Graphons arising from graphs definable over finite fields

Mirna Džamonja, Ivan Tomašić

COLLOQUIUM MATHEMATICUM

VOL. 169 2022 NO. 2

GRAPHONS ARISING FROM GRAPHS DEFINABLE OVER FINITE FIELDS

BY

MIRNA DŽAMONJA (Paris) and IVAN TOMAŠIĆ (London)

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 summary of results. Tao's algebraic 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 definable 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*. The result was proved by Tao [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 history. In a private correspondence to Tao, Hrushovski [10] gave another proof using the model-theoretic tools for studying the growth rates of definable sets over finite fields, as developed by Chatzidakis–van den Dries–Macintyre [5]. 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' structures studied in the the context of asymptotic classes of finite structures by Macpherson–Steinhorn [16] and Elwes–Macpherson [8]. García–Macpherson–Steinhorn [9] state, without proof, a

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

Key words and phrases: graphons, regularity lemma, finite fields, Frobenius automorphism, ACFA.

Received 8 September 2020; revised 23 December 2021.

Published online 9 March 2022.

version of the lemma for classes of finite structures controlled by suitable pseudofinite dimensions, which potentially generalise asymptotic classes.

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 by Tao in the revisited and simplified proof of his lemma 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 (see Section 1.4, Corollary 7.5 below) 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 the growth rates of definable sets are controlled by rational monomials; see Definition 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 the *stable* and [7] in the *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 regularis-

ing 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 statement of the result is uniform as discussed in Section 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 Section 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.
(4) While finite fields are extensively used in combinatorics, fields with powers of Frobenius touched upon in Section 1.4 are less known. We hope that our treatment will introduce them to a wider audience.
(5) As pointed out by Hrushovski in [10], as well as 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 Theorem 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 [12], such an analogue was found by Scott [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 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.

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, 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 $\tilde{\mathscr{W}}$ of *kernels* is the space of 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 $\tilde{\mathscr{W}}$ 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 equivalence classes of functions, by taking representatives and by using the notation $\mathscr{W}$ in place of $\tilde{\mathscr{W}}$. Sometimes we need to be more explicit, for example a suitable metric on $\tilde{\mathscr{W}}$ will be introduced below, but it is only a pseudometric on $\mathscr{W}$.
(2) The space of *graphons* is
$$\mathscr{W}_0 = \{W \in \mathscr{W} : 0 \le W \le 1\}.$$
(3) We write
$$\mathscr{W}_1 = \{W \in \mathscr{W} : -1 \le W \le 1\}.$$

The inequalities are given up to a measure 0 set (see Definition 2.3).

Remark 2.2. The classical literature on graphons [14] usually stipulates that kernels and graphons have to be symmetric functions, and we write $\mathscr{W}^{\rm sym}$, $\mathscr{W}_0^{\rm sym}$ and $\mathscr{W}_1^{\rm sym}$ 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

$$\mathscr{W} \to \mathscr{W}^{\mathrm{sym}},$$

assigning to every $W \in \mathscr{W}$ the symmetric 'bipartite kernel'

$$W^{\mathrm{sym}}(x,y) = \begin{cases} W(2x, 2y-1) & \text{when } (x,y) \in [0,1/2] \times [1/2,1], \\ W^*(2x-1, 2y) & \text{when } (x,y) \in [1/2,1] \times [0,1/2], \\ 0 & \text{elsewhere.} \end{cases}$$

Here W^* denotes the *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 Definition 2.13 and will preserve the membership in $\mathscr{W}$, $\mathscr{W}_0$ and $\mathscr{W}_1$, 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 *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 *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 $[0,1] = \coprod_{i=1}^{n} U_i$ and $[0,1] = \coprod_{j=1}^{m} V_j$ into measurable sets such 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* $\Gamma = (U, V, E, w)$ is 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 $[0,1] = \coprod_{i\in U} U_i$ and $[0,1] = \coprod_{j\in V} V_j$ 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

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

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

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

NOTATION 2.7. Let Γ 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 $W \in \mathscr{W}$ 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*

$$Tf = \sum_i \sigma_i \langle f, u_i\rangle v_i,$$

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

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

The spectrum of T is discrete, and the non-zero eigenvalues λ_i satisfy $\lim_i \lambda_i = 0$. If W is symmetric, the eigenvalues λ_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 λ_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

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

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 φ_i mapping I_i onto $[0,1]$, and ψ_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 $\mathscr{W}$.

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 $\mathscr{W}$ of kernels is defined by

$$\|W\|_\square = \sup_{S,T \subseteq [0,1]} \Big| \int_{S \times T} W(x,y)\,dx\,dy \Big|,$$

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 $W \in \mathscr{W}_1$, we have the inequalities*

$$\|W\|_\square \le \|W\|_1 \le \|W\|_2 \le \|W\|_\infty \le 1.$$

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 Γ is *ϵ-homogeneous of density $w \in [0,1]$* provided, for every $A \subseteq U$ and $B \subseteq V$,

$$\big||E \cap (A \times B)| - w|A|\,|B|\big| \leq \epsilon |U|\,|V|.$$

(2) We say that Γ is *ϵ-regular of density $w \in [0,1]$* if, for every $A \subseteq U$ with $|A| > \epsilon|U|$ and $B \subseteq V$ with $|B| > \epsilon|V|$,

$$\big||E \cap (A \times B)| - w|A|\,|B|\big| \leq \epsilon |A|\,|B|.$$

The well-known connection between regularity lemmas (the existence of a homogeneous partition) and the cut metric (see discussion following [14, 9.3]) 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 Γ is ϵ-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 $U = \coprod_{i=1}^n U_i$ and $V = \coprod_{j=1}^m V_j$ such that $\Gamma \restriction (U_i \times V_j)$ is ϵ-homogeneous with density w_{ij}. Let $\bar\Gamma$ be a weighted bipartite graph on U, 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 Γ 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 $W(\Gamma)$, then using (1), for measurable $A_i \subseteq \bar U_i$ and $B_j \subseteq \bar V_j$ we obtain

$$\Big| \int\limits_{A_i \times B_j} (W(\Gamma)(x,y) - w_{ij})\,dx\,dy \Big| \leq \epsilon \mu(\bar U_i)\mu(\bar V_j).$$

Hence, for measurable $A, B \subseteq [0,1]$,

$$\begin{aligned}\Big| \int\limits_{A \times B} (W(\Gamma)(x,y) - W(\bar\Gamma))\,dx\,dy \Big| &= \Big| \sum_{i,j} \int\limits_{A_i \times B_j} (W(\Gamma)(x,y) - w_{ij})\,dx\,dy \Big| \\ &\leq \sum_{i,j} \epsilon \mu(\bar U_i)\mu(\bar V_j) = \epsilon. \ \blacksquare\end{aligned}$$

2.7. The space of graphons. The cut distance is only a pseudometric on $\mathscr{W}_0$, so here we work with $\tilde{\mathscr{W}}_0$, the associated metric space of classes of graphons at zero cut distance from each other.

THEOREM 2.18. *The space* $(\tilde{\mathscr{W}}_0, \delta_\square)$ *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 Remark 2.2, which maps $\mathscr{W}_0$ onto a closed subset of the space of symmetric graphons. ∎

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

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

$$\tilde{\varphi} : \mathcal{C} \to \mathrm{Set}, \qquad \tilde{\varphi}(F) = \varphi(F) = \{a \in F^n : F \models \varphi(a)\},$$

mapping a structure F in $\mathcal{C}$ to the set of realisations of the formula φ in F.

If φ is a sentence in the language $\mathcal{L}$, then

$$\tilde{\varphi} : \mathcal{C} \to \mathrm{Set}$$

assigns to each $F \in \mathcal{C}$ the truth value of φ in F.

For $n > 0$, the 'affine space' $\mathbb{A}^n$ is the assignment

$$F \mapsto F^n,$$

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

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

$$F \mapsto \{\bot, \top\}.$$

For each formula φ, the restriction of $\tilde{\varphi}$ to $\mathcal{C}^\prec$ is a functor

$$\tilde{\varphi} : \mathcal{C}^\prec \to \mathrm{Set}.$$

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

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

DEFINITION 3.1. A subassignment $\mathbf{S}$ 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 $\mathbf{X}$ and $\mathbf{Y}$ are definable sets, their argumentwise cartesian product, denoted

$$\mathbf{X} \times \mathbf{Y},$$

is clearly definable.

If $\mathbf{X}, \mathbf{Y} \subseteq \mathbb{A}^n$ are definable, so are the sets

$$\mathbf{X} \cap \mathbf{Y}, \quad \mathbf{X} \cup \mathbf{Y}, \quad \mathbf{X} \setminus \mathbf{Y}.$$

DEFINITION 3.3. A *definable function*

$$\boldsymbol{\pi} : \mathbf{X} \to \mathbf{Y}$$

between definable sets $\mathbf{X}$ and $\mathbf{Y}$ is given through a definable subset $\boldsymbol{\Gamma} \subseteq \mathbf{X} \times \mathbf{Y}$ such that, for every $F \in \mathcal{C}$, $\boldsymbol{\Gamma}(F) \subseteq \mathbf{X}(F) \times \mathbf{Y}(F)$ defines a function

$$\boldsymbol{\pi}_F : \mathbf{X}(F) \to \mathbf{Y}(F).$$

REMARK 3.4. If $\boldsymbol{\pi} : \mathbf{X} \to \mathbf{Y}$ is a definable function, then the *image* $\boldsymbol{\pi}(\mathbf{X})$ defined by

$$\boldsymbol{\pi}(\mathbf{X})(F) = \boldsymbol{\pi}_F(\mathbf{X}(F))$$

is a definable subset of $\mathbf{Y}$.

DEFINITION 3.5. A *definable function*

$$\mathbf{f} : \mathbf{X} \to E$$

from a definable set $\mathbf{X}$ 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 $\mathbf{X}_i \subseteq \mathbf{X}$, $i = 1, \ldots, n$,

such that, for every $F \in \mathcal{C}$,

$$\mathbf{X}(F) = \coprod_i \mathbf{X}_i(F)$$

and $\mathbf{f}_F : \mathbf{X}(F) \to E$ is defined by

$$\mathbf{f}_F \restriction_{\mathbf{X}_i(F)} = e_i.$$

We stipulate that, for $\mathbf{X} = \mathbb{A}^0$, the $\mathbf{X}_i$ are associated with sentences φ_i such that, for every $F \in \mathcal{C}$,

$$\mathbf{f}_F = e_i \quad \text{if and only if} \quad F \models \varphi_i.$$

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

$$\tilde{\varphi}_c : \mathcal{C}_F \to \mathrm{Set},$$

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

DEFINITION 3.7. Let $\mathbf{S}$ be a definable set. A *definable set* $\mathbf{X}$ *over* $\mathbf{S}$ (or, *with parameters in* $\mathbf{S}$) is a definable function

$$\boldsymbol{\pi} : \mathbf{X} \to \mathbf{S}$$

in the sense of Definition 3.3. A *definable map of definable sets over* $\mathbf{S}$ is a commutative diagram

$$\begin{array}{ccc} \mathbf{X} & \longrightarrow & \mathbf{Y} \\ & \searrow \quad \swarrow & \\ & \mathbf{S} & \end{array}$$

DEFINITION 3.8. Given definable sets $\boldsymbol{\pi} : \mathbf{X} \to \mathbf{S}$ and $\boldsymbol{\pi}' : \mathbf{X}' \to \mathbf{S}$ over $\mathbf{S}$, their *fibre product* $\mathbf{X} \times_{\mathbf{S}} \mathbf{X}'$ is

$$\begin{aligned} \mathbf{X} \times_{\mathbf{S}} \mathbf{X}'(F) &= \mathbf{X}(F) \times_{\mathbf{S}(F)} \mathbf{X}'(F) \\ &= \{(x, x') \in \mathbf{X}(F) \times \mathbf{X}'(S) : \boldsymbol{\pi}(x) = \boldsymbol{\pi}'(x')\}, \end{aligned}$$

and it is again a definable set over $\mathbf{S}$.

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

$$\mathbf{X}_s = \mathbf{X} \times_{\mathbf{S}} \{s\} : \mathcal{C}_F \to \mathrm{Set},$$

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

$$\mathbf{X}_s(F') = \boldsymbol{\pi}_{F'}^{-1}(s)$$

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

REMARK 3.9. By the above, a definable set $\mathbf{X}$ over $\mathbf{S}$ gives rise to a family of definable sets $\mathbf{X}_s$ parametrised by parameters s from $\mathbf{S}$.

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 $\mathbf{X}$ is a definable set, it can naturally be considered as a definable set over $\mathbb{A}^0$. Indeed, we consider the definable map which takes $\mathbf{X}(F)$ to $\top$ if and only if $\mathbf{X}(F) \neq \emptyset$.

DEFINITION 3.10. Let $\mathbf{X}$ be a definable set over $\mathbf{S}$ and let E be a set. Then a *definable function*

$$\mathbf{f} : \mathbf{X} \to E$$

over $\mathbf{S}$ on a class of structures $\mathcal{C}$ is determined by a choice of

(1) finitely many definable functions $\mathbf{e}_1, \ldots, \mathbf{e}_n : \mathbf{S} \to E$, and
(2) finitely many definable sets $\mathbf{X}_1, \ldots, \mathbf{X}_n$ over $\mathbf{S}$

such that, for every $F \in \mathcal{C}$ and every $s \in \mathbf{S}(F)$, we have

$$\mathbf{X}_{1,s} \amalg \cdots \amalg \mathbf{X}_{n,s} = \mathbf{X}_s,$$

and

$$\mathbf{f}_s : \mathbf{X}_s \to E$$

is given on $\mathcal{C}_F$ by

$$\mathbf{f}_{s,F'} \restriction_{\mathbf{X}_{i,s}(F')} = \mathbf{e}_i(s).$$

3.2. Counting and asymptotic classes

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

(1) a definable function $\boldsymbol{\mu}_{\mathbf{X}} : \mathbf{S} \to \mathbb{Q}$,
(2) a definable function $\mathbf{d}_{\mathbf{X}} : \mathbf{S} \to \mathbb{N}$,
(3) a constant $C_{\mathbf{X}} > 0$

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

$$\big||\mathbf{X}_s(F)| - \boldsymbol{\mu}_{\mathbf{X}}(s)|F|^{\mathbf{d}_{\mathbf{X}}(s)}\big| \leq C_{\mathbf{X}}|F|^{\mathbf{d}_{\mathbf{X}}(s)-1/2}.$$

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

(1) a definable function $\boldsymbol{\mu}_{\mathbf{X}} : \mathbf{S} \to \mathbb{Q}$,
(2) a definable function $\mathbf{d}_{\mathbf{X}} : \mathbf{S} \to \mathbb{N}$

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

$$\big||\mathbf{X}_s(F)| - \boldsymbol{\mu}_{\mathbf{X}}(s)|F|^{\mathbf{d}_{\mathbf{X}}(s)}\big| \leq \epsilon|F|^{\mathbf{d}_{\mathbf{X}}(s)}.$$

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* χ if, for every definable set $\mathbf{X}$ over $\mathbf{S}$, there exist

(1) a definable function $\boldsymbol{\mu}_{\mathbf{X}} : \mathbf{S} \to \mathbb{Q} \cup \{\infty\}$,
(2) a definable function $\mathbf{d}_{\mathbf{X}} : \mathbf{S} \to \mathbb{N} \cup \{\infty\}$,
(3) a constant $C_{\mathbf{X}} > 0$

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

$$\left| |\mathbf{X}_s(F)| - \boldsymbol{\mu}_{\mathbf{X}}(s)\chi(F)^{\mathbf{d}_{\mathbf{X}}(s)} \right| \leq C_{\mathbf{X}}\chi(F)^{\mathbf{d}_{\mathbf{X}}(s)-1/2},$$

where we stipulate that $\boldsymbol{\mu}_{\mathbf{X}}(s) < \infty$ if and only if $\mathbf{d}_{\mathbf{X}}(s) < \infty$.

We define an *asymptotic class relative to* χ analogously, to reflect the error term from Definition 3.12.

Terminology 3.14. With notation of Definition 3.13, we may informally refer to the number $\boldsymbol{\mu}_{\mathbf{X}}(s)$ as a *measure*, and to the number $\mathbf{d}_{\mathbf{X}}(s)$ as a *dimension*. We say that $\mathbf{X} \to \mathbf{S}$ is of *finite relative dimension* provided $\mathbf{d}_{\mathbf{X}}$ maps into $\mathbb{N}$.

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

$$\boldsymbol{\mu}_{\mathbf{Y}/\mathbf{X}} : \mathbf{S} \to \mathbb{Q}, \qquad \boldsymbol{\mu}_{\mathbf{Y}/\mathbf{X}}(s) = \begin{cases} \frac{\boldsymbol{\mu}_{\mathbf{Y}}(s)}{\boldsymbol{\mu}_{\mathbf{X}}(s)} & \text{if } \mathbf{d}_{\mathbf{Y}}(s) = \mathbf{d}_{\mathbf{X}}(s), \\ 0 & \text{otherwise.} \end{cases}$$

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

$$\frac{|\mathbf{Y}_s(F)|}{|\mathbf{X}_s(F)|} - \boldsymbol{\mu}_{\mathbf{Y}/\mathbf{X}}(s)$$

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 χ, then the above expressions are $O(\chi(F)^{-1/2})$.

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

Remark 3.16. A (relative) CDM-class is clearly a (relative) asymptotic 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 [5] by Chatzidakis–van den Dries–Macintyre.

In Section 7, following a discussion of difference fields, we will explain in Subsection 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. This 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 $W^6 = W \circ W^* \circ W \circ W^* \circ W \circ W^*$, we have*

$$\|W^6 - W'\|_\infty \leq 2\epsilon^2.$$

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

$$(Tf)(v) = \int_{[0,1]} W(u,v) f(u)\, du$$

associated with the graphon W. Its adjoint T^* is given by

$$(T^*g)(u) = \int_{[0,1]} W(u,v) g(v)\, dv.$$

The Cauchy–Schwarz inequality implies that for all $f \in L^2([0,1])$,

$$\|Tf\|_2 \leq \|Tf\|_\infty \leq \|f\|_2, \tag{1}$$

and similarly, for $g \in L^2[0,1]$,

$$\|T^*g\|_2 \leq \|T^*g\|_\infty \leq \|g\|_2. \tag{2}$$

We can apply the singular value decomposition to the Hilbert–Schmidt operator T, which gives

$$Tf = \sum_i \sigma_i \langle f, u_i\rangle_2 y_i \quad \text{and} \quad T^*g = \sum_i \sigma_i \langle g, y_i\rangle_2 u_i$$

for some sequence σ_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

$$TT^*g = \sum_i \sigma_i^2 \langle g, y_i\rangle_2 y_i,$$

whence

$$\operatorname{tr}(TT^*) = \sum_i \sigma_i^2.$$

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

$$\sum_i \sigma_i^2 = \operatorname{tr}(TT^*) = \int |W(x,y)|^2 \, dx \, dy \le 1. \tag{3}$$

Using

$$y_i = \frac{1}{\sigma_i} T u_i, \qquad u_i = \frac{1}{\sigma_i} T^* y_i,$$

as well as (1), (2), we obtain

$$\|y_i\|_\infty \le \frac{1}{\sigma_i} \quad \text{and} \quad \|u_i\|_\infty \le \frac{1}{\sigma_i}. \tag{4}$$

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

$$TT^*TT^*TT^* g = \sum_i \sigma_i^6 \langle g, y_i \rangle_2 y_i. \tag{5}$$

Given an $\epsilon > 0$, we split

$$TT^*TT^*TT^* = A_\epsilon + B_\epsilon,$$

where $A = A_\epsilon$ is a low rank operator

$$Ag = \sum_{i : \sigma_i \ge \epsilon} \sigma_i^6 \langle g, y_i \rangle_2 y_i,$$

and $B = B_\epsilon$ is the error term

$$Bg = \sum_{i : \sigma_i < \epsilon} \sigma_i^6 \langle g, y_i \rangle_2 y_i.$$

Using the triangle inequality, Hölder's inequality (Fact A.2) and (4), (3), for any $g \in L^1[0,1]$ we have

$$\begin{aligned} \|Bg\|_\infty &\le \sum_{i : \sigma_i < \epsilon} \sigma_i^6 |\langle g, y_i \rangle_2| \frac{1}{\sigma_i} \le \sum_{i : \sigma_i < \epsilon} \sigma_i^6 \|g\|_1 \|y_i\|_\infty \frac{1}{\sigma_i} \\ &\le \sum_{i : \sigma_i < \epsilon} \sigma_i^6 \|g\|_1 \frac{1}{\sigma_i^2} \le \epsilon^2 \sum_i \sigma_i^2 \|g\|_1 \le \epsilon^2 \|g\|_1. \end{aligned}$$

Let

$$\delta = \epsilon^2 / 3.$$

Using (4), we discretise

$$y_i = y'_{i,\epsilon} + e_{i,\epsilon},$$

where $y'_i = y'_{i,\epsilon}$ takes at most $1/(\sigma_i\delta)$ values, $\|y'_i\|_\infty \leq \|y_i\|_\infty$, and e_i is bounded in magnitude by δ. We split $A_\epsilon = A'_\epsilon + E_\epsilon$, where

$$A'g = \sum_{i:\sigma_i\geq\epsilon} \sigma_i^6 \langle g, y'_i\rangle_2 y'_i, \qquad Eg = \sum_{i:\sigma_i\geq\epsilon} \sigma_i^6 (\langle g, y'_i\rangle_2 e_i + \langle g, e_i\rangle_2 y'_i + \langle g, e_i\rangle_2 e_i).$$

By the choice of δ, and arguments analogous to the above, we get

$$\begin{aligned}\|Eg\|_\infty &\leq \sum_{i:\,\sigma_i\geq\epsilon} \sigma_i^6 (\|g\|_1\|y'_i\|_\infty\|e_i\|_\infty + \|g\|_1\|e_i\|_\infty\|y'_i\|_\infty + \|g\|_1\|e_i\|_\infty^2)\\ &\leq \|g\|_1 \sum_{i:\,\sigma_i\geq\epsilon} \sigma_i^6 \left(2\frac{\delta}{\sigma_i} + \delta^2\right) \leq \|g\|_1 \sum_{i:\,\sigma_i\geq\epsilon} 2\sigma_i^5\delta + \sigma_i^6\delta^2 \leq \epsilon^2\|g\|_1.\end{aligned}$$

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 (3), we see that the number of summands in the definition of A is at most $1/\epsilon^2$. We partition

$$[0,1] = V_1 \cup \cdots \cup V_n,$$

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

$$n = N(\epsilon) \leq (1/(\epsilon\delta))^{1/\epsilon^2} = (3/\epsilon^3)^{1/\epsilon^2}.$$

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

$$\boldsymbol{\Gamma}_s = \boldsymbol{\Gamma}_s(F) = (\mathbf{U}_s(F), \mathbf{V}_s(F), \mathbf{E}_s(F)).$$

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

Definition 5.2. Let $\mathbf{U}$, $\mathbf{V}$ be definable sets over $\mathbf{S}$. A *definable step-function* on $\mathbf{U} \times \mathbf{V}$ over $\mathbf{S}$ is a definable function $\mathbf{W} : \mathbf{U} \times \mathbf{V} \to \mathbb{R}$ over $\mathbf{S}$

such that there exist definable sets $\mathbf{U}_1, \ldots, \mathbf{U}_m$ and $\mathbf{V}_1, \ldots, \mathbf{V}_n$ over $\mathbf{S}$ such that, for each $F \in \mathcal{C}$ and each $s \in \mathbf{S}(F)$,

$$\mathbf{U}_s = \coprod_{i=1}^{n} \mathbf{U}_{i,s} \quad \text{and} \quad \mathbf{V}_s = \coprod_{j=1}^{m} \mathbf{V}_{j,s},$$

and $\mathbf{W}_s$ is constant on each $\mathbf{U}_{i,s} \times \mathbf{V}_{j,s}$.

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

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

$$\{\mathbf{W}_s(F) : F \in \mathcal{C},\ s \in \mathbf{S}(F)\}$$

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

Proof. By definition, there are definable sets $\mathbf{U}_1, \ldots, \mathbf{U}_m$ and $\mathbf{V}_1, \ldots, \mathbf{V}_n$ over $\mathbf{S}$ partitioning $\mathbf{U}$ and $\mathbf{V}$ and definable functions $\mathbf{e}_{ij} : \mathbf{S} \to [0,1]$ with

$$\mathbf{W}_s \restriction_{\mathbf{U}_{i,s} \times \mathbf{V}_{j,s}} = \mathbf{e}_{ij}(s).$$

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

$$\boldsymbol{\mu}_{\mathbf{U}_i/\mathbf{U}} : \mathbf{S} \to \mathbb{Q} \quad \text{and} \quad \boldsymbol{\mu}_{\mathbf{V}_j/\mathbf{V}} : \mathbf{S} \to \mathbb{Q}$$

such that

$$\frac{|\mathbf{U}_{i,s}(F)|}{|\mathbf{U}_s(F)|} - \mu_{\mathbf{U}_i/\mathbf{U}}(s) = o(1),$$

and analogously for $\mathbf{V}_j$.

We can partition

$$\mathbf{S} = \mathbf{S}_1 \amalg \cdots \amalg \mathbf{S}_r$$

so that, for all $s \in \mathbf{S}_l(F)$,

$$\mathbf{e}_{ij}(s) = e_{ijl} \in [0,1], \qquad \boldsymbol{\mu}_{\mathbf{U}_i/\mathbf{U}}(s) = \mu_{il} \in \mathbb{Q}, \qquad \boldsymbol{\mu}_{\mathbf{V}_j/\mathbf{V}}(s) = \nu_{jl} \in \mathbb{Q}.$$

Thus, again by Remark 3.15, up to isomorphism, the stepfunction associated to $\mathbf{W}_s(F)$ through Definition 2.6 is within $o(1)$ from a stepfunction associated to a weighted bipartite graph with vertex weights $\{\mu_{il} : i\}$ and $\{\nu_{jl} : j\}$ 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 $\{\mathbf{W}_s(F) : s \in \mathbf{S}_l(F)\}$ with respect to the preorder induced by $\chi(F)$ has a stepfunction as a limit in the $\|\cdot\|_1$-norm. ∎

THEOREM 5.4 (Tao's algebraic regularity lemma for asymptotic classes). *Let $\boldsymbol{\Gamma} = (\mathbf{U}, \mathbf{V}, \mathbf{E})$ be a definable bipartite graph of finite relative dimension over a definable set $\mathbf{S}$ on an asymptotic class $\mathcal{C}$ relative to χ. Then there exists a definable set $\tilde{\mathbf{S}}$ over $\mathbf{S}$ and a definable stepfunction $\mathbf{W}$ over $\tilde{\mathbf{S}}$ such that for every $\epsilon > 0$, there exists an $M > 0$ such that for every $F \in \mathcal{C}$ with $\chi(F) \geq M$ and every $\tilde{s} \in \tilde{\mathbf{S}}(F)$ mapping onto $s \in \mathbf{S}(F)$,*

$$d_{\square}(\boldsymbol{\Gamma}_s(F), \mathbf{W}_{\tilde{s}}(F)) \leq \epsilon.$$

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

$$d_{\square}(\boldsymbol{\Gamma}_s(F), \mathbf{W}_{\tilde{s}}(F)) \leq M\chi(F)^{-1/12}.$$

We 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 of [22, proof of 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 works for a relative CDM-class by replacing every instance of $|F|$ by $\chi(F)$, for $F \in \mathcal{C}$.

The weak regularity result established in Lemma 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^6 up to $2\epsilon^2$ in the $\|\cdot\|_\infty$-norm.

The idea is to improve this result by using the definability of $\boldsymbol{\Gamma}$ and the constraints on the growth rates of the sets of realisations of definable sets over CDM-classes in the spirit of Remark 3.15.

Let us name the key objects. For $F \in \mathcal{C}$ and an $s \in \mathbf{S}(F)$, 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 $W_s = W(\boldsymbol{\Gamma}_s(F))$.

The proof consists of the following conceptual steps:

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

(2) We verify that, for any large enough $F \in \mathcal{C}$ and any $s \in \mathbf{S}(F)$, if we write $W_s = W(\boldsymbol{\Gamma}_s(F))$, then any (not necessarily definable) stepfunction with at most $N(\epsilon)$ steps that regularises W_s^6 and $(W_s^*)^6$ in view of Lemma 4.1 turns out to be close to the steps of a realisation of $\mathbf{W}$ constructed in (1) in the supremum norm.
(3) From (1) and (2), it follows that W_s is close to a realisation of $\mathbf{W}$ in the cut metric, which shows the required regularity.

Step (1). We will use the definability of $\boldsymbol{\Gamma}$ and decide on an appropriate choice of ϵ. The integral kernel $K_3(v,v') = K_{3,s}(v,v')$ of $TT^*TT^*TT^*$ for $T = T_s$ is explicitly given as

$$K_3(v,v')(F) = \frac{|\mathbf{G}_{v,v',s}(F)|}{|\mathbf{U}_s(F)|^3 |\mathbf{V}_s(F)|^2},$$

where $\mathbf{G}_{v,v',s} \subseteq \mathbf{U}_s \times \mathbf{V}_s \times \mathbf{U}_s \times \mathbf{V}_s \times \mathbf{U}_s$ is the definable set

$$\mathbf{G}_{v,v',s} = \{(u_1,v_2,u_2,v_3,u_3) \in \mathbf{U}_s \times \mathbf{V}_s \times \mathbf{U}_s \times \mathbf{V}_s \times \mathbf{U}_s : \\ (u_1,v),(u_1,v_2),(u_2,v_2),(u_2,v_3),(u_3,v_3),(u_3,v') \in \mathbf{E}_s\}.$$

By Definition 3.11, there exists a definable function $\mathbf{c} = \mathbf{c}(v,v',s)$ symmetric in v, v' such that, for all $F \in \mathcal{C}$, $s \in \mathbf{S}(F)$ and $v, v' \in \mathbf{V}_s(F)$,

$$K_{3,s}(v,v') = \mathbf{c}(v,v',s) + O(|F|^{-1/2}). \tag{6}$$

Replacing the role of T_s by T_s^*, we consider the integral kernel $K_{3,s}^*(u,u')$ of $T^*TT^*TT^*T$, and we similarly find a definable function $\mathbf{c}^* = \mathbf{c}^*(u,u',s)$ symmetric in u, u' such that for all $F \in \mathcal{C}$, $s \in \mathbf{S}(F)$ and $u, u' \in \mathbf{U}_s(F)$,

$$K_{3,s}^*(u,u') = \mathbf{c}^*(u,u',s) + O(|F|^{-1/2}). \tag{7}$$

The set of values of $\mathbf{c}$ and $\mathbf{c}^*$ 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 $\bar{n} \le n \le N(\epsilon) = (3/\epsilon^3)^{1/\epsilon^2}$, writing

$$\mathbf{V}_{i,(v_1,\dots,v_n,s)} = \Big\{ v \in \mathbf{V}_s : \bigwedge_j \mathbf{c}(v,v_j,s) = \mathbf{c}(v_i,v_j,s) \Big\},$$

we see that the formula

$$\begin{aligned} \psi_{\bar{n},n}(v_1,\dots,v_n,s) \equiv {} & \Big(\bigwedge_i v_i \in \mathbf{V}_s \Big) \wedge \Big(\mathbf{V}_s = \coprod_{i \le n} \mathbf{V}_{i,(v_1,\dots,v_n,s)} \Big) \\ & \wedge \bigwedge_{i \le \bar{n}} \dim(\mathbf{V}_{i,(v_1,\dots,v_n,s)}) = \dim(\mathbf{V}_s) \\ & \wedge \dim\Big(\mathbf{V}_s \setminus \coprod_{i \le \bar{n}} \mathbf{V}_{i,(v_1,\dots,v_n,s)} \Big) < \dim(\mathbf{V}_s) \\ & \wedge \bigwedge_i \bigwedge_j \forall v \in \mathbf{V}_{i,\gamma}\, \forall v' \in \mathbf{V}_{j,\gamma}\ \mathbf{c}(v,v',s) = \mathbf{c}(v_i,v_j,s) \end{aligned}$$

expresses that $s' = (v_1, \dots, v_n, s)$ can serve as parameters for a definable stepfunction on $\mathbf{V}_s$ with steps $\mathbf{V}_{1,s'}, \dots, \mathbf{V}_{n,s'}$ and values $\mathbf{c}(v_i, v_j, s)$ on $\mathbf{V}_{i,s'} \times \mathbf{V}_{j,s'}$, where $\mathbf{V}_{1,s'}, \dots, \mathbf{V}_{\bar{n},s'}$ are the top-dimensional steps, which we will call *large* in what follows.

Let $\mathbf{S}_{\bar{n},n} \to \mathbf{S}$ be the definable set associated with the formula $\psi_{\bar{n},n}$; let

$$\mathbf{S}_V \to \mathbf{S}$$

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

$$\mathbf{V}_{i,\bar{n},n} \to \mathbf{S}_{\bar{n},n} \to \mathbf{S}_V.$$

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

$$\mathbf{U}_{j,\bar{m},m} \to \mathbf{S}^*_{\bar{m},m} \to \mathbf{S}_U$$

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

Using Definition 3.11 again, for all $i, \bar{n}, n$ and $j, \bar{m}, m$ as above, we find a definable function $\mathbf{f}_{j,\bar{m},m,i,\bar{n},n} : \mathbf{S}^*_{\bar{m},m} \times_{\mathbf{S}} \mathbf{S}_{\bar{n},n} \to \mathbb{Q}$ such that for all $F \in \mathcal{C}$, and $s' \in \mathbf{S}^*_{\bar{m},m}(F)$, $s'' \in \mathbf{S}_{\bar{n},n}(F)$ mapping onto $s \in \mathbf{S}(F)$,

(8)
$$\frac{\left|\mathbf{E}_s(F) \cap (\mathbf{U}_{j,\bar{m},m,s'}(F) \times \mathbf{V}_{i,\bar{n},n,s''}(F))\right|}{\left|\mathbf{U}_{j,\bar{m},m,s'}(F) \times \mathbf{V}_{i,\bar{n},n,s''}(F)\right|} = \mathbf{f}_{j,\bar{m},m,i,\bar{n},n}(s', s'') + O(|F|^{-1/2}).$$

Let

$$\tilde{\mathbf{S}} = \mathbf{S}_U \times_{\mathbf{S}} \mathbf{S}_V.$$

We consider the definable stepfunction $\mathbf{W}$ over $\tilde{\mathbf{S}}$ which is given by

$$\mathbf{W} \restriction_{\mathbf{U}_{j,\bar{m},m} \times \mathbf{V}_{i,\bar{n},n}} = \mathbf{f}_{j,\bar{m},m,i,\bar{n},n}$$

over the component $\mathbf{S}^*_{\bar{m},m} \times_{\mathbf{S}} \mathbf{S}_{\bar{n},n}$ of $\tilde{\mathbf{S}}$.

Step (2). Let us choose M such that for $|F| \geq M$, the error terms in (6), (7) and (8) are less than $\delta/2$.

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

We apply Lemma 4.1 to W_s for the ϵ chosen in Step (1), and this yields a partition

$$\mathbf{V}_s(F) = V_1 \amalg \cdots \amalg V_n$$

into $n \leq N(\epsilon)$ sets V_i such 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 (6), and the choice of F, each value of $K_3(v, v')$ is within $\delta/2$ from a value of the definable function $\mathbf{c}$. Hence we can improve the near-constancy to the property

$$K_{3,s}(v, v') = \gamma_{ij} + O(|F|^{-1/2}), \tag{9}$$

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

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

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

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

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

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

$$V_i = \Big\{ v \in \mathbf{V}_s(F) : \bigwedge_j \mathbf{c}(v, v_j) = \gamma_{ij} \Big\}.$$

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

$$V_i = \mathbf{V}_{i,\bar{n},n,s'}(F).$$

Note that $\mathbf{V}_{i,\bar{n},n}$ for $i \le \bar{n}$ are large, so Definition 3.11 implies that for some positive $a_i \in \mathbb{Q}$,

$$\frac{|V_i|}{|\mathbf{V}_s(F)|} = a_i + O(|F|^{-1/2}), \tag{10}$$

while for $i > \bar{n}$ we have $\dim(\mathbf{V}_{i,\bar{n},n,s'}) < \dim(V)$ and

$$|V_i| \ll |F|^{-1} |\mathbf{V}_s(F)|. \tag{11}$$

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 $\boldsymbol{\Gamma}_s(F)$ via Definition 2.6. Suppose $g \in L^2[0,1]$ is supported on $\bar{V}_j$ with mean zero. Combining (5) and (9), we obtain

$$\begin{aligned}\sum_i \sigma_i^6 |\langle g, y_i \rangle|^2 = |\langle TT^*TT^*TT^* g, g \rangle| &= \Big| \int_0^1 \int_0^1 K_3(v, v') g(v) g(v') \, dv \, dv' \Big| \\ &\ll |F|^{-1/2} \|g\|_2^2.\end{aligned}$$

On the other hand, Bessel's inequality gives

$$\sum_i |\langle g, y_i \rangle|^2 \le \|g\|_2^2,$$

and so Hölder's inequality (combining these two inequalities via Lemma 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

$$\mathbf{U}_s(F) = U_1 \amalg \cdots \amalg U_m$$

with $m \leq N(\epsilon)$ such that, upon fixing a parameter $s'' = (u_1,\dots,u_m,s) \in \mathbf{S}^*_{\bar{m},m,s}(F)$, by relabelling U_j, we have

$$U_j = \mathbf{U}_{j,\bar{m},m,s''}(F).$$

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 $\mathbf{W}$ 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),

$$d_\square(W_s, \mathbf{W}_{\tilde{s}}(F)) \ll |F|^{-1/12}.$$

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

$$1_A = (1_A - \mu(A)/\mu(\bar{U}_i)1_{\bar{U}_i}) + \mu(A)/\mu(\bar{U}_i)1_{\bar{U}_i},$$
$$1_B = (1_B - \mu(B)/\mu(\bar{V}_j)1_{\bar{V}_j}) + \mu(B)/\mu(\bar{V}_j)1_{\bar{V}_j}$$

into sums of mean-zero and constant functions. Note that the mean-zero parts are bounded by 1. We obtain

$$\langle T1_A, 1_B\rangle_2 = \frac{\mu(A)}{\mu(\bar{U}_i)}\frac{\mu(B)}{\mu(\bar{V}_i)}\langle T1_{\bar{U}_i}, 1_{\bar{V}_j}\rangle_2 + \sqrt{\mu(\bar{U}_i)\mu(\bar{V}_j)}\, O(|F|^{-1/12}).$$

On the other hand, by equation (8),

$$\begin{aligned}\langle T1_{\bar{U}_i}, 1_{\bar{V}_j}\rangle_2 &= \int\limits_{[0,1]^2} W(u,v)1_{\bar{U}_i}(u)1_{\bar{V}_j}(v)\,du\,dv \\ &= \frac{|\mathbf{E}_s \cap (U_i \times V_j)|}{|U|\,|V|} = (\mathbf{f}_{ij}(\tilde{s}) + O(|F|^{-1/2}))\mu(\bar{U}_i)\mu(\bar{V}_j),\end{aligned}$$

hence

$$\langle T1_A, 1_B\rangle_2 = \mu(A)\mu(B)\mathbf{f}_{ij}(\tilde{s}) + \sqrt{\mu(\bar{U}_i)\mu(\bar{V}_j)}\, O(|F|^{-1/12}).$$

Thus, for measurable $A \subseteq \bar{U}_i$ and $B \subseteq \bar{V}_j$, we have

$$\int_{[0,1]^2} (W_s(u,v) - \mathbf{f}_{ij}(\tilde{s}))1_A(u)1_B(v) \ll \sqrt{\mu(\bar{U}_i)\mu(\bar{V}_j)}\,|F|^{-1/12}.$$

For arbitrary measurable $A, B \subseteq [0,1]$, if we write $A_i = A \cap \bar{U}_i$ and $B_j = B \cap \bar{V}_j$,

$$\begin{aligned} d_\square(W_s, \mathbf{W}_{\tilde{s}}(F)) &= \sup_{A,B\subseteq[0,1]} \int_{[0,1]^2} (W_s - \mathbf{W}_{\tilde{s}}(F))1_A 1_B \\ &= \sup_{A,B\subseteq[0,1]} \sum_{i,j} \int_{[0,1]^2} (W_s - \mathbf{f}_{ij}(\tilde{s}))1_{A_i}1_{B_j} + O(|F|^{-1}) \\ &\ll \sum_{i,j} |F|^{-1/12}\sqrt{\mu(\bar{U}_i)\mu(\bar{V}_j)} + O(|F|^{-1}) \\ &\le \sqrt{\bar{m}\bar{n}}\,|F|^{-1/12} + O(|F|^{-1}), \end{aligned}$$

where the summation is over the large blocks and, by (11), 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 is a consequence of the fact that $\bar{m}$ and $\bar{n}$ are bounded independently of F. ∎

Combining Lemma 5.3 and Theorem 5.4, we obtain the following.

COROLLARY 5.6. *Let $\boldsymbol{\Gamma}$ be a definable graph of finite relative dimension over $\mathbf{S}$ on an asymptotic class $\mathcal{C}$ relative to χ. The set of accumulation points of the family of finite graphs*

$$\{\boldsymbol{\Gamma}_s(F) : F \in \mathcal{C}, s \in \mathbf{S}(F)\}$$

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

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

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

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

We are interested in accumulation points in the graphon space of the set of graphs $\boldsymbol{\Gamma}(\mathbb{F}_q)$ as $\mathbb{F}_q$ ranges over all finite fields (these are in fact symmetric variants of the well-known *Paley graphs*). When $\mathrm{char}(\mathbb{F}_q) \neq 2$, being a square is an event of 'CDM probability' (in the sense of Remark 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 $\mathrm{char}(\mathbb{F}_q) = 2$, everything is a square, so the graphs accumulate around the constant graphon $W(1)$.

The definable stepfunction $\mathbf{W} : \mathbf{U} \times \mathbf{V} \to \mathbb{Q}$ that regularises $\boldsymbol{\Gamma}$ is

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

Hence, the set of accumulation points is

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

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

$$\mathbf{E}(x, y) \equiv \exists z\ xy = z^2.$$

The definable sets

$$\mathbf{U}_0(x) \equiv \exists z\ x = z^2 \quad \text{and} \quad \mathbf{U}_1 = \mathbb{A}^1 \setminus \mathbf{U}_0$$

partition $\mathbb{A}^1$. Let $\mathbf{W}$ be the definable function which returns 1 on $\mathbf{U}_0 \times \mathbf{U}_0 \cup \mathbf{U}_1 \times \mathbf{U}_1$ and 0 elsewhere. The definable stepfunction $\mathbf{W}$ regularises $\boldsymbol{\Gamma}$, and the set of accumulation points of $\boldsymbol{\Gamma}(\mathbb{F}_q)$ in the space of graphons is

$$\left\{ W\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, W(1) \right\},$$

where the former denotes the stepfunction associated to the graph with edge weight matrix $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ as in Notation 2.7, which is the accumulation point of $\boldsymbol{\Gamma}(\mathbb{F}_q)$ for odd q, and the latter is the constant graphon 1, which is the accumulation point of $\boldsymbol{\Gamma}(\mathbb{F}_q)$ 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 $\boldsymbol{\Gamma}$ on $\mathbb{A}^1$ through the edge relation

$$\mathbf{E}(x, y) \equiv \exists z\ x + y = z^3.$$

In fact, the set of cubes in $\mathbb{F}_q$ is of size $(q-1)/3 + 1$ if $3 \mid q - 1$ 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

$$\exists t\ t^3 = 1 \wedge t \neq 1$$

so $\boldsymbol{\Gamma}$ is definably regularised. To conclude, the graphs $\boldsymbol{\Gamma}(\mathbb{F}_q)$ accumulate around $W(1/3)$ when $3 \mid q - 1$, and around $W(1)$ otherwise.

The following example illustrates the need to introduce parameters.

EXAMPLE 6.4. Consider the graph $\boldsymbol{\Gamma}$ on $\mathbb{A}^1$ with the edge relation

$$\mathbf{E}(x,y) \equiv \exists z\ xy = z^3.$$

The graph is trivially considered over the base $\mathbf{S} = \mathbb{A}^0$. Let

$$\tilde{S}(z) \equiv z^3 = 1 \wedge z \neq 1,$$

and consider the definable sets $\mathbf{U}_i$ over $\tilde{\mathbf{S}}$, $i = 0, 1, 2$:

$$\mathbf{U}_i(x;z) \equiv \exists u\ x = z^i u^3.$$

We see that $\boldsymbol{\Gamma}$ is regularised by the definable stepfunction $\mathbf{W}(x,y;z)$ over $\tilde{\mathbf{S}}$ taking value 1 on $\mathbf{U}_0 \times \mathbf{U}_0 \cup \mathbf{U}_1 \times \mathbf{U}_2 \cup \mathbf{U}_2 \times \mathbf{U}_1$ and 0 elsewhere, and the graphs $\boldsymbol{\Gamma}(\mathbb{F}_q)$ accumulate around the stepfunction

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

when $3 \mid q-1$, 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, σ) consisting of a commutative ring R with identity 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 σ-Rng. A difference ring (R, σ) is a *transformal domain* if R is a domain and σ is injective.

NOTATION 7.1. For a prime power q, we write

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

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 [20], we established the following difference analogue of [5].

THEOREM 7.2. *Let $\mathbf{X} \to \mathbf{S}$ be a definable map in the language of difference rings. Then there exist*

(1) *a definable function $\boldsymbol{\mu}_{\mathbf{X}} : \mathbf{S} \to \mathbb{Q} \cup \{\infty\}$,*

(2) *a definable function* $\mathbf{d_X} : \mathbf{S} \to \mathbb{N} \cup \{\infty\}$,
(3) *a constant* $C_{\mathbf{X}} > 0$

such that, for every K_q *with* $q > C_{\mathbf{X}}$ *and every* $s \in \mathbf{S}(K_q)$,

$$\big||\mathbf{X}_s(K_q)| - \boldsymbol{\mu}_{\mathbf{X}}(s) q^{\mathbf{d_X}(s)}\big| \leq C_{\mathbf{X}}\, q^{\mathbf{d_X}(s)-1/2}.$$

COROLLARY 7.3. *The class* $\mathcal{C}$ *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 Corollary 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 (Corollary 7.3) that fields with Frobenius constitute a relative CDM-class and the regularity lemma (Theorem 5.5), we obtain the following.

COROLLARY 7.5. *Let* $\boldsymbol{\Gamma} = (\mathbf{U}, \mathbf{V}, \mathbf{E})$ *be a definable bipartite graph of finite relative dimension over a definable set* $\mathbf{S}$ *in the language of difference rings. There exists a constant* $M = M(\boldsymbol{\Gamma}) > 0$*, a definable set* $\tilde{\mathbf{S}}$ *over* $\mathbf{S}$ *and a definable stepfunction* $\mathbf{W}$ *over* $\tilde{\mathbf{S}}$ *such that for every* K_q *with* $q > M$ *and every* $\tilde{s} \in \tilde{\mathbf{S}}(K_q)$ *mapping onto* $s \in \mathbf{S}(K_q)$*,*

$$d_{\square}(\boldsymbol{\Gamma}_s(K_q), \mathbf{W}_{\tilde{s}}(K_q)) \leq M q^{-1/12}.$$

EXAMPLE 7.6. Let $\mathbf{U}(x) \equiv x\sigma(x)^2 = 1$, and let $\boldsymbol{\Gamma}$ be a graph on $\mathbf{U} \times \mathbf{U}$ with the edge relation

$$\mathbf{E}(x, y) \equiv \exists z \in \mathbf{U}\ xy = z^3.$$

Then $\mathbf{U}(K_q)$ is the group μ_{2q+1} of $(2q+1)$th roots of unity in $\bar{\mathbb{F}}_q$, 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 $\boldsymbol{\Gamma}(K_q)$ 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 $\mu_{2q+1} = \mathbf{U}(K_q)$ 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 (Definition 3.11) proved for finite fields in [5].

EXAMPLE 7.7. Let $\boldsymbol{\Gamma}$ be defined on the same $\mathbf{U}$ as in Example 7.6, but with

$$\mathbf{E}(x, y) \equiv \exists z \in \mathbf{U}\ x\sigma(y) = z^3.$$

In order to simplify the notation, write $\Gamma = (U, U, E) = \boldsymbol{\Gamma}(K_q)$, so that $U = \mu_{2q+1} \subseteq \bar{\mathbb{F}}_q$. Let $\zeta \in \bar{\mathbb{F}}_q$ 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., $3 \mid 2q+1$ and $3 \mid q-1$;
(3) $\zeta \in U$ and $\zeta^q = \zeta^2$, i.e., $3 \mid 2q+1$ 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, σ) is a difference ring, a *difference* (k, σ)*-algebra* is a difference ring (A, σ) endowed with a difference ring homomorphism $(k, \sigma) \to (A, \sigma)$. A morphism between (k, σ)-algebras (A, σ) and (B, σ) 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, σ)-Alg for the category of (k, σ)-algebras.

DEFINITION 8.2. Let (k, σ) 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 σ 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, σ)-algebra.

DEFINITION 8.3. Let (k, σ) be a difference ring. We say that a (k, σ)-algebra A is

(1) of *finite* σ*-type* if it admits an epimorphism $k[x_1, \ldots, x_n]_\sigma \to A$ for some n;
(2) of *finite* σ*-presentation* if it is a quotient

$$A \simeq k[x_1, \ldots, x_n]_\sigma / \langle f_1, \ldots, f_m \rangle_\sigma,$$

where $\langle f_1, \ldots, f_m \rangle_\sigma$ is the *difference ideal* (ideal closed under σ) generated by $f_1, \ldots, f_n \in k[x_1, \ldots, x_n]_\sigma$ for some m, n.

DEFINITION 8.4. Let (A, σ) be a difference ring. The *affine difference scheme* associated with A is the functor

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

A *morphism* between affine difference schemes is a natural transformation

$$\varphi : \mathbf{X} \to \mathbf{Y}.$$

We will only ever discuss affine difference schemes, so we omit the word 'affine' from our descriptions. The category of difference schemes is denoted σ-Sch.

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

REMARK 8.6. By Yoneda's lemma, if $\mathbf{X}$ is associated with a difference ring (A, σ) and $\mathbf{Y}$ with (B, σ), a morphism $\mathbf{X} \to \mathbf{Y}$ 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,

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

Moreover, if $\mathbf{S}$ corresponds to a difference ring (k, σ), then

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

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

(1) $\mathbf{X}$ is *transformally integral* if (A, σ) is a transformal domain;
(2) f is *of finite σ-type* if (A, σ) is a (k, σ)-algebra of finite σ-type;
(3) f is *of finite σ-presentation* over S if (A, σ) is a (k, σ)-algebra of finite σ-presentation;
(4) f is a *σ-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 of the equation*

$$f(x) = 0$$

in a difference ring (R, φ) extending (k, σ) is defined as

$$\{a \in R^n : F(a, \varphi a, \ldots, \varphi^r a) = 0\}.$$

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 $\mathbf{S}$ be a difference scheme associated with a difference ring (k, σ), and consider a system of difference polynomial equations

$$\begin{aligned} f_1(x_1, \ldots, x_n) &= 0, \\ &\vdots \\ f_m(x_1, \ldots, x_n) &= 0, \end{aligned}$$

where $f_i \in k[x_1, \ldots, x_n]_\sigma$, and consider the difference scheme $\mathbf{X}$ over $\mathbf{S}$ associated to the (k, σ)-algebra of finite σ-presentation

$$A = k[x_1, \ldots, x_n]_\sigma / \langle f_1, \ldots, f_m \rangle_\sigma.$$

For any (k, σ)-algebra (R, σ), the set of (R, σ)-*rational points*

$$\mathbf{X}_{/\mathbf{S}}(R, \sigma) = \mathrm{Hom}_{(k,\sigma)\text{-Alg}}(A, R)$$

can be identified with the set of solutions of the above system in (R, σ), i.e., the intersection of the solution sets of all equations $f_i = 0$ in (R, σ).

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

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

(1) f is of *finite relative dimension* if it is such when considered as a definable map over $\mathcal{C}$ in the sense of Terminology 3.14;
(2) f is *dominant* if the dimension of its image as a definable set is of full dimension in $\mathbf{S}$;
(3) f is *finite* if the map $f(K_q)$ between realisation sets has finite fibres for all $K_q \in \mathcal{C}$;
(4) f is *generically finite* if it fails to have finite fibres only over a lower-dimensional subset of $\mathbf{S}$.

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 Remark 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* $X \rightsquigarrow Y$ 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 σ-presentation over $\mathbb{Z}$ and let $X_1, \dots, X_n, Y$ be difference schemes of finite σ-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 $K_q = (\bar{\mathbb{F}}_p, \varphi_q)$, 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

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

ASSUMPTION 8.16 (Expansion dichotomy). *Consider difference schemes X, Y, Z of finite σ-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 $X \times_S X \times_S Y \times_S Y \to Z \times_S Z \times_S Z \times_S Z$ given by*

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

is not dominant.

(2) (Moderate expansion property) *There exists a σ-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 [23, proof of Theorem 38] and consider that they can be followed formally to derive the above statement from Corollary 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 Assumption* 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* $(A, *)$ *over* S'*, which is either* $\mathbb{G}_{\mathrm{a}}$*,* $\mathbb{G}_{\mathrm{m}}$ *or of an elliptic curve over* S'*, and correspondences* $\chi : X \rightsquigarrow A$*,* $\upsilon : Y \rightsquigarrow A$ *and* $\zeta : Z \rightsquigarrow A$ *over* S' *such that* f *is in correspondence with the group law on* A*,*

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

The proof below 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 to [2] and [25].
- Hrushovski [11] shows 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 [6].
- For groups definable in algebraically closed fields, we shall use [3], and for groups definable in ACFA we refer to [13].

Martin Bays suggested that our proof can benefit from field-theoretic arguments specific to ACFA. We provide a summary of properties we use, which can all be found in [6].

FACT 8.18. *Let* (Ω, σ) *be a model of ACFA.*

(a) *While ACFA is a simple theory, the underlying field* Ω *is algebraically closed, so* stability theory *applies to it.*
(b) *The model-theoretic algebraic closure of* $A \subset \Omega$*, denoted* $\mathrm{acl}(A)$*, is the field-theoretic algebraic closure of the inversive difference subfield of* Ω *generated by* A*.*
(c) *The forking independence is witnessed by algebraic independence in the sense that, for* $A, B, C \subset \Omega$*,* $A \mathop{\smile\!\!\!\!\!\!|}_C B$ *if and only if the fields* $\mathrm{acl}(AC)$ *and* $\mathrm{acl}(BC)$ *are linearly disjoint over* $\mathrm{acl}(C)$*.*
(d) *Let* $a \in \Omega$*, let* k *be a difference subfield of* Ω*, and let* $k(a)_\sigma$ *be the difference field generated by* a *over* k*. If* $\mathrm{tr.deg}(k(a)_\sigma/k) = 1$*, then* $\sigma(a) \in \mathrm{acl}(k(a))$ *and thus* $a \not\mathop{\smile\!\!\!\!\!\!|}_k \sigma(a)$*. 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 correspondence with its algebraic locus in the sense of Definition* 8.14*.*

(f) *Given $a \in \Omega$ and a difference subfield k of Ω, the* canonical base $\mathrm{Cb}(a/k)$ *is the smallest perfect difference field over which the locus of a over k is defined. It is equal, in the pure field language, to the canonical base of the type of $(a, \sigma(a), \dots)$ over k. It has the property that $a \mathop{\smile\!\!\!\!\!\!|}_{\mathrm{Cb}(a/k)} k$. In arbitrary simple theories, canonical bases exist only as 'hyperimaginaries', and one is forced to work with the* bounded closure bdd *as opposed to the algebraic closure. In ACFA, canonical bases exist in the real word, so any occurrence of* bdd *below can be replaced by* acl.

Proof of Theorem 8.17. 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 Fact 8.18.

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

Let η be a scheme-theoretic generic point of S, and write X_η, Y_η and Z_η 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 Ω be a generic point over k on the graph of f_η, i.e., $x_0 \in X_\eta$, $y_0 \in Y_\eta$ are rank 1 points over k, with $x_0 \mathop{\smile\!\!\!\!\!\!|} y_0$ and $z_{00} = f_\eta(x_0, y_0) \in Z_\eta$.

We may assume that $z_{00} \mathop{\smile\!\!\!\!\!\!|} x_0$ and $z_{00} \mathop{\smile\!\!\!\!\!\!|} y_0$, which implies that $x_0 \in \mathrm{bdd}(y_0, z_{00})$ and $y_0 \in \mathrm{bdd}(x_0, z_{00})$. Otherwise, we would have $z_{00} \in \mathrm{bdd}(x_0)$ or $z_{00} \in \mathrm{bdd}(y_0)$ and we are in a degenerate case where the function f_η 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}) such 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 $z \in \mathrm{dcl}(xy)$, $y \in \mathrm{bdd}(xz)$ and $x \in \mathrm{bdd}(yz)$. Thus, π is a *generic invertible multiaction* as in [1, 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 $y_j z_{ij} \models x_i$), the set $\{z_{00}, z_{01}, z_{10}, z_{11}\}$ is dependent.

We claim that $\overline{\pi \circ \pi^{-1}}$ is a generic multiaction in the following sense (cf. [1, 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 $y_0 z_{00} \models x_0$, $y_0 z_{10} \models x_1$, so that $z_{00} z_{10} \models x_1 \circ x_0^{-1}$. Let

$$h = \mathrm{Cb}(z_{00} z_{10} / x_0 x_1) \in \overline{\pi \circ \pi^{-1}}.$$

We need to verify that

$$h \mathop{\smile\!\!\!\!\!\!|} x_0 \quad \text{and} \quad h \mathop{\smile\!\!\!\!\!\!|} x_1.$$

Let us choose $y_1 \equiv_{x_0 x_1} y_0$ such that $y_1 \mathop{\smile\!\!\!\!\!\!|}_{x_0 x_1} y_0$. 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 $y_1 z_{01} z_{11} \equiv_{x_0 x_1} y_0 z_{00} z_{10}$, which implies that $y_1 z_{01} \models x_0$ and $y_1 z_{11} \models x_1$.

Using $y_1 \mathop{\smile\!\!\!\!\!\!|}_{x_0 x_1} y_0$, we find that $z_{01} \mathop{\smile\!\!\!\!\!\!|}_{x_0 x_1} z_{00} z_{10}$, which together with $z_{01} \mathop{\smile\!\!\!\!\!\!|} x_0 x_1$ yields $z_{01} \mathop{\smile\!\!\!\!\!\!|} x_0 x_1 z_{00} z_{10}$ and $z_{01} \mathop{\smile\!\!\!\!\!\!|}_{z_{00} z_{10}} x_0 x_1$. Using the non-dominance assumption, we get $z_{11} \in \mathrm{bdd}(z_{01} z_{00} z_{10})$, so

$$z_{01} z_{11} \mathop{\smile\!\!\!\!\!\!|}_{z_{00} z_{10}} x_0 x_1.$$

We conclude that

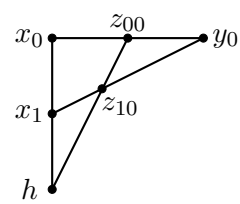
$$h = \mathrm{Cb}(z_{00} z_{10} / x_0 x_1) = \mathrm{Cb}(z_{01} z_{11} / x_0 x_1) = \mathrm{Cb}(z_{01} z_{11} / z_{00} z_{10}) \in \mathrm{bdd}(z_{00} z_{10}).$$

Thus, since $z_{00} z_{10} \mathop{\smile\!\!\!\!\!\!|} x_0$ and $z_{00} z_{10} \mathop{\smile\!\!\!\!\!\!|} x_1$, we confirm that $h \mathop{\smile\!\!\!\!\!\!|} x_0$ and $h \mathop{\smile\!\!\!\!\!\!|} x_1$.

In terminology of [1], [24], [2] and [25], it follows that the set of germs of the multiaction $\overline{\pi \circ \pi^{-1}}$ is a polygroup chunk, which entails that the 6-tuple

$$(x_0, x_1, h, y_0, z_{00}, z_{10})$$

forms a group configuration, conveniently depicted by the diagram



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 $\mathrm{Cb}(y_0 z_{00} / x_0)$, x_1 with $\mathrm{Cb}(y_0 z_{10} / x_1)$ and h with $\mathrm{Cb}(z_{00} z_{10} / h)$.

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.

By the fact that this is a rank 1 group configuration in (Ω, σ), the same tuple forms a group configuration over k in the underlying algebraically

closed field Ω. Indeed, the independence property (1) in the field reduct is clear, and the interalgebraicity property (2) follows from Fact 8.18(b, d). The property (3) in the reduct follows from the description 8.18(f) of canonical bases in the field language specialised to rank 1.

The (stable) group configuration for algebraically closed fields yields an ACF-definable group $(G, *)$ over the algebraic closure $\bar{k}$ such 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.

By Fact 8.18, the graph of f_η 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 [13, Lemma 3.3] on the 6-tuple above, that the graph of f_η is in correspondence with a difference definable group H over $\bar{k}$, and, by [13, 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, σ) be the difference extension field of (k, ς) generated by k'. Through standard constructibility arguments of σ-localising S to make it normal, and considering the normalisation of S in L, we obtain a generically σ-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 on A. ∎

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 Theorem 8.17.
(2) The morphism f corresponds to the addition law on an elliptic curve in a way described in Theorem 8.17.
(3) The morphism f is a moderate asymmetric expander over a σ-localisation of S.

A. Inequalities

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

(1) (Cauchy–Schwarz inequality) *For $x, y \in H$, we have*

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

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

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

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

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

Note that, for $p = q = 2$, the above yields the Cauchy–Schwarz inequality for $L^2(X, \mu)$.

LEMMA A.3. *If a_i, b_i, u, v are positive real numbers satisfying*

$$\sum_i a_i^3 b_i \le uv \quad \text{and} \quad \sum_i b_i \le v,$$

then

$$\sum_i a_i b_i \le u^{1/3} v.$$

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 $1/p + 1/q = 1$,

$$\sum_i |x_i y_i| \le \Big(\sum_i |x_i|^p\Big)^{1/p} \cdot \Big(\sum_i |y_i|^q\Big)^{1/q}.$$

We can now write

$$\begin{aligned}\sum_i a_i b_i = \sum_i a_i b_i^{1/3} \cdot b^{2/3} &\le \Big(\sum_i a_i^3 b_i\Big)^{1/3} \cdot \Big(\sum_i (b^{2/3})^{3/2}\Big)^{2/3} \\ &\le u^{1/3} v^{1/3} \cdot v^{2/3} = u^{1/3} v. \ \blacksquare\end{aligned}$$

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

Mirna Džamonja received funding from the European Union's Horizon 2020 research and innovation programme under the Maria Skłodowska-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, where she is a Visiting Professor.

REFERENCES

[1] I. Ben-Yaacov, *Group configurations and germs in simple theories*, J. Symbolic Logic 67 (2002), 1581–1600.
[2] I. Ben-Yaacov, I. Tomašić and F. O. Wagner, *Constructing an almost hyperdefinable group*, J. Math. Logic 4 (2004), 181–212.
[3] E. Bouscaren, *Model theoretic versions of Weil's theorem on pre-groups*, in: Model Theory of Groups, A. Nesin and A. Pillay (eds.), Univ. of Notre Dame Press, Notre Dame, IN, 1989, 177–185.
[4] E. Casanovas, *Simple Theories and Hyperimaginaries*, Lecture Notes in Logic 39, Assoc. Symbolic Logic, Chicago, IL, and Cambridge Univ. Press, Cambridge, 2011.
[5] Z. Chatzidakis, L. van den Dries and A. Macintyre, *Definable sets over finite fields*, J. Reine Angew. Math. 427 (1992), 107–135.
[6] Z. Chatzidakis and E. Hrushovski, *Model theory of difference fields*, Trans. Amer. Math. Soc. 351 (1999), 2997–3071.
[7] A. Chernikov and S. Starchenko, *Regularity lemma for distal structures*, J. Eur. Math. Soc. 20 (2018), 2437–2466.
[8] R. Elwes and D. Macpherson, *A survey of asymptotic classes and measurable structures*, in: Model Theory with Applications to Algebra and Analysis, Vol. 2, London Math. Soc. Lecture Note Ser. 350, Cambridge Univ. Press, Cambridge, 2008, 125–159.
[9] D. García, D. Macpherson and C. Steinhorn, *Pseudofinite structures and simplicity*, J. Math. Logic 15 (2015), no. 1, art. 1550002, 41 pp.
[10] E. Hrushovski, *Expanding comments*, private correspondence with T. Tao, 2013.
[11] E. Hrushovski, *The elementary theory of the Frobenius automorphisms*, arXiv:math/0406514 (2004).
[12] Y. Kohayakawa, *Szemerédi's regularity lemma for sparse graphs*, in: Foundations of Computational Mathematics (Rio de Janeiro, 1997), Springer, Berlin, 1997, 216–230.
[13] P. Kowalski and A. Pillay, *A note on groups definable in difference fields*, Proc. Amer. Math. Soc. 130 (2002), 205–212.
[14] L. Lovász, *Large Networks and Graph Limits*, Amer. Math. Soc. Colloq. Publ. 60, Amer. Math. Soc., Providence, RI, 2012.
[15] L. Lovász and B. Szegedy, *Szemerédi's lemma for the analyst*, Geom. Funct. Anal. 17 (2007), 252–270.
[16] D. Macpherson and C. Steinhorn, *One-dimensional asymptotic classes of finite structures*, Trans. Amer. Math. Soc. 360 (2008), 411–448.
[17] M. Malliaris and S. Shelah, *Regularity lemmas for stable graphs*, Trans. Amer. Math. Soc. 366 (2014), 1551–1585.
[18] J. Nešetřil and P. Ossona de Mendez, *A unified approach to structural limits and limits of graphs with bounded tree-depth*, Mem. Amer. Math. Soc. 263 (2020), no. 1272, v+108 pp.
[19] A. Pillay and S. Starchenko, *Remarks on Tao's algebraic regularity lemma*, arXiv: 1310.7538 (2013).
[20] M. Ryten and I. Tomašić, *ACFA and measurability*, Selecta Math. (N.S.) 11 (2005), 523–537.

[21] A. Scott, *Szemerédi's regularity lemma for matrices and sparse graphs*, Combin. Probab. Comput. 20 (2011), 455–466.
[22] T. 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] T. Tao, *Expanding polynomials over finite fields of large characteristic, and a regularity lemma for definable sets*, Contrib. Discrete Math. 10 (2015), no. 1, 22–98.
[24] I. Tomašić, *Geometric simplicity theory*, PhD thesis, Univ. of Edinburgh, 2001.
[25] I. Tomašić and F. O. Wagner, *Applications of the group configuration theorem in simple theories*, J. Math. Logic 3 (2003), 239–255.
[26] F. O. Wagner, *Simple Theories*, Math. Appl. (Dordrecht) 503, Kluwer, 2000.

Mirna Džamonja
IRIF (CNRS-Université de Paris)
Bâtiment Sophie Germain, Case courrier 7014
8 Place Aurélie Nemours
75205 Paris Cedex 13, France
E-mail: mdzamonja@irif.fr

Ivan Tomašić
School of Mathematical Sciences
Queen Mary University of London
London, E1 4NS, United Kingdom
E-mail: i.tomasic@qmul.ac.uk