GRAPHONS ARISING FROM GRAPHS DEFINABLE
OVER FINITE FIELDS

MIRNA DŽAMONJA AND IVAN TOMAŠIĆ

Abstract. We prove a version of Tao’s algebraic regularity lemma for
asymptotic classes in the context of graphons. We apply it to study
expander difference polynomials over fields with powers of Frobenius.

1. Introduction

1.1. Historical overview and the statement of results. Tao’s alge-
braic regularity lemma is a variant of the celebrated Szemerédi’s regularity
lemma that applies to graphs that can be defined by a first-order formula
over finite fields. It states that such a graph can be decomposed into defin-
able pieces which are roughly about the same size and such that the edges
between those pieces behave almost randomly. This process is referred to
as regularisation. It was proved by Tao in [23] in order to study expander
polynomials over finite fields, and initially it was formulated for fields of
large enough characteristics.

Further developments on Tao’s lemma have a somewhat complex his-
tory. In a private correspondence to Tao, Hrushovski [10] gave another
proof of the lemma using the model-theoretic tools for studying the growth
rates of definable sets over finite fields, as developed by Chatzidakis-van
den Dries-Macintyre in [6]. Independently, Pillay and Starchenko gave an
analogous proof in the preprint [19]. The advantage of these proofs is that
they remove the requirement of the large characteristics of the field. Pillay
and Starchenko state that their proof also works for ‘measurable’ struc-
tures studied in the the context of asymptotic classes of finite structures by
Macpherson-Steinhorn [16] and Elwes and Macpherson [8]. Garcia-Macpherson-
Steinhorn [9] state, without proof, a version of the lemma for classes of finite
structures controlled by suitable pseudofinite dimensions, which potentially
generalise asymptotic classes.

2020 Mathematics Subject Classification. Primary 03C60, 11G25. Secondary 14G10,
14G15.

Key words and phrases. graphons, regularity lemma, finite fields, Frobenius automor-
phism, ACFA.
In this paper we give a proof of Tao’s algebraic regularity lemma for finite fields of any characteristics using concepts from the theory of graphons, and we adapt the lemma to relative asymptotic classes. We apply it to study expander difference polynomials over fields with powers of Frobenius. As far as the algebraic regularity lemma for fields goes, our proof benefited from the strategy used in the revisited and simplified proof of his lemma by Tao in his blog article [22].

Upon writing a preliminary version of this article, we received a copy of [10]. We now understand that Hrushovski envisaged in that letter that Tao’s regularity lemma applies to fields with Frobenius [14, 7.5] and that Tao’s algebraic constraint is directly related to the existence of a model-theoretic group configuration.

1.2. Graphons. The space of graphons with the cut metric is known as one of the most suitable contexts for studying the limit behaviour of (dense) finite graphs, as detailed in [14].

It is also extremely useful for formulating and proving regularity lemmas. Indeed, to express that a graph can be regularised to a certain precision, it is enough to say that it is sufficiently close to a stepfunction in the space of graphons, a piecewise constant probability function on the unit square encoding a rectangular array of the edge densities of pieces forming a regular decomposition of the graph, see Definition 2.4.

The algebraic regularity lemma in this context states the following, see theorems 5.4, 5.5 and Corollary 5.6.

In the space of graphons, the set of accumulation points of the family of realisations of a definable bipartite graph over the structures ranging in an asymptotic class is a finite set of stepfunctions.

Asymptotic classes are defined as classes of finite structures such that growth rates of definable sets are controlled by rational monomials 3.12. Consequently, stepfunctions appearing in the above take arbitrary rational probability values. In contrast, several recent regularity lemmas from other contexts, including [17] in stable and [7] in distal context can be viewed as 0-1 laws, and the resulting stepfunctions only take values 0 and 1.

1.3. Definable regularisation. An important feature of this note is a uniform formulation of the lemma. We emphasise that realisations of a single definable stepfunction over finite structures from an asymptotic class can be used to regularise a definable graph. The proof is almost constructive, and
we carefully analyse the space of parameters needed to define the regularising stepfunction. It becomes apparent that there are finitely many possible asymptotic behaviours of its realisations as we vary the finite structures in an asymptotic class.

The strategy of proof is congenial to the context of graphons in view of the fact that we start with a general weak regularity Lemma [4.1] for graphons, and then use definability and Chatzidakis-van den Dries-Macintyre properties to dramatically improve the conclusion.

1.4. Fields with powers of Frobenius. We draw attention to the fact that the algebraic regularity lemma applies to finite-dimensional bipartite graphs definable in the language of difference rings over the class of algebraic closures of finite fields equipped with powers of the Frobenius automorphism, which was shown to be an asymptotic class in [20] (with a suitable modification to the definition of ‘asymptotic class’). This is a significant generalisation of the context of finite fields and a potential new source of interesting examples. We outline the possibility of studying difference expander polynomials and we describe the difference morphisms which are moderate expanders.

1.5. Novelty in the paper. The ‘bounded-complexity’ statements of Tao’s regularity lemma in various contexts have been (independently) known to experts. Our main proof loosely follows the strategy of the proof from [22].

The added value of this paper consists of the following contributions.

1. The uniform statement of the result as discussed in [1.3] for a given set of formulae defining the initial graph and without the use of the notion of bounded complexity of formulae, which makes the proof more constructive.

2. It is aesthetically convenient to state the regularity lemma in the space of graphons, as noted in [1.2].

3. The profound reason for working in the space of graphons is to fix the arguments from Tao’s proof [22] in a single space within the realm of classical functional analysis. While each step in his proof is formulated for operators on finite-dimensional normed spaces, the dimensions grow with each step so it is difficult to track the key objects through the changes, or to give a bound on the dimensions of relevant spaces a priori in terms of the complexity of the formulae defining the given bipartite graph.
While finite fields are extensively used in combinatorics, fields with powers of Frobenius touched upon in [1.4] are less known. We hope that our treatment will introduce them to a wider audience.

As pointed out by Hrushovski in [10] and Tao in a number of talks, the hidden group structure in ‘solving the algebraic constraint’-type theorems needed to discuss expanding properties of polynomials can be uncovered using the model-theoretic group configuration arguments. We present these ideas as explicitly as possible in [8.17] whilst proving the expanding dichotomy for ‘difference polynomials’.

1.6. Future work. Tao’s lemma, as much as the original Szemerédi’s lemma, is only interesting in the context of reasonably dense graphs. After much research on obtaining an analogue of Szemerédi’s lemma for sparse graphs, as reported by Kohayakawa in [13], such an analogue was found by Scott in [21]. In further work we hope to extend our ideas from graphons to the Nešetřil-Ossona de Mendez notion of measurings [18] in order to obtain local versions of Tao’s regularity lemma which would deal with sparse definable graphs. This seems like a viable programme in the view of Scott’s lemma [21] and the fact that measurings give a unified approach to combinatorial limits which has graphons as a special case, but also applies to classes of sparse graphs.

1.7. Acknowledgement. We would like to thank the anonymous referee for their dedicated work on the paper and numerous suggestions for improvements.

2. Background: Graphs and Graphons

The background on the space of graphons given in this section is based almost entirely on the book [14] by Lovász, presented with minor modifications needed to treat our ‘bipartite/non-symmetric’ version of kernels and graphons.

2.1. Kernels and graphons.

Definition 2.1. (1) The space of kernels $\mathcal{W}$ is the space of the equivalence classes of essentially bounded measurable functions $[0, 1]^2 \to \mathbb{R}$ modulo equivalence up to a measure 0 set, i.e., the underlying vector space of $\mathcal{W}$ is $L^\infty([0, 1]^2)$. Following the usual conventions, we often abuse the notation and speak of ‘functions’ or ‘kernels’ (and ‘graphons’ below) in terms of functions rather than the equivalence
classes of functions, by taking representatives and by using the notation \( W \) in place of \( \tilde{W} \). Sometimes we need to be more explicit, for example a suitable metric on \( \tilde{W} \) will be introduced below, but it is only a pseudo-metric on \( W \).

(2) The space of graphons is
\[
\mathcal{W}_0 = \{ W \in \mathcal{W} : 0 \leq W \leq 1 \}.
\]

(3) We write
\[
\mathcal{W}_1 = \{ W \in \mathcal{W} : -1 \leq W \leq 1 \}.
\]

The inequalities are given up to a measure zero set (see Definition 2.3).

Remark 2.2. The classical literature on graphons \cite{13} usually stipulates that kernels and graphons have to be symmetric functions, and we write \( \mathcal{W}^{\text{sym}}, \mathcal{W}_0^{\text{sym}} \) and \( \mathcal{W}_1^{\text{sym}} \) for the corresponding spaces.

Given that we are primarily interested in studying the limits of bipartite and directed graphs, we do not assume the symmetry, but our context is easily reconciled with the classical one through the map
\[
\mathcal{W} \to \mathcal{W}^{\text{sym}},
\]
assigning to every \( W \in \mathcal{W} \) the symmetric ‘bipartite kernel’
\[
W^{\text{sym}}(x, y) = \begin{cases} 
W(2x, 2y - 1) & \text{when } (x, y) \in [0, 1/2] \times [1/2, 1], \\
W^*(2x - 1, 2y) & \text{when } (x, y) \in [1/2, 1] \times [0, 1/2], \\
0 & \text{elsewhere.}
\end{cases}
\]

Here \( W^* \) denotes transpose kernel of \( W \), defined by
\[
W^*(x, y) = W(y, x).
\]

The above map will yield a closed embedding in the cut metric from \cite{2,13} and will preserve the membership in \( \mathcal{W}, \mathcal{W}_0 \) and \( \mathcal{W}_1 \), which allows us to resort to the classical theory of symmetric graphons whenever needed.

Definition 2.3. Let \( U, W \) be two kernels.

(1) We say that \( U \) and \( W \) are \textit{equal almost everywhere} if the set \( \{(x, y) \in [0, 1]^2 : U(x, y) \neq W(x, y)\} \) is a null set with respect to the Lebesgue measure.

(2) We say that \( U \) and \( W \) are \textit{isomorphic up to a null set} if there exist invertible measure-preserving maps \( \varphi, \psi : [0, 1] \to [0, 1] \) such that \( U \) and
\[
W^{\varphi, \psi}(x, y) = W(\varphi(x), \psi(y))
\]
are equal almost everywhere.
2.2. Graphs as graphons.

**Definition 2.4.** A *stepfunction* is a kernel $W$ such that there exist partitions $[0,1] = \bigsqcup_{i=1}^{n} U_i$ and $[0,1] = \bigsqcup_{j=1}^{m} V_j$ into measurable sets so that $W$ is constant on each $U_i \times V_j$.

We often informally refer to the number $\max(m,n)$ as the number of *steps* of the stepfunction as above.

**Definition 2.5.** A *weighted bipartite graph* $\Gamma = (U,V,E,w)$ consists of a bipartite graph $(U,V,E \subseteq U \times V)$ such that

1. for $i \in U$, we have a positive real weight $w_i$;
2. for $j \in V$, we have a positive real weight $w_j$;
3. for each edge $(i,j) \in E$, we have a weight $w_{ij} \in \mathbb{R}$.

The above assignment can be extended to $U \times V$ by stipulating that $w_{ij} = 0$ whenever $(i,j)$ is not an edge.

**Definition 2.6.** Let $\Gamma = (U,V,E \subseteq U \times V,w)$ be a finite weighted bipartite graph. We define the associated stepfunction

$$W(\Gamma)$$

by considering the partitions $[0,1] = \bigsqcup_{i \in U} U_i$ and $[0,1] = \bigsqcup_{j \in V} V_j$ with lengths $\mu(U_i) = w_i/w_U$ and $\mu(V_j) = w_j/w_V$, where $w_U = \sum_{i \in U} w_i$, $w_V = \sum_{j \in V} w_j$, and by letting

$$W(\Gamma)|_{U_i \times V_j} = w_{ij}$$

for $i \in U$, $j \in V$.

By a slight abuse of notation, whenever we wish to consider a graph $\Gamma$ in the space of graphons, we implicitly identify $\Gamma$ with $W(\Gamma)$.

**Notation 2.7.** Let $\Gamma$ be a finite edge-weighted bipartite graph, i.e., a weighted bipartite graph with all node weights 1, and let $A = (w_{ij})$ be its matrix of edge weights. We write

$$W(A) = W(\Gamma).$$

In particular, we use the notation

$$W(w)$$

for the constant kernel with value $w$. 
2.3. Kernel operators.

**Definition 2.8.** The kernel operator $T_W : L^1[0,1] \to L^\infty[0,1]$ associated to a kernel $W \in \mathcal{W}$ is defined by

$$ (T_W f)(x) = \int_0^1 W(x,y)f(y)dy. $$

**Remark 2.9.** Considered as an operator $L^2[0,1] \to L^2[0,1]$, $T = T_W$ is a Hilbert-Schmidt operator; it is a compact operator, with a singular value decomposition

$$ T(f) = \sum_i \sigma_i \langle f, u_i \rangle v_i, $$

where $\{u_i\}$ and $\{v_i\}$ are orthonormal sets and $\sigma_i$ are positive with $\sigma_i \to 0$ such that the Hilbert-Schmidt norm satisfies

$$ \|T\|_2^2 = \text{tr}(T^*T) = \sum_i \sigma_i^2 = \|W\|_{L^2([0,1]^2)}^2 < \infty. $$

The spectrum of $T$ is discrete, and the nonzero eigenvalues $\lambda_i$ satisfy $\lim_i \lambda_i = 0$. If $W$ is symmetric, the eigenvalues $\lambda_i$ are real and we have a spectral decomposition

$$ W(x,y) \sim \sum_k \lambda_k f_k(x)f_k(y), $$

where $f_k$ is the normalised eigenfunction corresponding to the eigenvalue $\lambda_k$.

2.4. Operations on kernels.

**Definition 2.10.** Let $W_i$ be a countable family of kernels, and let $a_i$ and $b_i$ be positive real numbers with $\sum_i a_i = 1$ and $\sum_i b_i = 1$. The direct sum of kernels $W_i$ with weights $(a_i, b_i)$, denoted

$$ W = \oplus_i (a_i, b_i)W_i, $$

is defined as follows. We partition the interval $[0,1]$ into intervals $I_i$ of lengths $a_i$ and also into intervals $J_i$ of lengths $b_i$. We consider the monotone affine maps $\varphi_i$ mapping $I_i$ onto $[0,1]$, and $\psi_i$ mapping $J_i$ onto $[0,1]$, and we let

$$ W(x,y) = \begin{cases} W_i(\varphi_i(x), \psi_i(y)), & \text{if } x \in I_i, y \in J_i \\ 0, & \text{otherwise.} \end{cases} $$

A kernel is said to be connected, if it is not isomorphic up to a null set to a non-trivial direct sum of kernels.

Apart from the obvious linear structure, we consider the following operations on $\mathcal{W}$. 


Definition 2.11. Let $U$ and $W$ be two kernels.

1. Their product is the kernel
   $$(UW)(x, y) = U(x, y)W(x, y).$$

2. Their operator product is the kernel
   $$(U \circ W)(x, y) = \int U(x, z)W(z, y) \, dz.$$

Remark 2.12. If $U$ and $W$ are kernels, then, considering the associated kernel operators as operators on $L^2[0, 1]$, we have
$$T^*_W = T_W^*, \quad \text{and} \quad T_{U \circ W} = T_U T_W.$$

2.5. The cut norm and distance.

Definition 2.13. The cut norm on the linear space $\mathcal{W}$ of kernels is defined by
$$\|W\|_\square = \sup_{S, T \subseteq [0, 1]} \left| \int_{S \times T} W(x, y) \, dx \, dy \right|,$$
where $S$ and $T$ vary over all measurable subsets of $[0, 1]$.

The associated cut metric is
$$d_\square(U, W) = \|U - W\|_\square.$$

Fact 2.14. For $W \in \mathcal{W}_1$, we have the inequalities
$$\|W\|_\square \leq \|W\|_1 \leq \|W\|_2 \leq \|W\|_\infty \leq 1.$$

Definition 2.15. Let $S_{[0, 1]}$ denote the group of all invertible measure-preserving maps $[0, 1] \to [0, 1]$. The cut distance between kernels $U$ and $W$ is
$$\delta_\square(U, W) = \inf_{\varphi, \psi \in S_{[0, 1]}} d_\square(U, W^{\varphi, \psi}).$$

2.6. Regularity/homogeneity and the cut distance.

Definition 2.16. Let $\Gamma = (U, V, E \subseteq U \times V)$ be a finite bipartite graph and $\epsilon > 0$.

1. We say that $\Gamma$ is $\epsilon$-homogeneous of density $w \in [0, 1]$ provided, for every $A \subseteq U$ and $B \subseteq V$,
   $$| |E \cap (A \times B)| - w|A||B| | \leq \epsilon|U||V|.$$

2. We say that $\Gamma$ is $\epsilon$-regular of density $w \in [0, 1]$ if, for every $A \subseteq U$ with $|A| > \epsilon|U|$ and $B \subseteq V$ with $|B| > \epsilon|V|$,
   $$| |E \cap (A \times B)| - w|A||B| | \leq \epsilon|A||B|. $$
The well-known connection between regularity lemmas (the existence of a homogeneous partition) and the cut metric (see the discussion following 9.3 in [14]) is explained in the following lemma.

**Lemma 2.17.** Let \( \Gamma = (U, V, E \subseteq U \times V) \) be a finite bipartite graph, \( \epsilon > 0 \).

1. The graph \( \Gamma \) is \( \epsilon \)-homogeneous with density \( w \in [0, 1] \) if and only if \( d_\square(W(\Gamma), W(w)) \leq \epsilon \).

2. Suppose that there exist partitions \( U = \bigsqcup_{i=1}^{n} U_i \) and \( V = \bigsqcup_{j=1}^{m} V_j \) so that \( \Gamma \upharpoonright (U_i \times V_j) \) is \( \epsilon \)-homogeneous with density \( w_{ij} \). Let \( \bar{\Gamma} \) be a weighted bipartite graph on \( U \cup V \) such that all vertices have weight 1 and all edges between \( U_i \) and \( V_j \) have weight \( w_{ij} \). Then

\[ d_\square(W(\Gamma), W(\bar{\Gamma})) \leq \epsilon. \]

**Proof.** Item (1) follows from definitions of homogeneity, associated graphon and cut distance. Suppose that \( \Gamma \) satisfies the assumptions of (2). If we write \( \bar{U}_i \) and \( \bar{V}_j \) for segments in \( [0, 1] \) corresponding to \( U_i \) and \( V_j \) in \( W(\Gamma) \), using (1), we obtain that, for measurable \( A_i \subseteq \bar{U}_i \) and \( B_j \subseteq \bar{V}_j \),

\[ \left| \int_{A_i \times B_j} W(\Gamma)(x, y) - w_{ij} \, dx \, dy \right| \leq \epsilon \mu(\bar{U}_i) \mu(\bar{V}_j). \]

Hence, for measurable \( A, B \subseteq [0, 1] \), we obtain

\[ \left| \int_{A \times B} W(\Gamma)(x, y) - W(\bar{\Gamma}) \, dx \, dy \right| = \left| \sum_{i,j} \int_{A_i \times B_j} W(\Gamma)(x, y) - w_{ij} \, dx \, dy \right| \leq \sum_{i,j} \epsilon \mu(\bar{U}_i) \mu(\bar{V}_j) = \epsilon. \]

\( \square \)

2.7. **The space of graphons.** The cut distance is only a pseudo-metric on \( \mathcal{W}_0 \), so here we work with \( \tilde{\mathcal{W}}_0 \), the metric space of classes of graphons at non-zero distance.

**Theorem 2.18.** The space \((\tilde{\mathcal{W}}_0, \delta_\square)\) is compact.

**Proof.** For symmetric graphons, this is a known theorem of Lovász and Szegedy [15, Theorem 5.1], which essentially follows from a variant of Szemerédi’s regularity lemma. Our version can be deduced from it through the closed embedding from 2.2, which maps \( \mathcal{W}_0 \) onto a closed subset of the space of symmetric graphons. \( \square \)
3. Background: Asymptotic classes of finite structures

3.1. The yoga of definable sets. Definability in classes of finite structures, in the absence of a monster model, is a somewhat delicate matter. While all this is standard for a logician, we hope to improve the exposition for a reader with a combinatorics background by adopting the following notation inspired by category theory.

Let \( C \) be an arbitrary category of structures for a fixed first-order language \( \mathcal{L} \) with (substructure) embeddings as morphisms, and let \( C^< \) be the subcategory in which the morphisms are elementary embeddings.

Given a first-order formula \( \varphi(x) \) in the language \( \mathcal{L} \) in variables \( x = x_1, \ldots, x_n \), we define the assignment

\[
\tilde{\varphi} : C \to \text{Set}, \quad \tilde{\varphi}(F) = \varphi(F) = \{ a \in F^n : F \models \varphi(a) \},
\]

mapping a structure \( F \) in \( C \) to the set of realisations of the formula \( \varphi \) in \( F \).

If \( \varphi \) is a sentence in the language \( \mathcal{L} \), then

\[
\tilde{\varphi} : C \to \text{Set}
\]

assigns to each \( F \in C \) the truth value of \( \varphi \) in \( F \).

For \( n > 0 \), the ‘affine space’ \( \mathbb{A}^n \) is the assignment

\[
F \mapsto F^n,
\]

associated with the trivial formula \( \land_{i=1}^n (x_i = x_i) \).

The space \( \mathbb{A}^0 \) is the constant assignment

\[
F \mapsto \{ \bot, \top \}.
\]

For each formula \( \varphi \), the restriction of \( \tilde{\varphi} \) to \( C^< \) is a functor

\[
\tilde{\varphi} : C^< \to \text{Set}.
\]

On the other hand, while \( \mathbb{A}^n : C \to \text{Set} \) is a functor, the assignment \( \tilde{\varphi} : C \to \text{Set} \) for an arbitrary formula \( \varphi(x) \) in variables \( x = x_1, \ldots, x_n \) (or a sentence) is at best a subassignment of \( \mathbb{A}^n \), i.e., for all \( F \in C \),

\[
\tilde{\varphi}(F) \subseteq \mathbb{A}^n(F) = F^n.
\]

**Definition 3.1.** A subassignment \( S \) of some \( \mathbb{A}^n \) with \( n \geq 0 \) is called a **definable set** if it is equivalent to the assignment \( \tilde{\varphi} \) associated with some first-order formula \( \varphi(x) \) in \( n \) variables.

We often emphasise that such sets are ‘parameter-free’, or ‘defined with no parameters’.
Remark 3.2. If $X$ and $Y$ are definable sets, their argument-wise cartesian product, denoted 
\[ X \times Y \]
is clearly definable.

If $X, Y \subseteq \mathbb{A}^n$ are definable, so are the following sets 
\[ X \cap Y, \quad X \cup Y, \quad X \setminus Y. \]

Definition 3.3. A definable function 
\[ \pi : X \to Y \]
between definable sets $X$ and $Y$ is given through a definable subset $\Gamma \subseteq X \times Y$ such that, for every $F \in \mathcal{C}$, $\Gamma(F) \subseteq X(F) \times Y(F)$ defines a function
\[ \pi_F : X(F) \to Y(F). \]

Remark 3.4. If $\pi : X \to Y$ is a definable function, the image $\pi(X)$ defined by
\[ \pi(X)(F) = \pi_F(X(F)) \]
is a definable subset of $Y$.

Definition 3.5. A definable function 
\[ f : X \to E \]
from a definable set $X$ to an arbitrary set $E$ is determined by a choice of 
1. finitely many values $e_1, \ldots, e_n \in E$, and 
2. finitely many definable subsets $X_i \subseteq X$, $i = 1, \ldots, n$,
so that, for every $F \in \mathcal{C}$,
\[ X(F) = \bigsqcup_i X_i(F) \]
and $f_F : X(F) \to E$ is defined by
\[ f_F \mid_{X_i(F)} = e_i. \]

We stipulate that, for $X = \mathbb{A}^n$, the $X_i$ are associated with sentences $\varphi_i$ such that, for every $F \in \mathcal{C}$,
\[ f_F = e_i \text{ if and only if } F \models \varphi_i \]

In the following, we discuss definable sets with parameters. We start by a construction allowing us to choose parameters from a fixed structure in our class, and continue onto constructions of definable parameter spaces.
Definition 3.6. Let $F \in \mathcal{C}$, and let $\mathcal{L}_F$ be the language obtained by adding the constant symbols for the elements of $F$ to $\mathcal{L}$. Let $\mathcal{C}_F$ be the subcategory of $\mathcal{C}$ consisting of superstructures of $F$.

If $\varphi(x;c)$ is a formula in the language $\mathcal{L}_F$ with parameters $c = c_1, \ldots, c_m \in F$, then we can define the assignment

$$\bar{\varphi}_c : \mathcal{C}_F \to \text{Set},$$

which to each $F' \supseteq F$ assigns the set of realisations $\varphi(F',c)$ in $F'$.

Definition 3.7. Let $S$ be a definable set. A definable set $X$ over $S$ (or, with parameters in $S$) is a definable function in the sense of 3.3

$$\pi : X \to S.$$ 

A definable map of definable sets over $S$ is a commutative diagram

$$\begin{array}{ccc}
X & \longrightarrow & Y \\
\downarrow & & \downarrow \\
S & \rightarrow & S
\end{array}$$

Definition 3.8. Given definable sets $\pi : X \to S$ and $\pi' : X' \to S$ over $S$, their fibre product $X \times_S X'$ is

$$X \times_S X'(F) = X(F) \times_{S(F)} X'(F) = \{(x,x') \in X(F) \times X'(S) : \pi(x) = \pi'(x')\},$$

and it is again a definable set over $S$.

Given $s \in S(F)$ for some $F \in \mathcal{C}$, we can consider the singleton $\{s\}$ as a definable set with parameters from $F$ over $S$, so we obtain a definable set

$$X_s = X \times_S \{s\} : \mathcal{C}_F \to \text{Set},$$

which is called the fibre of $X$ over $S$ with parameter $s$, and it is a definable set over $F$ in the sense of 3.6 via

$$X_s(F') = \pi_{F'}^{-1}(s),$$

for every $F'$ in $\mathcal{C}_F$.

Remark 3.9. By the above, a definable set $X$ over $S$ gives rise to a family of definable sets $X_s$ parametrised by parameters $s$ from $S$.

In model theory, this object is usually called a uniformly definable family of definable sets. The reason for our specific formulation is the additional precision needed to treat parameters over a family of structures, as opposed to working in a fixed ‘monster model’, which is the most familiar setting for model theory.
Note, if $X$ is a definable set, it can naturally be considered as a definable set over $A^0$. Indeed, we consider the definable map which takes $X(F)$ to $\top$ if and only if $X(F) \neq \emptyset$.

**Definition 3.10.** Let $X$ be a definable set over $S$ and let $E$ be a set. A *definable function*

$$f : X \to E$$

*over $S$ on a class of structures $C$ is determined by a choice of*

1. finitely many definable functions $e_1, \ldots, e_n : S \to E$, and
2. finitely many definable sets $X_1, \ldots, X_n$ over $S$

such that, for every $F \in C$, every $s \in S(F)$, we have

$$X_{1,s} \sqcup \cdots \sqcup X_{n,s} = X_s,$$

and

$$f_s : X_s \to E$$

is given on $C_F$ by

$$f_s,F : |_{X_{i,s}(F')} = e_i(s).$$

3.2. Counting and asymptotic classes.

**Definition 3.11.** Let $C$ be a class of finite structures (considered a category with substructure embeddings). We say that $C$ is a *CDM-class*, if, for every definable set $X$ over $S$, there exist

1. a definable function $\mu_X : S \to \mathbb{Q}$,
2. a definable function $d_X : S \to \mathbb{N}$,
3. a constant $C_X > 0$,

so that, for every $F \in C$ and every $s \in S(F)$,

$$\left| |X_s(F)| - \mu_X(s)|F|^{d_X(s)} \right| \leq C_X |F|^{d_X(s) - 1/2}.$$

**Definition 3.12.** Let $C$ be a class of finite structures (considered a category with substructure embeddings). We say that $C$ is an *asymptotic class* (in the sense of [16] and [8]), if, for every definable set $X$ over $S$, there exist

1. a definable function $\mu_X : S \to \mathbb{Q}$,
2. a definable function $d_X : S \to \mathbb{N}$,

so that, for every $\epsilon > 0$ there exists a constant $N > 0$ such that for every $F \in C$ with $|F| > N$ and every $s \in S(F)$,

$$\left| |X_s(F)| - \mu_X(s)|F|^{d_X(s)} \right| \leq \epsilon |F|^{d_X(s)}.$$
Definition 3.13. Let $\mathcal{C}$ be a class of structures with a given function $\chi : \mathcal{C} \to \mathbb{N}$.

We say that $\mathcal{C}$ is a CDM-class relative to $\chi$, if for every definable set $X$ over $S$, there exist

1. a definable function $\mu_X : S \to \mathbb{Q} \cup \{\infty\}$,
2. a definable function $d_X : S \to \mathbb{N} \cup \{\infty\}$,
3. a constant $C_X > 0$,

so that, for every $F \in \mathcal{C}$ and every $s \in S(F)$,

$$\left| |X_s(F)| - \mu_X(s)\chi(F)^{d_X(s)} \right| \leq C_X\chi(F)^{d_X(s)-1/2},$$

where we stipulate that $\mu_X(s) < \infty$ is and only if $d_X(s) < \infty$.

We define an asymptotic class relative to $\chi$ analogously, to reflect the error term from 3.12.

Terminology 3.14. With notation of 3.13, we may informally refer to the number $\mu_X(s)$ as measure, and to the number $d_X(s)$ as dimension.

We say that $X \to S$ is of finite relative dimension, provided $d_X$ maps into $\mathbb{N}$.

Remark 3.15. Suppose $X$ is of finite relative dimension over $S$, over a class $\mathcal{C}$ that is an asymptotic class relative to $\chi$. For each definable subset $Y$ of $X$ over $S$, we obtain a definable probability function

$$\mu_{Y/X} : S \to \mathbb{Q}, \quad \mu_{Y/X}(s) = \begin{cases} \frac{\mu_Y(s)}{\mu_X(s)} & \text{if } d_Y(s) = d_X(s) \\ 0 & \text{otherwise.} \end{cases}$$

By the above definitions, for $F$ ranging over $\mathcal{C}$ and $s \in S(F)$, the expressions

$$\frac{|Y_s(F)|}{|X_s(F)|} - \mu_{Y/X}(s)$$

are $o(1)$ in $\chi(F)$, i.e., converge to 0 as $\chi(F) \to \infty$. If $\mathcal{C}$ is a CDM-class relative to $\chi$, then the above expressions are $O(\chi(F)^{-1/2})$.

If $Y$ is lower-dimensional than $X$, then the above expressions are $O(\chi(F)^{-1})$.

Remark 3.16. A (relative) asymptotic class is clearly a (relative) CDM-class. A CDM-class is a relative CDM-class of finite structures with respect to the cardinality function such that all definable sets are finite dimensional. Similarly, an asymptotic class is a relative asymptotic class of finite structures with respect to the cardinality function such that all definable sets are finite dimensional.
Example 3.17. The class of finite fields was shown to be a CDM-class in the foundational paper [6] by Chatzidakis-van den Dries-Macintyre.

Following a discussion of difference fields in Section 7, we will explain in 7.3 how fields with powers of Frobenius constitute a relative CDM-class by the main theorem of [20].

For further examples of CDM-classes and a discussion of differences between CDM and asymptotic classes we refer the reader to [16].

4. A weak regularity lemma

In this section we show that an iterate of a graphon can be regularised in infinity norm. It is a general result for graphons and, although it does not use definability, it will serve as an important first step in the proof of our main theorem later.

Lemma 4.1. Let $W$ be a graphon. For every $\epsilon \in (0,1)$ there exists a stepfunction $W'$ with $N(\epsilon) \leq (3/\epsilon^3)^{(1/\epsilon^2)}$ steps such that, writing $W^6 = W \circ W^* \circ W \circ W^* \circ W \circ W^*$, \[
\|W^6 - W'\|_\infty \leq 2\epsilon^2.
\]

Proof. We reformulate the first part of the proof of [22, Lemma 3] in the language of graphons.

Let $T = T_W : L^2[0,1] \to L^2[0,1]$ be the kernel operator \[
T(f)(v) = \int_{[0,1]} W(u,v)f(u) \, du
\]
associated with the graphon $W$. Its adjoint $T^*$ is given by \[
T^*(g)(u) = \int_{[0,1]} W(u,v)g(v) \, dv.
\]

Cauchy-Schwarz inequality yields that for all $f \in L^2([0,1])$, \[
\|Tf\|_2 \leq \|Tf\|_\infty \leq \|f\|_2,
\]
and similarly, for $g \in L^2[0,1]$, \[
\|T^*g\|_2 \leq \|T^*g\|_\infty \leq \|g\|_2.
\]

We can apply the singular value decomposition to the Hilbert-Schmidt operator $T$, which gives \[
Tf = \sum_i \sigma_i \langle f, u_i \rangle_2 y_i
\]
and \[
T^*g = \sum_i \sigma_i \langle g, y_i \rangle_2 u_i
\]
for some sequence \( \sigma_i \) of singular values with \( \sigma_1 \geq \sigma_2 \geq \cdots > 0 \), where \( u_i \) and \( y_i \) are orthonormal systems in \( L^2[0,1] \).

The operator \( TT^* : L^2[0,1] \to L^2[0,1] \) can be diagonalised as

\[
TT^*g = \sum_i \sigma_i^2 \langle g, y_i \rangle y_i,
\]

whence

\[
\text{tr}(TT^*) = \sum_i \sigma_i^2.
\]

On the other hand, we obtain the Hilbert-Schmidt norm bound

\[
\begin{aligned}
\sum_i \sigma_i^2 &= \text{tr}(TT^*) = \int |W(x,y)|^2 \, dx \, dy \\
&\leq 1.
\end{aligned}
\]  

Using

\[
y_i = \frac{1}{\sigma_i} Tu_i, \quad u_i = \frac{1}{\sigma_i} T^* y_i,
\]

as well as (4.1), (4.2), we obtain

\[
\|y_i\|_\infty \leq \frac{1}{\sigma_i}, \quad \text{and} \quad \|u_i\|_\infty \leq \frac{1}{\sigma_i}.
\]

We use these bounds to find a low rank approximation to the sixth power

\[
TT^*TT^*TT^* : L^2[0,1] \to L^2[0,1].
\]

Intuitively, taking a high power ‘tames’ any unpredictable behaviour of \( T \) and produces a more manageable operator.

The above operator can be diagonalised as

\[
TT^*TT^*TT^*g = \sum_i \sigma_i^6 \langle g, y_i \rangle y_i.
\]

Given an \( \epsilon > 0 \), we split

\[
TT^*TT^*TT^* = A_\epsilon + B_\epsilon,
\]

where \( A = A_\epsilon \) is a low rank operator

\[
Ag = \sum_{i : \sigma_i \geq \epsilon} \sigma_i^6 \langle g, y_i \rangle y_i,
\]

and \( B = B_\epsilon \) is the error term

\[
Bg = \sum_{i : \sigma_i < \epsilon} \sigma_i^6 \langle g, y_i \rangle y_i.
\]
Using the triangle inequality, Hölder’s inequality \[ A.2 \] and \( 4.4 \), \( 4.3 \), for any \( g \in L^1[0,1] \),
\[
\|B g\|_\infty \leq \sum_{i: \sigma_i < \epsilon} \sigma_i^6 |\langle g, y_i \rangle| \frac{1}{\sigma_i} \leq \sum_{i: \sigma_i < \epsilon} \sigma_i^6 \|g\|_1 \|y_i\|_\infty \frac{1}{\sigma_i}
\]
\[
\leq \sum_{i: \sigma_i < \epsilon} \sigma_i^6 \|g\|_1 \frac{1}{\sigma_i} \leq \epsilon^2 \sum_i \sigma_i^2 \|g\|_1 \leq \epsilon^2 \|g\|_1.
\]
Let
\[
\delta = \frac{\epsilon^2}{3}.
\]
Using \( 4.4 \), we discretise
\[
y_i = y_i' + e_i;
\]
where \( y_i' \) takes at most \( 1/\sigma_i \delta \) values, \( \|y_i'\|_\infty \leq \|y_i\|_\infty \), and \( e_i \) is bounded in magnitude by \( \delta \). We split \( A_\epsilon = A_\epsilon' + E_\epsilon \), where
\[
A_\epsilon' = \sum_{i: \sigma_i \geq \epsilon} \sigma_i^6 \langle g, y_i' \rangle y_i'
\]
and
\[
E_\epsilon = \sum_{i: \sigma_i \geq \epsilon} \sigma_i^6 (\langle g, y_i' \rangle e_i + \langle g, e_i \rangle y_i' + \langle g, e_i \rangle e_i).
\]
By the choice of \( \delta \), and arguments analogous to the above, we get that
\[
\|E_\epsilon\|_\infty \leq \sum_{i: \sigma_i \geq \epsilon} \sigma_i^6 \left( \|g\|_1 \|y_i'\|_\infty \|e_i\|_\infty + \|g\|_1 \|e_i\|_\infty \|y_i'\|_\infty + \|g\|_1 \|e_i\|_\infty^2 \right)
\]
\[
\leq \|g\|_1 \sum_{i: \sigma_i \geq \epsilon} \sigma_i^6 \left( 2 \frac{\delta}{\sigma_i} + \delta^2 \right) \leq \|g\|_1 \sum_{i: \sigma_i \geq \epsilon} 2 \sigma_i^6 \delta + \sigma_i^6 \delta^2 \leq \epsilon^2 \|g\|_1.
\]
Thus, we have decomposed
\[
TT^*TT^*TT^* = A_\epsilon' + E_\epsilon',
\]
where \( A' \) is of ‘low rank’, and \( E' = E + B \) has integral kernel bounded pointwise by \( 2\epsilon^2 \).

Using \( 4.3 \), the number of summands in the definition of \( A \) is at most \( 1/\epsilon^2 \). We partition
\[
[0,1] = V_1 \cup \ldots \cup V_n,
\]
where \( V_j = V_j, \epsilon \) are the intersections of the level sets of the \( y_i' \) (removing any empty cells to ensure the \( V_j \) are all non-empty), and
\[
n = N(\epsilon) \leq \left( 1/\epsilon \delta \right)^{(1/\epsilon^2)} = \left( 3/\epsilon^3 \right)^{(1/\epsilon^2)}.
\]
The (sought-after) integral kernel \( W' \) of \( A' \) is constant on each \( V_j \times V_k \), so the integral kernel of \( TT^*TT^*TT^* \) fluctuates by at most \( 2\epsilon^2 \) on each \( V_j \times V_k \). □
5. Algebraic regularity lemma

Definition 5.1. A definable bipartite graph over $S$ is a triple $\Gamma = (U, V, E)$, where $U, V, E \subseteq U \times V$ are definable sets over $S$. For each $F \in \mathcal{C}$ and each point $s \in S(F)$, we obtain a bipartite graph

$$\Gamma_s = \Gamma_s(F) = (U_s(F), V_s(F), E_s(F)).$$

We are interested in describing the (limit) behaviour of the graphs $\Gamma_s(F)$ as $F$ and $s$ vary over an asymptotic class $\mathcal{C}$.

Definition 5.2. Let $U, V$ be definable sets over $S$. A definable stepfunction on $U \times V$ over $S$ is a definable function $W : U \times V \to \mathbb{R}$ over $S$ such that there exist definable sets $U_1, \ldots, U_m$ and $V_1, \ldots, V_n$ over $S$ so that, for each $F \in \mathcal{C}$ and each $s \in S(F)$,

$$U_s = \bigcup_{i=1}^m U_{i,s} \quad \text{and} \quad V_s = \bigcup_{j=1}^n V_{j,s},$$

and $W_s$ is constant on each $U_{i,s} \times V_{j,s}$.

For $F \in \mathcal{C}$ and $s \in S(F)$, by a slight abuse of notation, we will often identify the weighted graph $W_s(F)$ with its associated stepfunction $W(W_s(F)) \in \mathcal{W}_0$.

Lemma 5.3. Let $W$ be a definable stepfunction on $U \times V$ over $S$ of finite relative dimension on an asymptotic class $\mathcal{C}$ relative to $\chi$. The set of accumulation points of the set

$$\{W_s(F) : F \in \mathcal{C}, s \in S(F)\}$$

in the space $\mathcal{W}_0$ of graphons is a finite set of graphons represented by stepfunctions.

Proof. By definition, there exist definable sets $U_1, \ldots, U_m$ and $V_1, \ldots, V_n$ over $S$ partitioning $U$ and $V$ and definable functions $e_{ij} : S \to [0, 1]$ so that

$$W_s \mid_{U_{i,s} \times V_{j,s}} = e_{ij}(s).$$

Using 3.15 and the fact that we are dealing with the finite relative dimension, there exist definable functions

$$\mu_{U_i/U} : S \to \mathbb{Q} \quad \text{and} \quad \mu_{V_j/V} : S \to \mathbb{Q}$$

$$\frac{|U_{i,s}(F)|}{|U_s(F)|} - \mu_{U_i/U}(s) = o(1),$$

and analogously for $V_j$.

We can partition

$$S = S_1 \amalg \cdots \amalg S_r.$$
so that, for all \( s \in S_l(F) \),

\[
e_{ij}(s) = e_{ijl} \in [0, 1], \quad \mu_{U_i/U}(s) = \mu_{il} \in \mathbb{Q}, \quad \mu_{V_j/V}(s) = \nu_{jl} \in \mathbb{Q}.
\]

Thus, using 3.15, up to isomorphism, the stepfunction associated to \( W_s(F) \) through 2.6 is within \( o(1) \) from a stepfunction associated to a weighted bipartite graph with vertex weights \( \{ \mu_{il} : i \} \) and edge weights \( \{ e_{ijl} : i, j \} \) in \( \| \cdot \|_1 \)-norm, and hence, by Fact 2.14 in the cut norm.

In fact, the net of graphons associated to the graphs \( \{ W_s(F) : s \in S_l(F) \} \) with respect to the preorder induced by \( \chi(F) \) has a stepfunction as a limit in the \( \| \cdot \|_1 \)-norm. \( \square \)

**Theorem 5.4 (Tao’s algebraic regularity lemma for asymptotic classes).**

Let \( \Gamma = (U, V, E) \) be a definable bipartite graph of finite relative dimension over a definable set \( S \) on an asymptotic class \( C \) relative to \( \chi \). Then there exists a definable set \( \tilde{S} \) over \( S \) and a definable stepfunction \( W \) over \( \tilde{S} \) such that for every \( \varepsilon > 0 \), there exists an \( M > 0 \) such that for every \( F \in C \) with \( \chi(F) \geq M \), every \( \tilde{s} \in \tilde{S}(F) \) mapping onto \( s \in S(F) \),

\[
d(\Gamma_s(F), W_{\tilde{s}}(F)) \leq \varepsilon.
\]

**Theorem 5.5 (Tao’s algebraic regularity lemma).** With assumptions from 5.4, suppose that \( C \) is a CDM-class relative to \( \chi \). There exists a constant \( M = M(\Gamma) > 0 \), a definable set \( \tilde{S} \) over \( S \) and a definable stepfunction \( W \) over \( \tilde{S} \) such that for every \( F \in C \), every \( \tilde{s} \in \tilde{S}(F) \) mapping onto \( s \in S(F) \),

\[
d(\Gamma_s(F), W_{\tilde{s}}(F)) \leq M\chi(F)^{-1/12}.
\]

We will prove Theorem 5.5 because of the more interesting/precise analysis of the error term, and the proof of Theorem 5.4 follows along the same lines. We follow the ideas from the proof of [22, Lemma 3].

**Proof.** For simplicity of notation, we will write the proof for a CDM-class. Note that, because of the assumption of finite relative dimension, the same proof will work for a relative CDM-class by replacing every instance of \( |F| \) by \( \chi(F) \), for \( F \in C \).

The weak regularity result established in Proposition 4.1 states that for every graphon \( W \) and every \( \epsilon > 0 \), there exists a stepfunction with at most \( N(\epsilon) = (3/\epsilon^3)^{(1/\epsilon^2)} \) steps which approximates \( W^\epsilon \) up to \( 2\epsilon^2 \) in the \( \| \cdot \|_\infty \)-norm.

The idea is to improve this result by using the definability of \( \Gamma \) and the constraints on the growth rates of the sets of realisations of definable sets over CDM-classes in the spirit of 3.15.
Let us name the key objects. For \( F \in \mathcal{C} \) and an \( s \in S(F) \), let \( T = T_s : L^2[0, 1] \to L^2[0, 1] \) be the kernel operator

\[
T_s(f)(v) = \int_{[0,1]} W_s(u, v) f(u) \, du
\]

associated with the stepfunction \( W_s = W(\Gamma_s(F)) \).

The proof consists of the following conceptual steps.

(1) By the definability of \( \Gamma \), the set of relevant CDM-growth rates of certain definable invariants of \( \Gamma \) is finite and hence separated by some minimal distance \( \delta \). We choose an \( \epsilon > 0 \) small enough with respect to \( \delta \), and we construct a definable stepfunction \( W \) with at most \( N(\epsilon) \) steps anticipating all the possible behaviours that may occur in regularising each \( \Gamma_s(F) \) for varying \( F \in \mathcal{C} \) and \( s \in S(F) \).

(2) We verify that, for any large enough \( F \in \mathcal{C} \), any \( s \in S(F) \), writing \( W_s = W(\Gamma_s(F)) \), any (not necessarily definable) stepfunction with at most \( N(\epsilon) \) steps that regularises \( W_s^6 \) and \( (W_s^*)^6 \) in view of \( 4.1 \) turns out to be close to the steps of a realisation of \( W \) constructed in (1) in the supremum norm.

(3) From (1) and (2), it follows that \( W_s \) is close to a realisation of \( W \) in the cut metric, which shows the required regularity.

**Step 1.** We will use the definability of \( \Gamma \) and decide on an appropriate choice of \( \epsilon \).

The integral kernel \( K_3(v, v') = K_{3,s}(v, v') \) of \( TT^*TT^*TT^* \) for \( T = T_s \) is explicitly given as

\[
K_3(v, v')(F) = \frac{|G_{v,v',s}(F)|}{|U_s(F)|^2|V_s(F)|^2},
\]

where \( G_{v,v',s} \subseteq U_s \times V_s \times U_s \times V_s \times U_s \times V_s \times U_s \times V_s \times U_s \times V_s \times U_s \) is the definable set

\[
G_{v,v',s} = \{(u_1, v_2, v_3, u_3) \in U_s \times V_s \times U_s \times V_s : (u_1, v_1, (u_1, v_2), (u_2, v_3), (u_3, v_3), (u_3, v') \in E_s)\}.
\]

Using 3.11, we get that there exists a a definable function \( c = c(v, v', s) \) symmetric in \( v, v' \) such that, for all \( F \in \mathcal{C} \), \( s \in S(F) \) and \( v, v' \in V_s(F) \),

\[
K_{3,s}(v, v') = c(v, v', s) + O(|F|^{-1/2}).
\]

Replacing the role of \( T_s \) by \( T_s^* \), we consider the integral kernel

\[
K_{3,s}^*(u, u')
\]

of \( TT^*TT^*TT^*T \), and we similarly find a definable function \( c^* = c^*(u, u', s) \) symmetric in \( u, u' \) such that for all \( F, s \in S(F) \) and \( u, u' \in U_s(F) \),

\[
K_{3,s}^*(u, u') = c^*(u, u', s) + O(|F|^{-1/2}).
\]
The set of values of $c$ and $c^*$ is finite, so its elements are separated by some minimal distance $\delta > 0$.

Let us choose an $\epsilon \in (0, 1)$ such that

$$2\epsilon^2 < \delta/2.$$

For every $\tilde{n} \leq n \leq N(\epsilon) = (3/\epsilon^3)^{1/2}$, writing

$$V_i(v_1, \ldots, v_n, s) = \{ v \in V_s : \bigwedge_j c(v, v_j, s) = c(v_i, v_j, s) \},$$

the formula

$$\psi_{\tilde{n}, n}(v_1, \ldots, v_n, s) \equiv \left( \bigwedge_i v_i \in V_s \right) \wedge \left( V_s = \bigcap_{i \leq \tilde{n}} V_i(v_1, \ldots, v_n, s) \right)$$

expresses that $s' = (v_1, \ldots, v_n, s)$ can serve as parameters for a definable stepfunction on $V_s$ with steps $V_1, s', \ldots, V_n, s'$ and values $c(v_i, v_j, s)$ on $V_1, s' \times V_j, s'$, where $V_1, s', \ldots, V_n, s'$ are the top-dimensional steps that we will call large in the sequel.

Let $S_{\tilde{n}, n} \to S$ be the definable set associated with the formula $\psi_{\tilde{n}, n}$, and let

$$S_V \to S$$

be the disjoint union of all the $S_{\tilde{n}, n}$ for $\tilde{n} \leq n \leq N(\epsilon)$. For each $i \leq n$, we have a definable set

$$V_{i, \tilde{n}, n} \to S_{\tilde{n}, n} \to S_V.$$  

Analogously, using $c^*$ in place of $c$, for every $\tilde{m} \leq m \leq N(\epsilon)$ we construct definable sets

$$U_{j, \tilde{m}, m} \to S^*_{\tilde{m}, m} \to S_U,$$

for $j \leq m$, which form steps of a definable stepfunction with values $c^*$.

Using [3.11] again, for each $i, \tilde{n}, n$ and $j, \tilde{m}, m$ as above, there is a definable function $f_{j, \tilde{m}, m, i, \tilde{n}, n} : S^*_{\tilde{m}, m} \times S S_{\tilde{n}, n} \to \mathbb{Q}$ such that for all $F \in \mathcal{C}$, and $s' \in S^*_{\tilde{m}, m}(F)$, $s'' \in S_{\tilde{n}, n}(F)$ mapping onto $s \in S(F)$,

$$\left| \frac{E_s(F) \cap (U_{j, \tilde{m}, m, s'}(F) \times V_{i, \tilde{n}, n, s'}(F))}{|U_{j, \tilde{m}, m, s'}(F) \times V_{i, \tilde{n}, n, s'}(F)|} \right| = f_{j, \tilde{m}, m, i, \tilde{n}, n}(s', s'') + O(|F|^{1/2}).$$

Let

$$\tilde{S} = S_U \times S S_V.$$
We consider the definable stepfunction $W$ over $\tilde{S}$ which is given by

$$W \rvert_{u_j,m,m \times v_i,n,n} = f_{j,m,m,i,n,n}$$

over the component $S_{m,m}^* \times S_{n,n}$ of $\tilde{S}$.

**Step 2.** Let us choose $M$ such that for $|F| \geq M$, the error terms in 5.1, 5.2, 5.3 are less than $\delta/2$.

Let us consider $F \in \mathcal{C}$ with $|F| \geq M$, a parameter $s \in S(F)$ and the stepfunction $W_s = W(T_s(F))$.

We apply Proposition 4.1 to $W_s$ for the $\epsilon$ chosen in Step 1, and this yields a partition

$$V_s(F) = V_1 \cdots V_n$$

into $n \leq N(\epsilon)$ sets $V_i$ so that the integral kernel $K_3(v, v')$ fluctuates by less than $\delta/2$ on each $V_i \times V_j$. On the other hand, by (5.1), and the choice of $F$, each value of $K_3(v, v')$ is within $\delta/2$ from a value of the definable function $c$. Hence we can improve the near-constancy to the property

$$K_{3,s}(v, v') = \gamma_{ij} + O(|F|^{-1/2})$$

for some value $\gamma_{ij}$ of $c$, whenever $v \in V_i$, $v' \in V_j$ and $i, j \in \{1, \ldots, n\}$.

Note that $K_3$ is symmetric, so $\gamma_{ij}$ will be symmetric too. Moreover, by reducing the number of $V_i$, we can assume that $\gamma_{ij} = \gamma_{i'j}$ for all $j \in \{1, \ldots, n\}$ implies $i = i'$, i.e., that the array $\gamma$ is reduced.

Using this property, we see that, if $v_1, \ldots, v_n$ satisfy the definable condition

$$\bigwedge_i \bigwedge_j c(v_i, v_j, s) = \gamma_{ij},$$

then there exists a permutation $\sigma$ of $\{1, \ldots, n\}$ such that $\gamma_{\sigma(i), \sigma(j)} = \gamma_{ij}$ and $v_i \in V_{\sigma(i)}$.

Note, if $v_i \in V_i$ for $i \in \{1, \ldots, n\}$ are parameters, then $V_i$ is definable as

$$V_i = \{v \in V_s(F) : \bigwedge_j c(v, v_j) = \gamma_{ij}\}.$$ 

Hence, choosing parameters $s' = (v_1, \ldots, v_n, s) \in S_{n,n,s}(F)$ determines $V_i$ up to a permutation $\sigma$ preserving $\gamma_{i,j}$. Moreover, for a fixed $s'$ we may relabel the $V_i$ so that

$$V_i = V_{i,n,n,s'}(F).$$

Note that $V_{i,n,n}$ for $i \leq \bar{n}$ are large, so 3.11 gives that for some positive $a_i \in \mathbb{Q}$,

$$\frac{|V_i|}{|V_s(F)|} = a_i + O(|F|^{-1/2}),$$

(5.5)
while for $i > \bar{n}$, $\dim(V_{i,\bar{n},n,s'}) < \dim(V)$ and
\begin{equation}
|V_i| \ll |F|^{-1}|V_s(F)|.
\end{equation}

Let $\bar{V}_j \subseteq [0,1]$ denote the segment corresponding to a large step $V_j$ in the stepfunction $W_s$ associated to the graph $\Gamma_s(F)$ via 2.6. Suppose $g \in L^2[0,1]$ is supported on $\bar{V}_j$ with mean zero. Combining (4.5) and (5.4), we obtain
\[
\sum_i \sigma_i^6 |\langle g, y_i \rangle|^2 = |\langle TT^*TT^*TT^*g, g \rangle| = \left| \int_0^1 \int_0^1 K_3(v, v')g(v)g(v') dv dv' \right| \ll |F|^{-1/2} \|g\|_2^2.
\]

On the other hand, Bessel's inequality gives
\[
\sum_i |\langle g, y_i \rangle|^2 \leq \|g\|_2^2,
\]
and so Hölder’s inequality (combining these two inequalities via A.3) shows that
\[
\|T^*g\|_2^2 = \sum_i \sigma_i^2 |\langle g, y_i \rangle|^2 \ll |F|^{-1/6} \|g\|_2^2,
\]
whence
\[
\|T^*g\|_2 \ll |F|^{-1/12} \|g\|_2.
\]

Similarly, arguing for $K_3^*(u, u')$, we can find a partition
\[U_s(F) = U_1 \cup \cdots \cup U_m\]
with $m \leq N(\epsilon)$ so that, upon fixing a parameter $s'' = (u_1, \ldots, u_m, s) \in S_{m,m,s}(F)$, by relabeling $U_j$, we have
\[U_j = U_{j,m,m,s''}(F).
\]
Moreover,
\[
\|Tf\|_2 \ll |F|^{-1/12} \|f\|_2,
\]
when $f$ is supported on one of the segments $\bar{U}_i$ corresponding to a large step $U_i$ with mean zero.

**Step 3.** Let $W$ be the definable stepfunction over $\tilde{S}$ defined in Step 1.

We claim that for $\tilde{s} \in \tilde{S}(F)$ which maps to $s'$ and $s''$ from Step 2,
\[
d_{\square}(W_s, W_{\tilde{s}}(F))) \ll |F|^{-1/12}.
\]

Combining the estimates obtained in Step 2 with the Cauchy-Schwarz inequality, we deduce that
\[
|\langle Tf, g \rangle|_2 \ll |F|^{-1/12} \|f\|_2 \|g\|_2
\]
whenever $f$ and $g$ are supported on $\bar{V}_i$ and $\bar{U}_j$ with at least one of $f$, $g$ of mean zero.
For measurable $A \subseteq \bar{U}_i$ and $B \subseteq \bar{V}_j$, we decompose the characteristic functions
\[ 1_A = (1_A - \mu(A)/\mu(\bar{U}_i))\mu(\bar{U}_i) + \mu(A)/\mu(\bar{U}_i)1_{\bar{U}_i} \]
and
\[ 1_B = (1_B - \mu(B)/\mu(\bar{V}_j))\mu(\bar{V}_j) + \mu(B)/\mu(\bar{V}_j)1_{\bar{V}_j} \]
into sums of mean-zero and constant functions. Note that the mean-zero parts are bounded by 1. We obtain that
\[ \langle T1_A, 1_B \rangle_2 = \frac{\mu(A)\mu(B)}{\mu(\bar{U}_i)\mu(\bar{V}_j)}\langle T1_{\bar{U}_i}, 1_{\bar{V}_j} \rangle_2 + \sqrt{\mu(\bar{U}_i)\mu(\bar{V}_j)}O(|F|^{-1/12}). \]

On the other hand, using 5.3,
\[ \langle T1_{\bar{U}_i}, 1_{\bar{V}_j} \rangle_2 = \int_{[0,1]^2} W(u, v)1_{\bar{U}_i}(u)1_{\bar{V}_j}(v)\,dudv \]
\[ = \frac{|E_s \cap (U_i \times V_j)|}{|U||V|} = (f_{ij}(\bar{s}) + O(|F|^{-1/2}))\mu(\bar{U}_i)\mu(\bar{V}_j), \]
hence
\[ \langle T1_A, 1_B \rangle_2 = \mu(A)\mu(B)f_{ij}(\bar{s}) + \sqrt{\mu(\bar{U}_i)\mu(\bar{V}_j)}O(|F|^{-1/12}). \]
Thus, for measurable $A \subseteq \bar{U}_i$ and $B \subseteq \bar{V}_j$, we have
\[ \int_{[0,1]^2} (W_s(u, v) - f_{ij}(\bar{s}))1_A(u)1_B(v) \ll \sqrt{\mu(\bar{U}_i)\mu(\bar{V}_j)}|F|^{-1/12}. \]

For arbitrary measurable $A, B \subseteq [0, 1]$, writing $A_i = A \cap \bar{U}_i$ and $B_j = B \cap \bar{V}_j$,
\[ d_{\square}(W_s, W_{\bar{s}}(F)) = \sup_{A, B \subseteq [0,1]} \int_{[0,1]^2} (W_s - W_{\bar{s}}(F))1_A1_B \]
\[ = \sup_{A, B \subseteq [0,1]} \sum_{i,j} \int_{[0,1]^2} (W_s - f_{ij}(\bar{s}))1_{A_i}1_{B_j} + O(|F|^{-1}) \]
\[ \ll \sum_{i,j} |F|^{-1/12}\sqrt{\mu(\bar{U}_i)\mu(\bar{V}_j)} + O(|F|^{-1}) \]
\[ \leq \sqrt{m}\bar{m}|F|^{-1/12} + O(|F|^{-1}), \]
where the summation is over the large blocks and, by (5.6), and the behaviour over non-large/lower-dimensional blocks is absorbed by the $O(|F|^{-1})$ term. The last inequality follows from Jensen’s inequality.

The desired conclusion follows from the fact that $\bar{m}$ and $\bar{n}$ are bounded independently of $F$.

Combining 5.3 and 5.4, we obtain the following.
Corollary 5.6. Let $\Gamma$ be a definable graph of finite relative dimension over $S$ on an asymptotic class $C$ relative to $\chi$. The set of accumulation points of the family of finite graphs

$$\{\Gamma_s(F) : F \in C, s \in S(F)\}$$

in the space $\tilde{W}_0$ of graphons has a set of representatives which is a finite set of stepfunctions.

6. Finite fields

The original algebraic regularity lemma by Tao in [23] was formulated for graphs definable in the language of rings over finite fields. We give several examples of graphons arising in this context by along the lines of our graphon formulation 5.6.

Example 6.1. Consider the definable graph $\Gamma = (U, V, E)$ where $U = V = \mathbb{A}^1$, and the edge relation is

$$E(x, y) \equiv \exists z \ x + y = z^2.$$

We are interested in accumulation points in the graphon space of the set of graphs

$$\Gamma(\mathbb{F}_q),$$

as $\mathbb{F}_q$ ranges over all finite fields (these are in fact a symmetric variant of the well-known Paley graphs). When $\text{char}(\mathbb{F}_q) \neq 2$, being a square is an event of ‘CDM probability’ (in the sense of 3.15) $1/2$. Thus, the edge density is approximately $1/2$ and these graphs accumulate around the constant graphon $W(1/2)$. These graphs are interesting in model theory due to the fact that their ultraproduct limit is the random graph.

For $\text{char}(\mathbb{F}_q) = 2$, everything is a square, so the graphs accumulate around the constant graphon $W(1)$. The definable stepfunction $W : U \times V \to \mathbb{Q}$ that regularises $\Gamma$ is

$$W(x, y) = \begin{cases} 1/2 & \text{if } \exists z \ z + z \neq 0; \\ 1 & \text{if } \forall z \ z + z = 0. \end{cases}$$

Hence, the set of accumulation points is

$$\{W(1/2), W(1)\}.$$

Example 6.2. Consider the definable graph $\Gamma = (\mathbb{A}^1, \mathbb{A}^1, E)$ with the edge relation

$$E(x, y) \equiv \exists z \ xy = z^2.$$
The definable sets
\[ U_0(x) \equiv \exists z \ x = z^2 \quad \text{and} \quad U_1 = \mathbb{A}^1 \setminus U_0 \]
partition \( \mathbb{A}^1 \). Let \( W \) be the definable function which returns 1 on \( U_0 \times U_0 \cup U_1 \times U_1 \) and 0 elsewhere. The definable stepfunction \( W \) regularises \( \Gamma \), and the set of accumulation points of \( \Gamma(\mathbb{F}_q) \) in the space of graphons is
\[ \left\{ W \left( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right), W(1) \right\}, \]
where the former denotes the stepfunction associated to the graph with edge weights matrix \( \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \) as in [2,7], which is the accumulation point of \( \Gamma(\mathbb{F}_q) \) for odd \( q \), and the latter is the constant graphon 1, which is the accumulation point of \( \Gamma(\mathbb{F}_q) \) for even \( q \).

In the following example, unlike in Example 6.1, the limit behaviour is not determined by the characteristic only.

**Example 6.3.** Let us define the graph \( \Gamma \) on \( \mathbb{A}^1 \) through the edge relation
\[ E(x, y) \equiv \exists z \ x + y = z^3. \]
In fact, the set of cubes in \( \mathbb{F}_q \) is of size \( (q - 1)/3 + 1 \) if \( 3|q - 1 \) and of size \( q \) otherwise.

The first condition happens precisely when the primitive cube root of unity belongs to \( \mathbb{F}_q \), which is expressible by the first-order condition
\[ \exists t \ t^3 = 1 \land t \neq 1 \]
so \( \Gamma \) is definably regularised. To conclude, the graphs \( \Gamma(\mathbb{F}_q) \) accumulate around \( W(1/3) \) when \( 3|q - 1 \), and around \( W(1) \) otherwise.

The following example illustrates the need to introduce parameters.

**Example 6.4.** Consider the graph \( \Gamma \) on \( \mathbb{A}^1 \) with the edge relation
\[ E(x, y) \equiv \exists z \ xy = z^3. \]
The graph is trivially considered over the base \( S = \mathbb{A}^0 \). Let
\[ \tilde{S}(z) \equiv z^3 = 1 \land z \neq 1, \]
and consider the definable sets \( U_i \) over \( \tilde{S} \), \( i = 0, 1, 2 \):
\[ U_i(x; z) \equiv \exists u \ (x = z^i u^3). \]
We see that \( \Gamma \) is regularised by the definable stepfunction \( W(x, y; z) \) over \( \tilde{S} \) taking value 1 on \( U_0 \times U_0 \cup U_1 \times U_1 \cup U_2 \times U_1 \) and 0 elsewhere, and
the graphs $\Gamma(\mathbb{F}_q)$ accumulate around the stepfunction

$$W = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

when $3|q - 1$, and around $W(1)$ when $3 \nmid q - 1$.

7. Fields with powers of Frobenius

7.1. Difference algebra. A difference ring is a pair

$$(R, \sigma)$$

consisting of a commutative ring with identity $R$ and an endomorphism $\sigma : R \to R$. We call it a difference field if $R$ happens to be a field.

A homomorphism of difference rings

$$f : (R, \sigma_R) \to (S, \sigma_S)$$

is a ring homomorphism $f : R \to S$ satisfying

$$\sigma_S \circ f = f \circ \sigma_R.$$

The category of difference rings is denoted by

$$\sigma\text{-}\text{Rng}$$

A difference ring $(R, \sigma)$ is a transformal domain if $R$ is a domain and $\sigma$ is injective.

Notation 7.1. For a prime power $q$, we write

$$K_q = (\overline{\mathbb{F}}_q, \varphi_q)$$

for the difference field consisting of the algebraic closure of a finite field, together with a power of the Frobenius automorphism

$$\varphi_q(x) = x^q.$$

7.2. Counting points over fields with Frobenius. In \cite{20}, we established the following difference analogue of \cite{6}.

Theorem 7.2. Let $X \to S$ be a definable map in the language of difference rings. Then there exist

1. a definable function $\mu_X : S \to \mathbb{Q} \cup \{\infty\}$,
2. a definable function $d_X : S \to \mathbb{N} \cup \{\infty\}$,
3. a constant $C_X > 0$,

so that, for every $K_q$ with $q > C_X$ and every $s \in S(K_q)$,

$$| |X_s(K_q)| - \mu_X(s) q^{d_X(s)}| \leq C_X q^{d_X(s) - 1/2}.$$
Corollary 7.3. The class $C$ of fields $K_q$ for prime powers $q$ is a CDM-class relative to the function $$\chi(K_q) = q.$$ 

Remark 7.4. An alternative approach to [7.3] is to study finite fields with powers of Frobenius, which requires the study of $N$-dimensional asymptotic classes. The authors of [9] go even further, and note that Tao’s regularity lemma applies in classes of structures with pseudofinite dimension.

7.3. Algebraic regularity lemma for fields with Frobenius. Using the fact [7.3] that fields with Frobenius constitute a relative CDM-class and regularity lemma 5.5, we obtain the following.

Corollary 7.5. Let $\Gamma = (U, V, E)$ be a definable bipartite graph of finite relative dimension over a definable set $S$ in the language of difference rings. There exists a constant $M = M(\Gamma) > 0$, a definable set $\tilde{S}$ over $S$ and a definable stepfunction $W$ over $\tilde{S}$ such that for every $K_q$ with $q > M$, every $\tilde{s} \in \tilde{S}(K_q)$ mapping onto $s \in S(K_q)$, $$d_{\square}(\Gamma_s(K_q), W_{\tilde{s}}(K_q)) \leq M q^{-1/12}.$$ 

Example 7.6. Let $U(x) \equiv x\sigma(x)^2 = 1$, and let $\Gamma$ be a graph on $U \times U$ with the edge relation $$E(x, y) \equiv \exists z \in U \ xy = z^3.$$ Then $U(K_q)$ is the group $\mu_{2q+1}$ of $(2q + 1)$-th roots of unity in $\bar{F}_q$, and the size of the set of cubes depends on whether $\mu_3 \subseteq \mu_{2q+1}$, i.e., whether $3|2q + 1$. Henceforth we can follow the reasoning from Example 6.4 and we obtain the same limit stepfunctions.

On the other hand, let us point out that the graphs $\Gamma(K_q)$ cannot be obtained from a graph interpretable over finite fields, which shows that the difference context is genuinely richer. Indeed, already in even characteristic, for each $q$, the smallest finite field containing the set $\mu_{2q+1} = U(K_q)$ is $\mathbb{F}_{4q^2}$, so the size of the set of realisations grows roughly as the square root of the size of the corresponding finite field, which cannot happen by the CDM-property 3.11 proved for finite fields in [6].

Example 7.7. Let $\Gamma$ be defined on the same $U$ as in Example 7.6 but with $$E(x, y) \equiv \exists z \in U \ x\sigma(y) = z^3.$$ In order to facilitate the notation, write $\Gamma = (U, U, E) = \Gamma(K_q)$, so that $U = \mu_{2q+1} \subseteq \bar{F}_q$. Let $\zeta \in \bar{F}_q$ be a primitive cube root of unity, and write $U_i = \zeta^i U^3$, for $i = 0, 1, 2$. We split the fields $K_q$ into subclasses:
(1) \( \zeta \notin U \), i.e. \( 3 \nmid 2q + 1 \);
(2) \( \zeta \in U \) and \( \zeta^q = \zeta \), i.e., \( 3|2q + 1 \) and \( 3|q - 1 \);
(3) \( \zeta \in U \) and \( \zeta^q = \zeta^2 \), i.e., \( 3|2q + 1 \) and \( 3 \nmid q - 1 \).

The set of corresponding limits is
\[
\left\{ W(1), W \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, W \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \right\}.
\]

8. Expander difference polynomials

8.1. Difference schemes.

Definition 8.1. If \((k, \sigma)\) is a difference ring, a difference \((k, \sigma)\)-algebra
is a difference ring \((A, \sigma)\) endowed with a difference ring homomorphism
\((k, \sigma) \to (A, \sigma)\). A morphism between \((k, \sigma)\)-algebras \((A, \sigma)\) and \((B, \sigma)\) is a
difference ring homomorphism \((A, \sigma) \to (B, \sigma)\) which commutes with the
structure maps \((k, \sigma) \to (A, \sigma)\) and \((k, \sigma) \to (B, \sigma)\). We write
\((k, \sigma)\)-Alg
for the category of \((k, \sigma)\)-algebras.

Definition 8.2. Let \((k, \sigma)\) be a difference ring. The difference polynomial
ring in variables \(x_1, \ldots, x_n\) is the difference ring in infinitely many variables
\[
k[x_1, \ldots, x_n]_\sigma = k[x_{1,i}, \ldots, x_{n,i} : i \geq 0],
\]
with \(\sigma\) inferred from the rule \(\sigma(x_{j,i}) = x_{j,i+1}\) and its action on \(k\). Informally
speaking, if \(x = (x_1, \ldots, x_n)\), we can write
\[
k[x]_\sigma = k[x, \sigma x, \sigma^2 x, \ldots].
\]
Note that \(k[x]_\sigma\) is naturally a \((k, \sigma)\)-algebra.

Definition 8.3. Let \((k, \sigma)\) be a difference ring. We say that a \((k, \sigma)\)-algebra
\(A\) is

(1) of finite \(\sigma\)-type, if it admits an epimorphism \(k[x_1, \ldots, x_n]_\sigma \to A\), for
some \(n\);
(2) of finite \(\sigma\)-presentation, if it is a quotient
\[
A \cong k[x_1, \ldots, x_n]_\sigma / \langle f_1, \ldots, f_m \rangle_\sigma,
\]
where \(\langle f_1, \ldots, f_m \rangle_\sigma\) is the difference ideal (ideal closed under \(\sigma\))
generated by \(f_1, \ldots, f_m \in k[x_1, \ldots, x_n]_\sigma\), for some \(m, n\).
Definition 8.4. Let $(A, \sigma)$ be a difference ring. The affine difference scheme associated with $A$ is the functor

$$X : \sigma\text{-Rng} \to \text{Set}, \quad X(R, \sigma) = \text{Hom}_{\sigma\text{-Rng}}((A, \sigma), (R, \sigma)).$$

A morphism between affine difference schemes is a natural transformation $\varphi : X \to Y$.

We shall only ever discuss affine difference schemes, so we omit the word ‘affine’ from our descriptions. The category of difference schemes is denoted $\sigma\text{-}\text{Sch}$.

Definition 8.5. Given $S \in \sigma\text{-}\text{Sch}$, a difference scheme over $S$ is a morphism $X \to S$. A morphism between schemes $X \to S$ and $Y \to S$ is given by a morphism $X \to Y$ which commutes with the structure morphisms to $S$, as in the commutative diagram from 3.7. We obtain the category of difference schemes over $S$,

$$\sigma\text{-}\text{Sch}/S.$$

Remark 8.6. By Yoneda’s lemma, if $X$ is associated with a difference ring $(A, \sigma)$ and $Y$ with $(B, \sigma)$, a morphism $X \to Y$ always arises from a difference ring homomorphism $(B, \sigma) \to (A, \sigma)$, whence the category of difference schemes is opposite to that of difference rings,

$$\sigma\text{-}\text{Sch} \simeq (\sigma\text{-Rng})^{\text{op}}.$$

Moreover, if $S$ corresponds to a difference ring $(k, \sigma)$, then

$$\sigma\text{-}\text{Sch}/S \simeq ((k, \sigma)-\text{Alg})^{\text{op}}.$$  

Definition 8.7. Let $f : X \to S$ be a morphism of difference schemes associated with a difference ring homomorphism $(k, \sigma) \to (A, \sigma)$. We say that

1. $X$ is transformally integral, if $(A, \sigma)$ is a transformal domain;
2. $f$ is of finite $\sigma$-type if $(A, \sigma)$ is a $(k, \sigma)$-algebra of finite $\sigma$-type;
3. $f$ is of finite $\sigma$-presentation over $S$ if $(A, \sigma)$ is a $(k, \sigma)$-algebra of finite $\sigma$-presentation.
4. $f$ is a $\sigma$-localisation, if $A = k[1/a]_\sigma$ for some $a \in k$.

Definition 8.8. Let $f = F(x, \sigma x, \ldots, \sigma^r x) \in k[x]_\sigma$, with $x = (x_1, \ldots, x_n)$, and some ordinary polynomial $F \in k[t_0, t_1, \ldots, t_r]$. The set of solutions to the equation

$$f(x) = 0$$
in a difference ring \((R, \varphi)\) extending \((k, \sigma)\) is defined as
\[
\{ a \in R^n : F(a, \varphi a, \ldots, \varphi^n a) = 0 \}.
\]

**Example 8.9.** The set of solutions of the difference equation \(\sigma x = x\) in the difference field \(K_q = (\overline{\mathbb{F}}_p, \varphi_q)\) is the finite field \(\mathbb{F}_q\).

**Remark 8.10.** Let \(S\) be a difference scheme associated with a difference ring \((k, \sigma)\), and consider a system of difference polynomial equations
\[
f_1(x_1, \ldots, x_n) = 0
\]
\[
\vdots
\]
\[
f_m(x_1, \ldots, x_n) = 0,
\]
where \(f_i \in k[x_1, \ldots, x_n]_{\sigma}\), and consider the difference scheme \(X\) over \(S\) associated to the \((k, \sigma)\)-algebra of finite \(\sigma\)-presentation
\[
A = k[x_1, \ldots, x_n]_{\sigma}/\langle f_1, \ldots, f_m \rangle_{\sigma}.
\]
For any \((k, \sigma)\)-algebra \((R, \sigma)\), the set of \((R, \sigma)\)-rational points
\[
X/S(R, \sigma) = \text{Hom}_{(k, \sigma)\text{-Alg}}(A, R)
\]
can be identified with the set of solutions of the above system in \((R, \sigma)\), i.e., the intersection of the solution sets of all equations \(f_i = 0\) in \((R, \sigma)\).

This observation clarifies how a difference scheme of finite presentation can be considered a (quantifier-free) *definable set* as in [3.1].

**Definition 8.11.** Let \(f : X \to S\) be a morphism of difference schemes of finite \(\sigma\)-presentation over \(Z\), and let \(C\) be the relative CDM-class of fields with Frobenius \(K_q = (\overline{\mathbb{F}}_p, \varphi_q)\). We say that

(1) \(f\) is of *finite relative dimension*, if it is such when considered as a definable map over \(C\) in the sense of [3.14];

(2) \(f\) is *dominant*, if the dimension of its image as the definable set is of full dimension in \(S\);

(3) \(f\) is *finite*, if the map \(f(K_q)\) between realisation sets has finite fibres for all \(K_q \in C\);

(4) \(f\) is *generically finite*, if it fails to have finite fibres only over a lower-dimensional subset of \(S\).

**Remark 8.12.** All of the above properties of difference schemes and their morphisms have intrinsic formulations rooted in difference algebra, as given in [11] and [20]. In particular, the notion of *total dimension* of difference schemes is shown to agree with the CDM-dimension in [20]. We choose these equivalent formulations in order to simplify the presentation.
Terminology 8.13. Following the comparison of difference schemes to definable sets made in 8.10, we shall no longer use the boldface notation to denote difference schemes.

Definition 8.14. Let $X$ and $Y$ be difference schemes. A correspondence $X \leadsto Y$ is given by a difference scheme $W \subseteq X \times Y$ such that the projection $W \to X$ is dominant with generically finite fibres.

8.2. Algebraic constraint and group configuration. The results in this subsection parallel those of Tao in [23].

Definition 8.15. Let $S$ be a transformally integral difference scheme of finite $\sigma$-presentation over $\mathbb{Z}$ and let $X_1, \ldots, X_n, Y$ be difference schemes of finite $\sigma$-presentation and finite relative dimension over $S$, and let $f : X_1 \times_S \cdots \times_S X_n \to Y$ be a morphism of difference schemes over $S$. We say that $f$ is a moderate asymmetric expander if there exist constants $c, C > 0$ such that, for every difference field $K_q = (\mathbb{F}_p, \varphi_q)$, every $s \in S(K_q)$ and every choice of $A_i \subseteq X_{i,s}(K_q)$ with $|A_i| \geq C|X_{i,s}(K_q)|^{1-c}$, we have

$$|f_s(A_1, \ldots, A_n)| \geq C^{-1}|Y_s(K_q)|.$$

Assumption 8.16 (Expansion dichotomy). Let $X, Y, Z$ be difference schemes of finite $\sigma$-presentation over a transformally integral $S$ with geometrically transformally integral fibres of finite relative dimension, and let $f : X \times_S Y \to Z$ be a morphism of difference schemes over $S$. Then at least one of the following statements hold:

1. (Algebraic constraint). The morphism $X \times_S X \times_S Y \times_S Y \to Z \times_S Z \times_S Z$,

$$(x, x', y, y') \mapsto (f(x, y), f(x, y'), f(x', y), f(x', y'))$$

is not dominant.

2. (Moderate expansion property). There exists a $\sigma$-localisation $S'$ of $S$ and a constant $C > 0$ such that for every $K_q$, every $s \in S'(K_q)$, and every $A \subseteq X_s(K_q)$, $B \subseteq Y_s(K_q)$, if $|A||B| \geq Cq^{-1/8}|X_s(K_q)||Y_s(K_q)|$, then $|f_s(A, B)| \geq C^{-1}|Z_s(K_q)|$. In particular, the morphism $f$ is a moderate asymmetric expander above $S'$.

We have gone through the steps and methods used in the proof of [23, Theorem 38] and consider that they can be followed formally to derive the
above statement from 7.5. On the other hand, writing out the full proof would require us to repeat lengthy passages of Tao’s work, including the self-improvement trick from [22] to obtain the error term of $O(|F|^{-1/4})$ in 7.5 and the generalisation of the regularity lemma 7.5 to hypergraphs following [23, Theorem 35].

**Theorem 8.17** (Solving the algebraic constraint in dimension 1). With notation of 8.16, assume that $X$, $Y$ and $Z$ are of relative dimension 1 over $S$ and that the morphism $X \times_S X \times_S Y \times_S Y \to Z \times_S Z \times_S Z \times_S Z$ given by

$$(x, x', y, y') \mapsto (f(x, y), f(x, y'), f(x', y), f(x', y'))$$

is not dominant. Then there exists a generically finite morphism $S' \to S$ and an algebraic group scheme $(A, \ast)$ over $S'$, which is either $\mathbb{G}_a$, $\mathbb{G}_m$ or of an elliptic curve over $S'$, and correspondences $\chi : X \leadsto A$, $\upsilon : Y \leadsto A$ and $\zeta : Z \leadsto A$ over $S'$ such that $f$ is in correspondence with the group law on $A$,

$$\zeta(f(x, y)) = \chi(x) \ast \upsilon(y).$$

The proof requires a vast amount of model theory and it is no longer possible to keep our exposition self-contained, hence we provide a guide to the literature for the interested reader.

- Ultraproducts of CDM-classes are measurable structures in the sense of [16] and can be treated as simple theories, which are governed by a notion of forking independence. For background on simple theories and for the basics of forking calculus we refer the reader to [4] and [26].
- For group configuration in simple theories, we refer the reader to [2] and [25].
- Hrushovski shows in [11] that the limit theory of the relative CDM-class of fields with Frobenius is the theory ACFA of existentially closed difference fields. It is extensively studied in [5].
- For groups definable in algebraically closed fields, we shall use [3], and for groups definable in ACFA we refer to [12].

Martin Bays suggested that our proof can benefit from field-theoretic arguments specific to ACFA. We provide the summary of properties we use, which can all be found in [5].

**Fact 8.18.** Let $(\Omega, \sigma)$ be a model of ACFA.

(a) While ACFA is a simple theory, the underlying field $\Omega$ is algebraically closed, so stability theory applies to it.
(b) The model-theoretic algebraic closure of \( A \subset \Omega \), denoted \( \text{acl}(A) \), is the field-theoretic algebraic closure of the inversive difference subfield of \( \Omega \) generated by \( A \).

c) The forking independence is witnessed by algebraic independence in the sense that, for \( A, B, C \subset \Omega \), \( A \downarrow_C B \) if and only if the fields \( \text{acl}(AC) \) and \( \text{acl}(BC) \) are linearly disjoint over \( \text{acl}(C) \).

d) Let \( a \in \Omega \), let \( k \) be a difference subfield of \( \Omega \), and let \( k(a)_{\sigma} \) be the difference field generated by \( a \) over \( k \). If \( \text{tr.deg}(k(a)_{\sigma}/k) = 1 \), then \( \sigma(a) \in \text{acl}(k(a)) \) and thus \( a \not\equiv_k \sigma(a) \). This amounts to saying that the difference locus of \( a \) over \( k \) is of total dimension 1, or, equivalently, of CDM-dimension 1. Such an element also has SU-rank 1 over \( k \). Henceforth we shall simply refer to such elements as having rank 1.

e) The difference locus of a tuple \( a \in \Omega \) of rank 1 elements is dense in its algebraic locus, and hence it is in a correspondence with its algebraic locus in the sense of \( 8.14 \).

(f) Given \( a \in \Omega \) and a difference subfield \( k \) of \( \Omega \), the canonical basis \( Cb(a/k) \) is the smallest perfect difference field over which the locus of \( a \) over \( k \) is defined. It equals, in the pure field language, to the canonical base of the type of \( (a, \sigma(a), \ldots) \) over \( k \). It has the property that \( a \not\equiv_{Cb(a/k)} k \). In arbitrary simple theories, canonical bases exist only as ‘hyperimaginaries’, and one is forced to work with bounded closure \( \text{bdd} \) as opposed to the algebraic closure. In ACFA canonical bases exist in the real word, so any occurrence of \( \text{bdd} \) below can be replaced by \( \text{acl} \).

**Proof.** Our first objective is to extract a group configuration from the algebraic constraint. We will use the context and notation of partial generic multiactions familiar from [1], [24], [2] and [25], which works in a simple setting, suitable for dealing with arbitrary asymptotic classes whose non-principal ultraproducts are supersimple of finite SU-rank.

An interested reader can reformulate it in purely field-theoretic language using \( 8.18 \).

The base difference scheme \( S \) is associated with a transformal domain \((R, \varsigma)\), so we can take its fraction field \((k, \varsigma)\) and work in a large model \((\Omega, \sigma)\) of ACFA extending it.

Let \( \eta \) be a scheme-theoretic generic point of \( S \), and write \( X_\eta, Y_\eta \) and \( Z_\eta \) for the generic fibres of \( X, Y \) and \( Z \) over \( S \), considered as difference schemes
over $k$. For simplicity of notation, all the independences and bounded/algebraic closures and types we write below will be over $k$.

Let $(x_0, y_0, z_{00})$ in $\Omega$ be a generic point over $k$ on the graph of $f_\eta$, i.e., $x_0 \in X_\eta$, $y_0 \in Y_\eta$ are rank 1 points over $k$, with $x_0 \downarrow y_0$ and $z_{00} = f_\eta(x_0, y_0) \in Z_\eta$.

We may assume that $z_{00} \downarrow x_0$ and $z_{00} \downarrow y_0$, which implies that $x_0 \in bdd(y_0, z_{00})$ and $y_0 \in bdd(x_0, z_{00})$. Otherwise, we would have that $z_{00} \in bdd(x_0)$ or $z_{00} \in bdd(y_0)$ and we are in a degenerate case where the function $f_\eta$ defines a correspondence dependent on a single variable, which leads to the conclusion of the theorem in a trivial way.

Let $\pi \subseteq X_\eta \times Y_\eta \times Z_\eta$ be the partial type over $k$ refining the locus of $(x_0, y_0, z_{00})$ so that

$$\models \pi(x, y, z)$$

provided $f_\eta(x, y) = z$ and the points $x \in X_\eta$, $y \in Y_\eta$, $z \in Z_\eta$ are pairwise independent. It is handy to abbreviate this by writing $yz \models x$. Note, if $yz \models x$, then $z \in dcl(xy)$, $y \in bdd(xz)$ and $x \in bdd(yz)$. Thus, $\pi$ is a generic invertible multiaction as in [11 Definitions 2.1, 2.3] and [24 2.1].

The assumption of non-dominance in this context states that for all $x_0$, $x_1$, $y_0$, $y_1$, $z_{ij} = f_\eta(x_i, y_j)$ (or $y_jz_{ij} \models x_i$), the set $\{z_{00}, z_{01}, z_{10}, z_{11}\}$ is dependent.

We claim that

$$\pi \circ \pi^{-1}$$

is a generic multiaction in the following sense (cf. [11 Definition 2.5, Section 2.3], [24 2.2, 2.3], [25 1.2]). Starting with an independent triple $x_0$, $x_1$, $y_0$, let $y_0z_{00} \models x_0$, $y_0z_{10} \models x_1$, so that $z_{00}z_{10} \models x_1 \circ x_0^{-1}$. Let $h = Cb(z_{00}z_{10}/x_0x_1) \in \pi \circ \pi^{-1}$. We need to verify that $h \downarrow x_0$ and $h \downarrow x_1$.

Let us choose $y_1 \equiv_{x_0x_1} y_0$ such that $y_1 \downarrow_{x_0x_1} y_0$. It follows that the set $\{y_0, y_1, x_0, x_1\}$ is independent. We denote by $z_{01}$ and $z_{11}$ the elements satisfying $y_1z_{01}z_{11} \equiv_{x_0x_1} y_0z_{00}z_{10}$, which implies that $y_1z_{01} \models x_0$ and $y_1z_{11} \models x_1$.

Using $y_1 \downarrow_{x_0x_1} y_0$, we get that $z_{01} \downarrow_{x_0x_1} z_{00}z_{10}$, which, together with $z_{01} \downarrow x_0x_1$ yields $z_{01} \downarrow x_0x_1z_{00}z_{10}$ and $z_{01} \downarrow z_{00}z_{10} x_0x_1$. Using the non-dominance assumption, $z_{11} \in bdd(z_{01}z_{00}z_{10})$, so

$$z_{01}z_{11} \downarrow_{z_{00}z_{10}} x_0x_1.$$  

We conclude that

$$h = Cb(z_{00}z_{10}/x_0x_1) = Cb(z_{01}z_{11}/x_0x_1) = Cb(z_{01}z_{11}/z_{00}z_{10}) \in bdd(z_{00}z_{10}).$$

Thus, since $z_{00}z_{10} \downarrow x_0$ and $z_{00}z_{10} \downarrow x_1$, we confirm that $h \downarrow x_0$ and $h \downarrow x_1$. 


In terminology of [1], [24], [2] and [25], it follows that the set of germs of the multiaction \( \pi \circ \pi^{-1} \) is a polygroup chunk, which entails that the tuple

\[
(x_0, x_1, h, y_0, z_{00}, z_{10})
\]

forms a group configuration, conveniently depicted by the diagram

```
    z_{00}
   /   \
 x_0 -- y_0
   |  /   \
   x_1 -- z_{10}
     \  |
      h
```

where:

1. any three non-collinear points are independent;
2. in any triple of collinear points, any two are interbounded over the third;
3. \( x_0 \) is interbounded with \( Cb(y_0z_{00}/x_0) \), \( x_1 \) is interbounded with \( Cb(y_0z_{10}/x_1) \) and \( h \) is interbounded with \( Cb(z_{00}z_{10}/h) \).

Instead of using the group configuration for simple theories at this point, we prefer to use techniques specific to ACFA to construct a definable group.

Using the fact that this is a rank 1 group configuration in \( (\Omega, \sigma) \), the same tuple forms a group configuration over \( k \) in the underlying algebraically closed field \( \Omega \). Indeed, the independence property (1) in the field reduct is clear, and interalgebraicity property (2) follows from \( 8.18(b), (d) \). The property (3) in the reduct is follows from the description of canonical bases \( 8.18(f) \) in the field language specialised to rank 1.

The (stable) group configuration for algebraically closed fields yields an ACF-definable group \( (G, \ast) \) over the algebraic closure \( \bar{k} \) so that the algebraic locus of \( (x_0, y_0, z_{00}) \) over \( \bar{k} \) is in correspondence with the group law on \( G \). Moreover, by results of van den Dries and Hrushovski (described in [3]), there is a 1-dimensional algebraic group \( A_0 \) defined over \( \bar{k} \) whose group law is in correspondence with that of \( G \).

Using \( 8.18 \) the graph of \( f_\eta \) is in correspondence with the difference locus of \( (x_0, y_0, z_{00}) \) over \( k \), which is in correspondence with the algebraic locus of \( (x_0, y_0, z_{00}) \) over \( k \), which is in correspondence with the locus of the same tuple over \( \bar{k} \), which is in correspondence with the group law in \( A_0 \) over \( \bar{k} \).

Alternatively, we could have argued, using [12] Lemma 3.3] on the 6-tuple above, that the graph of \( f_\eta \) is in correspondence with a difference definable group \( H \) over \( \bar{k} \), and, by [12] Theorem 3.1], \( H \) is virtually definably embeddable into an algebraic group \( A \) of dimension 1 over \( \bar{k} \).
In either case, the theorem lists all three possibilities for the 1-dimensional group $A_0$. Moreover, $A_0$ is defined over a finite extension $k'$ of $k$, so let $(L, \sigma)$ be the difference extension field of $(k, \varsigma)$ generated by $k'$. Through standard constructibility arguments of $\sigma$-localising $S$ to make it normal, and considering the normalisation of $S$ in $L$, we obtain a generically $\sigma$-finite difference scheme $S' \to S$ and a group scheme $A$ over $S'$ whose generic fibre is $A_0$, and correspondences over $S'$ relating the morphism $f$ and the group law in $A$.

\[\square\]

**Remark 8.19.** Combining Assumption 8.16 and Theorem 8.17, we obtain the following.

Let $X, Y, Z$ be difference schemes of finite transformal type and relative total dimension 1 over $S$, and let

$$f : X \times_S Y \to Z$$

be a morphism of difference schemes over $S$. Then at least one of the following statements hold.

1. The morphism $f$ corresponds to the additive or multiplicative group law in a way described in 8.17.
2. The morphism $f$ corresponds to the addition law on an elliptic curve in a way described in 8.17.
3. The morphism $f$ is a moderate asymmetric expander over a $\sigma$-localisation of $S$.

**Appendix A. Inequalities**

**Fact A.1.** Let $H$ be a Hilbert space. We have the following inequalities.

1. (Cauchy-Schwartz inequality). For $x, y \in H$, we have

$$|\langle x, y \rangle| \leq \|x\| \|y\| .$$

2. (Bessel’s inequality). If $e_i$ is an orthonormal sequence in $H$, for every $x \in H$

$$\sum_i |\langle x, e_i \rangle|^2 \leq \|x\|^2 .$$

**Fact A.2** (Hölder’s inequality). Let $(X, \mu)$ be a measure space. Let $p, q \in [1, \infty]$ be Hölder conjugates, namely satisfying $1/p + 1/q = 1$. Then, for all measurable real or complex functions $f, g$ on $X$,

$$\|fg\|_1 \leq \|f\|_p \|g\|_q .$$

Note that, for $p = q = 2$, the above yields the Cauchy-Schwartz inequality for $L^2(X, \mu)$.
Lemma A.3. If \( a_i, b_i, u, v \) are positive real numbers satisfying
\[
\sum_i a_i^3 b_i \leq uv \quad \text{and} \quad \sum_i b_i \leq v,
\]
then
\[
\sum_i a_i b_i \leq u^{\frac{1}{3}} v.
\]

Proof. Hölder’s inequality, applied to the space \( X = \mathbb{N} \) with the counting measure, states that for real numbers \( x_i, y_i, i \in \mathbb{N} \), and \( p, q \) satisfying \( 1/p + 1/q = 1 \),
\[
\sum_i |x_i y_i| \leq \left( \sum_i |x_i|^p \right)^{\frac{1}{p}} \left( \sum_i |y_i|^q \right)^{\frac{1}{q}}.
\]
We can now write
\[
\sum_i a_i b_i = \sum_i a_i^{\frac{1}{3}} b_i^{\frac{2}{3}} \leq \left( \sum_i a_i^3 b_i \right)^{\frac{1}{3}} \cdot \left( \sum_i (b_i^2)^{\frac{3}{2}} \right)^{\frac{2}{3}} \leq u^{\frac{1}{3}} v^{\frac{1}{3}} \cdot v^{\frac{2}{3}} = u^{\frac{1}{3}} v.
\]

Acknowledgements. Mirna Džamonja received funding from the European’s Union Horizon 2020 research and innovation programme under the Maria Skołodowska-Curie grant agreement No 1010232. She also thanks the Institut d’Histoire et de Philosophie des Sciences et des Techniques, CNRS et Université Panthéon-Sorbonne, Paris, where she is an Associated Member and the University of East Anglia, Norwich, UK, where she is a Visiting Professor. The authors thank the referee for a very careful, detailed and helpful referee report.

References

[1] Itay Ben-Yaacov. Group configurations and germs in simple theories. J. Symbolic Logic 67(4):1581–1600, 2002.
[2] Itay Ben-Yaacov, Ivan Tomašić, and Frank O. Wagner. Constructing an almost hyperdefinable group. J. Math. Log., 4(2):181–212, 2004.
[3] Elisabeth Bouscaren. Model theoretic versions of Weil’s theorem on pre-groups. In Model Theory of Groups, (A. Nesin and A. Pillay, eds.), Notre Dame Press, 1989.
[4] Enrique Casanovas. Simple theories and hyperimaginaries. Lecture Notes in Logic, 39. Association for Symbolic Logic; Cambridge University Press, 2011.
[5] Zoé Chatzidakis and Ehud Hrushovski. Model theory of difference fields. Trans. Amer. Math. Soc., 351(8):2997–3071, 1999.
[6] Zoé Chatzidakis, Lou van den Dries, and Angus Macintyre. Definable sets over finite fields. *J. Reine Angew. Math.*, 427:107–135, 1992.

[7] Artem Chernikov and Sergei Starchenko. Regularity lemma for distal structures. *J. Eur. Math. Soc.*, 20(10):2437–2466, 2018.

[8] Richard Elwes and Dugald Macpherson. A survey of asymptotic classes and measurable structures. In *Model theory with applications to algebra and analysis. Vol. 2*, volume 350 of *London Math. Soc. Lecture Note Ser.*, pages 125–159. Cambridge Univ. Press, Cambridge, 2008.

[9] Dario Garcia, Dugald Macpherson, and Charles Steinhorn. Pseudofinite structures and simplicity. *J. Math. Log.*, 15(1), 2015.

[10] Ehud Hrushovski. Expanding comments. Private correspondence with Terence Tao, 2013.

[11] Ehud Hrushovski. The elementary theory of the Frobenius automorphisms. *arXiv:math/0406514*, 2004.

[12] Piotr Kowalski and A. Pillay. A note on groups definable in difference fields. *Proc. Amer. Math. Soc.*, 130(1):205–212, 2002.

[13] Yoshiharu Kohayakawa. Szemerédi’s regularity lemma for sparse graphs. In *Foundations of Computational Mathematics, Rio de Janeiro 1997*, pages 216–230, Springer-Berlin, 1997.

[14] László Lovász. *Large networks and graph limits*, volume 60 of *American Mathematical Society Colloquium Publications*. American Mathematical Society, Providence, RI, 2012.

[15] László Lovász and Balázs Szegedy. Szemerédi’s lemma for the analyst. *Geom. Funct. Anal.*, 17, 2007.

[16] Dugald Macpherson and Charles Steinhorn. One-dimensional asymptotic classes of finite structures. *Trans. Amer. Math. Soc.*, 360(1):411–448, 2008.

[17] Maryanthe Malliaris and Saharon Shelah. Regularity lemmas for stable graphs. *Trans. Amer. Math. Soc.*, 366(3):1551–1585, 2014.

[18] Jaroslav Nešetřil and Patrice Ossona de Mendez. A unified approach to structural limits and limits of graphs with bounded tree-depth. *Memoirs of the American Mathematical Society* 263(1272). American Mathematical Society, Providence, RI, 2020.

[19] Anand Pillay and Sergei Starchenko. Remarks on Tao’s algebraic regularity lemma. *arXiv:1310.7538*, 2013.

[20] Mark Ryten and Ivan Tomašić. ACFA and measurability. *Selecta Math. (N.S.),* 11(3-4):523–537, 2005.
[21] Alexander Scott. Szemerédi regularity lemma for matrices and sparse graphs. *Combin. Probab. Comput.*, 20(3):455–466, 2011.

[22] Terence Tao. A spectral theory proof of the algebraic regularity lemma. https://terrytao.wordpress.com/2013/10/29/a-spectral-theory-proof-of-the-algebraic-regularity-lemma/, 2013.

[23] Terence Tao. Expanding polynomials over finite fields of large characteristic, and a regularity lemma for definable sets. *Contrib. Discrete Math.*, 10(1):22–98, 2015.

[24] Ivan Tomašić. Geometric simplicity theory. PhD thesis, University of Edinburgh, 2001.

[25] Ivan Tomašić and Frank O. Wagner. Applications of the group configuration theorem in simple theories. *J. Math. Log.*, 3(2):239–255, 2003.

[26] Frank O. Wagner. Simple theories. *Mathematics and its Applications*, 503. Kluwer Academic Publishers, 2000.