will be considered a logical symbol and not a member of $\mathcal{L}$.

For any relation $E\in\mathcal{L}$, we let arity(E) denote the arity of E. We will also need two quantities related to the arities of relations in $\mathcal{L}$. We let

$$q_{\mathcal{L}} := \sup\{\mathrm{arity}(E) : E\in\mathcal{L}\}$$

and $n_{\mathcal{L}} := |\mathcal{L}|\cdot q_{\mathcal{L}}$.

We consider an n-element sequence $\overline{a}$ of elements of A to be a map of the form $\overline{a}\colon\{0,\dots n-1\}\to A$, and therefore $\emptyset$ is the empty sequence, and range$(\overline{a})$ is the set of elements occurring in the sequence $\overline{a}$. We also identify a finite sequence $\overline{a}$ with the tuple of its elements $\langle\overline{a}(0),\dots,\overline{a}(\mathrm{len}(\overline{a})-1)\rangle$.

For finite sequences $\langle a_i\rangle_{i\le n}$ and $\langle b_j\rangle_{j\le m}$, we say that $\langle a_i\rangle_{i\le n}$ is an *initial segment* of $\langle b_j\rangle_{j\le m}$, written

$$\langle a_i\rangle_{i\le n}\preceq\langle a_j\rangle_{j\le m},$$

when $n\le m$ and when $a_i=b_i$ for all $i\le n$. Given a tuple $\overline{a}=\langle a_1,\dots,a_k\rangle$ and an element b, we write $\overline{a}^\wedge b$ to denote the tuple $\langle a_1,\dots,a_k,b\rangle$.

We now introduce two special kinds of partitions. An *equitable* partition is one whose pieces differ in size by at most 1, and an *indivisible* partition is one for which whether or not a relation holds of a tuple depends only on the parts of the partition these elements are in.

Definition 1.1. *Suppose $\mathcal{M}$ is an $\mathcal{L}$-structure with underlying set M. We say that P is a partition of $\mathcal{M}$ if it is a partition of M. We say that P is* **equitable** *if for any $p_o, p_1 \in P$,*

$$\big|\,|p_0| - |p_1|\,\big| \leq 1$$

Definition 1.2. *We say that a partition P of an $\mathcal{L}$-structure $\mathcal{M}$ is* **indivisible** *if for each relation $E \in \mathcal{L}$, for all $p_0, \ldots, p_{\text{arity}(E)-1} \in P$, and for any pair of sequences $\langle a_i^0 \rangle_{i<\text{arity}(E)}, \langle a_i^1 \rangle_{i<\text{arity}(E)}$ such that $a_i^0, a_i^1 \in p_i$, where $0 \leq i < \text{arity}(E)$, we have*

$$\mathcal{M} \models E(a_0^0, \ldots, a_{\text{arity}(E)-1}^0) \leftrightarrow E(a_0^1, \ldots, a_{\text{arity}(E)-1}^1).$$

In other words, a partition is indivisible if we can quotient out by the equivalence relation that it induces and then assign a compatible $\mathcal{L}$-structure relation.

Definition 1.3. *Suppose $\mathcal{M}$ and $\mathcal{N}$ are $\mathcal{L}$-structures with underlying sets M and N respectively. A map $\alpha\colon M \to N$ is a* **full homomorphism** *from $\mathcal{M}$ to $\mathcal{N}$ if for each relation $E \in \mathcal{L}$ and all tuples $a_0, \ldots, a_{\text{arity}(E)-1} \in M$ of (distinct) elements of M,*

$$\mathcal{M} \models E(a_0, \ldots, a_{\text{arity}(E)-1}) \text{ if and only if } \mathcal{N} \models E(\alpha(a_0), \ldots, \alpha(a_{\text{arity}(E)-1})).$$

Note that that full homomorphisms are not necessarily injective.

Definition 1.4. *We say that an $\mathcal{L}$-structure $\mathcal{M}$ is a* **blow-up** *of a $\mathcal{L}$-structure $\mathcal{N}$ when there is a surjective full homomorphism $i\colon \mathcal{M} \to \mathcal{N}$. We call i the* **witness** *to the blow-up.*

If also the sets $i^{-1}(\{b_0\})$ and $i^{-1}(\{b_1\})$ differ in size by at most one, for all $b_0, b_1 \in \mathcal{N}$, then we say that $\mathcal{M}$ is an **equitable blow-up** *of $\mathcal{N}$.*

Our regularity lemmas can be seen as stating that certain types of structures are close in edit distance to a blow-up of a small finite structure.

The following easy lemma, whose proof we omit, makes precise the notion that an $\mathcal{L}$-structure with an indivisible partition can be thought of as blow-up of a smaller $\mathcal{L}$-structure.

Lemma 1.5. *For an $\mathcal{L}$-structure $\mathcal{M}$ and a partition P of $\mathcal{M}$ the following are equivalent.*

- *P is indivisible.*
- *There exists an $\mathcal{L}$-structure $\mathcal{N}$ such that $\mathcal{M}$ is a blow-up of $\mathcal{N}$ with witness i such that*

$$P = \{i^{-1}(\{b\}) \,:\, b \in N\}.$$

Furthermore, $\mathcal{M}$ is an equitable blow-up of $\mathcal{N}$ if and only if P is equitable.

Intuitively, $\mathcal{M}$ is a blow-up of $\mathcal{N}$ if it can be obtained by replacing each element of $\mathcal{N}$ with an indiscernible set, while $\mathcal{M}$ is an equitable blow-up of $\mathcal{N}$ if these indiscernible sets are all almost the same size.

1.4. **Stability.** We now recall some basic facts from stability theory. The notation in this section follows that of [MS14].

Definition 1.6. *Let* $\overline{\tau} \in \mathbb{N}$. *An* $\mathcal{L}$*-formula* $\varphi(\overline{x};\overline{y})$ *has the* $\overline{\tau}$**-order property** *in an* $\mathcal{L}$*-structure* $\mathcal{M}$ *when there exist sequences* $\langle \overline{a}_i \rangle_{i<\overline{\tau}} \subseteq \mathcal{M}$ *(with* $\mathrm{len}(\overline{a}_i) = \mathrm{len}(\overline{x}$ *for all* $i < \overline{\tau}$*) and* $\langle \overline{b}_j \rangle_{j<\overline{\tau}} \subseteq \mathcal{M}$ *(with* $\mathrm{len}(\overline{b}_j) = \mathrm{len}(\overline{y})$ *for all* $j < \overline{\tau}$*) such that for all* $i, j < \overline{\tau}$,

$$\mathcal{M} \models \varphi(\overline{a}_i; \overline{b}_j) \Leftrightarrow i < j.$$

We say that $\varphi(\overline{x};\overline{y})$ *has the* **non-**$\overline{\tau}$**-order property** *in* $\mathcal{M}$ *when it does not have the* $\overline{\tau}$*-order property in* $\mathcal{M}$.

Note that the $\overline{\tau}$-order property is defined for a formula along with a given partition of its free variables, not just for the formula alone.

We will in fact work with a combinatorial property that holds in a structure essentially whenever the $\overline{\tau}$-order property does.

Definition 1.7. *Let* $\widehat{\tau} \in \mathbb{N}$. *An* $\mathcal{L}$*-formula* $\varphi(\overline{x};\overline{y})$ *has the* $\widehat{\tau}$**-branching property** *in an* $\mathcal{L}$*-structure* $\mathcal{M}$ *when there exist sequences* $\langle \overline{a}_i \rangle_{i\in\{0,1\}^{\widehat{\tau}}} \subseteq \mathcal{M}$ *(with* $\mathrm{len}(\overline{a}_i) = \mathrm{len}(\overline{x})$ *for all* $i \in \{0,1\}^{\widehat{\tau}}$*) and* $\langle \overline{b}_j \rangle_{j\in\{0,1\}^{<\widehat{\tau}}} \subseteq \mathcal{M}$ *(with* $\mathrm{len}(\overline{b}_j) = \mathrm{len}(\overline{y})$ *for* $j \in \{0,1\}^{<\widehat{\tau}}$*) such that for all* $i \in \{0,1\}^{\widehat{\tau}}$, *for all* $j \in \{0,1\}^{<\widehat{\tau}}$, *and for each* $h \in \{0,1\}$, *we have that*

$$j^\wedge h \preceq i$$

implies

$$\mathcal{M} \models \varphi(\overline{a}_i; \overline{b}_j) \Leftrightarrow (h = 1).$$

We say that $\varphi(\overline{x};\overline{y})$ *has the* **non-**$\widehat{\tau}$**-branching property** *in* $\mathcal{M}$ *when it does not have the* $\widehat{\tau}$*-branching property in* $\mathcal{M}$.

We now note a connection between the non-$\overline{\tau}$-order property and the non-$\widehat{\tau}$-branching property for a structure $\mathcal{M}$.

Lemma 1.8. *If* $\varphi(\overline{x};\overline{y})$ *has the non-*$\overline{\tau}$*-order property in* $\mathcal{M}$ *then* $\varphi(\overline{x};\overline{y})$ *has the non-*$2^{\widehat{\tau}}$*-branching property in* $\mathcal{M}$, *where* $\widehat{\tau} = 2^{\overline{\tau}+2} - 2$. *On the other hand, if* $\varphi(\overline{x};\overline{y})$ *has the non-*$\widehat{\tau}$*-branching property in* $\mathcal{M}$ *then* $\varphi(\overline{x};\overline{y})$ *has the non-*$2^{\overline{\tau}}$*-order property in* $\mathcal{M}$, *where* $\overline{\tau} = 2^{\widehat{\tau}+1}$.

Proof. See [Hod93, Lemma 6.7.9]. □

We will be interested in the situation when, for each relation $E \in \mathcal{L}$, the $\mathcal{L}$-structure $\mathcal{M}$ has on the non-$\overline{\tau}_E$-property for some $\overline{\tau}_E \in \mathbb{N}$. This is equivalent to the following.

Definition 1.9. *Let* $\mathcal{M}$ *be an* $\mathcal{L}$*-structure. We say that* $\mathcal{M}$ *has the* **non-**$\widehat{\tau}$**-branching property** *(***non-**$\overline{\tau}$**-order property***) if for each relation* $E \in \mathcal{L}$ *and each* $0 \leq j < \mathrm{arity}(E) - 1$, *the formula* $E(x_0, \ldots, x_{\mathrm{arity}(E)-1})$ *with the partition of variables* $(x_j; x_0, \ldots, x_{j-1}, x_{j+1}, \ldots, x_{\mathrm{arity}(E)-1})$ *has the non-*$\widehat{\tau}$*-branching property (non-*$\overline{\tau}$*-order property) in* $\mathcal{M}$.

For the rest of this paper, fix $\widehat{\tau} \in \mathbb{N}$.

2. Excellence

From now on we assume $\mathcal{M}$ is a finite $\mathcal{L}$-structure with underlying set M. We will prove our regularity lemma by showing that, under appropriate stability assumptions, we can find a partition of any subset A of M that is *almost* a blow-up. To do this, we will need a notion called ε-*excellence*, which captures this idea of being almost a blow-up.

We begin by allowing the domain of relations to consist of both elements and sets. As such, we let $\widehat{M} := M \cup \wp(M)$ where $\wp(M)$ denotes the power set of M. We now define how to augment a relation on M to all of $\widehat{M}$ (according to a given ε tolerance).

Definition 2.1. *Let $0 < \varepsilon < \frac{1}{2}$, let $E \in \mathcal{L}$ be an n-ary relation, and let $\overline{m}$ be a sequence of distinct elements of $\{0, \dots, n-1\}$. We now define, inductively on the length of $\overline{m}$, the collection of ε-**partial relations for** E. Such a partial relation is a function $\widehat{E}^{\overline{m}}_{\varepsilon} \colon \widehat{M}^n \to \{\top, \bot, \uparrow\}$.*

Let $A_0, \dots, A_{n-1} \in \widehat{M}$ and let $S := \{i < n : A_i \in \widehat{M} \setminus M\}$. If $S \neq \mathrm{range}(\overline{m})$, then define

$$\widehat{E}^{\overline{m}}_{\varepsilon}(A_0, \dots, A_{n-1}) := \uparrow .$$

Otherwise, when $S = \mathrm{range}(\overline{m})$, we will define $\widehat{E}^{\overline{m}}_{\varepsilon}(A_0, \dots, A_{n-1})$ by induction on $\ell := \mathrm{len}(\overline{m})$, as follows.

Case $\ell = 0$: In this case, $\overline{m} = \emptyset$, and so $S = \emptyset$. In particular, $A_0, \dots, A_{n-1}$ are elements of M. Define

- $\widehat{E}^{\emptyset}_{\varepsilon}(A_0, \dots, A_{n-1}) := \top$ *if* $M \models E(A_0, \dots, A_{n-1})$, *and*
- $\widehat{E}^{\emptyset}_{\varepsilon}(A_0, \dots, A_{n-1}) := \bot$ *if* $M \models \neg E(A_0, \dots, A_{n-1})$.

Case $\ell \geq 1$:
Let $\overline{k}$ be the initial subtuple of $\overline{m}$ of length $\ell - 1$, and let $j := \overline{m}(\ell - 1)$ be the last element of $\overline{m}$, so that $\overline{m} = \overline{k}^\wedge j$. Because $\overline{m}$ is a tuple of distinct elements, observe that $\overline{k} \colon \{0, \dots, \ell - 2\} \to S \setminus \{j\}$ is a bijection. For each $\delta \in \{\top, \bot\}$, define

$$A^{\delta}_{\overline{k},j} := \{a \in A_j : \widehat{E}^{\overline{k}}_{\varepsilon}(A_0, \dots, A_{j-1}, a, A_{j+1}, \dots, A_{n-1}) = \delta\}.$$

- *If $\frac{|A^{\top}_{\overline{m}}|}{|A_j|} > 1 - \varepsilon$ then define $\widehat{E}^{\overline{m}}_{\varepsilon}(A_0, \dots, A_{n-1}) := \top$.*
- *If $\frac{|A^{\bot}_{\overline{m}}|}{|A_j|} > 1 - \varepsilon$ then define $\widehat{E}^{\overline{m}}_{\varepsilon}(A_0, \dots, A_{n-1}) := \bot$.*
- *Otherwise $\widehat{E}^{\overline{m}}_{\varepsilon}(A_0, \dots, A_{n-1}) := \uparrow$.*

In this definition, unless $\widehat{E}^{\overline{m}}_{\varepsilon}$ is undefined on all arguments, S is the collection of indices which are sets. To understand this definition, it is worth walking through the cases where $|S| \leq 2$. First consider the case where $|S| = 0$. We then have $\overline{m} = \emptyset$, and all A_i are elements of M, and so we let $\widehat{E}_\varepsilon^{\overline{m}}$ agree with the relation E on $\langle A_0, \dots, A_{n-1} \rangle$.

Next consider the case where $|S| = 1$, with say $S = \{j\}$, i.e., when there is a unique element A_j of $\widehat{M} \setminus M$ among the arguments $A_0, \dots, A_{n-1}$. In this case we let $\widehat{E}_\varepsilon^{\langle j \rangle}(A_0, \dots, A_{n-1})$ be $\top$ if, when we fix $A_1, \dots, A_{j-1}, A_{j+1}, \dots, A_{n-1}$ and let the jth entry vary among the elements of A_j, at least a $(1-\varepsilon)$-fraction of the elements return a value of $\top$; and similarly for $\bot$. If this doesn't happen, i.e., if there is no near-consensus among the elements of A_j, then we return $\uparrow$ signifying that its truth value is undefined.

Finally consider the case when $|S| = 2$, with say $S = \{j, p\}$. Suppose we have defined $\widehat{E}_\varepsilon^{\overline{m}}$ whenever $|\mathrm{range}(\overline{k})| = 1$. We would like to perform a similar sort of consensus-gathering to determine the value of $\widehat{E}_\varepsilon^{\overline{k}}$. Specifically, we would like to choose an element of $\widehat{M} \setminus M$ from among the arguments, fix all the other arguments, of which now there exists only one set, and then take a consensus (which we can do as we have defined $\widehat{E}_\varepsilon^{\overline{k}}$ in this case). However, we find that we first have to make a choice as to which set, A_j or A_p, we want to vary. In particular the outcome of this process might be different if we first vary A_j and then vary A_p or if instead we first vary A_p and then vary A_j. The purpose of the parameter $\overline{k}$ is to keep track of the order in which we are varying the parameters to calculate the truth value. As we will see, we will mainly be interested in sets which have a property called ε-excellence, which implies that the same truth value is returned no matter which order we consider the arguments (i.e., where $\widehat{E}_\varepsilon^{\overline{k}}$ is independent of the order of the range of $\overline{k}$).

In order to define the notion of ε-excellence, we first need to define a notion of (ε, ℓ, E)-*goodness* for each relation $E \in \mathcal{L}$ and each $\ell \in \mathbb{N}$ such that $0 \leq \ell < \mathrm{arity}(E)$.

Definition 2.2. *Fix $\varepsilon > 0$, let $E \in \mathcal{L}$ of arity n, and let $\ell < n$. We define the notion of (ε, ℓ, E)*-**goodness** *for any $A_0 \in \widehat{M}$ by induction on ℓ as follows.*

Case $\ell = 0$:
$A_0 \in \widehat{M}$ is $(\varepsilon, 0, E)$-good if and only if $A_0 \in M$.

Case $\ell \leq 1$:
$A_0 \in \widehat{M}$ is (ε, ℓ, E)-good if and only if A_0 is (ε, k, E)-good for $1 \leq k < \ell$ and for all

- *$A_1, \dots, A_{\ell-1} \in \widehat{M} \setminus M$ such that A_i is $(\varepsilon, \ell - i, E)$-good for every $1 \leq i < \ell$,*
- *$A_\ell, \dots, A_{n-1} \in M$, and*
- *permutations σ of $\{0, \dots, n-1\}$,*

we have

$$\widehat{E}_\varepsilon^{\langle\sigma(\ell-1),\ldots,\sigma(1),\sigma(0)\rangle}(A_{\sigma(0)}, A_{\sigma(1)}, \ldots, A_{\sigma(n-1)}) \in \{\top, \bot\}.$$

We say that A is ε-**excellent** *if A is $(\varepsilon, \mathrm{arity}(E), E)$-good for all relation symbols $E \in \mathcal{L}$.*

Note that our definition of $(\varepsilon, 1, E)$-goodness is the same, when $\mathcal{M}$ is a (symmetric) graph with edge relation E, as the notion of ε-goodness in [MS14], and the higher arity notions are motivated by what is needed to generalize their proof to arbitrary finite relational languages.

Intuitively, we think of a set A as being (ε, ℓ, E)-good provided that whenever we only fix sets for at most ℓ-many coordinates in the definition of $\widehat{E}_\varepsilon^{\overline{m}}$, i.e., $|\mathrm{range}(\overline{m})| \leq \ell$, with the first set fixed being (ε, ℓ, E)-good (and the others being sufficiently good) then $\widehat{E}_\varepsilon^{\overline{m}}$ returns a truth value.

Once again it will be instructive to walk through the cases when $\ell = 1$ or 2. First, a set A is $(\varepsilon, 1, E)$-good if $\widehat{E}_\varepsilon^{\overline{m}}$ is defined on any collection of arguments where A is the only argument that is not an element of M.

The case of $(\varepsilon, 2, E)$-good is somewhat more complicated. Namely, A_0 is $(\varepsilon, 2, E)$-good if whenever A_1 is $(\varepsilon, 1, E)$-good then any ε-partial relation which first varies A_0 and then varies A_1 will always return a truth value when the other arguments are in M. In particular, this holds no matter which place A_0 and A_1 take in the relation.

The notion of (ε, ℓ, E)-goodness generalizes this idea. A set A_0 is (ε, ℓ, E)-good if whenever we have a sequence $A_1, \ldots, A_{k-1}$, with $1 \leq k \leq \ell$, of decreasing E-goodness then any ε-partial relation which first varies A_0, then varies A_1, ..., will always return a truth value (no matter what the arguments are from M).

It is worth noting that if A is (ε, ℓ, E)-good and $1 \leq \ell^* < \ell$ then A is also (ε, ℓ^*, E)-good. So in particular, if A is ε-excellent then A is (ε, ℓ^*, E)-good for all $\ell^* \leq n$. This is important because it means that if $A_0, \ldots, A_{n-1}$ are all ε-excellent then $\widehat{E}_\varepsilon^{\overline{m}}(A_0, \ldots, A_{n-1})$ must have a truth value. We can further preserve goodness while weakening ε, leading to the following straightforward but crucial observation.

Lemma 2.3. *Let $E \in \mathcal{L}$, and suppose $1 \leq i^* \leq i$ and $0 < \varepsilon \leq \varepsilon^*$. If $A \in \widehat{M} \setminus M$ is (ε, i, E)-good, then A is (ε^*, i^*, E)-good.*

As we will see, a crucial property of ε-excellence is that given a collection of ε-excellent sets, all ε-partial relations have the same truth value on that collection.

In particular, the next proposition tells us that when we have a sequence $A_0, \ldots, A_{k-1}$ of (ε, k, E)-good sets, then whenever $A_0, \ldots, A_{k-1}$ are the only set arguments of $\widehat{E}_\varepsilon^{\overline{m}}$, it has a truth value, which is independent of the ordering of $\overline{m}$. In particular, Proposition 2.4 will imply that for any partial relation applied to ε-excellent sets, the augmented relation won't depend on the ordering in which the truth value is calculated.

Proposition 2.4. *Let $E \in \mathcal{L}$ have arity n, and let $0 < \varepsilon < \frac{1}{4}$. Suppose that $A_0, \dots, A_{k-1}$ be (ε, k, E)-good sets such that $A_k, \dots, A_{n-1} \in M$. For any two injective functions $\beta_0, \beta_1 \colon \{0, \dots, k-1\} \to \{0, \dots, k-1\}$ and any permutation σ of $\{0, \dots, n-1\}$,*

$$\widehat{E}_\varepsilon^{\sigma\circ\beta_0}(A_{\sigma(0)}, \dots, A_{\sigma(n-1)}) = \widehat{E}_\varepsilon^{\sigma\circ\beta_1}(A_{\sigma(0)}, \dots, A_{\sigma(n-1)}) \in \{\top, \bot\}.$$

Proof. Without loss of generality we may assume that $\sigma = \mathrm{id}$, as the proof of the general case is the same. Our proof proceeds by induction.

<u>*Case $k = 1$:*</u>
This case is immediate, as $\beta_0 = \beta_1$.

<u>*Case $k = i + 1$:*</u>
As every permutation of $\{0, \dots, k-1\}$ is equal to a composition of transpositions, it suffices to prove the result when β_0 is a transposition of β_1. Therefore, we may assume without loss of generality that $\beta_0 = \langle k-1, \dots, 2, 1, 0\rangle$ and $\beta_1 = \langle k-1, \dots, 2, 0, 1\rangle$.

Let $\beta^* = \langle k-1, \dots, 2\rangle$ and for distinct $i, j \in \{0, 1\}$, define the relation

$$\widehat{H}_\varepsilon^{ij}(B, B^*) = \widehat{E}_\varepsilon^{\beta^* {}^\wedge \langle i,j\rangle}(B, B^*, A_2, \dots, A_{n-1}).$$

Then our goal is to show that $\widehat{H}_\varepsilon^{01}(A_0, A_1) = \widehat{H}_\varepsilon^{10}(A_0, A_1)$.

Suppose $\widehat{H}_\varepsilon^{10}(A_0, A_1) = \top$. Then there are at most

$$(\varepsilon \cdot |A_0|) \cdot |A_1| + ((1-\varepsilon)|A_0|) \cdot (\varepsilon \cdot |A_1|)$$

many pairs $(a, b) \in A_0 \times A_1$ such that $\widehat{E}_\varepsilon^{\beta^*}(a, b, A_2, \dots, A_{n-1}) = \bot$.

Similarly, if $\widehat{H}_\varepsilon^{01}(A_0, A_1) = \bot$, then there are at most

$$(\varepsilon \cdot |A_1|) \cdot |A_0| + ((1-\varepsilon) \cdot |A_1|) \cdot (\varepsilon \cdot |A_0|)$$

many pairs $(a, b) \in A_0 \times A_1$ such that $\widehat{E}_\varepsilon^{\beta^*}(a, b, A_2, \dots, A_{n-1}) = \top$.

Hence if

$$\begin{aligned} |A_0||A_1| > {} & (\varepsilon \cdot |A_0|) \cdot |A_1| + ((1-\varepsilon)|A_0|) \cdot (\varepsilon \cdot |A_1|) + \\ & (\varepsilon \cdot |A_1|) \cdot |A_0| + ((1-\varepsilon)|A_1|) \cdot (\varepsilon \cdot |A_0|) \\ = {} & 2(2\varepsilon - \varepsilon^2)|A_0||A_1|, \end{aligned}$$

then $\widehat{H}_\varepsilon^{10}(A_0, A_1) = \top$ and $\widehat{H}_\varepsilon^{01}(A_0, A_1) = \bot$ cannot both hold simultaneously.

A similar calculation shows that if

$$|A_0||A_1| > 2(2\varepsilon - \varepsilon^2)|A_0||A_1|$$

then $\widehat{H}_\varepsilon^{01}(A_0, A_1) = \top$ and $\widehat{H}_\varepsilon^{10}(A_0, A_1) = \bot$ cannot both hold simultaneously.

Now, $\varepsilon < \frac{1}{4}$, and so $2(2\varepsilon - \varepsilon^2) < 1$. Hence $\widehat{H}_\varepsilon^{01}(A_0, A_1) = \widehat{H}_\varepsilon^{10}(A_0, A_1)$ (as they have truth values because A_0 and A_1 are (ε, k, E)-good). Therefore the result follows. $\square$

From now on we will assume that $\varepsilon < \frac{1}{4}$. Suppose $\{A_0, \dots, A_{n-1}\} \subseteq \widehat{M}$ is such that exactly k are (ε, k, E)-good and exactly $\operatorname{arity}(E) - k$ are in M. Then by Proposition 2.4, the value of $\widehat{E}^{\overline{m}}_{\varepsilon}(A_0, \dots, A_{n-1})$ is independent of $\overline{m}$. In this circumstance, we will refer to $\widehat{E}^{\overline{m}}_{\varepsilon}$ simply as $\widehat{E}_{\varepsilon}$. In particular, this holds when all arguments are ε-excellent. In summary, we have the following corollary.

Corollary 2.5. *For any ε-excellent elements $A_0, \dots, A_{n-1} \in \widehat{M}$, and any $E \in \mathcal{L}$, $\widehat{E}_{\varepsilon}(A_0, \dots, A_{n-1}) \in \{\top, \bot\}$.*

The following technical lemma tells us that, for a relation E and appropriately good sets, at most a small fraction of the tuples consistent with those sets disagree with the partial relation $\widehat{E}^{\overline{m}}_{\varepsilon}$ about the truth value of E.

Lemma 2.6. *Let $E \in \mathcal{L}$ have arity n and let σ be a permutation of $\{0, \dots, n-1\}$. Suppose $A_0, \dots, A_{k-1}$ are sets such that A_i is $(\varepsilon, k-i, E)$-good for $0 \le i < k$. Further suppose that $A_k, \dots, A_{n-1} \in M$. Define*

$$Z := \{(a_0, \dots, a_{n-1}) : a_i \in A_{\sigma^{-1}(i)} \text{ when } A_{\sigma^{-1}(i)} \in \widehat{M} \setminus M, \text{ and } a_i = A_{\sigma^{-1}(i)} \text{ when } A_{\sigma^{-1}(i)} \in M\}.$$

Then the following hold.

(a) *If $\widehat{E}^{\sigma^{-1}(\langle k-1, \dots, 0\rangle)}_{\varepsilon}(A_{\sigma(0)}, \dots, A_{\sigma(n-1)}) = \top$ then*

$$\left|\{(a_0, \dots, a_{n-1}) \in Z : M \models \neg E(a_0, \dots, a_{n-1})\}\right| \le k \cdot \varepsilon \cdot \prod_{0 \le i < k} |A_i|.$$

(b) *If $\widehat{E}^{\sigma^{-1}(\langle k-1, \dots, 0\rangle)}_{\varepsilon}(A_{\sigma(0)}, \dots, A_{\sigma(n-1)}) = \bot$ then*

$$\left|\{(a_0, \dots, a_{n-1}) \in Z : M \models E(a_0, \dots, a_{n-1})\}\right| \le k \cdot \varepsilon \cdot \prod_{0 \le i < k} |A_i|.$$

Proof. The proofs of (a) and (b) are essentially identical so we will only prove (a). Further we can assume without loss of generality that $\sigma = \mathrm{id}$. To simplify notation we will will omit the superscript of the partial relation and refer to $\widehat{E}^{\langle k-1, \dots, 0\rangle}_{\varepsilon}$ by $\widehat{E}_{\varepsilon}$.

For simplicity of notation let $F(x_0, \dots, x_{k-1}) := E(x_0, \dots, x_{k-1}, A_k, \dots, A_{n-1})$. Note that as A_i is $(\varepsilon, k-i, E)$-good, A_i is also $(\varepsilon, k-i, F)$-good.

We will proceed in stages.

Stage 0:

Let $B^{\emptyset}_0 := \{b \in A_0 : \widehat{F}_{\varepsilon}(b, A_1, \dots, A_{k-1}) = \bot\}$, and let $C^{\emptyset}_0 := \{(a_0, \dots, a_{k-1}) \in \prod_{i<k} A_i : a_0 \in B^{\emptyset}_0\}$.

Let $F^0 := F \cup C^{\emptyset}_0$. Note that whenever $a_0 \in B^{\emptyset}_0$ and $a_i \in A_i$ for $1 \le i < k$, we have $F^0(a_0, \dots, a_{k-1}) = \top$.

In particular, for every $b \in A_0$ we have $F^0(b, A_1, \dots, A_{k-1}) = \bot$.

Stage $\ell + 1$:

Suppose, for $\overline{c} \in \prod_{i<\ell} A_i$, that the sets F^{ℓ}, $B^{\overline{c}}_{\ell}$ have been defined so that

(1) whenever $a_i \in A_i$ for $0 \leq i < k$ and $a_j \in B_j^{\langle a_p \rangle_{p<j}}$ for some $j \leq \ell$ then $F^\ell(a_0, \ldots, a_{k-1}) = \top$, and
(2) whenever $a_i \in A_i$ for $0 \leq i \leq \ell$ and $a_j \notin B_j^{\langle a_p \rangle_{p<j}}$ for all $j \leq \ell$ then $\widehat{F}_\varepsilon(a_0, \ldots, a_\ell, A_{\ell+1}, \ldots, A_{k-1}) = \top$.

We now show how to appropriately define sets for parameters of length $\ell + 1$.

Fix $\overline{a} = \langle a_0, \ldots, a_\ell \rangle$ so that (2) holds. Let

$$B_{\ell+1}^{\overline{a}} := \big\{ b \in A_{\ell+1} \,:\, \widehat{F}_\varepsilon(a_0, \ldots, a_\ell, b, A_{\ell+2}, \ldots, A_{k-1}) = \bot \big\},$$

and let

$$C_{\ell+1}^{\overline{a}} := \big\{ \overline{a}^\wedge \langle a_{\ell+1}, \ldots, a_{k-1}) \rangle \,:\, a_{\ell+1} \in B_{\ell+1}^{\overline{a}} \text{ and } a_{\ell+2} \in A_{\ell+2}, \ldots, a_{k-1} \in A_{k-1} \big\}$$

If $\overline{a} = \langle a_0, \ldots, a_\ell \rangle$ are such that (1) holds then let $C_{\ell+1}^{\langle a_p \rangle_{p \leq \ell}} = \emptyset$.

Let $F^{\ell+1} := F_\ell \cup \bigcup \{ C_{\ell+1}^{\overline{a}} \,:\, (2) \text{ holds} \}$. In particular the inductive hypothesis holds.

When we reach the kth inductive stage we therefore have $F^{k-1}(a_0, \ldots, a_{k-1}) = \top$ whenever $(a_0, \ldots, a_{k-1}) \in \prod_{i<k} A_i$.

$$|\{ (a_0, \ldots, a_{n-1}) \in Z \,:\, M \models \neg E(a_0, \ldots, a_{n-1}) \}| \leq |F^{k-1} \setminus F|.$$

But $F^{k-1} \setminus F \subseteq \bigcup_{i<k} \bigcup \{ C_i^{\overline{a}} \,:\, \overline{a} \in \prod_{p<i} A_p \}$.

But as each A_i is (ε, k, E)-good, for each $\overline{a} \in \prod_{p<i} A_p$ we have

$$|C_i^{\overline{a}}| \leq \varepsilon \cdot \prod_{i \leq p < k} |A_p|,$$

and so

$$|\bigcup \{ C_i^{\overline{a}} \,:\, \overline{a} \in \prod_{p<i} A_p \}| \leq \varepsilon \cdot \prod_{p<k} |A_p|.$$

But then $|F^{k-1} \setminus F| \leq k \cdot \varepsilon \cdot \prod_{p<k} |A_p|$, as desired. □

In particular, if we have a partition of our graph into ε-excellent pieces, then we can define a *consensus* truth value to any relation and tuple of pieces of the partition that makes the partition almost indivisible in the following sense.

Proposition 2.7. *Suppose P is an equitable partition of M such that each element of P is ε-excellent. Then for each relation $E \in \mathcal{L}$ there is a relation E^* on M, with the same arity as E, such that*

- $|(E \Delta E^*) \cap \prod_{i<\text{arity}(E)} p_i| \leq \text{arity}(E) \cdot \varepsilon \cdot \prod_{i<n} |p_i|$ *for all tuples $\langle p_i \rangle_{i<n}$ from P, and*
- *P is an indivisible partition of (M, E^*).*

Proof. Fix $E \in \mathcal{L}$ and suppose arity$(E) = n$. Let $\widehat{p} = (p_0, \ldots, p_{n-1}) \subseteq P$ and let $E_{\widehat{p}} = E \cap \prod_{i<n} p_i$. Without loss of generality let $\widehat{E}_\varepsilon(p_0, \ldots, p_{n-1}) = \top$. By

Lemma 2.6 we know that

$$\left|\prod_{i<n} p_i \setminus E_{\widehat{p}}\right| \leq n \cdot \varepsilon \cdot \prod_{i<n} |p_i|.$$

Now let E^* be such that for any $(p_0, \ldots, p_{n-1}) \in P$, if $\widehat{E}_\varepsilon(p_0, \ldots, p_{n-1}) = \top$ then $E^* \cap \prod_{i<n} p_i = \prod_{i<n} p_i$, and if $\widehat{E}_\varepsilon(p_0, \ldots, p_{n-1}) = \bot$ then $E^* \cap \prod_{i<n} p_i = \emptyset$. It is then clear that (M, E^*) is indivisible. □

3. Obtaining excellent sets

In this section, we will show how to use the fact that $\mathcal{G}$ has the non-$\widehat{\tau}$-branching property to get large excellent sets. Specifically, we start with a set A and try and build a binary-branching tree of subsets of A, where the set at a child node has size at least ε times the size of the set at the parent node, and where the sets at any two children disagree on some "question" that excellent sets "decide". If this process of building a tree terminates, then there must be some set which we could not divide into two pieces each of size an ε fraction of the set, each of which gives a different answer to a question that ε-sets can answer. Hence we will deduce that such a set must itself be ε-excellent. We will then show that such a tree must have a height bounded by a term definable from $\widehat{\tau}$, which will give us a bound on how large (as a fraction of our original set) an ε-excellent set we can find.

In addition, when such a tree branches we will further require the subsets at the children nodes to be not merely "sufficiently large", but also one of a given predetermined set of sizes. In this way we will ensure that the sizes of all ε-excellent sets we create have a large greatest common divisor. This will be important when, in Section 4, we wish to divide our partition of ε-excellent sets into an equitable partition of ε-excellent sets.

Definition 3.1. *A* **rock** *is a tuple* $\langle A, Q, \ell, (B^0, \ldots, B^{\ell-1}, B^{\ell+1}, \ldots, B^{\text{arity}(Q)-1}), \beta\rangle$, *where*

- $A \in \wp(G) \setminus \emptyset$,
- *Q is a relation symbol in $\mathcal{L}$,*
- $\ell \in \mathbb{N}$ *such that* $\ell < \text{arity}(Q)$,
- *each* $B^t \in \wp(G) \setminus \emptyset$, *and*
- $\beta\colon \{1, \ldots, \text{arity}(Q)-1\} \to \{0, \ldots, \text{arity}(Q)-1\} \setminus \{\ell_i\}$ *is an injection (and hence a bijection).*

We say that such a rock **covers** *the set A.*

Definition 3.2. *Let $k \in \mathbb{N}$ and $\varepsilon > 0$. A finite sequence $\langle m_j\rangle_{j\leq k}$ of positive integers is a* **staircase** *if $\frac{m_{j+1}}{m_j} \leq \varepsilon$ for all $j < k$.*

Definition 3.3. *Let $k \in \mathbb{N}$ and $\varepsilon > 0$, and suppose $\boldsymbol{m} := \langle m_j \rangle_{j \le k}$ is a staircase. Define an* $(\varepsilon, \boldsymbol{m})$**-mesa** *of height k to consist of a tree of rocks*

$$\left\langle \left(A_i, Q_i, \ell_i, (B_i^0, \dots, B_i^{\ell-1}, B_i^{\ell+1}, \dots, B_i^{\operatorname{arity}(Q)-1}), \beta_i\right) \right\rangle_{i \in \{0,1\}^{<k}}$$

along with a collection of sets (called **pre-caps***) $\langle A_i \rangle_{i \in \{0,1\}^k}$ indexed by the children of the leaves, that satisfy, for each $i \in \{0,1\}^{<k}$,*

- *B_i^j is $(\varepsilon, \beta_i^{-1}(j), Q_i)$-good for each $j \in \{0, \dots, \operatorname{arity}(Q_i) - 1\} \setminus \{\ell_i\}$.*
- *$|A_{i^\wedge s}| \in \boldsymbol{m}$ and $|A_{i^\wedge s}| \ge |\varepsilon| \cdot |A_i|$ for each $s \in \{0, 1\}$.*
- *$\left(\widehat{Q_i}\right)_\varepsilon^{\beta_i(\langle 1, \dots, \operatorname{arity}(Q_i)-1 \rangle)}(B_i^0, \dots, B_i^{\ell_i - 1}, a, B_i^{\ell_i+1}, \dots, B_i^{\operatorname{arity}(Q_i)-1}) = \bot$ for all $a \in A_{i^\wedge 0}$.*
- *$\left(\widehat{Q_i}\right)_\varepsilon^{\beta_i(\langle 1, \dots, \operatorname{arity}(Q_i)-1 \rangle)}(B_i^0, \dots, B_i^{\ell_i - 1}, a, B_i^{\ell_i+1}, \dots, B_i^{\operatorname{arity}(Q_i)-1}) = \top$ for all $a \in A_{i^\wedge 1}$.*

The set $A_{\langle\rangle}$ is called the **head***.*

Consider an $(\varepsilon, \boldsymbol{m})$-mesa as above, suppose m_{k+1} is such that $\frac{m_{k+1}}{m_k} \le \varepsilon$, and let A_p be a pre-cap such that $\varepsilon \cdot |A_p| \le m_{k+1}$. Then A_p is an m_{k+1}**-cap** *if there is no rock $\langle A_p, Q, \ell, (B^0, \dots, B^{\ell-1}, B^{\ell+1}, \dots, B^{\operatorname{arity}(Q)-1}), \beta \rangle$ covering it such that*

$$m_{k+1} \le \{a \in A_p \,:\, \widehat{Q}_\varepsilon^{\beta(\langle 1 \dots \operatorname{arity}(Q)-1 \rangle)}(B^0, \dots, B^{\ell-1}, a, B^{\ell+1}, \dots B^{\operatorname{arity}(Q)-1}) = \bot\}$$

and

$$m_{k+1} \le \{a \in A_p \,:\, \widehat{Q}_\varepsilon^{\beta(\langle 1 \dots \operatorname{arity}(Q)-1 \rangle)}(B^0, \dots, B^{\ell-1}, a, B^{\ell+1}, \dots B^{\operatorname{arity}(Q)-1}) = \top\}.$$

A **cap** *of an $(\varepsilon, \boldsymbol{m})$-mesa is an m_{k+1}-cap of the mesa for some $m_{k+1} \le \varepsilon m_k$.*

An $(\varepsilon, \boldsymbol{m})$-mesa has **constant location** *ℓ if $\ell_i = \ell$ for all $i \in \{0,1\}^{<k}$, and has* **constant relation** *Q if $Q_i = Q$ for all $i \in \{0,1\}^{<k}$.*

Let Y be an $(\varepsilon, \boldsymbol{m})$-mesa, and suppose $\boldsymbol{m'}$ has $\boldsymbol{m}$ as an initial segment. Then an $(\varepsilon, \boldsymbol{m'})$-mesa Z is an **extension** *of Y if (i) Z extends Y (as a tree of rocks), and (ii) Z at the level after the height of Y contains, for each pre-cap of Y, a rock that covers that pre-cap.*

Suppose m_{k+1} is such that $\frac{m_{k+1}}{m_k} \le \varepsilon$. An $(\varepsilon, \boldsymbol{m})$-mesa is m_{k+1}**-maximal** *if it has no extensions which are $(\varepsilon, \boldsymbol{m}^\wedge m_k)$-mesas.*

Note that if C is the cap of a mesa, then every rock covering C determines the truth value of its relation symbol (with its arguments and its ordering), in the sense that there is only one truth value that a large fraction of C agrees with.

Lemma 3.4. *Let Y be an $(\varepsilon, \boldsymbol{m})$-mesa with notation as in Definition 3.3. Let $m_{k+1} \le \varepsilon m_k$, and suppose that Y is m_{k+1}-maximal.*

(a) *Let $p \in \{0,1\}^k$. If the pre-cap A_p is an m_{k+1}-cap of Y, then A_p is ε-excellent.*
(b) *There is a (not necessarily unique) m_{k+1}-cap of Y.*

Proof. (a) This follows immediately from the definition of m_{k+1}-cap and the fact that $\frac{m_{k+1}}{m_k} \leq \varepsilon$.

(b) If there is no m_{k+1}-cap for any $p \in \{0,1\}^k$, then by the definition of an $(\varepsilon, \boldsymbol{m})$-mesa we can find an extension of Y to an $(\varepsilon, \boldsymbol{m}{}^\wedge m_{k+1})$-mesa, contradicting the assumption that Y was m_{k+1}-maximal. □

In fact, an $(\varepsilon, \boldsymbol{m})$-mesa is m_{k+1}-maximal if and only if it has some m_{k+1}-cap.

We will eventually want to obtain a bound on the height of an $(\varepsilon, \boldsymbol{m})$-mesa based on the underlying $\mathcal{L}$-structure G having the non-$\widehat{\tau}$-branching property. To do this, we will need an $(\varepsilon, \boldsymbol{m})$-mesa with constant relation and constant location.

We first define what it means for a mesa to be a substructure of another.

Definition 3.5. *Let $k, k* \in \mathbb{N}$, let $\varepsilon > 0$, and suppose $\boldsymbol{m} := \langle m_j \rangle_{j \leq k}$ and $\boldsymbol{m^*} := \langle m^*_j \rangle_{j \leq k^*}$ are staircases. Let Y be an $(\varepsilon, \boldsymbol{m})$-mesa and Y^* an $(\varepsilon, \boldsymbol{m^*})$-mesa.*

Then Y^* is a **substructure** *of Y if there are injective maps $\alpha\colon \{0,1\}^{\leq k^*} \to \{0,1\}^{\leq k}$ and $\gamma\colon \{0, \dots, k^*-1\} \to \{0, \dots, k-1\}$ such that, for all $i, i' \in \{0,1\}^{\leq k^*}$,*

- *$m^*_h = m_{\gamma(h)}$ for all $h \leq k^*$,*
- *$\mathrm{len}\big(\alpha(i)\big) = \gamma\big(\mathrm{len}(i)\big)$,*
- *if i is an initial segment of i' then $\alpha(i)$ is an initial segment of $\alpha(i')$,*
- *if $\mathrm{len}(i) < k^*$, then the rock of Y at node $\alpha(i)$ equals the rock of Y^* at node i,*
- *if $\mathrm{len}(i) = k^*$ and $\gamma(k^*) = k$, then the pre-cap of Y at node $\alpha(i)$ equals the pre-cap of Y^* at node i, and*
- *if $\mathrm{len}(i) = k^*$ and $\gamma(k^*) < k$, then the rock of Y at node $\alpha(i)$ covers the the pre-cap of Y^* at node i.*

We will soon show the key fact that for every $k^* \in \mathbb{N}$ there is some $k \geq k^*$, depending only on k^*, such that every $(\varepsilon, \boldsymbol{m})$-mesa of height at least k has some substructure that is a $(\varepsilon, \boldsymbol{m^*})$-mesa with constant location. We will use the following Ramsey-theoretic result about colored trees.

Lemma 3.6 ([PST12, Theorem 2 (i)])**.** *Let $p, q \geq 2$. Suppose T is a binary branching tree of height at least $H > 5 \cdot q \cdot p \cdot \log p$ along with a map ι from the nodes of the tree to $\{0, \dots, q-1\}$. Then there is a binary branching tree T^* and an injection $\alpha\colon T^* \to T$ such that*

- *T^* has height p,*
- *α preserves the partial ordering of nodes in the tree, and preserves when two nodes are on the same level, and*
- *$\iota \circ \alpha\colon T^* \to \{0, \dots, q-1\}$ is constant.*

Lemma 3.7. *Suppose $n_{\mathcal{L}}, k^* \geq 2$, and suppose Y is an $(\varepsilon, \boldsymbol{m})$-mesa of height $k > 5 \cdot n_{\mathcal{L}} \cdot k^* \cdot \log k^*$. Then there is some staircase $\boldsymbol{m^*}$ of length k^* and some substructure Y^* of Y that is an $(\varepsilon, \boldsymbol{m^*})$-mesa which has constant location and constant relation.*

Proof. This follows immediately from Lemma 3.6. □

Our next step is to show how to get from a mesa having constant location and constant relation to a witness to the k-branching property.

For $E \in \mathcal{L}$ and $0 \leq \ell \leq \mathrm{arity}(E) - 1$, write

$$\mathcal{E}^\ell(x_\ell, x_0, \ldots, x_{\ell-1}, x_{\ell+1}, \ldots, x_{\mathrm{arity}(E)-1}) := E(x_0, \ldots, x_{\mathrm{arity}(E)-1}),$$

so that we may easily isolate x_ℓ from the other variables when talking about stability.

Lemma 3.8. *Suppose there is an $(\varepsilon, \boldsymbol{m})$-mesa Y of height k with constant location ℓ and constant relation Q, and suppose $2^k \cdot (\mathrm{arity}(E) - 1) \cdot \varepsilon < 1$. Then $(G, \mathcal{E}^\ell(x_\ell; x_0, \ldots, x_{\ell-1}, x_{\ell+1}, \ldots, x_{n-1}))$ has the k-branching property.*

Proof. We use the notation for the components of Y as in Definition 3.3. Without loss of generality, we may assume $\ell = 0$. For each $\eta \in \{0,1\}^k$ let $a_\eta \in A_\eta$. Now for each $\eta \in \{0,1\}^k$ and each $\nu \in \{0,1\}^{<k}$ define

$$U_{\nu,\eta} := \Big\{(b_1, \ldots, b_{n-1}) \in \prod_{1 \leq j < n} B^j_\nu : \widehat{E}^{\beta_\nu(\langle 1 \cdots n-1 \rangle)}_\varepsilon(a_\eta, b_1, \ldots, b_{n-1}) \neq \widehat{E}^{\beta_\nu(\langle 1 \cdots n-1 \rangle)}_\varepsilon(a_\eta, B^1_\nu, \ldots, B^{n-1}_\nu)\Big\}.$$

Now by Lemma 2.6, we have $|U_{\nu,\eta}| < (n-1) \cdot \varepsilon \cdot \prod_{1 \leq j < n} |B^j_\nu|$ for every $\eta \in \{0,1\}^k$ and $\nu \in \{0,1\}^{<k}$. Hence

$$\Big|\bigcup_{\nu \preceq \eta} U_{\nu,\eta}\Big| < 2^k \cdot (n-1) \cdot \varepsilon \cdot \prod_{1 \leq j < n} |B^j_\nu|$$

for every $\nu \in \{0,1\}^{<k}$.

But we assumed $2^k \cdot (n-1) \cdot \varepsilon < 1$, and so for any ν we can find some $\boldsymbol{b}_\nu := (b^1_\nu, \ldots, b^{n-1}_\nu) \in \prod_{1 \leq j < n-1} B^j_\nu \setminus \bigcup_{\nu \preceq \eta} U_{\nu,\eta}$.

But then by construction, $\langle \boldsymbol{b}_\nu \rangle_{\nu \in \{0,1\}^{<k}}$ and $\langle a_\eta \rangle_{\eta \in \{0,1\}^k}$ witness that $\mathcal{E}^0(x_0; x_1, \ldots, x_{n-1})$ has the k-branching property. □

Putting all of these together we get the following crucial proposition.

Proposition 3.9. *Suppose $0 < \varepsilon < 2^{-\widehat{\tau}} \cdot n_{\mathcal{L}}^{-1}$. Further suppose $\mathcal{G}$ does not have the $\widehat{\tau}$-branching property, and let $g = \lceil 5 \cdot n_{\mathcal{L}} \cdot \widehat{\tau} \cdot \log \widehat{\tau} \rceil$. Finally, suppose $\boldsymbol{m} := \langle m_i \rangle_{i \leq g}$ is a staircase, and that A is a set with $|A| \geq m_0$. Then A contains an ε-excellent subset A' of size m_i for some $i \leq g$.*

Proof. By Lemma 1.8, for any $E \in \mathcal{L}$ and $\ell < \mathrm{arity}(E)$ the structure

$$(G, \mathcal{E}^\ell(x_\ell; x_0, \ldots, x_{\ell-1}, x_{\ell+1}, \ldots, x_{\mathrm{arity}(E)-1}))$$

has the non-$\widehat{\tau}$-branching property. By our assumption on ε, we may apply Lemma 3.8, and so any $(\varepsilon, \boldsymbol{m})$-mesa of constant location and constant relation E

can have height at most $\widehat{\tau}$. But then by Lemma 3.7, the height of any $(\varepsilon, \boldsymbol{m})$-mesa is at most $5 \cdot n \cdot \widehat{\tau} \cdot \log \widehat{\tau}$.

In particular there must be some $j \leq g$ and $(\varepsilon, \langle m_i \rangle_{i \leq g})$-mesa which is m_{j+1}-maximal. But then by Lemma 3.4 this mesa must have a cap, which has size m_j for some $j \leq g$. Further, by Lemma 3.4 this cap is ε-excellent. □

Having developed a method to find a large ε-excellent subset of any sets, we now aim to find a partition of $\mathcal{G}$ such that (1) all but one element of the partition is ε-excellent and (2) for any two elements of the partition, the size of one divides the size of the other, along with a bound on the size of the non-ε-excellent element.

Proposition 3.10. *Suppose* $0 < \varepsilon < 2^{-\widehat{\tau}} \cdot n_{\mathcal{L}}^{-1}$, *and that*

- $\mathcal{G}$ *does not have the* $\widehat{\tau}$*-branching property,*
- $n = |\mathcal{L}| \cdot q_{\mathcal{L}}$,
- $g = \lceil 5 \cdot n_{\mathcal{L}} \cdot \widehat{\tau} \cdot \log \widehat{\tau} \rceil$,
- $r = \lfloor \frac{1}{\varepsilon} \rfloor$, *and*
- $\boldsymbol{m} := \langle m_i \rangle_{i \leq g}$ *is a staircase such that*
  - $\frac{m_i}{m_{i+1}} = r$ *for all* $0 \leq i < g$, *and*
  - $|G| \geq m_0$.

Then there is a subset $G^* \subseteq G$ *and a partition* P *of* G^* *such that*

- $|G \setminus G^*| < m_0$,
- *each element of* P *is* ε*-excellent, and*
- $|p| \in \boldsymbol{m}$ *for all* $p \in P$.

Proof. We define the partition by induction. For the base case, let $G_0 := G$ and let P_0 be an ε-excellent subset of G_0 with $|P_0| \in \boldsymbol{m}$, as guaranteed by Proposition 3.9.

For the inductive step, suppose we that have already defined G_n and $\langle P_j \rangle_{j \leq n}$, where each P_j is ε-excellent and whose size is in $\boldsymbol{m}$. Let $G_{n+1} := G_n \setminus P_n$.

If $|G_{n+1}| < m_0$ then let $G^* := G \setminus G_{n+1}$ and let $P := \{P_i\}_{i \leq n}$; then G^* and P have the desired properties.

Otherwise let P_{n+1} be an ε-excellent subset of G_{n+1} with $|P_{n+1}| \in \boldsymbol{m}$, as guaranteed by Proposition 3.9, and proceed to the next step of the induction. □

4. Equitable partitions of excellent sets

We have just seen, in Proposition 3.10, that a large subset of a sufficiently large structure $\mathcal{G}$ may be partitioned into ε-excellent sets. In this section, we show, in Proposition 4.5, how to refine this into an equitable partition of $\mathcal{G}$ into $(\varepsilon + \zeta)$-excellent sets, for some $\zeta > 0$.

Then, in the main results of this section, Proposition 4.6 and Theorem 4.7, we show how to uniformly distribute the elements of our structure not in this large subset, obtaining an equitable partition of the entire structure which witnesses that it is close in edit distance to an equitable blow-up.

Our first lemma immediately implies that if a set agrees with an ε-excellent set on the truth values of all edge relations in $\mathcal{L}$ with respect to all parameters that are elements of G, then the set itself must be ε-excellent.

Lemma 4.1. *Let $E \in \mathcal{L}$ and let n be the arity of E. Suppose that A is (ε, k, E)-good and that A' is such that for all elements $b_1, \ldots, b_{n-1} \in G$ and every permutation σ of n,*

$$\widehat{E}_\varepsilon(x_{\sigma(0)}, \ldots, x_{\sigma(n-1)}) = \widehat{E}_\varepsilon(y_{\sigma(0)}, \ldots, y_{\sigma(n-1)})$$

where $x_0 = A$ and $y_0 = A'$, and $x_i = y_i = b_i$ whenever $1 \le i < n$. Then A' is (ε, k, E)-good.

Proof. We will prove the following statement $(*_k)$ by induction on k:
$(*_k)$: For all $b_{k-1}, \ldots, b_n \in G$ and permutations σ of n, if B_i is $(\varepsilon, k-i, E)$-good for all $1 \le i < k$, then

$$\widehat{E}_\varepsilon^{\sigma^+}(x_{\sigma(0)}, \ldots, x_{\sigma(n-1)}) = \widehat{E}_\varepsilon^{\sigma^+}(y_{\sigma(0)}, \ldots, y_{\sigma(n-1)})$$

where $x_0 = A$ and $y_0 = A'$, where $\sigma^+ := \sigma|_{\{0,\ldots,k-1\}}$, and $x_i = y_i = B_i$ whenever $1 \le i < k$, and $x_i = y_i = b_i$ whenever $k \le i < n$.

<u>*Case $k = 1$:*</u>
This is immediate by our assumption.

<u>*Case $k > 1$:*</u>
By the inductive assumption, A' is $(\varepsilon, k-1, E)$-good. We must show that it is (ε, k, E)-good.

Now suppose $A_1, \ldots, A_{k-1} \subseteq G$ and $a_k, \ldots, a_{n-1} \in G$, where A_i is $(\varepsilon, k-i, E)$-good whenever $1 \le i < k$. Without loss of generality, it suffices to show that

$$\widehat{E}_\varepsilon^{\mathrm{id}}(A, A_1, \ldots, A_{k-1}, a_k, \ldots, a_{n-1}) = \widehat{E}_\varepsilon^{\mathrm{id}}(A', A_1, \ldots, A_{k-1}, a_k, \ldots, a_{n-1}),$$

where id is the identity map on $\{0, \ldots, k-1\}$. But we know that

$$\widehat{E}_\varepsilon^{\mathrm{id}}(A, A_1, \ldots, A_{k-1}, a_k, \ldots, a_{n-1}) \in \{\top, \bot\}.$$

Suppose that $\widehat{E}_\varepsilon^{\mathrm{id}}(A, A_1, \ldots, A_{k-1}, a_k, \ldots, a_{n-1}) = \top$. Then

$$\frac{|\{a \in A_{k-1} \,:\, \widehat{E}_\varepsilon^{\mathrm{id}}(A, A_1, A_2, \ldots, A_{k-2}, a, a_k, \ldots, a_{n-1}) = \top\}|}{|A_{k-1}|} \ge 1 - \varepsilon.$$

But then by the inductive hypothesis we also have

$$\frac{|\{a \in A_{k-1} \,:\, \widehat{E}_\varepsilon^{\mathrm{id}}(A', A_1, A_2, \ldots, A_{k-2}, a, a_k, \ldots, a_{n-1}) = \top\}|}{|A_{k-1}|} \ge 1 - \varepsilon.$$

Hence $\widehat{E}_\varepsilon^{\mathrm{id}}(A', A_1, \ldots, A_{k-1}, a_k, \ldots, a_{n-1}) = \top$.

The case when $\widehat{E}_\varepsilon^{\mathrm{id}}(A, A_1, \ldots, A_{k-1}, a_k, \ldots, a_{n-1}) = \bot$ is identical. ∎

Now we want to show that if our ε-excellent set is sufficiently large then a uniformly random equitable partition will be $(\varepsilon+\zeta)$-excellent with high probability, for some ζ.

Lemma 4.2. *If $\varphi(\overline{x};\overline{y})$ has the non-$\overline{\tau}$-order property in a structure $\mathcal{M}$ then for any finite $A \subseteq \mathcal{M}$ with $|A| > 2$,*

$$|\{\{\overline{a} \in A : \varphi(\overline{a}, \overline{b})\} : \overline{b} \in \mathcal{M}\}| \leq |A|^{\overline{\tau}}.$$

Proof. This is immediate from [She90, Theorem II.4.10(4)]. □

The following result provides an upper bound on the probability that the fraction of elements satisfying property S will be more than the expected value by an additive constant t.

Proposition 4.3 ([Ska13])**.** *Suppose we have N elements of which K have a property S. Let $H(n, N, K)$ be the random variable which selects without replacement s elements and returns the number which have property S. Then for any $t > 0$ we have*

$$\mathbb{P}\left[\frac{H(s,N,K)}{s} \geq \frac{K}{N} + t\right] \leq e^{-2t^2 s}.$$

For our purposes we will have an ε-excellent set A and we will want to sample a random partition P of A. We will then want to ask the following question, for a given partition element $p \in P$, a given relation E and a given collection of good sets $B_1, \ldots, B_{\text{arity}(E)-1}$: What is the probability that the statement "the fraction of elements of p which disagree with A on the value of E with respect to $B_1, \ldots, B_{\text{arity}(E)-1}$ is greater than $\varepsilon + \zeta$" is true?

Now, Proposition 4.3 tells us that not only is this probability small, but even if we were to ask polynomially many such questions, the probability that any of them would hold is (asymptotically) small. But we also know by Lemma 4.2 that there exist only polynomially many such questions, hence the probability that any of them hold is (asymptotically) small. But if none of the questions holds of p then we know p is $(\varepsilon + \zeta)$-excellent, which was our goal. We will now make this precise.

Proposition 4.4. *Consider a population with N elements. Let $M_0, \ldots, M_k$ be subsets of the population where $k = CN^\ell$ for constants C and ℓ, and suppose that r divides N. Then for any $t > 0$, so long as $r \log r + \log C < 2t^2 N - r\ell \log N$, there is an equitable partition of N into r pieces such that for each element X of the partition, we have*

$$\frac{|M_i \cap X|}{|X|} \leq \frac{|M_i|}{N} + t$$

whenever $0 \leq i \leq k$.

Proof. By Proposition 4.3,

$$\mathbb{P}\left[\bigvee_{i\leq k}\left(\frac{H(N/r,N,M_i)}{N/r}\geq\frac{|M_i|}{N}+t\right)\right]\leq C\cdot N^\ell\cdot e^{-2t^2N/r}.$$

If P is a uniformly random partition then for any $p\in P$ and $i\leq k$, the probability that p contains at least h many elements in M_i is $\mathbb{P}[H(N/r,N,M_i)\geq h]$. Hence we have

$$\mathbb{P}\left[\bigvee_{p\in P}\bigvee_{i\leq k}\left(\frac{|p\cap M_i|}{|p|}\geq\frac{|M_i|}{N}+t\right)\right]\leq r\cdot C\cdot N^\ell\cdot e^{-2t^2N/r}.$$

But if $r\log r+\log C<2t^2N-r\ell\log N$, we then have

$$\mathbb{P}\left[\bigvee_{p\in P}\bigvee_{i\leq k}\left(\frac{|p\cap M_i|}{|p|}\geq\frac{|M_i|}{N}+t\right)\right]<1,$$

and so there must be some such partition P of N. □

Putting these all together we have the following.

Proposition 4.5. *Let ε, $\zeta>0$. Suppose A is an ε-excellent class, and $r\in\mathbb{N}$ is such that r divides $|A|$. Further, suppose*

$$r\log r+\log(2|\mathcal{L}|(q_{\mathcal{L}}!))<2\zeta^2|A|-r2^{\widehat{\tau}+1}\log|A|.$$

Then there is an equitable partition of A into r pieces, each of which is $(\varepsilon+\zeta)$-excellent.

Proof. Let $M_0,\dots,M_k$ be sets of the form

$$\{a_0\in A\,:\,G\models E(a_{\sigma(0)},\dots,a_{\sigma(\ell-1)})\}$$

or of the form

$$\{a_0\in A\,:\,G\models\neg E(a_{\sigma(0)},\dots,a_{\sigma(\ell-1)})\}$$

for some $E\in\mathcal{L}$, some $a_1,\dots,a_{\ell-1}\in G$, and some permutation σ of $\{0,\dots,\ell-1\}$, where $\ell:=\operatorname{arity}(E)$. Then by Lemma 1.8, $\mathcal{G}$ has the non-$2^{\widehat{\tau}+1}$ order property. Hence by Lemma 4.2, we have $k\leq 2|\mathcal{L}|\cdot q_{\mathcal{L}}!\cdot|A|^{2^{\widehat{\tau}+1}}$. The result then follows immediately from Lemma 4.1 and Proposition 4.4. □

Proposition 4.6. *Let $\zeta>0$. Suppose $0<\varepsilon<2^{-\widehat{\tau}}\cdot n_{\mathcal{L}}^{-1}$, and that*

(a) *$\mathcal{G}$ does not have the $\widehat{\tau}$-branching property,*
(b) *$g:=\lceil 5\cdot n_{\mathcal{L}}\cdot\widehat{\tau}\cdot\log\widehat{\tau}\rceil$,*
(c) *m is a positive natural number such that $m\cdot\lfloor\frac{1}{\varepsilon}\rfloor^g\leq|G|$, and*
(d) *$2\zeta^2m-\frac{|G|}{m}2^{\widehat{\tau}+1}\log m>\frac{|G|}{m}\log\frac{|G|}{m}+\log(2|\mathcal{L}|(q_{\mathcal{L}}!))$.*

Then there is a subset $G^+\subseteq G$ and a partition P of G^+ such that

(i) *$|G\setminus G^+|<m\cdot\lfloor\frac{1}{\varepsilon}\rfloor^g$,*
(ii) *each element of P is $(\varepsilon+\zeta)$-excellent,*

(iii) *P is equitable, and*
(iv) *each element of P has size m.*

Proof. Let $m_g = m$ and let $m_{i-1} = m_i \cdot \lfloor \frac{1}{\varepsilon} \rfloor$ for $1 \le i \le g$. By assumption (c) we have that $|G| \ge m_0$. Using assumptions (a) and (b) we can apply Proposition 3.10 to get a $G^+ \subseteq G$ and P^+ which satisfies (i), where each element of P^+ is ε-excellent, and where m divides the size of each element of P^+. Note that the size r of the partition P^+ is bounded above by $\frac{G}{m}$ and the size of any such partition is bounded below by m. Hence by applying (d), we obtain

$$2\zeta^2|p| - r2^{\widehat{\tau}+1}\log|p| > r\log r + \log(2|\mathcal{L}|(q_{\mathcal{L}}!))$$

for any element $p \in P^+$, and so we can apply Proposition 4.5 to find a refinement P of P^+ which is equitable and where every element is $(\varepsilon + \zeta)$-excellent. □

Finally, now that we have an equitable partition of a large subset of our graph, each of whose elements are appropriately excellent, we are able to prove our main theorem.

Theorem 4.7. *Let $\zeta, \eta > 0$ and let $m := \lceil |G| \cdot \eta \rceil > 2$. Suppose $0 < \varepsilon < 2^{-\widehat{\tau}} \cdot n_{\mathcal{L}}^{-1}$, and that*

(a) *$\mathcal{G}$ does not have the $\widehat{\tau}$-branching property,*
(b) *$g := \lceil 5 \cdot n_{\mathcal{L}} \cdot \widehat{\tau} \cdot \log \widehat{\tau} \rceil$,*
(c) *$\beta := \varepsilon^g - (\eta + \frac{1}{|G|}) > 0$, and*
(d) *$2\zeta^2\eta m - 2^{\widehat{\tau}+1}\log m > \eta \log(2|\mathcal{L}|(q_{\mathcal{L}}!)) - \log \eta$.*

Then there is an $\mathcal{L}$-structure $\mathcal{H}$ with the same underlying set G as $\mathcal{G}$ and an equitable partition P^* of $\mathcal{H}$ such that for all $E \in \mathcal{L}$,

- *for all $\langle p_i^* \rangle_{i<\ell} \subseteq P^*$,*
$$\left| (E^{\mathcal{G}} \triangle E^{\mathcal{H}}) \cap \prod_{i<\ell} p_i^* \right| \le \ell \cdot \left(\frac{(\varepsilon+\zeta)\cdot\beta + \eta}{\beta} \right) \cdot \prod_{i<\ell} |p_i|,$$
- *P^* is indivisible, and*
- *$\frac{\beta}{\varepsilon^g \cdot \eta} \le |P^*| \le \frac{1}{\eta} + 1$,*

where $\ell := \mathrm{arity}(E)$.

Proof. First note that by (d) and the fact that $\frac{|G|}{m} \le \eta^{-1}$, condition (d) of Proposition 4.6 holds. Next, $m \cdot \lfloor \frac{1}{\varepsilon} \rfloor^g \le \lceil |G| \cdot \eta \rceil \frac{1}{\varepsilon^g} \le (|G| \cdot \eta + 1)\frac{1}{\varepsilon^g} = |G| \cdot \frac{\eta + \frac{1}{|G|}}{\varepsilon^g} \le |G|$ and we so we can find a subset G^+ and an equitable partition P^+ of G^+ as in Proposition 4.6 where $|G \setminus G^+| < m \cdot \lfloor \frac{1}{\varepsilon} \rfloor^g$ and each element of P^+ has size m.

As each element of P^+ is $(\varepsilon + \zeta)$-excellent, by Proposition 2.7 there is a structure (G^+, E^{**}) on the same underlying set as G^+ such that P^+ is indivisible and $|(E^{\mathcal{G}}|_{G^+} \triangle E^{**}) \cap \prod_{i<\ell} p_i| \le \ell \cdot (\varepsilon + \zeta) \cdot \prod_{i<\ell} |p_i|$ for all $p_0, \ldots, p_{\ell-1} \in P$.

Finally, we can extend P^+ to an equitable partition P^* of G by adding elements of $G \setminus G^+$ arbitrarily while preserving the appropriate sizes of the elements of P.

As $|G \setminus G^+| < m \cdot \lfloor \frac{1}{\varepsilon} \rfloor^g$, we have

$$
\begin{aligned}
|P^*| &\geq \frac{|G| - \lceil |G| \cdot \eta \rceil \cdot \lfloor \frac{1}{\varepsilon} \rfloor^g}{\lceil |G| \cdot \eta \rceil} \geq \frac{|G| - (|G| \cdot \eta + 1) \cdot \lfloor \frac{1}{\varepsilon} \rfloor^g}{|G| \cdot \eta} \\
&= \frac{1 - (\eta + \frac{1}{|G|}) \cdot \lfloor \frac{1}{\varepsilon} \rfloor^g}{\eta} \geq \frac{1 - (\eta + \frac{1}{|G|}) \cdot (\frac{1}{\varepsilon})^g}{\eta} \\
&= \frac{\varepsilon^g - (\eta + \frac{1}{|G|})}{\varepsilon^g \cdot \eta} = \frac{\beta}{\varepsilon^g \cdot \eta}.
\end{aligned}
$$

Also note that each element of P^* has size at least m, and so $|P^*| \leq \frac{|G|}{m} \leq \frac{1}{\eta} + 1$.

Further note that by an appropriate assignment of edge relations on $G \setminus G^+$, we can extend E^{**} to an edge relation $E^{\mathcal{H}}$ such that P^* is also an indivisible partition of $\mathcal{H}$. Let

$$
k^* := \sup \left\{ \frac{|p^* \setminus p|}{|p|} \ : \ p \in P^+, \ p^* \in P^*, \text{ and } p \subseteq p^* \right\}.
$$

Then we have

$$
k^* \leq \frac{\frac{m \cdot \lfloor \frac{1}{\varepsilon} \rfloor^g}{|P^*|}}{m} = \frac{\lfloor \frac{1}{\varepsilon} \rfloor^g}{|P^*|} \leq \frac{(\frac{1}{\varepsilon})^g}{\frac{\beta}{\varepsilon^g \cdot \eta}} = \frac{\eta}{\beta}.
$$

Let X_0 be the collection of ℓ-tuples at least one element of which is contained in $G \setminus G^+$. Suppose $p_0^*, \ldots, p_{\ell-1}^* \in P^*$. We then have

$$
\begin{aligned}
\left| \big((E^{\mathcal{G}} \cap X_0) \triangle (E^{\mathcal{H}} \cap X_0) \big) \cap \prod_{i<\ell} p_i^* \right| &\leq \big| X_0 \cap \prod_{i<\ell} p_i^* \big| \\
&\leq \ell \cdot k^* \cdot \prod_{i<\ell} |p_i^*| \\
&\leq \ell \cdot \frac{\eta}{\beta} \cdot \prod_{i<\ell} |p_i^*|.
\end{aligned}
$$

Putting this together we get

$$
\begin{aligned}
\left| (E^{\mathcal{G}} \triangle E^{\mathcal{H}}) \cap \prod_{i<\ell} p_i^* \right| &\leq \ell \cdot (\varepsilon + \zeta) \cdot \prod_{i<\ell} |p_i^*| + \ell \cdot \frac{\eta}{\beta} \cdot \prod_{i<\ell} |p_i^*| \\
&\leq \ell \cdot \left(\varepsilon + \zeta + \frac{\eta}{\beta} \right) \cdot \prod_{i<\ell} |p_i^*| \\
&\leq \ell \cdot \left(\frac{(\varepsilon + \zeta) \cdot \beta + \eta}{\beta} \right) \cdot \prod_{i<\ell} |p_i^*|.
\end{aligned}
$$

□

There is a tension among the three parameters ε, η, and ζ. Namely, as η becomes smaller, the potential size of the partition becomes larger, but at the same time, the fraction of elements that we need to change becomes smaller. On the other hand, as ε becomes smaller, both the potential partition size and the number of elements we need to change become larger. Finally, ζ must be chosen to as to be consistent with the other two parameters in (d); in particular, as η becomes smaller, ζ must get larger.

While Theorem 4.7 provides precise lower bounds on how large a structure we need in order for stable regularity to come into play, these bounds can be unwieldy. If instead we are willing to simply consider "sufficiently large" structures then the result has a much cleaner form.

Theorem 4.8 (Stable regularity for relational structures)**.** *For every $\varepsilon > 0$ there is a k_ε such that whenever*

(a) *$\mathcal{G}$ is a $\mathcal{L}$-structure with $|G| \geq k_\varepsilon$,*
(b) *$\mathcal{G}$ does not have the $\widehat{\tau}$-branching property,*
(c) *$g := \lceil 5 \cdot n_{\mathcal{L}} \cdot \widehat{\tau} \cdot \log \widehat{\tau} \rceil$, and*
(d) *$\varepsilon < 2^{-(g+1)(g+2)}$,*

then there is an $\mathcal{L}$-structure $\mathcal{H}$ with the same underlying set G as $\mathcal{G}$ and an equitable partition P of $\mathcal{H}$ such that for all $E \in \mathcal{L}$,

- *for all $\langle p_i \rangle_{i<\ell)} \subseteq P$,*
$$\Big|(E^{\mathcal{G}} \triangle E^{\mathcal{H}}) \cap \prod_{i<\ell} p_i\Big| \leq \ell \cdot \varepsilon \cdot \prod_{i<\ell} |p_i|,$$
- *P is indivisible, and*
- *$|P| \leq \varepsilon^{-g-2}$,*

where $\ell :=$ arity(E).

Proof. Suppose $\varepsilon > 0$ satisfies (c) and (d). We will choose ε_1, ζ_1, $\eta_1 > 0$ and $k_\varepsilon \in \mathbb{N}$ in terms of ε such that for all $\mathcal{G}$ satisfying (a) and (b) we may apply Theorem 4.7 to ε_1, ζ_1, and η_1 to produce an $\mathcal{L}$-structure $\mathcal{H}$ and equitable partition P, which we will verify have the desired properties.

Choose γ_1 such that $1 < \gamma_1 < 2$ and let $p > 4$ be such that $\gamma_1^g(1+\varepsilon) < p < 2^{g+1}$ (which is possible as $g > 1$, as $\gamma_1 < 2$, and as $\varepsilon < 1$). Therefore
$$\gamma_1^g < p - \varepsilon \cdot \gamma_1^g$$
and so
$$\frac{\gamma_1^g}{p - \varepsilon \cdot \gamma_1^g} < 1.$$
But then we also have have
$$\frac{\gamma_1^g}{1 - \frac{\varepsilon}{p+1} \cdot \gamma_1^g} < \frac{\gamma_1^g}{1 - \frac{\varepsilon}{p} \cdot \gamma_1^g} = p\frac{\gamma_1^g}{p - \varepsilon \cdot \gamma_1^g} < p. \tag{A}$$

Choose $\varepsilon_1 = \frac{\varepsilon}{(p+2)\gamma_1}$. In particular, we have $\varepsilon_1 \cdot \gamma_1 < \frac{\varepsilon}{p+1} < 1$. Further, as $p > 1$ and $\gamma_1 > 1$, we have $\varepsilon_1 < \frac{1}{1+\gamma_1^{g+1}}$, and so $\varepsilon_1(1 + \gamma_1^{g+1}) < 1$.

Let $\zeta_1 := \varepsilon_1 \cdot (\gamma_1 - 1)$, so that $\gamma_1 = 1 + \frac{\zeta_1}{\varepsilon_1}$.
Let $\eta_1 := \varepsilon_1^{g+1} \cdot \gamma_1^{g+1} = (\varepsilon_1 + \zeta_1)^{g+1}$.
Let $\beta := \varepsilon_1^g - (\eta_1 + \frac{1}{|G|})$.
Let k_ε be large enough that

(1) $k_\varepsilon \eta_1 > 2$,
(2) $2k_\varepsilon \zeta_1^2 \eta_1^2 - 2^{\widehat{\tau}+1} \log\big(k_\varepsilon \eta_1\big) > \eta_1 \log(2|\mathcal{L}|(q_{\mathcal{L}}!)) - \log \eta_1$,
(3) $k_\varepsilon > \frac{2^{\widehat{\tau}+1}}{2\zeta_1^2 \eta_1^2}$,
(4) $\frac{\gamma_1^g}{1 - \frac{\varepsilon}{p+1}\gamma_1^g - \frac{1}{\varepsilon_1^g k_\varepsilon}} < p$, and
(5) $k_\varepsilon > \varepsilon_1^{-g-1}$.

(Any sufficiently large k_ε satisfies (4) by (A), and clearly (1), (2), (3), and (5) hold for all sufficiently large k_ε.)

Let $m := \lceil |G| \cdot \eta_1 \rceil$ and let $\ell := \mathrm{arity}(E)$. We have assumed that $\mathcal{G}$ does not have the $\widehat{\tau}$-branching property. We now show that $m > 2$, that $\beta > 0$, and that $2\zeta_1^2 \eta_1 m - 2^{\widehat{\tau}+1} \log m > \eta_1 \log(2|\mathcal{L}|(q_{\mathcal{L}}!)) - \log \eta_1$ (so that we may apply Theorem 4.7).

Note that (1) ensures that $m = \lceil |G| \cdot \eta_1 \rceil > 2$. The function $2\zeta_1^2 \eta_1 x - 2^{\widehat{\tau}+1} \log x$ is increasing for $x > \frac{2^{\widehat{\tau}+1}}{2\zeta_1^2 \eta_1}$, and so (2) and (3) imply that

$$2\zeta_1^2 \eta_1 m - 2^{\widehat{\tau}+1} \log m > \eta_1 \log(2|\mathcal{L}|(q_{\mathcal{L}}!)) - \log \eta_1$$

holds.

Now

$$\beta = \varepsilon_1^g - (\varepsilon_1\gamma_1)^{g+1} - \frac{1}{|G|} \geq \varepsilon_1^g - (\varepsilon_1\gamma_1)^{g+1} - \frac{1}{k_\varepsilon} > \varepsilon_1^g - (\varepsilon_1\gamma_1)^{g+1} - \varepsilon_1^{g+1},$$

where the last inequality follows from (5). But

$$\varepsilon_1^g - (\varepsilon_1\gamma_1)^{g+1} - \varepsilon_1^{g+1} = (\varepsilon_1)^g(1 - \varepsilon_1(\gamma_1^{g+1} + 1)).$$

Recall that $1 > \varepsilon_1(\gamma_1^{g+1}+1)$, and so $\beta > 0$. Also note that (iv) implies $\varepsilon < 2^{-\widehat{\tau}} \cdot n_{\mathcal{L}}^{-1}$, and so $\varepsilon_1 < 2^{-\widehat{\tau}} \cdot n_{\mathcal{L}}^{-1}$.

Hence we may apply Theorem 4.7 to ε_1, ζ_1, and η_1 to obtain an $\mathcal{L}$-structure $\mathcal{H}$ with the same underlying set G as $\mathcal{G}$ and an equitable partition P of $\mathcal{H}$ such that for all $E \in \mathcal{L}$,

- for all $\langle p_i^* \rangle_{i<\ell} \subseteq P$,
$$\left| (E^{\mathcal{G}} \triangle E^{\mathcal{H}}) \cap \prod_{i<\ell} p_i^* \right| \leq \ell \cdot \left(\frac{(\varepsilon_1 + \zeta_1) \cdot \beta + \eta_1}{\beta} \right) \cdot \prod_{i<\ell} |p_i|,$$
- P is indivisible, and
- $|P| \leq \frac{1}{\eta_1} + 1$.

We must show that $\frac{(\varepsilon_1+\zeta_1)\cdot\beta+\eta_1}{\beta} \leq \varepsilon$ and that $\frac{1}{\eta_1} + 1 \leq \varepsilon^{-g-2}$.

Recall that $\varepsilon_1 + \zeta_1 = \varepsilon_1\gamma_1$. Observe that

$$\begin{aligned}\frac{\eta_1}{\beta} &= \frac{(\varepsilon_1\gamma_1)^{g+1}}{\beta} \\ &= \frac{(\varepsilon_1\gamma_1)^{g+1}}{\varepsilon^g - (\varepsilon_1\gamma_1)^{g+1} - \frac{1}{|G|}} \\ &\leq \frac{(\varepsilon_1\gamma_1)^{g+1}}{\varepsilon^g - (\varepsilon_1\gamma_1)^{g+1} - \frac{1}{k_\varepsilon}} \\ &= \varepsilon_1 \cdot \gamma_1 \frac{\gamma_1^g}{1 - (\varepsilon_1\gamma_1)\gamma_1^g - \frac{1}{\varepsilon_1^g k_\varepsilon}} \\ &< \varepsilon_1\gamma_1 p,\end{aligned}$$

where the last inequality follows from (4). Hence $\frac{(\varepsilon_1+\zeta_1)\cdot\beta+\eta_1}{\beta} = \varepsilon_1\gamma_1 + \frac{\eta_1}{\beta} < \varepsilon_1\gamma_1 + \varepsilon_1\gamma_1 p < \varepsilon$.

Now, we have

$$\frac{1}{\eta_1} = \frac{1}{(\varepsilon_1\gamma_1)^{g+1}} = \big(\frac{p+2}{\varepsilon}\big)^{g+1} < (2p)^{g+1}\varepsilon^{-g-1} - 1$$

as $p > 4$. Finally, we have

$$\frac{1}{\eta_1} + 1 < (2p)^{g+1}\varepsilon^{-g-1} < (2 \cdot 2^{g+1})^{g+1}\varepsilon^{-g-1} \leq 2^{(g+1)(g+2)}\varepsilon^{-g-1} < \varepsilon^{-g-2},$$

where the last inequality follows from (d). □

Note that the corresponding counting and removal lemmas follow immediately from Theorem 4.8.

5. Almost stable regularity for relational structures

We now consider structures that are not stable, but which have very few witnesses to their non-stability. In this "almost stable" situation we will show that there is also a highly structured regularity lemma, in which a modification of the original structure arises as a finite blow-up. However, in this almost stable case, we merely get a *global* regularity lemma, rather than a *local* one.

More precisely, instead of obtaining a blow-up by changing a small fraction of the relations across each tuple of elements of the partition (of appropriate length), we can instead obtain a blow-up only by changing a small fraction of the relations across the entire structure. The key difference is that the vertices corresponding to these modified relations might be concentrated in certain regions of the structure, in which they make up a large fraction of the vertices.

This distinction between local and global regularity is often referred to as the distinction between *regularity* and *weak regularity*.

Definition 5.1. *Let $\mathcal{F}$ and $\mathcal{G}$ be finite $\mathcal{L}$-structures with underlying sets F and G respectively, and set $n = |G|$ and $k = |F|$. Define the* **induced homomorphism density** *of $\mathcal{F}$ in $\mathcal{G}$ to be*

$$t_{\text{ind}}(\mathcal{F}, \mathcal{G}) = \frac{\big|\text{ind}(\mathcal{F}, \mathcal{G})\big|}{n(n-1)\cdots(n-k+1)},$$

where $\text{ind}(\mathcal{F}, \mathcal{G})$ is the number of embeddings from $\mathcal{F}$ to $\mathcal{G}$, in other words, injective homomorphisms that yields an induced substructure (i.e., which preserve all relations and all negations of relations).

For more details on induced homomorphism densities in the case of graphs, see [Lov12, §5.2]; for a more general setting, see [AC14, §2] and [Kru16, Chapter 1].

Definition 5.2. *Let $\widehat{\tau} \in \mathbb{N}$. An $\mathcal{L}$-structure $\mathcal{M}$* **minimally** *has the $\widehat{\tau}$-branching property for a quantifier-free formula $\varphi(\overline{x}; \overline{y})$ if $\mathcal{M}$ has the $\widehat{\tau}$-branching property for $\varphi(\overline{x}; \overline{y})$ and no induced substructure of $\mathcal{M}$ has the $\widehat{\tau}$-branching property for $\varphi(\overline{x}; \overline{y})$.*

Lemma 5.3. *If $\mathcal{M}$ minimally has the $\widehat{\tau}$-branching property for $\varphi(\overline{x}; \overline{y})$ then $|\mathcal{M}| \leq 2^{\widehat{\tau}} \cdot (|\overline{x}| + |\overline{y}|)$.*

Proof. Suppose $\mathcal{M}$ has the $\widehat{\tau}$-branching property for $\varphi(\overline{x}; \overline{y})$ but $|\mathcal{M}| > 2^{\widehat{\tau}} \cdot (|\overline{x}| + |\overline{y}|)$. Let $M_0 \subseteq M$ consist of all tuples in a witness to the $\widehat{\tau}$-branching property for $\varphi(\overline{x}; \overline{y})$. Then $|M_0| \leq 2^{\widehat{\tau}} \cdot (|\overline{x}| + |\overline{y}|)$, and so $\mathcal{M}_0$, the induced substructure of $\mathcal{M}$ with underlying set M_0, is a proper substructure of $\mathcal{M}$. Hence $\mathcal{M}_0$ also has the $\widehat{\tau}$-branching property for $\varphi(\overline{x}; \overline{y})$, and so $\mathcal{M}$ was not minimal. □

Definition 5.4. *Let $\widehat{\tau} \in \mathbb{N}$ and $\delta > 0$. An $\mathcal{L}$-structure $\mathcal{M}$ has the $(\delta, \widehat{\tau})$-branching property for a quantifier-free formula $\varphi(\overline{x}; \overline{y})$ if there is a structure $\mathcal{N}$ which minimally has the $\widehat{\tau}$-branching property and for which $t_{\text{ind}}(\mathcal{N}, \mathcal{M}) \geq \delta$.*

We say an $\mathcal{L}$-structure $\mathcal{M}$ has the $(\delta, \widehat{\tau})$-branching property if it has the $(\delta, \widehat{\tau})$-branching property for some relation $E \in \mathcal{L}$ with some partition of the variables where one part is a singleton.

Note that a structure $\mathcal{M}$ has the $\widehat{\tau}$-branching property for a quantifier-free formula $\varphi(\overline{x}; \overline{y})$ exactly when there is a structure $\mathcal{N}$ which minimally has the $\widehat{\tau}$-branching property for $\varphi(\overline{x}; \overline{y})$ and for which there exists at least one embedding from $\mathcal{N}$ into $\mathcal{M}$. This motivates the idea that a structure not having the $(\delta, \widehat{\tau})$-branching property is a sign that it has very few witnesses to non-stability.

The next result follows from [AC14, Theorem 2].

Proposition 5.5 ([AC14, Theorem 2])**.** *Suppose $\langle \mathcal{F}_i \rangle_{i \leq \ell}$ is a finite collection of finite $\mathcal{L}$-structures. Then for every $\varepsilon > 0$ there is an $n_\varepsilon \in \mathbb{N}$ and a $\delta > 0$ such that whenever*

- *$\mathcal{M}$ is a finite $\mathcal{L}$-structure with $|M| > n_\varepsilon$ and*
- *$t_{\text{ind}}(\mathcal{F}_i, \mathcal{M}) < \delta$ for all $i \leq \ell$,*

then there is an $\mathcal{L}$*-structure* $\mathcal{M}^*$ *with the same underlying set as* $\mathcal{M}$ *such that*

- $t_{\text{ind}}(\mathcal{F}_i, \mathcal{M}^*) = 0$ *for all* $i \leq \ell$ *and*
- $|E^{\mathcal{M}} \triangle E^{\mathcal{M}^*}| \leq \varepsilon \cdot |M|^{\text{arity}(E)}$ *for all* $E \in \mathcal{L}$.

Note that [AC14, Theorem 2] was originally stated in terms of quantities of the form $p(\mathcal{F}_i, \mathcal{M})$ (and analogously for $\mathcal{M}^*$), which equals $t_{\text{ind}}(\mathcal{F}_i, \mathcal{M})/t_{\text{ind}}(\mathcal{F}_i, \mathcal{F}_i)$ (by their Fact 1). Note that when $\mathcal{F}_i$ minimally has the $\widehat{\tau}$-branching property for all $E \in \mathcal{L}$ with partitions of the variables where one part is a singleton, then the denominator $t_{\text{ind}}(\mathcal{F}_i, \mathcal{F}_i)$ is bounded by $(2^{\widehat{\tau}} \cdot q_{\mathcal{L}})^{2^{\widehat{\tau}} \cdot q_{\mathcal{L}}}$ by Lemma 5.3. Hence one can check that the removal lemma Proposition 5.5 is essentially equivalent to theirs.

Theorem 5.6 (Almost stable regularity for relational structures)**.** *Let* $g := \lceil 5 \cdot n_{\mathcal{L}} \cdot \widehat{\tau} \cdot \log \widehat{\tau} \rceil$. *For all* $\varepsilon > 0$ *there is a* k_ε *and* $\delta > 0$ *such that if*

- $\mathcal{M}$ *is a* $\mathcal{L}$*-structure with* $|\mathcal{M}| \geq k_\varepsilon$*, and*
- $\mathcal{M}$ *does not have the* $(\delta, \widehat{\tau})$*-branching property,*

then there is a structure $\mathcal{H}$ *with the same underlying set as* $\mathcal{M}$ *and an equitable partition* P *of* $\mathcal{H}$ *such that*

(i) $|E^{\mathcal{M}} \triangle E^{\mathcal{H}}| \leq \text{arity}(E) \cdot \varepsilon \cdot |M|^{\text{arity}(E)}$ *for all* $E \in \mathcal{L}$,
(ii) P *is indivisible, and*
(iii) $|P| \leq \left(\frac{\varepsilon}{2}\right)^{-g-2}$.

Proof. First apply Proposition 5.5 with $\frac{\varepsilon}{2}$ to get a structure $\mathcal{M}^*$ without the $\widehat{\tau}$-branching property such that $|E^{\mathcal{M}} \triangle E^{\mathcal{M}^*}| \leq \text{arity}(E) \cdot \frac{\varepsilon}{2} \cdot |M|^{\text{arity}(E)}$ for all $E \in \mathcal{L}$. Then apply Theorem 4.8 with $\mathcal{M}^*$ and $\frac{\varepsilon}{2}$ to get a structure $\mathcal{H}$ and partition P such that (ii) and (iii) hold and $|E^{\mathcal{M}^*} \triangle E^{\mathcal{H}}| \leq \text{arity}(E) \cdot \frac{\varepsilon}{2} \cdot |M|^{\text{arity}(E)}$ for all $E \in \mathcal{L}$. Then condition (i) follows by considering the symmetric difference of $E^{\mathcal{M}}$ and $E^{\mathcal{H}}$. □

6. Borel stable regularity for relational structures

We now consider ways of extending the almost stable regularity lemma from finite relational structures to Borel relational structures. Somewhat analogously for the case of graphs, Lovász and Szegedy [LS07] have developed analytic versions of the graph regularity lemma, expressed in terms of *graphons* and measurable partitions of their domains.

In this section we provide an almost stable regularity lemma for Borel structures, which shows that every Borel structure that is almost stable (in a sense we make precise) is close in L^1 to a Borel blow-up of a finite structure.

We will define Borel structures to have underlying set $[0, 1]$, and we will mostly deal with Lebesgue measure λ on $[0, 1]$. Note that whenever (P, μ) is a standard probability space, there is a measure preserving map from $([0, 1], \lambda)$ onto (P, μ). Hence the main arguments of this section go through with $([0, 1], \lambda)$ replaced by an arbitrary standard probability space.

We begin with definitions of Borel structures and the notions of L^1-distance, blow-ups, and induced homomorphism densities for them. These can be seen as analogous to the corresponding notions for the theory of graphons [Lov12, Chapter 7].

Definition 6.1. *A* **Borel** *$\mathcal{L}$-structure $\mathcal{G}$ is an $\mathcal{L}$-structure with underlying set $[0,1]$ such that for all $E \in \mathcal{L}$, the relation $E^{\mathcal{G}}$ interpreting the relation symbol E is Borel.*

It will often be convenient to work with characteristic functions instead of relations.

Definition 6.2. *Let $\mathcal{H}$ be an $\mathcal{L}$-structure (with arbitrary underlying set). For each $E \in \mathcal{L}$, define $\widetilde{E}^{\mathcal{H}}\colon [0,1]^{\mathrm{arity}(E)} \to \{0,1\}$ to be the characteristic function of the relation $E^{\mathcal{H}}$. Note that these functions are Borel when H is a Borel L-structure.*

The L^1-distance plays a key role in our arguments in this section.

Definition 6.3. *Suppose $\mathcal{G}_0, \mathcal{G}_1$ are Borel $\mathcal{L}$-structures. We define the L^1*-**distance** *between $\mathcal{G}_0$ and $\mathcal{G}_1$, written $d_1(\mathcal{G}_0, \mathcal{G}_1)$, to be*

$$\sum_{E\in\mathcal{L}} \int_{[0,1]^{\mathrm{arity}(E)}} \left| \widetilde{E}^{\mathcal{G}_0}(\mathbf{x}) - \widetilde{E}^{\mathcal{G}_1}(\mathbf{x}) \right| \mathrm{d}\mathbf{x},$$

where $\mathbf{x}$ is a tuple of variables of length $\mathrm{arity}(E)$.

We now consider finite structures, and their relationship to Borel structures via Borel blow-ups. All finite structures in this section will have underlying set an initial segment of $\mathbb{N}$.

Every finite $\mathcal{L}$-structure with counting measure induces a Borel $\mathcal{L}$-structure, by taking its *Borel blow-up.* For each k such that $0 \le k < r-1$, define $\iota_r(k) := [\frac{k}{r}, \frac{k+1}{r})$ and $\iota_r(r-1) := [\frac{r-1}{r}, 1]$.

Definition 6.4. *Suppose $\mathcal{G}$ is a finite $\mathcal{L}$-structure. Define its* **Borel blow-up**, *$\overline{\mathcal{G}}$, to be the Borel $\mathcal{L}$-structure such that for all $E \in \mathcal{L}$ and $i < \mathrm{arity}(E)$, whenever $x_i \in \iota_r(k_i)$ for all $k_i < |G|$ we have*

$$\overline{\mathcal{G}} \models E(x_0, \ldots, x_{\mathrm{arity}(E)-1}) \quad \textit{if and only if} \quad \mathcal{G} \models E(k_0, \ldots, k_{\mathrm{arity}(E)-1}).$$

Observe that the Borel blow-up of a finite structure is a particular kind of blow-up, in the sense of Definition 1.4.

By a standard argument, every Borel $\mathcal{L}$-structure is close in L^1 to the Borel blow-up of some finite $\mathcal{L}$-structure.

Lemma 6.5. *Let $\mathcal{G}$ be a Borel $\mathcal{L}$-structure. For all $\varepsilon > 0$ and all $n_0 \in \mathbb{N}$, there is an $n > n_0$ and an $\mathcal{L}$-structure $\mathcal{H}$ with underlying set n such that $d_1(\mathcal{G}, \overline{\mathcal{H}}) < \varepsilon$.*

Proof. There is some $\ell \in \mathbb{N}$ such that for every $E \in \mathcal{L}$, some set $S_{E,\varepsilon} \subseteq [0,1]^{\mathrm{arity}(E)}$ that is a finite union of sets of the form $\prod_{s<\mathrm{arity}(E)} \iota_\ell(k_s)$ satisfies $\lambda(E \triangle S_{E,\varepsilon}) < \varepsilon/|\mathcal{L}|$.

Let $\mathcal{H}$ be the $\mathcal{L}$-structure with underlying set $\{0,\ldots,\ell-1\}$ satisfying

$$\mathcal{H} \models E(k_0,\ldots,k_{\mathrm{arity}(E)-1}) \quad \text{if and only if} \quad \textstyle\prod_{s<\mathrm{arity}(E)} \iota_\ell(k_s) \subseteq S_{E,\varepsilon}.$$

for $E \in \mathcal{L}$ and $k_0,\ldots,k_{\mathrm{arity}(E)-1} < \ell$. By construction of $\mathcal{H}$, by summing over all relation symbols $E \in \mathcal{L}$, we have $d_1(\mathcal{G},\overline{\mathcal{H}}) < \varepsilon$. □

For finite structures of the same size (hence on the same underlying set, an initial segment of $\mathbb{N}$) with a single relation, their normalized edit distance is the same as their L^1-distance. This fact follows immediately from Definitions 6.3 and 6.4 of L^1-distance and Borel blow-up.

Lemma 6.6. *Suppose $\mathcal{G}$ and $\mathcal{G}^*$ are finite $\mathcal{L}$-structures on the same underlying set G. Then*

$$d_1(\overline{\mathcal{G}},\overline{\mathcal{G}^*}) = \sum_{E\in\mathcal{L}} \frac{|E^{\mathcal{G}} \triangle E^{\mathcal{G}^*}|}{|G^{\mathrm{arity}(E)}|}.$$

We will later need finite blow-ups to make a structure large enough so as to apply the results of earlier sections. A finite blow-up can also be seen as an instance of Definition 1.4.

Definition 6.7. *Let $\mathcal{G}$ be a finite $\mathcal{L}$-structure and let $p \in \mathbb{N}$ be positive. The* p**-fold blow-up** *of $\mathcal{G}$ is defined to be the structure $\mathcal{G}_p$ of size $p \cdot |G|$ such that for each relation $E \in \mathcal{L}$ and $x_0,\ldots,x_{\mathrm{arity}(E)-1} \in G_p$*

$$\mathcal{G}_p \models E(x_0,\ldots,x_{\mathrm{arity}(E)-1}) \quad \textit{if and only if} \quad \mathcal{G} \models E(\lfloor \tfrac{x_0}{p} \rfloor,\ldots,\lfloor \tfrac{x_{\mathrm{arity}(E)-1}}{p} \rfloor).$$

We call $\mathcal{G}_p$ a **finite blow-up** *of $\mathcal{G}$.*

It is immediate that replacing a finite structure by a finite blow-up does not change its Borel blow-up.

Lemma 6.8. *Suppose $\mathcal{G}_p$ is the p-fold blow-up of a finite $\mathcal{L}$-structure $\mathcal{G}$. Then $\overline{\mathcal{G}_p} = \overline{\mathcal{G}}$.*

We may define induced homomorphism densities for Borel $\mathcal{L}$-structures, similarly to Definition 5.1. For more details on an analogous notion for graphons, see [Lov12, §7.2].

Definition 6.9. *Suppose $\mathcal{F}$ is a finite $\mathcal{L}$-structure with underlying set $\{0,\ldots,|\mathcal{F}|-1\}$ and $\mathcal{G}$ is a Borel $\mathcal{L}$-structure. We define the* **induced homomorphism density** *of $\mathcal{F}$ in $\mathcal{G}$ to be*

$$t_{\mathrm{ind}}(\mathcal{F},\mathcal{G}) := \int_{I(\mathcal{F},\mathcal{G})} \mathrm{d}\mathbf{x}$$

where $I(\mathcal{F},\mathcal{G})$ is the set of embeddings from $\mathcal{F}$ to $\mathcal{G}$, considered as a Borel subset of $[0,1]^{|F|}$.

The following lemma is immediate.

Lemma 6.10. *Let $\mathcal{F}$ and $\mathcal{G}$ be finite $\mathcal{L}$-structures. Then*

$$t_{\text{ind}}(\mathcal{F}, \mathcal{G}) \leq t_{\text{ind}}(\mathcal{F}, \overline{\mathcal{G}}).$$

In the case of Borel structures, we only ever care about a structure up to measure-zero sets. However, any stable Borel structure can be modified on a set of measure 0 to make it unstable, and so we need to consider a weaker notion of stability for Borel structures. We use Lemma 6.10 to extend the definition of the $(\delta, \widehat{\tau})$-branching property to Borel $\mathcal{L}$-structures.

Definition 6.11. *Let $\widehat{\tau} \in \mathbb{N}$ and $\delta > 0$. A Borel $\mathcal{L}$-structure $\mathcal{G}$ has the $(\delta, \widehat{\tau})$-branching property for a quantifier-free formula $\varphi(\overline{x}; \overline{y})$ if there is a structure $\mathcal{N}$ which minimally has the $\widehat{\tau}$-branching property for $\varphi(\overline{x}; \overline{y})$ and for which $t_{\text{ind}}(\mathcal{N}, \mathcal{G}) \geq \delta$.*

We can obtain a bound on the differences of induced homomorphism densities obtained from a bound on the L^1-distances of two structures.

Lemma 6.12. *Let $\mathcal{F}$ be a finite $\mathcal{L}$-structure and let $\mathcal{G}$ and $\mathcal{H}$ be Borel $\mathcal{L}$-structures. If $d_1(\mathcal{G}, \mathcal{H}) \leq \varepsilon$ then*

$$\left| t_{\text{ind}}(\mathcal{F}, \mathcal{G}) - t_{\text{ind}}(\mathcal{F}, \mathcal{H}) \right| \leq |\mathcal{L}||F|^{q_{\mathcal{L}}} \cdot \varepsilon.$$

Proof. Let $\mathbf{1}_X$ denote the indicator function of a set X. Observe that

$$\int_{[0,1]^{|F|}} \left| \mathbf{1}_{I(\mathcal{F},\mathcal{G})}(\mathbf{x}) - \mathbf{1}_{I(\mathcal{F},\mathcal{H})}(\mathbf{x}) \right| \mathrm{d}\mathbf{x} \leq \sum_{E \in \mathcal{L}} |F|^{\text{arity}(E)} \cdot \varepsilon,$$

where $|\mathbf{x}| = |F|$. But $q_{\mathcal{L}}$ is the maximum arity of a relation symbol in $\mathcal{L}$, and so $\sum_{E \in \mathcal{L}} |F|^{\text{arity}(E)} \leq |\mathcal{L}||F|^{q_{\mathcal{L}}}$, as desired. □

Definition 6.13. *A partition of $[0,1]$ is* **Borel** *if it is a countable partition each element of which is Borel. A Borel partition is* **equitable** *if every element has the same Lebesgue measure.*

A Borel $\mathcal{L}$-structure with an equitable finite partition can be thought of as a Borel blow-up of a finite structure (up to measure-preserving isomorphism).

Definition 6.14. *Suppose $\mathcal{G}$ is a Borel $\mathcal{L}$-structure. A Borel partition P of $[0,1]$ is* **indivisible** *with respect to $\mathcal{G}$ if for all relations $E \in \mathcal{L}$, for all $p_0, \ldots, p_{\text{arity}(E)-1} \in P$, and for any pair of sequences $\langle a_i^0 \rangle_{i<\text{arity}(E)}, \langle a_i^1 \rangle_{i<\text{arity}(E)}$ such that $a_i^0, a_i^1 \in p_i$ for $i < \text{arity}(E)$, we have*

$$\widetilde{E}^{\mathcal{G}}(a_0^0, \ldots, a_{\text{arity}(E)-1}^0) = \widetilde{E}^{\mathcal{G}}(a_0^1, \ldots, a_{\text{arity}(E)-1}^1).$$

Whereas in equitable partitions of finite structures, the size of the parts can differ by up to 1 (when the partition size does not divide the structure size), in the Borel case the Lebesgue measure of any two parts must be be equal. The following lemma relates these two notions.

Lemma 6.15. *Let $\mathcal{G}$ be a finite $\mathcal{L}$-structure. Suppose P is an indivisible partition of $\mathcal{G}$. Then there is a Borel $\mathcal{L}$-structure $\mathcal{G}^+$ and an equitable partition P^+ of $[0,1]$ such that*

- *P^+ is indivisible with respect to $\mathcal{G}^+$ and*
- *$d_1(\overline{\mathcal{G}}, \mathcal{G}^+) \leq \sum_{E\in\mathcal{L}} \frac{|P|-1}{|G|}$.*

Proof. Let $r := \min\{|p| : p \in P\}$. Let A contain exactly r elements from each $p \in P$. Note that $|G \setminus A| \leq |P| - 1$ as P is equitable. For each $p \in P$ let $p^* := \bigcup_{a\in p\cap A} \iota_{|G|}(a)$.

Let S be a partition of $[0,1] - \bigcup_{p\in P} p^*$ into $|P|$-many pieces $\langle s_p\rangle_{p\in P}$ of equal Lebesgue measure. For each $p \in P$, let $p^+ := p^* \cup s_p$. Define $P^+ := \{p^+ : p \in P\}$. It is then immediate that P^+ is an equitable partition.

For the remainder of this proof, consider $E \in \mathcal{L}$, and let $\ell := \mathrm{arity}(E)$; the result will follow by summing over all relation symbols in $\mathcal{L}$. For every $p \in P$ choose $x_p \in p$. For every $p_0^+, \ldots, p_{\ell-1}^+ \in P^+$, and for every $y_0, \ldots, y_{\ell-1} \in [0,1]$ such that $y_i \in p_i^+$ for all $i < \ell$, let

$$\mathcal{G}^+ \models E(y_0, \ldots, y_{\ell-1}) \quad \text{if and only if} \quad \mathcal{G} \models E(x_{p_0}, \ldots, x_{p_{\ell-1}}).$$

Note that P^+ is indivisible with respect to $\mathcal{G}^+$.

Because P was indivisible with respect to $\mathcal{G}$, the definition of $\mathcal{G}^+$ does not depend on the choice of the elements x_p. In particular this means $\widetilde{E}^{\mathcal{G}^+}|_{\iota_{|G|}(A)^\ell} = E^{\overline{\mathcal{G}}}|_{\iota_{|G|}(A)^\ell}$. Finally, we have

$$\lambda^\ell([0,1]^\ell \setminus \iota_{|G|}(A)^\ell) \;\leq\; \lambda([0,1] \setminus \iota_{|G|}(A)) \;\leq\; \frac{|P|-1}{|G|},$$

which completes the argument for this particular $E \in \mathcal{L}$. □

Theorem 6.16 (Almost stable regularity for Borel structures)**.** *Suppose $\varepsilon > 0$. There is a $\delta > 0$ such that whenever*

(a) *$\mathcal{G}$ is a Borel $\mathcal{L}$-structure that does not have the $(\delta, \widehat{\tau})$-branching property and*
(b) *$g = \lceil 5 \cdot n_{\mathcal{L}} \cdot \widehat{\tau} \cdot \log \widehat{\tau}\rceil$,*

there is a Borel $\mathcal{G}^+$ and an equitable partition P of $\mathcal{G}^+$ such that

(i) *$d_1(\mathcal{G}, \mathcal{G}^+) \leq \varepsilon$,*
(ii) *P is indivisible with respect to $\mathcal{G}^+$, and*
(iii) *$|P| \leq (\frac{\varepsilon}{6q_{\mathcal{L}}|\mathcal{L}|})^{-(g+1)(g+2)}$.*

Proof. Let $\varepsilon_1 > 0$, and let δ be as determined by Theorem 5.6 (with ε_1 as its ε).

Suppose $\mathcal{G}$ satisfies condition (a). Then there must be some $\delta_0 < \delta$ such that $\mathcal{G}$ also satisfies condition (a) with respect to δ_0. Let ε_0 be such that $\delta_0 + |\mathcal{L}| \cdot (2^{\widehat{\tau}} \cdot q_{\mathcal{L}})^{q_{\mathcal{L}}} \cdot \varepsilon_0 < \delta$ and $\varepsilon_0 < \varepsilon/3$. By Lemma 6.5 (with ε_0 as its ε) we can find a finite $\mathcal{H}$ such that $d_1(\overline{\mathcal{H}}, \mathcal{G}) < \varepsilon_0$.

Suppose, towards a contradiction, that $\mathcal{H}$ has the $(\delta, \widehat{\tau})$-branching property. Then there is some finite $\mathcal{M}$ that minimally has the $\widehat{\tau}$-branching property such that $t_{\mathrm{ind}}(\mathcal{M}, \mathcal{H}) \geq \delta$. By Lemma 6.10, we then have $t_{\mathrm{ind}}(\mathcal{M}, \overline{\mathcal{H}}) \geq \delta$. We also have $|\mathcal{M}| \leq 2^{\widehat{\tau}} q_{\mathcal{L}}$ by Lemma 5.3. Then by Lemma 6.12, we know that

$$\left|t_{\mathrm{ind}}(\mathcal{M}, \overline{\mathcal{H}}) - t_{\mathrm{ind}}(\mathcal{M}, \mathcal{G})\right| \leq |\mathcal{L}||\mathcal{M}|^{q_{\mathcal{L}}} \cdot \varepsilon_0 \leq |\mathcal{L}|(2^{\widehat{\tau}} q_{\mathcal{L}})^{q_{\mathcal{L}}} \cdot \varepsilon_0$$

which implies that

$$t_{\mathrm{ind}}(\mathcal{M}, \mathcal{G}) \geq \delta - |\mathcal{L}|(2^{\widehat{\tau}} q_{\mathcal{L}})^{q_{\mathcal{L}}} \cdot \varepsilon_0 > \delta_0.$$

Hence $\mathcal{G}$ has the $(\delta_0, \widehat{\tau})$-branching property, contradicting our choice of δ_0. Therefore $\mathcal{H}$ must not have the $(\delta, \widehat{\tau})$-branching property.

By Lemma 6.8, we may replace $\mathcal{H}$ by a finite blow-up so that $|H|$ is large enough to apply Theorem 5.6 (with $\frac{\varepsilon_1}{2}$ as its ε). We thereby obtain an $\mathcal{L}$-structure $\mathcal{H}^*$ with the same underlying set as $\mathcal{H}$ and an equitable partition P_H such that

- P_H is indivisible with respect to $\mathcal{H}^*$,
- $|P_H| \leq (\frac{\varepsilon_1}{2})^{-(g+1)(g+2)}$, and
- $|E^{\mathcal{H}} \triangle E^{\mathcal{H}^*}| \leq \mathrm{arity}(E) \cdot |H|^{\mathrm{arity}(E)} \cdot (\frac{\varepsilon_1}{2})$ for all $E \in \mathcal{L}$.

But then by Lemma 6.6 we have

$$d_1(\overline{\mathcal{H}}, \overline{\mathcal{H}^*}) \leq \sum_{E \in \mathcal{L}} \mathrm{arity}(E) \cdot \varepsilon_1 \leq |\mathcal{L}| \cdot q_{\mathcal{L}} \cdot \varepsilon_1$$

and so $d_1(\overline{\mathcal{H}^*}, \mathcal{G}) \leq |\mathcal{L}| \cdot q_{\mathcal{L}} \cdot \varepsilon_1 + \varepsilon_0$.

We may similarly replace $\mathcal{H}^*$ by a finite blow-up so as to apply Lemma 6.15 to find a Borel $\mathcal{L}$-structure G^+ and an equitable partition P such that

- $|P| = |P_H|$,
- P is indivisible for E^+, and
- $d_1(\mathcal{G}^+, \overline{\mathcal{H}^*}) \leq |\mathcal{L}| \cdot \frac{|P|-1}{k} \leq \varepsilon_1$.

So we have $d_1(\mathcal{G}^+, \mathcal{G}) \leq \varepsilon_1 + |\mathcal{L}| \cdot q_{\mathcal{L}} \cdot \varepsilon_1 + \varepsilon_0$. Hence for $\varepsilon_1 := \frac{\varepsilon}{3q_{\mathcal{L}}|\mathcal{L}|}$, we have $d_1(\mathcal{G}^+, \mathcal{G}) \leq \varepsilon$. Further, as $|P| = |P_H|$ we have $|P| \leq (\frac{\varepsilon_1}{2})^{-(g+1)(g+2)} = (\frac{\varepsilon}{6q_{\mathcal{L}}|\mathcal{L}|})^{-(g+1)(g+2)}$. □

Acknowledgements

The authors would like to thank M. Malliaris for helpful conversations.

The main results of this paper were presented at the North American Annual Meeting of the Association for Symbolic Logic [ASL17] at the University of Connecticut on May 26, 2016, and at the Workshop on Model Theory of Finite and Pseudofinite Structures at the University of Leeds on July 29, 2016.

This research was facilitated by the the Trimester Program on Universality and Homogeneity of the Hausdorff Research Institute for Mathematics at the University of Bonn (2013), the Program on Model Theory, Arithmetic Geometry

http://www.maths.leeds.ac.uk/fps/programme.html

and Number Theory of the Mathematical Sciences Research Institute (2014), and the Lorentz Center workshop on Logic and Random Graphs (2015).

Work on this publication by C. F. was made possible through the support of ARO grant W911NF-13-1-0212 and a grant from Google.

References

[AC14] A. Aroskar and J. Cummings, *Limits, regularity and removal for finite structures*, To appear in J. Symb. Logic. ArXiv e-print 1412.8084 (2014).

[ADL+94] N. Alon, R. A. Duke, H. Lefmann, V. Rödl, and R. Yuster, *The algorithmic aspects of the regularity lemma*, J. Algorithms **16** (1994), no. 1, 80–109.

[ASL17] *2016 North American Annual Meeting of the Association for Symbolic Logic: University of Connecticut Storrs, CT, USA May 23–26, 2016*, Bull. Symb. Log. **23** (2017), no. 3, 345–373.

[CPT17] G. Conant, A. Pillay, and C. Terry, *A group version of stable regularity*, ArXiv e-print 1710.06309 (2017).

[CS15] A. Chernikov and S. Starchenko, *Regularity lemma for distal structures*, To appear in J. Eur. Math. Soc. ArXiv e-print 1507.01482 (2015).

[CS16] ———, *Definable regularity lemmas for NIP hypergraphs*, ArXiv e-print 1607.07701 (2016).

[ES12] G. Elek and B. Szegedy, *A measure-theoretic approach to the theory of dense hypergraphs*, Adv. Math. **231** (2012), no. 3-4, 1731–1772.

[Gow97] W. T. Gowers, *Lower bounds of tower type for Szemerédi's uniformity lemma*, Geom. Funct. Anal. **7** (1997), no. 2, 322–337.

[Gow07] ———, *Hypergraph regularity and the multidimensional Szemerédi theorem*, Ann. of Math. (2) **166** (2007), no. 3, 897–946.

[Gre05] B. Green, *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geom. Funct. Anal. **15** (2005), no. 2, 340–376.

[Hod93] W. Hodges, *Model theory*, Encyclopedia of Mathematics and its Applications, vol. 42, Cambridge University Press, Cambridge, 1993.

[Kru16] A. Kruckman, *Infinitary limits of finite structures*, Ph.D. thesis, University of California, Berkeley, 2016.

[Lov12] L. Lovász, *Large networks and graph limits*, American Mathematical Society Colloquium Publications, vol. 60, American Mathematical Society, Providence, RI, 2012.

[LS07] L. Lovász and B. Szegedy, *Szemerédi's lemma for the analyst*, Geom. Funct. Anal. **17** (2007), no. 1, 252–270.

[LS10] ———, *Regularity partitions and the topology of graphons*, An irregular mind, Bolyai Soc. Math. Stud., vol. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 415–446.

[MP16] M. Malliaris and A. Pillay, *The stable regularity lemma revisited*, Proc. Amer. Math. Soc. **144** (2016), no. 4, 1761–1765.

[MS14] M. Malliaris and S. Shelah, *Regularity lemmas for stable graphs*, Trans. Amer. Math. Soc. **366** (2014), no. 3, 1551–1585.

[PST12] J. Pach, J. Solymosi, and G. Tardos, *Remarks on a Ramsey theory for trees*, Combinatorica **32** (2012), no. 4, 473–482.

[RS07] V. Rödl and M. Schacht, *Regular partitions of hypergraphs: regularity lemmas*, Combin. Probab. Comput. **16** (2007), no. 6, 833–885.

[RS10] ———, *Regularity lemmas for graphs*, Fete of combinatorics and computer science, Bolyai Soc. Math. Stud., vol. 20, János Bolyai Math. Soc., Budapest, 2010, pp. 287–325.

[She90] S. Shelah, *Classification theory and the number of nonisomorphic models*, second ed., Studies in Logic and the Foundations of Mathematics, vol. 92, North-Holland Publishing Co., Amsterdam, 1990.

[Ska13] M. Skala, *Hypergeometric tail inequalities: ending the insanity*, ArXiv e-print 1311.5939 (2013).

[Tao06] T. Tao, *Szemerédi's regularity lemma revisited*, Contrib. Discrete Math. **1** (2006), no. 1, 8–28.

[TW17] C. Terry and J. Wolf, *Stable arithmetic regularity in the finite-field model*, ArXiv e-print 1710.02021 (2017).

Harvard University, Cambridge, MA 02138, USA
E-mail address: `nate@math.harvard.edu`

Remine, Fairfax, VA 22031, USA
E-mail address: `cameron@remine.com`

Wheaton College, Norton, MA 02766, USA
E-mail address: `rrpatel@cantab.net`