

DEFINABLE REGULARITY LEMMAS FOR NIP HYPERGRAPHS

by ARTEM CHERNIKOV[†]

(Department of Mathematics, University of California Los Angeles, 520 Portola Plaza, Los Angeles, CA 90095-1555, USA)

and SERGEI STARCHENKO[‡]

(Department of Mathematics, University of Notre Dame, 255 Hurley, Notre Dame, IN 46556, USA)

[Received 27 September 2018]

Abstract

We present a systematic study of the regularity phenomena for NIP hypergraphs and connections to the theory of (locally) generically stable measures, providing a model-theoretic hypergraph version of the results of Alon-Fischer-Newman and Lovász-Szegedy for graphs of bounded VC-dimension. We also consider the two extremal cases of regularity for stable and distal hypergraphs, improving and generalizing the corresponding results for graphs in the literature. Finally, we consider a related question of the existence of large (approximately) homogeneous definable subsets of NIP hypergraphs and provide some positive results and counterexamples, in particular for graphs definable in the p -adics.

1. Introduction

Szemerédi’s regularity lemma is a fundamental result in (hyper-)graph combinatorics with numerous applications in extremal combinatorics, number theory and computer science (see [29] for a survey). We recall it in a simplified form. By a graph $G = (V, E)$ we mean a set V with a symmetric subset $E \subseteq V^2$. For $A, B \subseteq V$, we denote by $E(A, B)$ the set of edges between A and B , that is, $E(A, B) = E \cap (A \times B)$. Given $M \in \mathbb{N}$, we write $[M]$ to denote the set $\{1, 2, \dots, M\}$.

FACT 1.1 (Szemerédi regularity lemma) *For every real $\varepsilon > 0$ there exists some constant $M = M(\varepsilon) \in \mathbb{N}$ satisfying the following. Let $G = (V, E)$ be an arbitrary finite graph. Then there is a partition $V = V_1 \cup \dots \cup V_M$ into disjoint sets, real numbers δ_{ij} , $i, j \in [M]$, and an exceptional set of pairs $\Sigma \subseteq [M] \times [M]$ such that*

$$\sum_{(i,j) \in \Sigma} |V_i| |V_j| \leq \varepsilon |V|^2$$

and for each $(i, j) \in [M] \times [M] \setminus \Sigma$ we have

$$||E(A, B)| - \delta_{ij}|A||B|| < \varepsilon |V_i| |V_j|$$

[†]Corresponding author. E-mail: chernikov@math.ucla.edu

[‡]E-mail: Starchenko.1@nd.edu

for all $A \subseteq V_i$, $B \subseteq V_j$.

The bounds on $M(\varepsilon)$ are known to be extremely bad: Gowers had demonstrated that it grows as an exponential tower of height polynomial in $(\frac{1}{\varepsilon})$ (see, for example [36]).

Several recent results demonstrate that better bounds and stronger regularity can be obtained for certain families of hypergraphs satisfying additional combinatorial restrictions. For example, in [15, 16] it is shown that when the edge relation is semialgebraic, of bounded description complexity, then the size of the partition can be bounded by a polynomial in terms of $\frac{1}{\varepsilon}$, all good pairs are actually homogeneous, and the sets in the partition can be chosen to be semialgebraic, of bounded complexity. Similar polynomial bounds were obtained by Tao [52] for algebraic hypergraphs of bounded description complexity in large finite fields and by Alon, Fischer, Newman [3] and Lovász, Szegedy [31] for graphs of bounded VC-dimension.

These results can be naturally viewed as results about hypergraphs with the edge relation definable, in the sense of first-order logic, in certain tame structures, and the restrictions on the complexity of the edge relation in all of the results above are surprisingly well aligned with generalized stability and classification in model theory. For example, as demonstrated in [10] (see also [48]), the results in [15, 16] can be generalized to graphs definable in arbitrary *distal* structures (see Section 4.2), and that moreover this strong form of regularity characterizes distality. Here ‘semialgebraic graphs’ correspond to the special case of ‘graphs definable in the field of reals’, but the result also applies to graphs definable in the p -adics, for example. Similarly, the result in [52] can be viewed as a result about graphs definable in *pseudofinite fields*, and admits a natural model-theoretic proof and generalizations [20, 24, 39]. Another very important example is given by the regularity lemma for *stable graphs* [33] (model-theoretic stability is the notion of tameness at the core of Shelah’s classification [44], see Section 4.1). Similarly, the results in [31] can be interpreted as results about graphs definable in *NIP structures* (see below).

Another point of view on the hypergraph regularity phenomenon is through the prism of probability theory. Namely, the existence of a regular partition can be viewed as a finitary version of the existence of the conditional expectation. There are several proofs of the hypergraph regularity lemma in the literature making this precise by reducing working with a family of finite graphs to working with some kind of an analytic ‘limit object’ equipped with a probability measure (see [14, 30, 51]).

Similarly, regularity for restricted families of graphs can be viewed as the study of (finitely additive) probability measures on certain restricted families of Boolean algebras. Such measures in the model-theoretic setting of Boolean algebras of definable sets were introduced by Keisler [27], and recently the study of Keisler measures has attracted a lot of attention, especially the study of *generically stable measures* in *NIP structures* [8, 25, 26, 50]. The class of NIP structures was introduced by Shelah in his work on the classification program [44]. It contains all stable and $\mathcal{o}$ -minimal structures, along with other important algebraic examples, and we refer to [2, 47] for an introduction to the area (see also Section 3.3 for the definition and some examples). The study of Keisler measures in NIP structures can be viewed as a model-theoretic counterpart of the Vapnik–Chervonenkis theory [54], and generically stable measure are those Keisler measures that satisfy a form of the VC-theorem for all uniformly definable families (see Section 3.3).

The connection between the study of generically stable measures in model theory and regularity lemmas for definable hypergraphs was pointed out in the distal case in [10], and the aim of this article is to systematically develop these connections for the general (local) NIP setting.

In Section 2, we give a decomposition result for products of finitely additive probability measures that are well-approximated by counting measures (which we call *finitely approximated measures*, see Section 2.4), with and without the assumption of finite VC-dimension. Namely, assume we are given some sets $V_1, \dots, V_k$ equipped with Boolean algebras $\mathcal{B}_1, \dots, \mathcal{B}_k$ of subsets and finitely additive probability measures $\mu_1, \dots, \mu_k$ on them. Let $R \subseteq V_1 \times \dots \times V_k$ be an edge relation such that all of its fibers are measurable. It then follows from the finite approximation assumption that there is a Boolean algebra $\mathcal{B}$ of subsets of $V_1 \times \dots \times V_k$ extending the product Boolean algebra $\mathcal{B}_1 \otimes \dots \otimes \mathcal{B}_k$ with $R \in \mathcal{B}$, and such that $\mathcal{B}$ can be equipped with a natural product measure μ satisfying a Fubini property (Section 2.4). Moreover, relatively to μ , the set R can be approximated by a union of boxes (that is, sets of the form $A_1 \times \dots \times A_k$ with $A_i \in \mathcal{B}_i$) up to measure ε , for any real $\varepsilon > 0$, and in the finite VC-dimension case the number of boxes needed is polynomial in $\frac{1}{\varepsilon}$ (Theorem 2.19). On the one hand, this can be viewed as a version of the results for graphons from [31] in a setting better suited for the model-theoretic applications, and generalized to hypergraphs. On the other hand, this result can also be viewed as developing elements of the *local* theory of generically stable measures, and refining some of the results in [26] for such measures. In our setting, instead of working with Borel measures on the space of types, we use directly the (equivalent) theory of integration for finitely additive measures (sometimes called the theory of charges [7]), and we give some details for the sake of exposition. Note that we are only assuming bounded VC-dimension on R -definable sets, and our definition of a finitely approximated measure is weaker than the definition of fin measures in [26] (see Remark 3.7), so we have to redefine the product of finitely approximated measures.

In Section 3, we apply these results to obtain a *definable* regularity lemma for hypergraphs of bounded VC-dimension, in particular for hypergraphs definable in an NIP structure, uniformly over all generically stable measures. For a general k -ary hypergraph (V, E) with $E \subseteq \binom{V}{k}$ for some $k \geq 2$, with V a large finite set, the hypergraph regularity lemma [21, 37, 42] allows to represent the characteristic function $\chi_E : \binom{V}{k} \rightarrow \{0, 1\}$ of the hyperedge relation in the form

$$\chi_E = f_{k-1} + \dots + f_1 + f^\perp$$

where: f_1 has the form $\sum_{i_1, \dots, i_k} \alpha_{i_1, \dots, i_k} \prod_j \chi_{V_{i_j}}(x_j)$ for some partition $V = \bigcup_{i \leq n} V_i$ and some real numbers $\alpha_{i_1, \dots, i_k}$ (that is, a partition of the vertices with weights $\alpha_{i_1, \dots, i_k}$ indicating the density of edges on $V_{i_1} \times \dots \times V_{i_k}$); the f_j in general are sums of j -ary cylinder sets (for instance, f_2 is, roughly speaking, the portion of χ_E which can be described using directed graphs, f_3 using directed 3-hypergraphs, etc. up to $k-1$); and $f^\perp$ is the quasi-random k -ary function (representing the random determination of which hyperedges of E are actually present). In Theorem 3.3, we show that if E has small VC-dimension then all summands of this decomposition except for f_1 are small, hence E is approximated by a union of boxes. More precisely, for each k, d and $\varepsilon > 0$, there is a bound $N = O_{k,d}((\frac{1}{\varepsilon})^{4(k-1)d^2})$ so that whenever (E, V) is a k -ary hypergraph with VC-dimension at most d , there is a partition of V into N parts so that E is given, up to symmetric difference of measure $< \varepsilon$, by a union of boxes of the form $V_{i_1} \times \dots \times V_{i_k}$. Moreover, each of the sets V_i in the partition is E -definable, that is, given by a Boolean combination of the fibers of E of size bounded in terms of d, k, ε .

In Section 4, we discuss regularity in two extreme opposite special cases of the NIP hypergraphs. Namely, we generalize and improve the aforementioned stable [33, 34] and distal [10] regularity lemmas in our setting. The (global) model-theoretic implications of these results can be summarized as follows.

THEOREM 1.2

(1) (Corollary 3.8) *Let $\mathcal{M}$ be an NIP structure and $k \geq 2$. For every definable relation $E(x_1, \dots, x_k)$ there is some $c = c(E)$ such that for any $\varepsilon > 0$ and any generically stable Keisler measures μ_i on $M^{|x_i|}$ there are partitions $M^{|x_i|} = \bigcup_{j < K} A_{i,j}$ and a set $\Sigma \subseteq \{1, \dots, K\}^k$ such that:*

- (a) $K \leq \left(\frac{1}{\varepsilon}\right)^c$.
- (b) $\mu\left(\bigcup_{(i_1, \dots, i_k) \in \Sigma} A_{1,i_1} \times \dots \times A_{k,i_k}\right) \leq \varepsilon$, where $\mu = \mu_1 \otimes \dots \otimes \mu_k$,
- (c) for all $\vec{i} = (i_1, \dots, i_k) \notin \Sigma$ we have

$$\left| \mu(E \cap (A_{1,i_1} \times \dots \times A_{k,i_k})) - \delta_{\vec{i}} \mu(A_{1,i_1} \times \dots \times A_{k,i_k}) \right| < \varepsilon \mu(A_{1,i_1} \times \dots \times A_{k,i_k})$$

for some $\delta_{\vec{i}} \in \{0, 1\}$.

- (d) each $A_{i,j}$ is defined by an instance of an E -formula depending only on E and ε .
- (2) (Corollary 4.14) *Assume that $\mathcal{M}$ is stable. Then, in addition:*
- (a) we can take the μ_i 's to be arbitrary Keisler measures (as all measures are automatically generically stable in this case),
 - (b) we may assume that $\Sigma = \emptyset$, that is, all tuples in the partition are ε -regular.
- (3) (Theorem 4.16) *Assume that $\mathcal{M}$ is distal. Then in addition we have:*
- (a) for all $(i_1, \dots, i_k) \notin \Sigma$, either $(A_{1,i_1} \times \dots \times A_{k,i_k}) \cap E = \emptyset$ or $A_{1,i_1} \times \dots \times A_{k,i_k} \subseteq E$,
 - (b) if the relation E is defined by an instance of a formula θ , then we can take each $A_{i,j}$ to be defined by an instance of a formula $\psi_i(x_i, z_i)$ which only depends on θ (and not on ε).

Finally, in Section 5, we consider a related question of the existence of large (approximately) homogeneous definable subsets of definable NIP hypergraphs (that is, the measure theoretic versions of the results of Erdős, Hajnal and Rödl, see, for example [19]). As a corollary of the regularity lemma, we show that for every d and $\alpha, \varepsilon > 0$ there is some $\delta = \delta(d, \alpha, \varepsilon) > 0$ such that the following holds. Let a hypergraph $R \subseteq V_1 \times \dots \times V_k$ of VC-dimension at most d be given, and let μ_i be measures on V_i which are all finitely approximated on R . Assume that the density of R on $V_1 \times \dots \times V_k$ (relatively to the product measure) is at least α . Then it is possible to find R -definable sets $A_i \subseteq V_i$ such that $\mu_i(A_i) \geq \delta$ and such that the density of R on $A_1 \times \dots \times A_k$ is at least $1 - \varepsilon$ (Theorem 5.1). The situation is quite different in the non-partitioned case. Namely, when $V = V_1 = \dots = V_k$, $\mu = \mu_1 = \dots = \mu_k$ and R is a symmetric relation, we would like to find a definable subset A of V of positive measure, such that the density of R on A is ε -close to 0 or 1 (the result above applied to this situation would typically produce *disjoint* sets $A_1, \dots, A_k$). A classical theorem of Rödl (see Fact 5.4) implies that this is indeed possible for pseudofinite counting measures, with all internal sets added to the language. We provide an example of a definable graph in the p -adics which does not admit uniformly definable sets of positive measure with this property, relatively to the additive Haar measure (Section 5.2.1) (hence demonstrating that an analogue of Rödl's theorem does not hold for finitely approximated measures in general).

1.1. Related work

The results of this paper originally appeared in a July 2016 preprint [9]. In a preprint [17] from October 2017 (which has later appeared as [18]), its authors obtain a stronger bound on the size of the partition for hypergraph regularity with finite VC-dimension in the case of uniform measures on finite

spaces (corresponding to Corollary 3.6) and demonstrate its optimality. Independently from our work in the present paper, in a preprint [1] its authors obtain a version of the stable case of Theorem 1.2 for finite hypergraphs (and more generally, finite relational structures). The NIP hypergraph regularity lemma is further generalized to n -dependent hypergraphs in [11].

Recall that a partition is *equitable* if any two sets in it have the same measure (possibly up to a rounding error). We remark that it is not always possible to choose an *E-definable* equipartition with the additional properties discussed above in the NIP, stable and distal cases. However, restricting to finite hypergraphs with uniform finitely supported measures and giving up definability, this can be achieved in each of the three cases. See [18] for the finite VC-dimension/NIP hypergraphs; [33] for stable graphs and [1] for stable hypergraphs and [10, Section 5.3] (along with [48, Section 3]) for distal hypergraphs (where in fact a definable equitable partition can be chosen in many cases).

1.2. Notation

Given $r, s, \delta \in \mathbb{R}$ with $\delta \geq 0$, we write $r \approx^\delta s$ if $|r - s| \leq \delta$.

2. Decomposing product measures

In this section, we present some general results on decomposing products of finitely additive probability measures that can be locally approximated by frequency measures.

2.1. Notation

We will use the following notation:

- For $k \in \mathbb{N}$, we will denote by $[k]$ the set $\{1, \dots, k\}$.
- For an integer k and $I \subseteq [k]$, we will denote by I^c the complement $I^c = [k] \setminus I$. For $i \in [k]$ instead of $\{i\}^c$ we write i^c .
- For sets $V_1, \dots, V_k$ and $I \subseteq [k]$, we denote by V_I the product $V_I = \prod_{i \in I} V_i$.
- Let $R \subseteq V_1 \times \dots \times V_k$ be a k -ary relation and $I \subseteq [k]$. Viewing R as a binary relation on $V_I \times V_{I^c}$, for $b \in V_{I^c}$ we denote by R_b the fiber

$$R_b = \{a \in V_I : (a, b) \in R\}.$$

DEFINITION 2.1 Let $V_1, \dots, V_k$ be sets, $R \subseteq V_1 \times \dots \times V_k$ and $I \subseteq [k]$.

- (1) We say that a subset $X \subseteq V_I$ is *R-definable over a set* $D \subseteq V_{I^c}$ if it is a finite Boolean combination of sets of the form R_b with $b \in D$, and say that X is *R-definable* if it is *R-definable over* V_{I^c} .
- (2) We say that a set $A \subseteq V_1 \times \dots \times V_k$ is *$R_\otimes$ -definable* if A can be written as a finite union of sets of the form $X_1 \times \dots \times X_k$, such that each $X_i \subseteq V_i$ is *R-definable*.
- (3) In addition, given a tuple $\vec{D} = (D_1, \dots, D_k)$ with $D_i \subseteq V_{I^c}$, we say that A is *$R_\otimes$ -definable over* $\vec{D}$ if every X_i above is *R-definable over* D_i . Note that the subsets of $V_1 \times \dots \times V_k$ which are *$R_\otimes$ -definable over* $\vec{D}$ form a Boolean algebra. For such a tuple $\vec{D}$, we define $\|\vec{D}\| := \max\{|D_i| : i \in [k]\}$.

We recall the notion of VC-dimension (see, for example [35, Chapter 10]). Let V be a set, finite or infinite, and let $\mathcal{F}$ be a family of subsets of V . Given $A \subseteq V$, we say that it is *shattered* by $\mathcal{F}$ if

for every $A' \subseteq A$ there is some $S \in \mathcal{F}$ such that $A \cap S = A'$. The *VC-dimension* of $\mathcal{F}$, that we will denote by $VC(\mathcal{F})$, is the smallest integer d such that no subset of V of size $d+1$ is shattered by $\mathcal{F}$. For a set $B \subseteq V$, let $\mathcal{F} \cap B = \{A \cap B : A \in \mathcal{F}\}$. The *shatter function* of $\mathcal{F}$ is defined as $\pi_{\mathcal{F}}(n) = \max \{|\mathcal{F} \cap B| : B \subseteq V, |B| = n\}$.

FACT 2.2 (Sauer–Shelah lemma) *If $VC(\mathcal{F}) \leq d$ then for $n \geq d$ we have $\pi_{\mathcal{F}}(n) \leq \sum_{i \leq d} \binom{n}{i} = O(n^d)$.*

DEFINITION 2.3 For sets $V_1, \dots, V_k$ and a set $R \subseteq V_1 \times \dots \times V_k$ we say that R has *VC-dimension at most d* if for every $i \in [k]$ the family $\{R_a : a \in V_{[k] \setminus \{i\}}\}$ of subsets of V_i is a family with VC-dimension at most d .

The next fact follows from the Sauer–Shelah lemma.

FACT 2.4 *For every $d \in \mathbb{N}$ there is a constant C_d such that for any relation $R \subseteq V \times W$ of VC-dimension at most d and any finite $D \subseteq V$, the number of atoms in the Boolean algebra of subsets of W which are R -definable over D is at most $C_d |D|^d$.*

2.2. Basics on Boolean algebras and measures

Recall that for a set V , a *field on V* is a Boolean algebra of subsets of V .

For sets $V_1, \dots, V_k$ and fields $\mathcal{B}_i$ on V_i , $i \in [k]$, as usual, we denote by $\mathcal{B}_1 \otimes \dots \otimes \mathcal{B}_k$ the field on $V_1 \times \dots \times V_k$ generated by the sets $X_1 \times \dots \times X_k$ with $X_i \in \mathcal{B}_i$. It is not hard to see that every set in $\mathcal{B}_1 \otimes \dots \otimes \mathcal{B}_k$ is a disjoint union of sets of the form $X_1 \times \dots \times X_k$ with $X_i \in \mathcal{B}_i$. Given $I = \{i_1, \dots, i_n\} \subseteq [k]$, we let $\mathcal{B}_I := \bigotimes_{i \in I} \mathcal{B}_i = \mathcal{B}_{i_1} \otimes \dots \otimes \mathcal{B}_{i_n}$.

Finitely additive probability measures

DEFINITION 2.5 Let V be a set and $\mathcal{B}$ be a field on V . In this paper, a *measure on $\mathcal{B}$* is a finitely additive probability measure on $\mathcal{B}$, that is, a function $\mu : \mathcal{B} \rightarrow \mathbb{R}^{\geq 0}$ such that $\mu(\emptyset) = 0$, $\mu(V) = 1$ and $\mu(A \cup B) = \mu(A) + \mu(B) - \mu(A \cap B)$ for all $A, B \in \mathcal{B}$.

Let $V_1, \dots, V_k$ be sets and $\mathcal{B}_i$ be fields on V_i , $i \in [k]$. Assume we have a measure μ_i on $\mathcal{B}_i$ for each $i \in [k]$. It is not hard to see that there is a unique measure μ on $\mathcal{B}_1 \otimes \dots \otimes \mathcal{B}_k$ with $\mu(A_1 \times \dots \times A_k) = \prod_{i=1}^k \mu_i(A_i)$ for all $A_i \in \mathcal{B}_i$, $i \in [k]$. We will denote this measure μ by $\mu_1 \times \dots \times \mu_k$.

Integration with respect to finitely additive measures

We will need some basic facts about integration relatively to finitely additive measures (we refer to [7] for a detailed account).

As usual for a set V and a subset $X \subseteq V$ we will denote by 1_X the indicator function of X on V .

We fix a set V and a field $\mathcal{B}$ on V .

We say that a function $f : V \rightarrow \mathbb{R}$ is *$\mathcal{B}$ -simple* if there are $X_1, \dots, X_n \in \mathcal{B}$ and $r_1, \dots, r_n \in \mathbb{R}$ with $f = \sum_{i=1}^n r_i 1_{X_i}$. Obviously, the set of all $\mathcal{B}$ -simple functions forms an $\mathbb{R}$ -algebra.

For a measure μ on $\mathcal{B}$ and a $\mathcal{B}$ -simple function $f = \sum_{i=1}^n r_i 1_{X_i}$ we define

$$\int_V f d\mu = \sum_{i=1}^n r_i \mu(X_i).$$

It is easy to see that the above integral does not depend on a representation of f as a simple function. If a subset $A \subseteq V$ is in $\mathcal{B}$ then we also define

$$\int_A f d\mu = \int_V (1_A f) d\mu = \sum_{i=1}^n r_i \mu(A \cap X_i).$$

REMARK 2.6 Clearly for $A \in \mathcal{B}$ we have $\mu(A) = \int_V 1_A d\mu$.

We say that a function $f: V \rightarrow \mathbb{R}$ is $\mathcal{B}$ -integrable, or just *integrable*, if it is in the closure of the set of $\mathcal{B}$ -simple functions with respect to the L_∞ -norm, that is, for all $\varepsilon > 0$ there is a $\mathcal{B}$ -simple function g with $|f(x) - g(x)| < \varepsilon$ for all $x \in V$. The following claim is obvious.

CLAIM 2.7 A function $f: V \rightarrow \mathbb{R}$ is $\mathcal{B}$ -integrable if and only if for any $\varepsilon > 0$ there are $Y_1, \dots, Y_n \in \mathcal{B}$ covering V such that for any $i \in [n]$ and any $c, c' \in Y_i$ we have $|f(c) - f(c')| < \varepsilon$.

If f is $\mathcal{B}$ -integrable and μ is a measure on $\mathcal{B}$ then the integral of f with respect to μ is defined as

$$\int_V f d\mu = \lim_{n \rightarrow \infty} \int_V g_n d\mu,$$

where $(g_n)_{n \in \mathbb{N}}$ is a sequence of $\mathcal{B}$ -simple functions convergent to f . It is easy to see that this integral does not depend on the choice of a convergent sequence. Also for a $\mathcal{B}$ -integrable function f and a set $A \in \mathcal{B}$ we define

$$\int_A f d\mu = \int_V (1_A f) d\mu.$$

2.3. On ε -nets

Let V be a set, $\mathcal{B}$ a field on V and μ a measure on $\mathcal{B}$. Let $\mathcal{F}$ be a family of subsets of V with $\mathcal{F} \subseteq \mathcal{B}$. As usual, for $\varepsilon > 0$ we say that a subset $T \subseteq V$ is an ε -net for $\mathcal{F}$ with respect to μ if for every $F \in \mathcal{F}$ we have $\mu(F) \geq \varepsilon \implies F \cap T \neq \emptyset$.

The following is a well-known consequence of the classical VC-theorem (see [54, 28] and also [31]).

FACT 2.8 Let V be a set, $\mathcal{B}$ a field on V and μ a measure on $\mathcal{B}$ with a finite support (that is, there exists a finite set $A \in \mathcal{B}$ with $\mu(A) = 1$). If $\mathcal{F} \subseteq \mathcal{B}$ is a VC-family with VC-dimension at most d then for any $\varepsilon > 0$ it admits an ε -net T with $|T| \leq 8d \frac{1}{\varepsilon} \log \frac{1}{\varepsilon}$.

2.4. Finitely approximated measures

Throughout this section we let V be a set and $\mathcal{B}_V$ a field on V .

DEFINITION 2.9 Let μ be a measure on $\mathcal{B}_V$, and let $\mathcal{F} \subseteq \mathcal{B}_V$ be a family of subsets of V . We say that μ is *finitely approximated* on the family $\mathcal{F}$ if for every $\varepsilon > 0$ there are $p_1, \dots, p_n \in V$ (possibly with

repetitions) with

$$|\mu(F) - \text{Av}(p_1, \dots, p_n; F)| < \varepsilon \text{ for every } F \in \mathcal{F},$$

where $\text{Av}(p_1, \dots, p_n; F) = \frac{1}{n} |\{i \in [n] : p_i \in F\}|$. We say that $p_1, \dots, p_n$ is an ε -approximation of μ on $\mathcal{F}$.

Notation

Let $\mathcal{F}$ be a family of subsets of V . For $m \in \mathbb{N}$, we let $\mathcal{F}^m$ be the family of all subsets of V given by the Boolean combinations of at most m sets from $\mathcal{F}$.

DEFINITION 2.10 Let now W be another set, and let $R \subseteq V \times W$ be a relation such that $R_b \in \mathcal{B}_V$ for all $b \in W$.

- (1) Let $\mathcal{R}_V := \{R_b : b \in W\}$, then $\mathcal{R}_V \subseteq \mathcal{B}_V$ by assumption.
- (2) We say that a measure μ on $\mathcal{B}_V$ is *finitely approximated on the relation R* if it is finitely approximated on each of the (infinitely many) families of sets $\mathcal{R}_V^m$, $m \in \mathbb{N}$.

REMARK 2.11

- (1) In particular, if μ is finitely approximated on the relation R , then it is finitely approximated on the family of sets $\mathcal{R}_V^\Delta = \{R_b \Delta R_{b'} : b, b' \in W\}$.
- (2) Note that μ being finitely approximated on $\mathcal{R}_V^\Delta$ does not imply that μ is finitely approximated on $\mathcal{R}_V$. For example, let $V = \mathbb{R}$, let $\mathcal{B}_V$ be the Boolean algebra generated by all intervals in V , and let $\mathcal{R}_V$ be the family of all intervals unbounded from above. Let μ be the 0–1 measure on $\mathcal{B}_V$ such that the measure of a set is 1 if and only if it is unbounded from above. Then all sets in $\mathcal{R}_V^\Delta$ have measure 0, so we can take the empty set as an ε -approximation for μ on $\mathcal{R}_V^\Delta$, for any $\varepsilon > 0$. But there are no finite ε -approximations for μ on $\mathcal{R}_V$, for any $\varepsilon < 1$, as any finite set can be avoided by some unbounded interval of measure 1. Similarly, if $\mathcal{R}_V$ is the family of all intervals bounded from above, then μ is trivially finitely approximated on $\mathcal{R}_V$. However, it is not finitely approximated on the family of all complements of the sets in $\mathcal{R}_V$.

See Example 3.11 for many examples of finitely approximated measures.

CLAIM 2.12 Let $R \subseteq V \times W$ be as in Definition 2.10, μ a measure on $\mathcal{B}_V$, and let $\mathcal{B}_W$ be a Boolean algebra on W such that $R_a \in \mathcal{B}_W$ for all $a \in V$.

Assume that μ is finitely approximated on the family $\mathcal{R}_V^\Delta$. Then for any set $A \in \mathcal{B}_V$, the function

$$h_{R,A} : W \rightarrow \mathbb{R}, \text{ given by } h_{R,A}(b) = \mu(R_b \cap A)$$

is $\mathcal{B}_W$ -integrable.

Proof. Let $\varepsilon > 0$. By assumption we can choose $p_1, \dots, p_n \in V$ such that

$$|\mu(R_b \Delta R_{b'}) - \text{Av}(p_1, \dots, p_n; R_b \Delta R_{b'})| < \varepsilon$$

for every $b, b' \in W$. For $I \subseteq [n]$ let $C_I \subseteq W$ be the set

$$C_I = \{b \in W : p_i \in R_b \Leftrightarrow i \in I\}.$$

Clearly each $C_I \in \mathcal{B}_W$, the sets $C_I, I \subseteq [n]$, cover W and for every $I \subseteq [n]$ and $b, b' \in C_I$ we have $\mu(R_b \Delta R_{b'}) < \varepsilon$. Hence, for any $b, b' \in C_I$ we have

$$|h_{R,A}(b) - h_{R,A}(b')| \leq \mu(A \cap (R_b \Delta R_{b'})) \leq \mu(R_b \Delta R_{b'}) < \varepsilon.$$

By Claim 2.7 the function $h_{R,A}$ is $\mathcal{B}_W$ -integrable. □

2.5. Products of finitely approximated measures

In this section, we let V, W, Z be sets, $R \subseteq V \times W \times Z$ a relation, and let $\mathcal{B}_V, \mathcal{B}_W$ be fields on V, W , respectively. We assume that $\mathcal{R}_V = \{R_{(b,c)} : (b,c) \in W \times Z\} \subseteq \mathcal{B}_V$ and $\mathcal{R}_W = \{R_{(a,c)} : (a,c) \in V \times Z\} \subseteq \mathcal{B}_W$. Let μ be a measure on $\mathcal{B}_V$ and ν a measure on $\mathcal{B}_W$.

Under the above assumptions, we will denote by $\mathcal{B}_{V \times W}[R]$ the field on $V \times W$ generated by $\mathcal{B}_V \otimes \mathcal{B}_W$ and the sets $R_c, c \in Z$. Notice that $\mathcal{B}_{V \times W}[R]$ contains all R -definable subsets of $V \times W$.

Given measures μ and ν on $\mathcal{B}_V$ and $\mathcal{B}_W$, respectively, in general there are many different measures on $\mathcal{B}_{V \times W}[R]$ extending the product measure $\mu \times \nu$. In this section, in the case when at least one of μ, ν is finitely approximated, we construct a certain canonical measure extending $\mu \times \nu$.

Let, for example, μ be a measure on $\mathcal{B}_V$ that is finitely approximated on the relation $R \subseteq V \times (W \times Z)$. By Claim 2.12, if E is an arbitrary R -definable subset of $V \times W$ and $A \in \mathcal{B}_V, B \in \mathcal{B}_W$, then the function $h_{E,A} : B \rightarrow \mathbb{R}$, given by $h_{E,A}(b) = \mu(A \cap E_b) = \int_A 1_E(x, b) d\mu$, is $\mathcal{B}_W$ -integrable. Hence the double integral

$$\int_B \left(\int_A 1_E(x, y) d\mu \right) d\nu$$

is well defined for any $A \in \mathcal{B}_V, B \in \mathcal{B}_W$.

Similarly, if ν is a measure on $\mathcal{B}_W$ that is finitely approximated on the relation R viewed as a relation on $W \times (V \times Z)$, then for any $A \in \mathcal{B}_V, B \in \mathcal{B}_W$ the double integral

$$\int_A \left(\int_B 1_E(x, y) d\nu \right) d\mu$$

is well defined as well.

PROPOSITION 2.13

- (1) *Let μ be a measure on $\mathcal{B}_V$ that is finitely approximated on the relation $R \subseteq V \times (W \times Z)$. There is a unique measure ω on $\mathcal{B}_{V \times W}[R]$ whose restriction to $\mathcal{B}_V \otimes \mathcal{B}_W$ is $\mu \times \nu$ and such that*

$\omega(E \cap (A \times B)) = \int_B \int_A 1_E(x, y) d\mu d\nu$ for every R -definable $E \subseteq V \times W$, $A \in \mathcal{B}_V, B \in \mathcal{B}_W$. We denote this measure by $\mu \times \nu$.

- (2) If ν is a measure on $\mathcal{B}_W$ that is finitely approximated on the relation $R \subseteq W \times (V \times Z)$, then there is a unique measure, denoted by $\mu \times \nu$, on $\mathcal{B}_{V \times W}[R]$ such that $\mu \times \nu(E \cap (A \times B)) = \int_A \int_B 1_E(x, y) d\nu d\mu$ for every R -definable $E \subseteq V \times W$, $A \in \mathcal{B}_V, B \in \mathcal{B}_W$.
- (3) If μ and ν are measures on $\mathcal{B}_V$ and $\mathcal{B}_W$, respectively, both finitely approximated on the relation R , then $\mu \times \nu$ is also finitely approximated on the relation $R \subseteq (V \times W) \times Z$ and $\mu \times \nu(E) = \nu \times \mu(E)$ for any R -definable $E \subseteq U \times V$.

Proof. (1) It is easy to see that every set Y in $\mathcal{B}_{V \times W}[R]$ is a finite disjoint union of sets of the form $E_i \cap (A_i \times B_i)$ where $E_i \subseteq V \times W$ is R -definable and $A_i \in \mathcal{B}_V, B_i \in \mathcal{B}_W$. We define

$$\omega(Y) = \sum_i \int_{B_i} \int_{A_i} 1_{E_i}(x, y) d\mu d\nu.$$

It is easy to check that ω is well defined and is a finitely additive probability measure on $\mathcal{B}_{V \times W}[R]$ satisfying the requirements. Uniqueness is straightforward from the definition of ω .

(2) is identical to (1).

(3) Given $m \in \mathbb{N}$ and viewing R as a binary relation on $(V \times W) \times Z$, we must show that $\mu \times \nu$ is finitely approximated on the family $\mathcal{R}_{V \times W}^m$ (see Definition 2.10). Notice that for any R -definable $E \in \mathcal{R}_{V \times W}^m$ and $a \in V, b \in W$, we have $E_a \in \mathcal{R}_V^m$ and $E_b \in \mathcal{R}_W^m$ (for R viewed as the corresponding binary relation).

Fix an arbitrary $\varepsilon > 0$. Let $p_1, \dots, p_n \in V$ be such that $\mu(F) \approx^\varepsilon \text{Av}(p_1, \dots, p_n; F)$ for all $F \in \mathcal{R}_V^m$, and let $q_1, \dots, q_m \in W$ be such that $\nu(F') \approx^\varepsilon \text{Av}(q_1, \dots, q_m; F')$ for all $F' \in \mathcal{R}_W^m$.

We claim that the set $\{(p_i, q_j) : 1 \leq i \leq n, 1 \leq j \leq m\}$ gives a 2ε -approximation for $\mu \times \nu$ on $\mathcal{R}_{V \times W}^m$. Let $E \in \mathcal{R}_{V \times W}^m$. Using linearity of integration, we have

$$\begin{aligned} \mu \times \nu(E) &= \int_W \left(\int_V 1_E(v, w) d\mu \right) d\nu \\ &\approx^\varepsilon \int_W \left(\frac{1}{n} \sum_{i=1}^n 1_{E_w}(p_i) \right) d\nu = \frac{1}{n} \sum_{i=1}^n \left(\int_W 1_{E_w}(p_i) d\nu \right) \\ &= \frac{1}{n} \sum_{i=1}^n \left(\int_W 1_{E_{p_i}}(w) d\nu \right) \approx^\varepsilon \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{m} \sum_{j=1}^m 1_{E_{p_i}}(q_j) \right) \\ &= \frac{1}{nm} \sum_{1 \leq i \leq n, 1 \leq j \leq m} 1_E(p_i, q_j), \end{aligned}$$

so $\mu \times \nu(E) \approx^{2\varepsilon} \text{Av}(\{(p_i, q_j) : 1 \leq i \leq n, 1 \leq j \leq m\}; E)$.

The fact that $\mu \ltimes \nu(E) = \nu \ltimes \mu(E)$ follows since, by the above, for any $\varepsilon > 0$ we have

$$\begin{aligned}
 \mu \ltimes \nu(E) &\approx^{2\varepsilon} \frac{1}{n} \sum_{i=1}^n \left(\frac{1}{m} \sum_{j=1}^m 1_{E_{p_i}}(q_j) \right) \\
 &= \frac{1}{m} \sum_{j=1}^m \left(\frac{1}{n} \sum_{i=1}^n 1_{E_{q_j}}(p_i) \right) \approx^\varepsilon \frac{1}{m} \sum_{j=1}^m \left(\int_V 1_{E_{q_j}}(v) d\mu \right) \\
 &= \int_V \left(\frac{1}{m} \sum_{j=1}^m 1_{E_v}(q_j) \right) d\mu \approx^\varepsilon \int_V \left(\int_W 1_E(v, w) d\nu \right) d\mu \\
 &= \nu \ltimes \mu(E),
 \end{aligned}$$

hence $\mu \ltimes \nu(E) \approx^{4\varepsilon} \nu \ltimes \mu(E)$ for an arbitrary $\varepsilon > 0$. $\square$

REMARK 2.14 If $R \subseteq U \times V$ is a relation, then everything above can be applied to R by viewing it as $R \subseteq V \times U \times Z$, where Z is any one-element set.

It is not hard to see that a product of finitely approximated measures satisfies the following weak Fubini property.

LEMMA 2.15 Let $R \subseteq V \times U$. Let μ be a measure on $\mathcal{B}_V$ which is finitely approximated on the relation $R \subseteq V \times U$, and let ν be a measure on $\mathcal{B}_U$. For any $\varepsilon > 0$, if $\mu(R_a) < \varepsilon$ for all $a \in U$ then $(\mu \ltimes \nu)(R) < \varepsilon$.

We extend products of finitely approximated measures to an arbitrary number of sets.

DEFINITION 2.16 Given $k \in \mathbb{N}$, we say that $(R; V_i, \mathcal{B}_i, \mu_i : 1 \leq i \leq k)$ is a *compatible finitely approximated system* if for all i we have:

- (1) $R \subseteq V_1 \times \dots \times V_k$,
- (2) $\mathcal{B}_i$ is a field on V_i ,
- (3) μ_i is a measure on $\mathcal{B}_i$,
- (4) $R_a \in \mathcal{B}_i$ for each $a \in V_i^c$,
- (5) μ_i is finitely approximated on the relation R (viewed as a binary relation on $V_i \times V_i^c$).

DEFINITION 2.17 Assume that $(R; V_i, \mathcal{B}_i, \mu_i : 1 \leq i \leq k)$ is a compatible finitely approximated system. By induction on $2 \leq n \leq k$ we define the field $\mathcal{B}_1 \times \dots \times \mathcal{B}_n$ on $V_1 \times \dots \times V_n$ and the measure $\mu_1 \ltimes \dots \ltimes \mu_n$ on $\mathcal{B}_1 \times \dots \times \mathcal{B}_n$ as follows.

On the induction step $n+1$, we set $V = V_1 \times \dots \times V_n$, $W = V_{n+1}$, $Z = V_{n+2} \times \dots \times V_k$. Viewing R as $R \subseteq V \times W \times Z$, we set $\mathcal{B}_1 \times \dots \times \mathcal{B}_{n+1} = \mathcal{B}_{U \times V}[R]$, and $\mu_1 \ltimes \dots \ltimes \mu_{n+1} := (\mu_1 \ltimes \dots \ltimes \mu_n) \ltimes \mu_{n+1}$.

Note that in particular $R \in \mathcal{B}_1 \times \dots \times \mathcal{B}_k$ and $E \in \mathcal{B}_1 \times \dots \times \mathcal{B}_k$ for every $R_\otimes$ -definable $E \subseteq V_1 \times \dots \times V_k$.

2.6. Approximations by rectangular sets

PROPOSITION 2.18 *Let V, W be sets, $R \subseteq V \times W$ a subset, μ a measure on V which is finitely approximated on the relation R . Then for any $\varepsilon > 0$ there are R -definable subsets $X_1, \dots, X_m \subseteq W$ partitioning W such that for every $i \in [m]$ and any $a, a' \in X_i$ we have $\mu(R_a \Delta R_{a'}) < \varepsilon$.*

In addition, if the family $\mathcal{R} = \{R_a : a \in W\}$ has VC-dimension at most d then we can choose $D \subseteq V$ of size at most $320d(\frac{1}{\varepsilon})^2$ such that every X_i is an atom in the Boolean algebra of sets R -definable over D , and $m \leq C_d(320d)^d (\frac{1}{\varepsilon})^{2d}$ for some constant C_d depending only on d .

Proof. Let $\mathcal{R}^\Delta = \{R_a \Delta R_{a'} : a, a' \in W\}$. Since μ is finitely approximated on R , there are $p_1, \dots, p_n \in V$ with $|\mu(F) - \text{Av}(p_1, \dots, p_n; F)| < \varepsilon$ for any $F \in \mathcal{R}^\Delta$.

For each $I \subseteq [n]$ let $X_I = \{a \in W : p_i \in R_a \Leftrightarrow i \in I\}$. It is easy to see that the sets $X_I, I \subseteq [n]$ partition W , every X_I is R -definable and for every $I \subseteq [n]$ and $a, a' \in X_I$ we have $\mu(R_a \Delta R_{a'}) < \varepsilon$.

Assume in addition that $\mathcal{R}$ is a VC-family with VC-dimension at most d . As above we choose $p_1, \dots, p_n \in V$ with

$$|\mu(F) - \text{Av}(p_1, \dots, p_n; F)| < \varepsilon/2$$

for any $F \in \mathcal{R}^\Delta$.

Let ω be a measure on $\mathcal{B}_V$ given by $\omega(X) = \text{Av}(p_1, \dots, p_n; X)$. Since $\mathcal{R}$ has VC-dimension at most d , the family $\mathcal{R}^\Delta$ had dimension at most $10d$ (see [31, Lemma 4.5]), and by Fact 2.8 we can choose an $\varepsilon/2$ -net D for $\mathcal{R}^\Delta$ and ω with $|D| \leq 80d^2 \log \frac{2}{\varepsilon}$. Clearly

$$80d^2 \log \frac{2}{\varepsilon} \leq 80d \left(\frac{2}{\varepsilon}\right)^2 = 320d \left(\frac{1}{\varepsilon}\right)^2.$$

For each $I \subseteq D$ let $X_I = \{a \in W : R_a \cap D = I\}$. It is easy to see that the sets $X_I, I \subseteq D$, partition W and every X_I is an atom in the Boolean algebra of all sets R -definable over D . Let $I \subseteq D$ and $a, a' \in X_I$. Then $R_a \cap D = R_{a'} \cap D$, hence $\omega(R_a \Delta R_{a'}) \leq \varepsilon/2$, and $\mu(R_a \Delta R_{a'}) < \varepsilon$. Finally, by Fact 2.4, the number of different atoms X_I 's is at most $C_d |D|^d \leq C_d (320d)^d (\frac{1}{\varepsilon})^{2d}$. $\square$

THEOREM 2.19 *Let $(R; V_i, \mathcal{B}_i, \mu_i : 1 \leq i \leq k)$ be a compatible finitely approximated system. Then for every $\varepsilon > 0$ there is an $R_\otimes$ -definable $A \subseteq V_1 \times \dots \times V_k$ with*

$$(\mu_1 \times \dots \times \mu_k)(R \Delta A) < \varepsilon.$$

In addition, if R has VC-dimension at most d (see Definition 2.3) then we can choose A to be $R_\otimes$ -definable over some $\vec{D}$ with $\|\vec{D}\| \leq C_{k,d} (\frac{1}{\varepsilon})^{2(k-1)d}$, where $C_{k,d}$ is a constant that depends on k and d only.

Proof. We proceed by induction on k .

The case $k = 2$. Let V_1, V_2 and $R \subseteq V_1 \times V_2$ be given. Using Proposition 2.18, we can find R -definable sets $X_1, \dots, X_m$ partitioning V_2 such that for every $i \in [m]$ and any $a, a' \in X_i$ we have $\mu_1(R_a \Delta R_{a'}) < \varepsilon$.

For each $i \in [m]$ we pick $a_i \in X_i$ and let $A = \bigcup_{i \in [m]} R_{a_i} \times X_i$. Obviously A is $R_\otimes$ -definable. It is not hard to see that for every $a \in V_2$ we have $\mu_1(R_a \Delta A_a) < \varepsilon$, hence, by Lemma 2.15, $(\mu_1 \times \mu_2)(R \Delta A) < \varepsilon$.

Assume in addition that R has VC-dimension at most d . Then by Proposition 2.18, we can assume that for some $D_2 \subseteq V_1$ with $|D_2| \leq 320d(\frac{1}{\varepsilon})^2$ every X_i is an R -definable atom over D_2 and $m \leq C_d(320d)^d(\frac{1}{\varepsilon})^{2d}$. Let $D_1 = \{a_1, \dots, a_m\}$, and $\vec{D} = (D_1, D_2)$. Obviously A is $R_\otimes$ -definable over $\vec{D}$ and we can take $C_{2,d} = C_d(320d)^d$.

Inductive step $k + 1$. Let $V_1, \dots, V_{k+1}$ and $R \subseteq V_1 \times \dots \times V_{k+1}$ be given.

Viewing $V_1 \times \dots \times V_{k+1}$ as $V_{[k]} \times V_{k+1}$ and using the case of $k = 2$ we obtain R -definable $X_1, \dots, X_m$ partitioning V_{k+1} and points $a_i \in X_i$, $i \in [m]$, such that for the set $A' = \bigcup_{i \in [m]} R_{a_i} \times X_i$ we have $(\mu_1 \times \dots \times \mu_{k+1})(R\Delta A') < \varepsilon/2$.

For each $i \in [m]$ let $R^i = R_{a_i}$. It is an R -definable subset of $V_1 \times \dots \times V_k$. Applying induction hypothesis to each R^i we obtain $R_\otimes^i$ -definable sets $A_i \subseteq V_1 \times \dots \times V_k$ such that $(\mu_1 \times \dots \times \mu_k)(R^i \Delta A_i) < \varepsilon/2$. Let $A = \bigcup_{i \in [m]} A_i \times X_i$. It is an $R_\otimes$ -definable set and using Proposition 2.13 and Lemma 2.15, it is not hard to see that

$$(\mu_1 \times \dots \times \mu_{k+1})(A' \Delta A) < \varepsilon/2,$$

hence $(\mu_1 \times \dots \times \mu_{k+1})(R\Delta A) < \varepsilon$, as required.

Assume in addition that R has VC-dimension at most d . As in the case $k = 2$ we can assume that every X_i is R -definable over $D_{k+1} \subseteq V_1 \times \dots \times V_k$ with $|D_{k+1}| \leq 320d(\frac{2}{\varepsilon})^2$ and also assume that

$$m \leq C_d |D_{k+1}|^d \leq C_d \left[320d(\frac{2}{\varepsilon})^2 \right]^d = C_d(1280d)^d(\frac{1}{\varepsilon})^{2d}.$$

It is easy to see that each R^i has VC-dimension at most d . Applying induction hypotheses we can assume that each A_i above is $R_\otimes^i$ -definable over $\vec{D}^i = (D_1^i, \dots, D_k^i)$ with $\|\vec{D}^i\| \leq C_{k,d}(\frac{2}{\varepsilon})^{2(k-1)d}$, where $D_j^i \subseteq \prod_{l \in [k] \setminus \{j\}} V_l$.

For each $i \in [m]$ and $j \in [k]$ let $\bar{D}_j^i = \{(c, a_i) : c \in D_j^i\}$, $D_j = \bigcup_{i \in [m]} \bar{D}_j^i$, and $\vec{D} = (D_1, \dots, D_{k+1})$.

It is not hard to see that A above is $R_\otimes$ -definable over $\vec{D}$ and

$$\|\vec{D}\| \leq m C_{k,d}(\frac{2}{\varepsilon})^{2(k-1)d} \leq C_d(1280d)^d(\frac{1}{\varepsilon})^{2d} C_{k,d} 2^{2(k-1)d}(\frac{1}{\varepsilon})^{2(k-1)d} = C_{k+1,d}(\frac{1}{\varepsilon})^{2kd}.$$

□

3. Definable regularity lemma for hypergraphs of bounded VC dimension

In this section, we apply the product measure decomposition results from Section 2 to regularity of definable hypergraphs. Our goal is to prove a stronger version of Fact 1.1 for hypergraphs of bounded VC-dimension.

3.1. Regularity lemmas for hypergraphs

In this paper, a k -hypergraph $G = (V_1, \dots, V_k; R)$ consists of sets $V_1, \dots, V_k$ and a subset $R \subseteq V_1 \times \dots \times V_k$. We do not assume that the sets V_i , $i \in [k]$ are pairwise distinct.

A k -uniform hypergraph $G = (V; R)$ is a set V with a symmetric subset $R \subseteq V^k$. Of course, every k -uniform hypergraph $G = (V; R)$ can be also viewed as a k -hypergraph $(V, \dots, V; R)$ that we will denote by $\tilde{G}$.

For a k -hypergraph $G = (V_1, \dots, V_k; R)$ and $A_1 \subseteq V_1, \dots, A_k \subseteq V_k$, we let $R(A_1, \dots, A_k) := R \cap A_1 \times \dots \times A_k$.

Let $G = (V_1, \dots, V_k; R)$ be a k -hypergraph. By a *rectangular partition* of G , we mean a k -tuple $\vec{\mathcal{P}} = (\mathcal{P}_1, \dots, \mathcal{P}_k)$ where each $\mathcal{P}_i$ is a finite partition of V_i . For a rectangular partition $\vec{\mathcal{P}} = (\mathcal{P}_1, \dots, \mathcal{P}_k)$, we define $\|\vec{\mathcal{P}}\| = \max\{|\mathcal{P}_i| : i \in [k]\}$, and for a set $X \subseteq V_1 \times \dots \times V_k$ we write $X \in \vec{\mathcal{P}}$ if $X = X_1 \times \dots \times X_k$ for some $X_i \in \mathcal{P}_i, i \in [k]$. We will also write $\Sigma \subseteq \vec{\mathcal{P}}$ to indicate that Σ consists of subsets $X \subseteq V_1 \times \dots \times V_k$ with $X \in \vec{\mathcal{P}}$.

We say that $\vec{\mathcal{P}}$ is *R-definable* if each $\mathcal{P}_i$ consists of R -definable sets. For a tuple $\vec{D} = (D_1, \dots, D_k)$ as in Definition 2.1 we say that $\vec{\mathcal{P}}$ is *R-definable over $\vec{D}$* if for each $i \in [k]$ every $X \in \mathcal{P}_i$ is R -definable over D_i .

A k -hypergraph $G = (V_1, \dots, V_k; R)$ has *VC-dimension at most d* if R has VC-dimension at most d in the sense of Definition 2.3. A k -uniform hypergraph $G = (V; R)$ has VC-dimension at most d if the corresponding k -hypergraph $\tilde{G}$ is NIP with VC-dimension at most d .

DEFINITION 3.1 Let $(R; V_i, \mathcal{B}_i, \mu_i : 1 \leq i \leq k)$ be a compatible finitely approximated system. Let $\mu := \mu_1 \times \dots \times \mu_k$. Given $\varepsilon > 0$, we say that an R -definable rectangular partition $\vec{\mathcal{P}}$ of $V_1 \times \dots \times V_k$ is *ε -regular with 0-1-densities* if there is $\Sigma \subseteq \vec{\mathcal{P}}$ such that

$$\sum_{X \in \Sigma} \mu(X) \leq \varepsilon,$$

and for every $X_1 \times \dots \times X_k \in \vec{\mathcal{P}} \setminus \Sigma$ either

$$\mu(X_1 \times \dots \times X_k) - \mu(R(X_1, \dots, X_k)) < \varepsilon \mu(X_1 \times \dots \times X_k)$$

or

$$\mu(R(X_1, \dots, X_k)) < \varepsilon \mu(X_1 \times \dots \times X_k).$$

REMARK 3.2 Note that in Fact 1.1 the condition on the density of the edges is stated not just for the sets of the form $V_i \times V_j$ with V_i, V_j from the partition, but also for arbitrary sets of the form $A \times B$ with $A \subseteq V_i, B \subseteq V_j$. However, in the case of regular partitions with 0-1-densities this strengthening follows for free: if, as in Definition 3.1, we have $|\mu(X_1 \times \dots \times X_k) - \delta \mu(R(X_1, \dots, X_k))| < \varepsilon \mu(X_1 \times \dots \times X_k)$ for some $\delta \in \{0, 1\}$, then for arbitrary sets $Y_i \subseteq X_i$ with $Y_i \in \mathcal{B}_i$ for $i \in [k]$ we also have $|\mu(Y_1 \times \dots \times Y_k) - \delta \mu(R(Y_1, \dots, Y_k))| < \varepsilon \mu(X_1 \times \dots \times X_k)$.

The next theorem demonstrates how existence of an approximation by rectangular sets for the product measure proved in Section 2 can be used to obtain a regular partition.

THEOREM 3.3 Let $(R; V_i, \mathcal{B}_i, \mu_i : 1 \leq i \leq k)$ be a compatible finitely approximated system and $\mu = \mu_1 \times \dots \times \mu_k$. Then for any $\varepsilon > 0$ there is an R -definable ε -regular partition $\vec{\mathcal{P}}$ with 0-1-densities.

In addition, if R is NIP with VC dimension at most d we can choose $\vec{\mathcal{P}}$ with $\|\vec{\mathcal{P}}\| \leq C_d(C_{k,d})^d \left(\frac{1}{\varepsilon}\right)^{4(k-1)d^2}$, where C_d and $C_{k,d}$ are the constants from Fact 2.4 and Theorem 2.19.

Proof. Using Theorem 2.19, there is a set A which is $R_\otimes$ -definable over some finite set $\vec{D} = (D_1, \dots, D_k)$ and $\mu(A \Delta R) < \varepsilon^2$. Say $A = \bigcup_{j \in [m]} A_1^j \times \dots \times A_k^j$ where each $A_i^j \subseteq V_i$ is R -definable.

For each $i \in [k]$, let $\mathcal{P}_i$ be the set of all atoms in the Boolean algebra generated by all R -definable over D_i subsets of V_i ; so each $\mathcal{P}_i$ consists of R -definable sets partitioning V_i . We claim that $\vec{\mathcal{P}}$ is ε -regular with 0–1-densities.

Let

$$\Sigma := \{X \in \vec{\mathcal{P}} : \mu(X \cap (A \Delta R)) \geq \varepsilon \mu(X)\}.$$

Since $\mu(A \Delta R) < \varepsilon^2$ and μ is finitely additive we obtain that

$$\sum_{X \in \Sigma} \mu(X) \leq \varepsilon.$$

Let $X = X_1 \times \cdots \times X_k \in \vec{\mathcal{P}} \setminus \Sigma$. We have

$$\mu(X \cap (A \Delta R)) < \varepsilon \mu(X).$$

By definition of $\vec{\mathcal{P}}$, either $X \subseteq A$ or $X \cap A = \emptyset$.

Assume first $X \subseteq A$, then $X \cap (A \Delta R) = X \setminus R(X_1, \dots, X_k)$ and

$$\mu(X \setminus R(X_1, \dots, X_k)) = \mu(X) - \mu(R(X_1, \dots, X_k)),$$

hence

$$\mu(X_1 \times \cdots \times X_k) - \mu(R(X_1, \dots, X_k)) \leq \varepsilon \mu(X_1 \times \cdots \times X_k).$$

If $X \cap A = \emptyset$ a similar argument shows that

$$\mu(R(X_1, \dots, X_k)) < \varepsilon \mu(X_1 \times \cdots \times X_k).$$

Assume in addition that R is NIP with VC-dimension at most d . Then using Theorem 2.19 we can assume that $|D_i| \leq C_{k,d} \left(\frac{1}{\varepsilon}\right)^{4(k-1)d}$ for $i \in [k]$, and by Fact 2.4,

$$|\mathcal{P}_i| \leq C_d |D_i|^d \leq C_d \left(C_{k,d} \left(\frac{1}{\varepsilon}\right)^{4(k-1)d}\right)^d = C_d (C_{k,d})^d \left(\frac{1}{\varepsilon}\right)^{4(k-1)d^2}.$$

□

REMARK 3.4 In the case when each V_i is finite the above theorem without the NIP part is trivial, since we can take $\mathcal{P}_i$ to be the set of all atoms in the Boolean algebra of **all** R -definable subsets of V_i .

This immediately gives an analogous theorem for k -uniform hypergraphs. We state it only in the NIP case.

THEOREM 3.5 *Let $G = (V; R)$ be a k -uniform hypergraph with VC-dimension at most d , $\mathcal{B}$ a Boolean algebra on V containing all of the fibers of R , and let μ be a measure on $\mathcal{B}$ which is finitely approximated on $R \subseteq V \times V^{k-1}$. Then for any $\varepsilon > 0$ there is an R -definable partition $\mathcal{P}$ of V such that*

$\vec{\mathcal{P}} = (\mathcal{P}, \dots, \mathcal{P})$ is an ε -regular partition of the k -hypergraph $\tilde{G} = (V, \dots, V; R)$ with 0-1-densities relatively to the measure $\mu^k = \mu \times \dots \times \mu$, and $|\vec{\mathcal{P}}| \leq C_d(kC_{k,d})^d \left(\frac{1}{\varepsilon}\right)^{4(k-1)d^2}$.

Proof. By assumption $(R; V_i, \mathcal{B}_i, \mu_i : 1 \leq i \leq k)$ is a compatible finitely approximated system, with $V_i := V, \mathcal{B}_i := \mathcal{B}, \mu_i := \mu$, and R viewed as a relation $R \subseteq V_1 \times \dots \times V_k$. Let $\mu^k = \mu_1 \times \dots \times \mu_k$.

Let $\vec{\mathcal{P}}$ be an ε -regular partition with 0–1-densities as in the proof of Theorem 3.3, that is, $\mathcal{P}_i$ is the set of all atoms in the Boolean algebra generated by all R -definable over $D_i \subseteq V^{k-1}$ subsets of V_i .

Let $D = \bigcup_{i \in [k]} D_i$. We take $\mathcal{P}$ to be the set of all atoms in the Boolean algebra of all R -definable over D subsets of V . Then $(\mathcal{P}, \dots, \mathcal{P})$ is also a ε -regular partition with 0–1-densities as it refines $\vec{\mathcal{P}}$, and by Fact 2.4,

$$|\mathcal{P}| \leq C_d |D|^d \leq C_d \left(kC_{k,d} \left(\frac{1}{\varepsilon} \right)^{4(k-1)d} \right)^d = C_d (kC_{k,d})^d \left(\frac{1}{\varepsilon} \right)^{4(k-1)d^2}.$$

□

Now we give some examples where Theorems 3.3 and 3.5 apply.

3.2. The finite case

Let $G = (V_1, \dots, V_k; R)$ be a finite k -hypergraph. For each $i \in [k]$ let μ_i be the uniform counting measure on V_i , that is, $\mu_i(X) = \frac{|X|}{|V_i|}$, and let μ be the uniform counting measure on $V_1 \times \dots \times V_k$. Then all μ_i and μ are finitely approximated measures with $\mu = \mu_1 \times \dots \times \mu_k$. Hence all the results of the previous section can be applied to finite k -hypergraphs with respect to the counting measures.

COROLLARY 3.6 *Let $G = (V; R)$ be a finite k -uniform hypergraph. Assume that R has VC-dimension at most d , as a relation on V^k .*

There is a partition $V = V_1 \dot{\cup} \dots \dot{\cup} V_M$ for some $M \leq C_d(kC_{k,d})^d \left(\frac{1}{\varepsilon}\right)^{4(k-1)d^2}$, numbers $\delta_{\vec{i}} \in \{0, 1\}$ for $\vec{i} \in [M]^k$, and an exceptional set $\Sigma \subseteq [M]^k$ such that

$$\sum_{(i_1, \dots, i_k) \in \Sigma} |V_{i_1}| \cdots |V_{i_k}| \leq \varepsilon |V|^k$$

and for each $\vec{i} = (i_1, \dots, i_k) \in [M]^k \setminus \Sigma$ we have

$$| |R(V_{i_1}, \dots, V_{i_k})| - \delta_{\vec{i}} |V_{i_1}| \cdots |V_{i_k}| | < \varepsilon |V_{i_1}| \cdots |V_{i_k}|.$$

So the size of the partition M depends only on ε , k and the VC-dimension d , polynomially in terms of $\frac{1}{\varepsilon}$, and not on the size of the graph.

3.3. Hypergraphs definable in NIP structures

Now we discuss the model-theoretic setting, which is the main motivating example for this article. For a detailed account of this setting, we refer to the introduction in [10] and to [50].

Let $\mathcal{M}$ be a first-order structure. Recall that a *Keisler measure* on M^n is a finitely additive probability measure on the Boolean algebra of all definable subsets of M^n . Given a formula $\phi(x)$ with parameters from M and a Keisler measure μ on $M^{|x|}$, we will write $\mu(\phi(x))$ to denote $\mu(\phi(M^{|x|}))$. Let us fix a definable relation $E(x_1, \dots, x_k)$, let $V_i = M^{|x_i|}$ and let $\mathcal{B}_i$ be the Boolean algebra of all definable subsets of $M^{|x_i|}$. Let μ_i be a Keisler measure on $M^{|x_i|}$, equivalently a measure on $\mathcal{B}_i$.

Recall that a structure $\mathcal{M}$ is an *NIP structure* if for every formula $\phi(x, y)$ the family of all ϕ -definable sets $\mathcal{F}_\phi = \{\phi(M, a) : a \in M^{|y|}\}$ has finite VC-dimension. In particular, if $\mathcal{M}$ is NIP, then any definable relation $E(x_1, \dots, x_k) \subseteq M^{|x_1|} \times \dots \times M^{|x_k|}$ has finite VC dimension (in the sense of Definition 2.3). Recall that, in an NIP structure $\mathcal{M}$, a Keisler measure μ on $M^{|x|}$ is *generically stable* if it is finitely approximated on *all* definable relations $\phi(x, y) \subseteq M^{|x|} \times M^{|y|}$, in particular on E .

REMARK 3.7 There are several equivalent characterizations of generically stable measures in NIP structures. Our definition of finitely approximated measures only requires the *existence* of an ε -approximation for every ε . A stronger notion of a *fin measure* is given in [26] requiring that in fact for every ε , there is some sufficiently large n such that *almost all* n -tuples (in the sense of the product measure $\mu^{(n)}$) give an ε -approximation. While μ is finitely approximated on all formulas if and only if it is fin on all formulas under the NIP assumption (by the results in [26]), it is not clear if the equivalence holds in general. It is known that μ is finitely approximated on all formulas if and only if it is fin on all formulas under the NIP assumption, by the results in [26]. On the other hand, examples separating the two notions outside of the NIP context were provided in [12].

Now, the semidirect product $\mu = \mu_1 \times \dots \times \mu_k$ corresponds to the non-forking product $\mu_1 \otimes \dots \otimes \mu_k$. Hence Theorem 3.3 translates into the following.

COROLLARY 3.8 *Let $\mathcal{M}$ be NIP. For every definable relation $E(x_1, \dots, x_k)$ there is some $c = c(E)$ such that: for any $\varepsilon > 0$ and any generically stable Keisler measures μ_i on $M^{|x_i|}$ there are partitions $M^{|x_i|} = \bigcup_{j < K} A_{i,j}$ and a set $\Sigma \subseteq [K]^k$ such that:*

- (1) $K \leq \left(\frac{1}{\varepsilon}\right)^c$.
- (2) $\mu\left(\bigcup_{(i_1, \dots, i_k) \in \Sigma} A_{1,i_1} \times \dots \times A_{k,i_k}\right) \leq \varepsilon$, where $\mu = \mu_1 \otimes \dots \otimes \mu_k$,
- (3) for all $(i_1, \dots, i_k) \notin \Sigma$ we have

$$|\mu(E \cap (A_{1,i_1} \times \dots \times A_{k,i_k})) - \delta_i \mu(A_{1,i_1} \times \dots \times A_{k,i_k})| < \varepsilon \mu(A_{1,i_1} \times \dots \times A_{k,i_k})$$

for some $\delta_i \in \{0, 1\}$.

- (4) each $A_{i,j}$ is defined by an instance of an E -formula, with this formula depending only on E and ε .

Theorem 3.3 is more general however as both NIP and finite approximability are only assumed locally for R , and can be applied outside of the context of NIP structures.

EXAMPLE 3.9 Let $\mathcal{M}$ be a pseudo-finite field, viewed as a structure in the ring language (for example, an ultraproduct of finite fields modulo some non-principal ultrafilter). Then the ultralimit of the counting measures gives a measure on the definable sets in $\mathcal{M}$. This measure is finitely approximable on all quantifier-free definable relations (by Lemma 4.3, as it is well known that all quantifier-free formulas in $\mathcal{M}$ are stable), but not finitely approximable for general definable relations (for example,

because the random graph is definable). Still, Theorem 3.3 can be applied to any quantifier-free definable relation in this situation.

We list some specific structures and Keisler measures for which Corollary 3.8 applies to all definable relations (again, see introduction in [10] for more details).

EXAMPLE 3.10 Examples of NIP structures:

- (1) Abelian groups and modules (see, for example [53]),
- (2) $(\mathbb{C}, +, \times, 0, 1)$ (see, for example [53]),
- (3) differentially closed fields (see, for example [53]),
- (4) free groups (in the pure group language $(\cdot, {}^{-1}, 0)$, see [43]),
- (5) planar graphs (in the language with a single binary relation corresponding to the edges, see [40]),
- (6) (weakly) o-minimal structures, for example, $M = (\mathbb{R}, +, \times, e^x)$ (see [10]),
- (7) Presburger arithmetic, that is, the ordered group of integers (see [10]),
- (8) p -minimal structures with Skolem functions, for example, $(\mathbb{Q}_p, +, \times)$ for each prime p ,
- (9) the (valued differential) field of transseries ([6, 4]),
- (10) algebraically closed valued fields (see, for example [47]).

EXAMPLE 3.11 Examples of generically stable Keisler measures (see for example, the introduction in [10] for more details on why these measures are generically stable):

- (1) Any Keisler measure concentrated on a finite set (as it is clearly finitely approximable).
- (2) Let λ_n be the Lebesgue measure on the unit cube $[0, 1]^n$ in $\mathbb{R}^n$. Let $\mathcal{M}$ be an o-minimal structure expanding the field of real numbers. If $X \subseteq \mathbb{R}^n$ is definable in $\mathcal{M}$, then, by o-minimal cell decomposition, $X \cap [0, 1]^n$ is Lebesgue measurable, hence λ_n induces a Keisler measure on M^n .
- (3) Similarly to (2), for every prime p a (normalized) Haar measure on a compact ball in $\mathbb{Q}_p$ induces a Keisler measure on $\mathbb{Q}_p^n$.

4. Stable and distal cases

Next we consider two extreme opposite special cases of NIP hypergraphs: stable and distal ones. Stable theories are at the cornerstone of Shelah's classification theory [44], and we refer to for example, [53, 38] for a general exposition of stability. Examples (1)–(5) in Example 3.10 are stable. Distal theories were introduced more recently in [46] aiming to capture ‘purely unstable’ structures in NIP theories. Examples (6)–(9) in Example 3.10 are distal. Example (10) gives a combination of these two cases: it has a stable part (the algebraically closed residue field) and a distal part (the value group), and the theory developed in [22] demonstrates that the whole structure can be analysed in terms of these two parts. There are certain generalizations of this decomposition principle for arbitrary NIP theories [45, 49].

4.1. Stable hypergraph regularity

A regularity lemma for stable graphs was proved in [33] for counting measures. Later, [34] provides a proof for general measures. However, the proof in [34] does not give any bounds on the size of the

partition. In this section, we combine these two approaches and prove a regularity lemma for stable hypergraphs relatively to arbitrary measures, bounding the size of the partition by a polynomial in $\frac{1}{\varepsilon}$.

DEFINITION 4.1

- (1) A binary relation $R \subseteq V \times W$ is *d-stable*, $d \in \mathbb{N}$, if there is no tree of parameters $(b_\eta : \eta \in 2^{<d})$ in W such that for any $\eta \in 2^d$ there is some $a_\eta \in V$ such that $a_\eta \in R_{b_\nu} \iff \nu \frown 1 \leq \eta$ (where $\leq$ is the tree order).
- (2) A relation $R \subseteq V_1 \times \dots \times V_k$ is *d-stable* if for every $I \subseteq [k]$, viewed as a binary relation on $V_I \times V_{I^c}$, it is *d-stable*.
- (3) A relation R is *stable* if it is *d-stable* for some d .

Note that if R is stable, then it has finite VC-dimension.

REMARK 4.2 Alternatively, stability of a relation can be defined in terms of the so called *order property*. Namely, $R \subseteq V \times W$ has the *e-order property*, $e \in \mathbb{N}$, if there are some elements a_i in V and b_i in W , $i = 1, \dots, e$, such that $a_i \in R_{b_j} \iff i < j$ for all $1 \leq i \neq j \leq e$. It is a standard fact in basic stability theory that R is stable (in the sense of Definition 4.1) if and only if it does not have the *e-order property* for some e (but the relation between the corresponding parameters e and d is exponential, see, for example [23, Lemma 6.7.9]).

LEMMA 4.3 *Let $R \subseteq V \times W$ be a stable relation, and $\mathcal{B}_V$ be a Boolean algebra on V such that $R_b \in \mathcal{B}_V$ for all $b \in W$. Then any measure μ on $\mathcal{B}_V$ is finitely approximable on R .*

Proof. Assume that R is *d-stable*. By Definition 2.10, we must show that μ is finitely approximable on the family $\mathcal{R}_V^m$ of subsets of V , for every $m \in \mathbb{N}$. Fix m .

Claim 1. For any $\varepsilon > 0$ there is some $t = t(\varepsilon, d, m)$ and some 0–1 measures $\delta_1, \dots, \delta_t$ on $\mathcal{B}_V$ (possibly with repetitions) such that $\mu(S) \approx^\varepsilon \frac{1}{t} \sum_{i=1}^t \delta_i(S)$ for all $S \in \mathcal{R}_V^m$.

Proof. As R is stable, it follows that the family $\mathcal{R}_V^m$ has finite VC-dimension, and it depends only on d, m . Fix $\varepsilon > 0$. By the VC-theorem there is some $t = t(\varepsilon, d, m)$ such that for every finite (or countable) $\mathcal{F} \subseteq \mathcal{R}_V^m$ there are some $a_1, \dots, a_t \in V$ such that $\mu(S) \approx^\varepsilon \frac{1}{t} \sum_{i=1}^t \mathbf{1}_S(a_i)$ for all $S \in \mathcal{F}$. For $1 \leq i \leq t$ and a finite $\mathcal{F}$, define a 0–1 measure $\delta_i^\mathcal{F}$ on $\mathcal{B}_V$ by $\delta_i^\mathcal{F}(S) := \mathbf{1}_S(a_i)$ for all $S \in \mathcal{B}_V$, then $\mu(S) \approx^\varepsilon \frac{1}{t} \sum_{i=1}^t \delta_i^\mathcal{F}(S)$ for all $S \in \mathcal{F}$. The claim now follows by compactness of the space of all 0–1 measures on $\mathcal{B}_V$ (see [25, Lemma 4.8] for a more detailed account).

Claim 2. Every 0–1 measure δ on $\mathcal{B}_V$ is finitely approximable on $\mathcal{R}_V^m$.

Proof. This is a straightforward consequence of the explicit form of the definability of types in local stability. Namely, consider a binary relation $E \subseteq V \times \mathcal{R}_V^m$ given by $E := \{(a, S) : a \in S\}$. Then E is *r-stable* for some r (as R is stable, and stability is preserved under Boolean combinations). We can identify our 0–1 measure δ restricted to E with a complete *E-type*. Then (see, for example, the

proof of [38, Chapter 1, Lemma 2.2]) for every $t \in \mathbb{N}$ we can choose some $c_1, \dots, c_t \in V$ such that for every $S \in \mathcal{R}_V^m$:

- if $|\{i : c_i \in S\}| > r$, then $\delta(S) = 1$;
- if $|\{i : c_i \notin S\}| > r$, then $\delta(S) = 0$.

Hence if t is large enough so that $\frac{r}{t} < \varepsilon$, then $c_1, \dots, c_t$ give an ε -approximation of δ on $\mathcal{R}_V^m$.

Now, let $\varepsilon > 0$ and m be arbitrary, and let $\delta_1, \dots, \delta_t$ be as given by Claim 1. By Claim 2, let A_i be a multiset in V giving an ε -approximation for δ_i on $\mathcal{R}_V^m$. It is straightforward to verify that $A = \bigcup_{i=1}^t A_i$ is a 2ε -approximation for μ on $\mathcal{R}_V^m$. $\square$

From now on we work in the same setting as in Section 2. Throughout the section we let the sets $V_1, \dots, V_k$ and a stable relation $R \subseteq V_1 \times \dots \times V_k$ be given, let $\mathcal{B}_i$ be a field on V_i , and let μ_i be a measure on $\mathcal{B}_i$. Assume moreover that for every $i \in [k]$, $R_b \in \mathcal{B}_i$ for all $b \in V_{I^c}$.

In view of Lemma 4.3, if $R \subseteq V_1 \times \dots \times V_k$ is a stable relation, then for every $I = \{i_1, \dots, i_n\} \subseteq [k]$ we have a semi-direct product measure $\mu_I = \mu_{i_1} \times \dots \times \mu_{i_n}$ on $\mathcal{B}_I = \mathcal{B}_{i_1} \times \dots \times \mathcal{B}_{i_n}$ (see Definition 2.17) which is finitely approximable on R (Proposition 2.13).

DEFINITION 4.4 For any $I \subseteq [k]$ and $\varepsilon > 0$, a set $A \in \mathcal{B}_I$ is ε -good if for any $b \in V_{I^c}$, either $\mu_I(A \cap R_b) < \varepsilon \mu_I(A)$ or $\mu_I(A \cap R_b) > (1 - \varepsilon) \mu_I(A)$.

REMARK 4.5 Notice that if a set is ε -good, $\varepsilon > 0$, then it has measure greater than 0.

LEMMA 4.6 Assume that μ_{I^c} is finitely approximable on R . For any $\varepsilon > 0$, consider the set

$$A = \{a \in V_I : \mu_{I^c}(R_a) < \varepsilon\}.$$

Then there is an R -definable set $A' \supseteq A$ such that $\mu_{I^c}(R_a) < 2\varepsilon$ for all $a \in A'$.

Proof. Let $b_1, \dots, b_n \in V_{I^c}$ be such that $\mu_{I^c}(R_a) \approx_{\frac{\varepsilon}{2}} \text{Av}(b_1, \dots, b_n; R_a)$ for all $a \in V_I$. Let $\mathcal{J} = \{J \subseteq [n] : \frac{|J|}{n} < \frac{3}{2}\varepsilon\}$, and let $A' = \bigcup_{J \in \mathcal{J}} \left(\bigcap_{j \in J} R_{b_j} \cap \bigcap_{j \notin J} (R_{b_j})^c \right)$. It is easy to check that A' satisfies the requirements. $\square$

LEMMA 4.7 Fix some $I \subseteq [k]$ and some $J \subseteq [k] \setminus I$. Let $\varepsilon > 0$ and $B \in \mathcal{B}_J$ be an ε -good set, and let $A \in \mathcal{B}_I$ and $c \in V_{[k] \setminus (I \cup J)}$ be arbitrary, such that A is of positive measure (note that B is of positive measure by Remark 4.5). Then (by Definition 4.4) A is a disjoint union of the sets

$$A_{B,c,\varepsilon}^0 = \{a \in A : \mu_J(R_{a,c} \cap B) < \varepsilon \mu_J(B)\}$$

and

$$A_{B,c,\varepsilon}^1 = \{a \in A : \mu_J(R_{a,c} \cap B) > (1 - \varepsilon) \mu_J(B)\}.$$

Assume that $\varepsilon < \frac{1}{4}$. Then $A_{B,c,\varepsilon}^0, A_{B,c,\varepsilon}^1 \in \mathcal{B}_I$.

Proof. Indeed, let μ'_I be given by conditioning μ_I on A (that is, $\mu'_I(X) = \frac{\mu(X \cap A)}{\mu(A)}$ for all X) and let μ'_J be given by conditioning μ_J on B . As R is stable, by Lemma 4.3 both μ'_I, μ'_J are finitely approximable on R . Hence, by Lemma 4.6 we can find some R -definable $A'_0 \supseteq A^0_{B,c,\varepsilon}, A'_1 \supset A^1_{B,c,\varepsilon}$ such that $\mu'_{I^c}(R_{a,c}) < 2\varepsilon$ for all $a \in A'_0$ and $\mu'_{I^c}(R_{a,c}) > (1 - 2\varepsilon)$ for all $a \in A'_1$ (in this case we are applying it to the complement R^c , which is also stable). As $\varepsilon < \frac{1}{4}$, it follows that in fact $A^0_{B,c,\varepsilon} = A'_0 \cap A, A^1_{B,c,\varepsilon} = A'_1 \cap A$. $\square$

In particular, it makes sense to speak of the μ_I -measure of $A^0_{B,c,\varepsilon}, A^1_{B,c,\varepsilon}$.

DEFINITION 4.8 Let $0 < \varepsilon \leq \delta < \frac{1}{4}$ be arbitrary, and let $I \subseteq [k]$. We say that a set $A \in \mathcal{B}_I$ is (ε, δ) -*excellent* if it is ε -good and for every $J \subseteq [k] \setminus I$, every δ -good $B \in \mathcal{B}_J$ and every $c \in V_{[k] \setminus (I \cup J)}$, either $\mu_I(A^0_{B,c,\delta}) < \varepsilon \mu_I(A)$ or $\mu_I(A^1_{B,c,\delta}) < \varepsilon \mu_I(A)$ (in the notation from Lemma 4.7).

REMARK 4.9 Note that if A is (ε, δ) -excellent, then it is also (ε', δ') -excellent for any $\varepsilon' > \varepsilon$ and $\varepsilon' \leq \delta' < \delta$. However, this need not be true if we take $\delta' > \delta$ since δ' -good sets need not be δ -good.

The following lemma is a generalization of [33, Claim 5.4], with an additional observation that the proof can be performed ‘definably’ and with respect to an arbitrary measure.

LEMMA 4.10 Let $R \subseteq V_1 \times \cdots \times V_k$ be d -stable and let $0 < \varepsilon \leq \delta < \frac{1}{2d}$ be arbitrary. Let $n \in [k]$. Assume that $A \in \mathcal{B}_n$ and $\mu_n(A) > 0$. Then there is an (ε, δ) -excellent R -definable set $A' \in \mathcal{B}_n$ with $\mu_n(A' \cap A) \geq \varepsilon^d \mu_n(A)$.

Proof. We will need the following claim.

Claim. Assume that $0 < \varepsilon \leq \delta < \frac{1}{4}$ and $A \in \mathcal{B}_n$ is not (ε, δ) -excellent. Then there are disjoint $A^0, A^1 \subseteq A$ with $A_i \in \mathcal{B}_n$ and $\mu(A_i) \geq \varepsilon \mu(A)$ for $i \in \{0, 1\}$, and such that for any finite $S^0 \subseteq A^0, S^1 \subseteq A^1$ with $|S^0| + |S^1| \leq \frac{1}{\delta}$ there is some $c \in V_{n^c}$ such that $a \in R_c$ for all $a \in S^1$ and $a \notin R_c$ for all $a \in S^0$. $\square$

Proof. If A is not ε -good, there is some $c \in V_{n^c}$ such that $\mu_n(A \cap R_c) \geq \varepsilon \mu_n(A)$ and $\mu_n(A \cap (R_c)^c) \geq \varepsilon \mu_n(A)$. We let $A^1 = A \cap R_c$ and $A^0 = A \cap (R_c)^c$.

If A is ε -good, as it is not (ε, δ) -excellent, there are some $J \subseteq [k] \setminus \{n\}$, some set $B \in \mathcal{B}_J$ which is δ -good, and some $c' \in V_{[k] \setminus (n \cup J)}$ such that A is a disjoint union of the sets $A^0 := A^0_{B,c',\delta}, A^1 := A^1_{B,c',\delta}$ (in the notation from Lemma 4.7) and $\mu_n(A^t) \geq \varepsilon \mu_n(A)$ for both $t \in \{0, 1\}$. Now given S^0, S^1 as in the claim, we have $\mu_J(B \cap R_{a,c'}) \leq \delta \mu_J(B)$ for all $a \in S^0$ and $\mu_J(B \cap (R_{a,c'})^c) \leq \delta \mu_J(B)$ for all $a \in S^1$. Let

$$B' = B \cap \left(\bigcup_{a \in S^0} R_{a,c'} \cup \bigcup_{a \in S^1} (R_{a,c'})^c \right).$$

As $|S^0| + |S^1| < \frac{1}{\delta}$, it follows that $\mu_J(B') < \frac{1}{\delta} \delta \mu_J(B) = \mu_J(B)$. In particular there is some $b' \in B \setminus B'$, and taking $c = b' \frown c'$ satisfies the claim.

Assume now that the conclusion of the lemma fails. By induction we choose sets $(A_\eta : \eta \in 2^{<d})$ in $\mathcal{B}_n$ such that $A_\emptyset = A$ and given $\eta \in 2^{<d}$, we take $A_{\eta \frown 0} := (A_\eta)^0, A_{\eta \frown 1} := (A_\eta)^1$ as given by the claim applied to A_η . For every $\eta \in 2^d$, pick some $a_\eta \in A_\eta$ (possible as $\mu_n(A_\eta) \geq \varepsilon^d \mu_n(A) > 0$). For every $\nu \in 2^{<d}$ there is some $c_\nu \in V_{n^c}$ such that $a_\eta \in R_{c_\nu}$ if and only if $\nu \frown 1 \leq \eta$ – which gives

contradiction to the d -stability of R . Namely we can take c given by the claim for A_ν and $S^0 = \{a_\eta : \eta \in 2^d, \nu \frown 0 \sqsubseteq \eta\}, S^1 = \{a_\eta : \eta \in 2^d, \nu \frown 1 \sqsubseteq \eta\}$ (note that $|S^0| + |S^1| \leq 2^d < \frac{1}{\delta}$ by assumption). $\square$

LEMMA 4.11 *Let $R \subseteq V_1 \times \cdots \times V_k$ be d -stable, and let $0 < \varepsilon \leq \delta < \frac{1}{2^d}$ be arbitrary. For any $n \in [k]$, there is a partition of V_n into (ε, δ) -excellent sets from $\mathcal{B}_n$, and the size of the partition can be bounded by a polynomial of degree $d+1$ in $\frac{1}{\varepsilon}$.*

Proof. Repeatedly applying Lemma 4.10, we let A_{m+1} be an $(\frac{\varepsilon}{2}, \delta)$ -excellent subset of $B_m := V_n \setminus (\bigcup_{1 \leq i \leq m} A_i)$ with $\mu_n(A_{m+1}) \geq (\frac{\varepsilon}{2})^d \mu_n(B_m)$. Then $\mu_n(B_{m+1}) \leq \mu_n(B_m) - (\frac{\varepsilon}{2})^d \mu_n(B_m) \leq (1 - (\frac{\varepsilon}{2})^d) \mu_n(B_m)$, hence $\mu_n(B_m) \leq (1 - (\frac{\varepsilon}{2})^d)^m$ for all m . Hence we have $\mu_n(B_m) \leq \frac{\varepsilon}{2} \mu_n(A_1)$ assuming $(1 - (\frac{\varepsilon}{2})^d)^m \leq (\frac{\varepsilon}{2})^{d+1}$. In this case, letting $A'_1 = A_1 \cup B_m$, it is easy to check that A'_1 is an (ε, δ) -excellent set, and $A'_1, A_2, \dots, A_m$ is a partition of V_n .

Finally, for the size of the partition, we have $(1 - (\frac{\varepsilon}{2})^d)^m \leq (\frac{\varepsilon}{2})^{d+1} \iff m \ln(1 - (\frac{\varepsilon}{2})^d) \leq (d+1) \ln \frac{\varepsilon}{2}$, and taking Taylor expansion this inequality holds provided $-m(\frac{\varepsilon}{2})^d \leq -(d+1)\frac{1}{(\frac{\varepsilon}{2})}$.

Hence we can take $m \leq c(\frac{1}{\varepsilon})^{d+1}$, for some $c = c(d)$. $\square$

Finally, we can use the partition in Lemma 4.11 to obtain a regular partition for $R \subseteq V_1 \times \cdots \times V_k$.

LEMMA 4.12 *Let $0 < \varepsilon \leq \delta < \frac{1}{2^d}$ be arbitrary. If $A \subseteq V_n$ is (ε, δ) -excellent and $B \subseteq V_{[n-1]}$ is δ -good then $B \times A$ is $(\varepsilon + \delta)$ -good.*

Proof. Let $c \in V_{[n]^c}$ be arbitrary. As B is δ -good and A is (ε, δ) -excellent, by Definition 4.8 we have $A = A_{B,c,\delta}^0 \cup A_{B,c,\delta}^1$ and either $\mu_n(A_{B,c,\delta}^0) < \varepsilon \mu_n(A)$ or $\mu_n(A_{B,c,\delta}^1) < \varepsilon \mu_n(A)$. Assume we are in the first case. Then, using the definition of $\mu_{[n]}$ and Lemma 2.15, we have

$$\begin{aligned} \mu_{[n]}((B \times A) \cap R_c) &= \int_A (\mu_{[n-1]}(R_{a,c} \cap B)) d\mu_n \\ &\geq \int_{A_{B,c,\delta}^1} (\mu_{[n-1]}(R_{a,c} \cap B)) d\mu_n \geq \int_{A_{B,c,\delta}^1} (1 - \delta) \mu_{[n-1]}(B) d\mu_n \\ &\geq (1 - \varepsilon) \mu_n(A) (1 - \delta) \mu_{[n-1]}(B) > (1 - (\varepsilon + \delta)) \mu_{[n]}(A \times B). \end{aligned}$$

Similarly, in the second case we obtain that $\mu_{[n]}((B \times A) \cap R_c) \leq (\varepsilon + \delta) \mu_{[n]}(A \times B)$. $\square$

THEOREM 4.13 *Let $R \subseteq V_1 \times \cdots \times V_k$ be d -stable, and let $0 < \varepsilon < \frac{1}{2^d}$ be arbitrary. Then there is an R -definable ε -regular partition $\vec{\mathcal{P}}$ of $V_1 \times \cdots \times V_k$ with 0-1-densities (see Definition 3.1) without any bad k -tuples in the partition (that is, $\Sigma = \emptyset$) and such that the size of the partition $\|\vec{\mathcal{P}}\|$ is bounded by a polynomial of degree $d+1$ in $\frac{1}{\varepsilon}$.*

Proof. For each $n \leq k$, let $\mathcal{P}_n$ be a partition of V_n into $(\frac{\varepsilon}{2^{k+1}}, \frac{\varepsilon}{2})$ -excellent sets as given by Lemma 4.11 (in particular, $\mathcal{P}_n$ has size polynomial in $\frac{1}{\varepsilon}$ as k is fixed), and let $\mathcal{P} := \{X_1 \times \cdots \times X_k : X_n \in \mathcal{P}_n\}$. We claim that $\mathcal{P}$ is ε -regular with $\Sigma = \emptyset$. Indeed, let $X = X_1 \times \cdots \times X_k \in \mathcal{P}$ be arbitrary, and let $X' := X_1 \times \cdots \times X_{k-1}$. Applying Lemma 4.12 k times, the set X' is $\frac{\varepsilon}{2}$ -good, and X_k is $(\frac{\varepsilon}{2}, \frac{\varepsilon}{2})$ -excellent

(by construction and Remark 4.9). Then, by Definition 4.8, X_k is a disjoint union of the sets $X_k^0 := (X_k)_{X', \frac{\varepsilon}{2}}^0$, $X_k^1 := (X_k)_{X', \frac{\varepsilon}{2}}^1 \in \mathcal{B}_k$ and $\mu_k(X_k^t) < \frac{\varepsilon}{2} \mu_k(X_k)$ for one of $t \in \{0, 1\}$. We have

$$\mu_{[k]}(R \cap X) = \int_{X_k} \mu_{[k-1]}(R_c \cap X') d\mu_k(c).$$

As X_k is a disjoint union of X_k^0, X_k^1 and $\mu(X_k^t) \leq \frac{\varepsilon}{2} \mu_k(X_k)$ for some $t \in \{0, 1\}$, we have

$$\left| \mu_{[k]}(R \cap X) - \int_{X_k^t} \mu_{[k-1]}(R_c \cap X') d\mu_k(c) \right| \leq$$

$$\frac{\varepsilon}{2} \mu_k(X_k) \mu_{[k-1]}(X_1 \times \cdots \times X_{k-1}) \leq \frac{\varepsilon}{2} \mu_{[k]}(X_1 \times \cdots \times X_k)$$

for some $t \in \{0, 1\}$.

Assume that $t = 0$. Then for all $c \in X_k^0$ we have $\mu_{[k-1]}(R_c \cap X') < \frac{\varepsilon}{2} \mu_{[k-1]}(X')$. Hence

$$\int_{X_k^0} \mu_{[k-1]}(R_c \cap X') d\mu_k(c) \leq \mu(X_k^0) \frac{\varepsilon}{2} \mu_{[k-1]}(X') \leq \frac{\varepsilon}{2} \mu_{[k]}(X_1 \times \cdots \times X_k),$$

and so $\mu_{[k]}(R \cap X) \leq \varepsilon \mu_{[k]}(X)$.

If $t = 1$, applying the same argument to the complement R^c we obtain $\mu_{[k]}(R^c \cap X) \leq \varepsilon \mu_{[k]}(X)$, hence $|\mu_{[k]}(R^c \cap X) - \mu_{[k]}(X)| \leq \varepsilon \mu_{[k]}(X)$. $\square$

Similarly to Corollary 3.8, Theorem 4.13 gives the following in the definable case. Recall that a structure $\mathcal{M}$ is *stable* if every binary relation definable in it is stable.

COROLLARY 4.14 *Let $\mathcal{M} = (M, \dots)$ be a stable structure and $k \geq 2$. For every definable $E(x_1, \dots, x_k)$ there is some $c = c(E)$ such that: for any $\varepsilon > 0$ and any Keisler measures μ_i on $M^{|x_i|}$ there are partitions $M^{|x_i|} = \bigcup_{j < K} A_{i,j}$ satisfying:*

- (1) $K \leq \left(\frac{1}{\varepsilon}\right)^c$;
- (2) for all $\vec{i} = (i_1, \dots, i_k) \in [K]^k$ we have

$$|\mu(E \cap (A_{1,i_1} \times \cdots \times A_{k,i_k})) - \delta_{\vec{i}} \mu(A_{1,i_1} \times \cdots \times A_{k,i_k})| < \varepsilon \mu(A_{1,i_1} \times \cdots \times A_{k,i_k})$$

for some $\delta_{\vec{i}} \in \{0, 1\}$ (where $\mu = \mu_1 \otimes \cdots \otimes \mu_k$);

- (3) each $A_{i,j}$ is defined by an instance of an E -formula, with this formula depending only on E and ε .

4.2. Distal case

The class of distal theories is defined and studied in [46], with the aim to isolate the class of purely unstable NIP theories (as opposed to the class of stable theories, see also [47]). For completeness of the exposition, we recall the distal regularity lemma established in [10], pointing out a stronger

form of definability for the regular partition than the one stated there. First we recall the definition of distality (and refer to the introduction in [10] for more details).

DEFINITION 4.15 [10] An NIP structure $\mathcal{M}$ is *distal* if and only if for every definable family $\{\phi(x, b) : b \in M^d\}$ of subsets of $M^{|x|}$ there is some $t \in \mathbb{N}$ and a definable family $\{\psi(x, c) : c \in M^{td}\}$ such that for every $a \in M$ and every finite set $B \subset M^d$ there is some $c \in B^t$ such that $a \in \psi(x, c)$ and for every $a' \in \psi(x, c)$ we have $a' \in \phi(x, b) \Leftrightarrow a \in \phi(x, b)$, for all $b \in B$.

THEOREM 4.16 Let $\mathcal{M}$ be distal and $k \geq 2$. For every definable $E(x_1, \dots, x_k)$, defined by an instance of some formula $\theta(x_1, \dots, x_k; z)$, there is some $c = c(\theta)$ such that: for any $\varepsilon > 0$ and any generically stable Keisler measures μ_i on $M^{|x_i|}$ there are partitions $M^{|x_i|} = \bigcup_{j < K} A_{i,j}$ and a set $\Sigma \subseteq \{1, \dots, K\}^k$ such that

- (1) $K \leq \left(\frac{1}{\varepsilon}\right)^c$.
- (2) $\mu\left(\bigcup_{(i_1, \dots, i_k) \in \Sigma} A_{1,i_1} \times \dots \times A_{k,i_k}\right) \leq \varepsilon$, where $\mu = \mu_1 \otimes \dots \otimes \mu_k$.
- (3) for all $(i_1, \dots, i_k) \notin \Sigma$, either $(A_{1,i_1} \times \dots \times A_{k,i_k}) \cap E = \emptyset$ or $A_{1,i_1} \times \dots \times A_{k,i_k} \subseteq E$.
- (4) Each $A_{i,j}$ is defined by an instance of a formula $\psi_i(x_i, z_i)$ which only depends on θ (and **not** on ε !).

Proof. This is proved in [10, Section 5.2], except for the fact that in (4) the formulas $\psi_i(x_i, z_i)$ can be chosen independently of ε —and we explain how to modify the proof there to obtain it. Namely, the proof of [10, Proposition 5.3] shows that, under the assumptions of the lemma, for each $i = 1, \dots, k$ we can find a finite set of formulas Δ_i and a constant $c \in \mathbb{N}$ depending *only* on θ (in view of [10, Corollary 4.6]), a finite set of parameters A_N depending on θ and ε with $|A_N| \leq \left(\frac{1}{\varepsilon}\right)^c$, and partitions $\mathcal{P}_i = \{A_{i,j} : j < K\}$ of $M^{|x_i|}$ satisfying the conclusion of the lemma, except for the bold font part, such that each $A_{i,j}$ is Δ_i -definable over A_N .

Let $\mathcal{Q}_i$ be a partition of $M^{|x_i|}$ into the sets of realizations of complete Δ_i -types over A_N . By distality of $\mathcal{M}$, let Δ'_i be a finite set of formulas such that for every $\phi \in \Delta_i$ it contains a formula ψ as in Definition 4.15. Let $\psi_i(x_i, z_i)$ be a conjunction of all formulas in Δ'_i . Then for every $a \in M^{|x_i|}$ there is a single instance $\psi_i(x_i, e)$ such that its parameters e are all from A_N and such that $\psi_i(x_i, e)$ isolates the complete Δ_i -type of a over A_N . Using this, we can choose a partition $\mathcal{Q}'_i$ of $M^{|x_i|}$ which refines $\mathcal{Q}_i$ (and so also refines $\mathcal{P}_i$) and such that every set in $\mathcal{Q}'_i$ is defined by an instance of $\phi_i(x_i, z_i)$ over A_N . Then the size of $\mathcal{Q}'_i$ is bounded by $|A_N|^{|z_i|} \leq \left(\frac{1}{\varepsilon}\right)^{c'}$ where $c' = c|z_i|$ only depends on θ . Hence $\mathcal{Q}'_i, i = 1, \dots, k$ give the desired partition. $\square$

5. Definable variants of the Erdős–Hajnal and Röd1 theorems

In this section, we are concerned with the question of finding a ‘large’ ‘approximately homogeneous’ definable subset of a definable hypergraph. ‘Large’ here refers to positive measure, relatively to a finitely approximable measure, and ‘approximately homogeneous’ means that the edge density on the set is close to 0 or 1 (see below for precise definitions). We consider two very different situations—(k -partite) k -hypergraphs and k -uniform hypergraphs (in the sense of Section 3.1).

5.1. Partitioned hypergraphs

First we consider the ‘partite’ situation. We are working in the same setting as in Section 3.1.

THEOREM 5.1 *Let $E \subseteq V_1 \times \cdots \times V_k$ be a k -hypergraph of VC-dimension at most d . Then for every $\alpha, \varepsilon > 0$ there is some $\delta = \delta(k, d, \alpha, \varepsilon) > 0$ such that the following holds.*

Let $\mathcal{B}_i$ be a field on V_i , and let μ_i be a measure on $\mathcal{B}_i$ which is finitely approximable on E , for $i = 1, \dots, k$. Let $\mu = \mu_1 \times \cdots \times \mu_k$. Assume that $\mu(E) \geq \alpha$. Then there are some E -definable sets $A_i \subseteq V_i$ such that $\mu_i(A_i) > \delta$ for all $i = 1, \dots, k$ and $d_E(A_1, \dots, A_k) > 1 - \varepsilon$ (where $d_E(A_1, \dots, A_k) = \frac{\mu(E \cap (A_1 \times \cdots \times A_k))}{\mu_1(A_1) \cdots \mu_k(A_k)}$ denotes the E -density).

Proof. This follows from the regularity lemma for NIP hypergraphs (Theorem 3.3). Let $\alpha, \varepsilon > 0$ and $d \in \mathbb{N}$ be given. Let $\varepsilon' = \frac{\min\{\alpha, \varepsilon\}}{4} > 0$. By Theorem 3.3 there exist $c_1, c_2 \in \mathbb{R}$ depending only on k, d , E -definable partitions $V_i = \bigcup_{j=1, \dots, n} A_{i,j}$ for each $i = 1, \dots, k$ with $n \leq c_1 (\frac{1}{\varepsilon'})^{c_2}$, $\delta_{\vec{j}} \in \{0, 1\}$ for each $\vec{j} \in [n]^k$ and $\Sigma \subseteq [n]^k$ such that $\sum_{(j_1, \dots, j_k) \in \Sigma} \mu_1(A_{1,j_1}) \cdots \mu_k(A_{k,j_k}) < \varepsilon'$ and

$$\begin{aligned} & \left| \mu(E \cap (A_{1,j_1} \times \cdots \times A_{k,j_k})) - \delta_{\vec{j}} \mu(A_{1,j_1} \times \cdots \times A_{k,j_k}) \right| \\ & < \varepsilon' \mu(A_{1,j_1} \times \cdots \times A_{k,j_k}) \end{aligned}$$

for all $\vec{j} = (j_1, \dots, j_k) \in [n]^k \setminus \Sigma$. In particular, if $\vec{j} \in [n]^k \setminus \Sigma$, $\delta_{\vec{j}} = 1$ and $\mu(A_{1,j_1} \times \cdots \times A_{k,j_k}) > 0$ then $d_E(A_{1,j_1}, \dots, A_{k,j_k}) > 1 - \varepsilon' > 1 - \varepsilon$.

Let $\delta := \frac{\varepsilon'}{c_1^k (\frac{1}{\varepsilon'})^{kc_2}} > 0$, it only depends on $k, d, \alpha, \varepsilon$. To prove the theorem, it is thus sufficient to show that there exists $\vec{j} \in [n]^k \setminus \Sigma$ so that $\delta_{\vec{j}} = 1$ and $\mu(A_{1,j_1} \times \cdots \times A_{k,j_k}) > \delta$ (which automatically implies $\mu_i(A_{i,j_i}) > \delta$ for all $i \in [k]$ as each μ_i takes values in $[0, 1]$). Assume that this fails. Then we have:

$$\begin{aligned} \mu(E) &= \sum_{\vec{j} \in [n]^k} \mu(E \cap (A_{1,j_1} \times \cdots \times A_{k,j_k})) \\ &\leq \sum_{\vec{j} \in \Sigma} \mu(A_{1,j_1} \times \cdots \times A_{k,j_k}) \\ &\quad + \sum_{\vec{j} \in [n]^k \setminus \Sigma, \mu(A_{1,j_1} \times \cdots \times A_{k,j_k}) \leq \delta} \mu(A_{1,j_1} \times \cdots \times A_{k,j_k}) \\ &\quad + \sum_{\vec{j} \in [n]^k \setminus \Sigma, \mu(A_{1,j_1} \times \cdots \times A_{k,j_k}) > \delta} \mu(A_{1,j_1} \times \cdots \times A_{k,j_k}) \varepsilon' \\ &\leq \varepsilon' + n^k \delta + \varepsilon' \leq 2\varepsilon' + \left(c_1 \left(\frac{1}{\varepsilon'} \right)^{c_2} \right)^k \delta, \end{aligned}$$

which by the choice of δ is at most $3\varepsilon'$ (in order to bound the third summand by ε' we use that μ is a probability measure and the sets in the sum come from a partition of $V_1 \times \cdots \times V_k$). But this contradicts the assumption that $\mu(E) \geq \alpha \geq 4\varepsilon'$. $\square$

REMARK 5.2 In the special case when μ is an ultraproduct of counting measures concentrated on finite sets, this gives a density version of the well-known lemma of Erdős and Hajnal, see, for example [19, Lemma 2.1]

In particular, the result holds when E is a definable relation in an NIP structure (see Section 3.3), giving uniform definability of the sets A_i in terms of E, α, ε .

In the case when E is definable in a distal structure we have the following strengthening proved in [10, Corollary 4.6].

FACT 5.3 *Let $\mathcal{M}$ be a distal structure and $\theta(x_1, \dots, x_k, y)$ a formula. Given $\alpha > 0$ there is $\delta > 0$ such that: for any relation $E(x_1, \dots, x_k)$ defined by an instance of θ and any generically stable measures μ_i on $M^{|x_i|}$, if $\mu(R) \geq \alpha$ (where $\mu = \mu_1 \otimes \dots \otimes \mu_k$), then there are definable sets $A_i \subseteq M^{|x_i|}$ with $\mu_i(A_i) \geq \delta$ for all $i = 1, \dots, k$ and $\prod_{i=1}^k A_i \subseteq R$. Moreover, each A_i can be defined by an instance of a formula $\psi_i(x_i, z_i)$ that depends only on θ and α .*

5.2. Non-partitioned case

In the non-partite case, however, it is much harder to find a large homogeneous subset (that is, a clique or an anti-clique), as it is well known in combinatorics, and we give some examples in the definable setting illustrating it.

The following is a classical result of Rödl.

FACT 5.4 ([41], see also [19, Theorem 1.1]) *For each $\varepsilon \in (0, \frac{1}{2})$ and finite graph H there is some $\delta = \delta(H, \varepsilon) > 0$ such that every H -free graph on n vertices contains an induced subgraph on at least δn vertices with edge density either at most ε or at least $1 - \varepsilon$.*

We consider a generalization of this property to finitely approximable measures.

DEFINITION 5.5 Let $\mathcal{M}$ be a structure and let $\mathfrak{M}$ be a class of Keisler measures. Let $\mathcal{E}$ be a collection of definable (symmetric) (hyper-)graphs in (some powers of) $\mathcal{M}$.

- (1) We will say that $\mathcal{E}$ satisfies the *Rödl property* with respect to $\mathfrak{M}$ if for every $E \subseteq (M^n)^k$ in $\mathcal{E}$ and every $\varepsilon > 0$ there is some $\delta > 0$ such that for every $\mu \in \mathfrak{M}$, a Keisler measure on M^n which is finitely approximable on E , there is some definable $A \subseteq M^n$ such that $\mu(A) \geq \delta$ and the $\mu^{(k)}$ -density of E on A is either $< \varepsilon$ or $> 1 - \varepsilon$.
- (2) If in addition such an A can be defined by an instance of some formula that depends only on E , and not on ε , then we say that $\mathcal{E}$ satisfies the *uniform Rödl property* with respect to $\mathfrak{M}$.
- (3) We will say that $\mathcal{E}$ satisfies the *strong Rödl property* with respect to $\mathfrak{M}$ if in (1) we can find a definable E -homogeneous subset of positive μ -measure.

Fact 5.4 implies that if $\mathcal{E}$ is a family of pseudofinite hypergraphs of bounded VC-dimension, then it satisfies the Rödl property with respect to the class $\mathfrak{M}$ of pseudofinite counting measures, in the language of set theory. We give some examples showing that there is little hope in generalizing this to arbitrary generically stable measures.

EXAMPLE 5.6 The strong Rödl property does not hold for graphs definable in the field of reals, with respect to the Lebesgue measure. To see this, consider the relation $E \subseteq \mathbb{R}^2 \times \mathbb{R}^2$ defined by $(a, b)E(a', b') \iff |a - a'| < |b - b'|$, and let μ be the generically stable measure on $\mathbb{R}^2$ given by restricting the Lebesgue measure on $[0, 1]^2$ to the definable sets. We claim that there is no definable E -homogeneous subset of $\mathbb{R}^2$ of positive measure. Indeed, any such set $A \subseteq [0, 1]^2$ would have to contain an E -homogeneous square, and it is easy to see that this is impossible by the definition of E (one can check, however, that the uniform Rödl property is satisfied as for any $\varepsilon > 0$ we can choose a sufficiently thin vertical stripe of positive measure such that the E -density on it is ε -close to 1).

It may be tempting to use the NIP regularity lemma as in the partitioned case (Theorem 5.1) to establish the Rödl property (applying it for a symmetric relation $R \subseteq V_1 \times V_2$ with $V = V_1 = V_2$, $\mu = \mu_1 = \mu_2$). However, it does not work. The reason is that, given an ε -regular partition $A_1, \dots, A_n$ of V , it is perfectly possible that all of the pairs on the diagonal (A_i, A_i) , $1 \leq i \leq n$ are bad simultaneously. Namely, if Σ is the collection of all bad pairs, we have that $\sum_{(i,j) \in \Sigma} \mu(A_i)\mu(A_j) < \varepsilon$. On the other hand, if let us say $(A_i : 1 \leq i \leq n)$ is an equipartition, we have $\sum_{1 \leq i \leq n} \mu(A_i)^2 \leq n \frac{1}{n^2} \leq \frac{1}{n}$, which can be smaller than ε when n is sufficiently large. In fact, this observation suggests an idea of a counterexample to the uniform Rödl property, which we present in the next subsection.

A counterexample to the uniform Rödl property

Throughout this section we are working in the field of 2-adics $\mathbb{Q}_2$, viewed as a first-order structure

$$\mathcal{M} = (\mathbb{Q}_2, 0, 1, +, \cdot, v(x) \leq v(y), (P_n(x))_{n \in \mathbb{N}_{\geq 2}})$$

with the universe $M = \mathbb{Q}_2$ in the Macintyre language (so $v(x) \leq v(y)$ is a binary predicate comparing the 2-adic valuations of x and y , and $P_n(x) \iff \exists y(x = y^n)$). Let μ be the Haar measure on $\mathbb{Q}_2$ normalized on the compact ball $\mathbb{Z}_2$ (restricted to definable sets). Then $\mathbb{Q}_2$ is a distal structure, and μ is a generically stable measure (for example, see the introduction in [10]). We think of elements in $\mathbb{Q}_2$ as branches of a binary tree and define $E \subseteq M^2$ by saying that $E(x, y)$ holds if and only if $v(x - y)$ is odd (that is, if the branches x and y split at an odd level). This is a symmetric relation definable in the Macintyre's language. We estimate the $\mu^{(2)}$ -density of E on certain definable sets.

LEMMA 5.7 *Assume that A is a (valuational) ball, then the density $d_E(A)$ is either $\frac{1}{3}$ or $\frac{2}{3}$ (depending on the radius of the ball).*

Proof. Let $A = \{b \in \mathbb{Q}_2 : v(c - b) \geq r\}$ be a (clopen) ball with center c of radius r , for some $r \in \mathbb{Z}$. We think of the elements of $\mathbb{Q}_2$ as bi-infinite binary sequences (with the set of non-zero entries containing the least element). Then we can write $A = \{\tau_0 \frown \tau : \tau \in 2^\omega\}$ for some $\tau_0 \in 2^{\mathbb{Z}_{\leq r}}$, where ‘ $\frown$ ’ denotes the concatenation of sequences. For each $n \in \omega$, consider the partition $A = \bigcup_{\sigma \in 2^n} A_\sigma$, where $A_\sigma = \{\tau_0 \frown \sigma \frown \tau : \tau \in 2^\omega\}$. By translation-invariance of the Haar measure we have $\mu(A_\sigma) = \frac{1}{2^n} \mu(A)$ for all $\sigma \in 2^n$, $n \in \omega$.

We would like to calculate the density $d_E(A) = \frac{\mu^{(2)}(E \cap A^2)}{\mu(A)^2}$. For $n \in \mathbb{N}_{\geq 1}$, we consider the definable set

$$F_n := \left\{ (a, a') \in A^2 : \bigvee_{\sigma \in 2^{n-1}} \bigvee_{t \in \{0,1\}} (a \in A_{\sigma \frown t} \wedge a' \in A_{\sigma \frown (t-1)}) \right\}.$$

From the definition we have:

- if $(a, a') \in F_n$ then $v(a - a') = r + n - 1$;
- in particular, if $r + n$ is even then $F_n \subseteq (E \cap A^2)$;
- if $r + n$ is odd then $F_n \subseteq (A^2 \setminus E)$;
- $F_n \cap F_{n'} = \emptyset$ for any $n \neq n' \in \mathbb{N}_{\geq 1}$;
- $\mu^{(2)}(F_n) = \frac{1}{2^n} \mu(A)^2$.

(Indeed, $\mu^{(2)}(F_n) = \sum_{\sigma \in 2^{n-1}} (\mu^{(2)}((A_{\sigma \frown 0} \times A_{\sigma \frown 1}) \cup (A_{\sigma \frown 1} \times A_{\sigma \frown 0})))$, and using that $(A_\sigma)_{\sigma \in 2^n}$ is a partition, $= 2^{n-1} \cdot 2 \cdot \left(\frac{1}{2^n} \mu(A)\right)^2 = \frac{1}{2^n} \mu(A)^2$.)

Using these observations we obtain the following estimates.

(1) If r is odd, then

$$\sum_{n \geq 1 \text{ odd}} \mu^{(2)}(F_n) \leq \mu^{(2)}(E \cap A) = d_E(A) \mu(A)^2 \leq \mu(A)^2 - \left(\sum_{n \geq 1 \text{ even}} \mu^{(2)}(F_n) \right).$$

We have:

$$\begin{aligned} \sum_{n \geq 1 \text{ odd}} \mu^{(2)}(F_n) &= \sum_{n \geq 1 \text{ odd}} \frac{1}{2^n} \mu(A)^2 = \mu(A)^2 \sum_{m \geq 0} \frac{1}{2^{2m+1}} \\ &= \mu(A)^2 \sum_{m \geq 0} \frac{1}{2} \frac{1}{4^m} = \mu(A)^2 \frac{1/2}{1 - 1/4} = \frac{2}{3} \mu(A)^2, \end{aligned}$$

$$\begin{aligned} \sum_{n \geq 1 \text{ even}} \mu^{(2)}(F_n) &= \sum_{n \geq 1 \text{ even}} \frac{1}{2^n} \mu(A)^2 = \mu(A)^2 \sum_{m \geq 1} \frac{1}{2^{2m}} \\ &= \mu(A)^2 \left(\sum_{m \geq 0} \frac{1}{4^m} - 1 \right) = \left(\frac{1}{1 - 1/4} - 1 \right) \mu(A)^2 = \frac{1}{3} \mu(A)^2. \end{aligned}$$

Combining we get $d_E(A) = \frac{2}{3}$.

(2) If r is even, then

$$\sum_{n \geq 1 \text{ even}} \mu^{(2)}(F_n) \leq \mu^{(2)}(E \cap A) = d_E(A) \mu(A)^2 \leq \mu(A)^2 - \left(\sum_{n \geq 1 \text{ odd}} \mu^{(2)}(F_n) \right),$$

and a similar computation shows that $d_E(A) = \frac{1}{3}$.

□

LEMMA 5.8 *Fix a formula $\phi(x, y)$. Then there is some $\gamma \in (0, 1)$ such that: for any tuple of parameters $b \in M^{|y|}$, if $\mu(\phi(x, b)) > 0$, then $\phi(x, b)$ contains some ball B with $\mu(B) \geq \gamma \mu(\phi(x, b))$.*

Proof. If $\mu(\phi(x, b)) > 0$, then $\phi(x, b)$ has to be infinite. As demonstrated in the original paper of Macintyre [32, Theorem 2], every infinite definable subset of M in the 2-adics has non-empty interior. In particular, it must contain some open ball of positive Haar measure.

However, to prove the claim we need a slightly more careful analysis. We recall a couple of facts about the 2-adic cell decomposition (see, for example, [5, Section 7]). Let $\phi(x, y)$ be fixed. Then there is some $N \in \mathbb{N}$, definable functions f_i, g_i and elements $\lambda_i \in M$ for $i \leq N$ such that for every $b \in M^{[y]}$, the set $\phi(M, b)$ is a union of at most N cells of the form

$$U_i(b) = \{x \in M; v(f_i(b)) \leq v(x - c_i(b)) < v(g_i(b)) \wedge P_{n_i}(\lambda_i(x - c_i(b)))\}.$$

Besides, we have the following fact.

FACT 5.9 (see, for example, [5, Lemma 7.4]) *Suppose $n > 1$, and let $x, y, a \in M$ be such that $v(y - x) > 2v(n) + v(y - a)$. Then $x - a$ and $y - a$ are in the same coset of P_n .*

Assume now that $\mu(\phi(x, b)) > 0$. Then for at least one $i \leq N$, the corresponding cell $U_i(b) \subseteq \phi(M, b)$ satisfies $\mu(U_i(b)) \geq \frac{1}{N}\mu(\phi(x, b))$. Let $U(b) := U_i(b)$, $f := f_i$, $g := g_i$, $n := n_i$, $c := c_i$, $\lambda := \lambda_i$.

We claim that there is some element $a \in U(b)$ with $v(a - c(b)) \leq v(f(b)) + n$. First, as the valuation group of $\mathbb{Q}_2$ is $\mathbb{Z}$, there must exist some $\beta \in \mathbb{Z}$ such that $v(f(b)) \leq \beta \leq v(f(b)) + n$ and $v(\lambda) + \beta = n\alpha$ for some $\alpha \in \mathbb{Z}$. Let $e \in M$ be arbitrary with $v(e) = \alpha$, and let $a := \frac{e^n}{\lambda} + c(b) \in M$. Then:

- (1) $\lambda(a - c(b)) = e^n$ (in particular $P_n(\lambda(a - c(b)))$ holds),
- (2) $v(a - c(b)) = v(\frac{e^n}{\lambda}) = v(e^n) - v(\lambda) = nv(e) - v(\lambda) = n\alpha - v(\lambda) = \beta$.

Now either $\beta < v(g(b))$, in which case $a \in U(b)$, or $\beta \geq v(g(b))$, in which case any element in $U(b)$ satisfies the claim.

Now we consider the ball $B := B_{\geq m}(a)$ for $m := 2v(n) + v(f(b)) + n + 1$. We claim that $B \subseteq U(b)$. Indeed, for any $x \in B$ we have $v(a - x) > 2v(n) + v(a - c(b))$, hence by Fact 5.9, $x - c(b)$ and $a - c(b)$ are in the same coset of P_n , and of course $v(x - c(b)) = v(a - c(b))$, so $x \in U(b)$.

Finally, as all balls of a given positive radius have the same Haar measure, we have $\mu(B) \geq \frac{1}{2^{2v(n)+n+1}}\mu(B_{\geq v(f(b))}(c(b)))$ and, as $U(b) \subseteq B_{\geq v(f(b))}(c(b))$, we have $\mu(U(b)) \leq \mu(B_{\geq v(f(b))}(c(b)))$. Hence

$$\mu(B) \geq \frac{1}{2^{2v(n)+n+1}}\mu(U(b)) \geq \left(\frac{1}{2^{2v(n)+n+1}} \cdot \frac{1}{N} \right) \mu(\phi(M, b)).$$

Note that the coefficient only depends on $\phi(x, y)$, and not on the choice of the parameter b . $\square$

We show that the uniform Rödl property fails for E . Assume toward contradiction that we can find some $\phi(x, y)$ such that for every $\varepsilon > 0$ there is some set $A \subseteq M$ definable by an instance of $\phi(x, y)$ and satisfying $\mu(A) > 0$ and $d_E(A) \in [0, \varepsilon) \cup (1 - \varepsilon, 1]$. Let us say $d_E(A) > 1 - \varepsilon$ (if $d_E(A) < \varepsilon$, we work with the complement of E instead). Let $\gamma > 0$ be as given by Lemma 5.8 for $\phi(x, y)$, and fix some $\varepsilon < \gamma$.

Now A contains some ball B with $\mu(B) = \delta\mu(A)$ for some $0 < \gamma \leq \delta \leq 1$, and we estimate the number of edges on A using Lemma 5.7.

$$\begin{aligned}
\mu^{(2)}(E \cap A^2) &= \mu^{(2)}(E \cap B^2) + \mu^{(2)}(E \cap (A \setminus B)^2) + 2\mu^{(2)}(E(A \setminus B, B)) \\
&\leq \frac{2}{3}\mu(B)^2 + \mu(A \setminus B)^2 + 2\mu(A \setminus B)\mu(B) \\
&= \frac{2}{3}\delta^2\mu(A)^2 + (1 - \delta)^2\mu(A)^2 + 2(1 - \delta)\delta\mu(A)^2 \\
&= \left(\frac{2}{3}\delta^2 + 1 - 2\delta + \delta^2 + 2\delta - 2\delta^2\right)\mu(A)^2 \\
&= \left(1 - \frac{1}{3}\delta^2\right)\mu(A)^2 \leq \left(1 - \frac{1}{3}\gamma^2\right)\mu(A)^2.
\end{aligned}$$

But as we have assumed $\varepsilon < \gamma \in (0, 1)$, this contradicts the assumption that $d_E(A) > 1 - \varepsilon$.

Uniform Rödl property fails for semialgebraic hypergraphs

It is well known that Fact 5.4 fails for hypergraphs (see the example at the very end of [41]). We observe that the uniform Rödl property fails already in the case of 3-hypergraphs in the *semialgebraic* setting.

For this, let $E(x_1, x_2, x_3) \subseteq \mathbb{R}^3$ be the relation given by $(x_1 < x_2 < x_3) \wedge (x_1 + x_3 - 2x_2 \geq 0)$, it is definable in the field of reals (it is considered in [13, Section 3.1]). We claim that it does not satisfy the uniform Rödl property relatively to the class of measures concentrated on finite sets. If we assume that it holds, then by o -minimality for every $\varepsilon > 0$ there is some $\delta > 0$ such that for any finite set $A \subseteq \mathbb{R}$ there is some interval $B \subseteq \mathbb{R}$ such that for $C = A \cap B$ we have $d_E(C) > 1 - \varepsilon$ or $d_E(C) < \varepsilon$. We observe that in fact the E -density tends to be $\frac{1}{2}$. Let arbitrary $\varepsilon < \frac{1}{2}$ and $\delta > 0$ be fixed. Let us take $A = \{1, 2, 3, \dots, N\}$ for some $N \in \mathbb{N}$ large enough (such that δN is also large), and let $C \subseteq A$, $C = \{p_1, \dots, p_n\}$ be an arbitrary interval of integers in A , $p_1 < \dots < p_n$, $|C| \geq \delta N$.

Assume that $E(p_i, p_j, p_k)$ does not hold for some $p_1 < i < j < k < p_n$. Let us define $q_i := p_n - p_{n-i+1} + p_1$. Then we have $p_1 < q_i < q_j < q_k < p_n$, and q_i, q_j, q_k are all in C since C is an interval. Moreover it's easy to see that $E(q_i, q_j, q_k)$ holds. This establishes a bijection between edges and non-edges in C , showing that the density on C is arbitrary close to $1/2$ for N large enough.

Acknowledgement

We would like to thank the referee, Matthias Aschenbrenner and Roland Walker for their very helpful comments on the earlier versions of the article.

Funding

Chernikov was supported by the NSF Research Grant DMS-1600796, by the NSF CAREER grant DMS-1651321 and by an Alfred P. Sloan Fellowship. Starchenko was supported by the NSF Research Grant DMS-1500671.

References

1. N. Ackerman, C. Freer and R. Patel, *Stable regularity for relational structures*, preprint, arXiv:1712.09305, 2017.
2. H. Adler, An introduction to theories without the independence property, preprint 2008, <http://www.logic.univie.ac.at/~adler/docs/nip.pdf>
3. N. Alon, E. Fischer and I. Newman, Efficient testing of bipartite graphs for forbidden induced subgraphs, *SIAM J. Comput.* **37** no. 3 (2007), 959–976.
4. M. Aschenbrenner, A. Chernikov, A. Gehret and M. Ziegler, *Distality in valued fields and related structures*, preprint, arXiv:2008.09889, 2020.
5. M. Aschenbrenner, A. Dolich, D. Haskell, D. Macpherson and S. Starchenko, Vapnik-Chervonenkis density in some theories without the independence property, I, *Trans. Amer. Math. Soc.* **368** no. 8 (2015), 5889–5949.
6. M. Aschenbrenner, L. van den Dries and J. van der Hoeven, *Asymptotic Differential Algebra and Model Theory of Transseries*, Vol. 195, Annals of Mathematics Studies, Princeton University Press, Princeton, NJ, 2017, xxi+849.
7. K. P. S. Bhaskara Rao and M. Bhaskara Rao, *Theory of Charges*, Pure and Applied Mathematics, Vol. 109, Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], New York, 1983, A study of finitely additive measures, With a foreword by D. M. Stone.
8. A. Chernikov, Model theory, Keisler measures, and groups, *Bull. Symb. Log.* **24** no. 3 (2018), 336–339.
9. Artem Chernikov and Sergei Starchenko, *Definable regularity lemmas for NIP hypergraphs*, preprint, arXiv:1607.07701v1, 2016.
10. A. Chernikov and S. Starchenko, Regularity lemma for distal structures, *J. Eur. Math. Soc.* **20** no. 10 (2018), 2437–2466.
11. A. Chernikov and H. Towsner, *Hypergraph regularity and higher arity VC-dimension*, preprint, arXiv:2010.00726, 2020.
12. G. Conant and K. Gannon, Remarks on generic stability in independent theories, *Ann. Pure Appl. Logic* **171** no. 2 (2020), 102736.
13. D. Conlon, J. Fox, J. Pach, B. Sudakov and A. Suk, Ramsey-type results for semi-algebraic relations, *Trans. Am. Math. Soc.* **366** no. 9 (2014), 5043–5065.
14. G. Elek and B. Szegedy, *Limits of hypergraphs, removal and regularity lemmas. a non-standard approach*, preprint, arXiv:0705.2179, 2007.
15. J. Fox, M. Gromov, V. Lafforgue, A. Naor and János Pach, Overlap properties of geometric expanders, *J. für die reine und angewandte Mathematik (Crelles J.)* **2012** no. 671 (2012), 49–83.
16. J. Fox, J. Pach and A. Suk, A polynomial regularity lemma for semialgebraic hypergraphs and its applications in geometry and property testing, *SIAM J. Comput.* **45** no. 6 (2016), 2199–2223.
17. J. Fox, J. Pach and A. Suk, *Erdős-Hajnal conjecture for graphs with bounded VC-dimension*, preprints, arXiv:1710.03745v1, 2017.
18. J. Fox, J. Pach and A. Suk, Erdős–Hajnal conjecture for graphs with bounded VC-dimension, *Discrete Comput. Geom.* **61** no. 4 (2019), 809–829.
19. J. Fox and B. Sudakov, Induced Ramsey-type theorems, *Adv. Math.* **219** no. 6 (2008), 1771–1800.
20. D. García, Dugald Macpherson, and Charles Steinhorn. Pseudofinite structures and simplicity, *J. Math. Log.* **15** no. 1 (2015), 1550002.

21. W. T. Gowers, Hypergraph regularity and the multidimensional Szemerédi theorem, *Ann. Math.* **166** no. 3 (2007), 897–946.
22. D. Haskell, E. Hrushovski and D. Macpherson, *Stable Domination and Independence in Algebraically Closed Valued Fields*, Lecture Notes in Logic, 30, Association for Symbolic Logic, Chicago, IL, Cambridge University Press, Cambridge, 2007, xii+182.
23. W. Hodges, *Model Theory*, Vol. 42, Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1993, xiv+772.
24. E. Hrushovski, *Notes on Tao's algebraic regularity lemma*, Unpublished, 2013.
25. E. Hrushovski and A. Pillay, On NIP and invariant measures, *J. Eur. Math. Soc.* **13** no. 4 (2011), 1005–1061.
26. E. Hrushovski, A. Pillay and P. Simon, Generically stable and smooth measures in NIP theories, *Trans. Am. Math. Soc.* **365** no. 5 (2013), 2341–2366.
27. H. J. Keisler, Measures and forking, *Ann. Pure Appl. Logic* **34** no. 2 (1987), 119–169.
28. J. Komlós, J. Pach and G. Woeginger, Almost tight bounds for ϵ -nets, *Discrete Comput. Geom.* **7** no. 2 (1992), 163–173.
29. J. Komlós and M. Simonovits, *Szemerédi's regularity lemma and its applications in graph theory*, 1996.
30. L. Lovász and B. Szegedy, Szemerédi's lemma for the analyst, *GAFA Geom. Funct. Anal.* **17** no. 1 (2007), 252–270.
31. L. Lovász and B. Szegedy, *An Irregular Mind*, Regularity partitions and the topology of graphons, Bolyai Soc. Math. Stud. Vol. 21, János Bolyai Math. Soc., Budapest, 2010, 415–446.
32. A. Macintyre, On definable subsets of p -adic fields, *J. Symb. Log.* **41** no. 3 (1976), 605–610.
33. M. Malliaris and S. Shelah, Regularity lemmas for stable graphs, *Trans. Am. Math. Soc.* **366** no. 3 (2014), 1551–1585.
34. M. Malliaris and A. Pillay, The stable regularity lemma revisited, *Proc. Am. Math. Soc.* **144** no. 4 (2016), 1761–1765.
35. J. Matoušek, *Lectures on Discrete Geometry* Vol. 108, Springer, New York, 2002.
36. G. Moshkovitz and A. Shapira, A short proof of Gower' lower bound for the regularity lemma, *Combinatorica* **36** no. 2 (2013), 187–194.
37. B. Nagle, V. Rödl and M. Schacht, The counting lemma for regular k -uniform hypergraphs, *Rand. Struct. Algorithms* **28** no. 2 (2006), 113–179.
38. A. Pillay, *Geometric Stability Theory*, Oxford Logic Guides, Vol. 32, Oxford Science Publications, The Clarendon Press, Oxford University Press, New York, 1996, x+361.
39. A. Pillay and S. Starchenko, *Remarks on Tao's algebraic regularity lemma*, Preprint, arXiv:1310.7538, 2013.
40. K.-P. Podewski and M. Ziegler, Stable graphs, *Fund. Math.* **100** no. 2 (1978), 101–107.
41. V. Rödl, On universality of graphs with uniformly distributed edges, *Discrete Math.* **59** no. 1 (1986), 125–134.
42. V. Rödl and J. Skokan, Regularity lemma for k -uniform hypergraphs, *Rand. Struct. Algorithms* **25** no. 1 (2004), 1–42.
43. Z. Sela, Diophantine geometry over groups VIII: stability, *Ann. Math.* **177** no. 3 (2013), 787–868.
44. S. Shelah, *Classification Theory and the Number of Nonisomorphic Models*, Studies in Logic and the Foundations of Mathematics, 2nd edn, Vol. 92, North-Holland Publishing Co., Amsterdam, 1990.
45. S. Shelah, *Dependent dreams: recounting types*, arXiv preprint arXiv:1202.5795, 2012.

46. P. Simon, Distal and non-distal NIP theories, *Ann. Pure Applied Log.* **164** no. 3 (2013), 294–318.
47. P. Simon, *A guide to NIP Theories*, Lecture Notes in Logic, Vol, 44, Association for Symbolic Logic, Chicago, IL, Cambridge Scientific Publishers, Cambridge, 2015, vii+156.
48. P. Simon, A note on “regularity lemma for distal structures”, *Proc. Am. Math. Soc.* **144** no. 8 (2016), 3573–3578.
49. P. Simon, *Type decomposition in NIP theories*, arXiv preprint arXiv:1604.03841, 2016.
50. S. Starchenko, (2016), *NIP, Keisler measures and combinatorics*, *Séminaire BOURBAKI*, 68eme année, 2015-2016, no 1114, <http://www.bourbaki.ens.fr/TEXTES/1114.pdf>.
51. T. Tao, Szemerédi’s regularity lemma revisited, *Contrib. Discrete Math.* **1** no. 1 (2006), 8–28.
52. T. Tao, Expanding polynomials over finite fields of large characteristic, and a regularity lemma for definable sets, *Contrib. Discrete Math.* **10** no. 1 (2015), 22–98.
53. K. Tent and M. Ziegler, *A Course in Model Theory*, Lecture Notes in Logic, Vol, 40, Association for Symbolic Logic, La Jolla, CA, Cambridge University Press, Cambridge, 2012, x+248.
54. V. N. Vapnik and A. Y. Chervonenkis, On the uniform convergence of relative frequencies of events to their probabilities, *Theory Probab. Appl.* **16** no. 2 (1971), 264–280.