STRUCTURE AND REGULARITY FOR SUBSETS OF GROUPS WITH FINITE VC-DIMENSION

G. CONANT, A. PILLAY, AND C. TERRY

Abstract. Suppose G is a finite group and $A \subseteq G$ is such that $\{gA : g \in G\}$ has VC-dimension strictly less than k. We find algebraically well-structured sets in G which, up to a chosen $\epsilon > 0$, describe the structure of A and behave regularly with respect to translates of A. For the subclass of groups with uniformly fixed finite exponent r, these algebraic objects are normal subgroups with index bounded in terms of k, r, and ϵ. For arbitrary groups, we use Bohr neighborhoods, of bounded dimension and width, inside normal subgroups of bounded index. Our proofs are largely model theoretic, and heavily rely on a structural analysis of compactifications of pseudofinite groups as inverse limits of Lie groups. The introduction of Bohr neighborhoods into the nonabelian setting uses model theoretic methods related to the work of Breuillard, Green, and Tao [7] and Hrushovski [21] on approximate groups, as well as a result of Alekseev, Glebskiĭ, and Gordon [1] on approximate homomorphisms.

1. Introduction and statement of results

Szemerédi's Regularity Lemma [39] is a fundamental result about graphs, which has found broad applications in graph theory, computer science, and arithmetic combinatorics. Roughly speaking, the regularity lemma partitions large graphs into few pieces such that almost all pairs of pieces have uniform edge density. In 2005, Green [18] proved the first *arithmetic regularity lemma*, which uses discrete Fourier analysis to define arithmetic notions of regularity for subsets of finite abelian groups. For groups of the form $\mathbb{F}_p^n$, Green's result states that given $A \subseteq \mathbb{F}_p^n$, there is $H \leq \mathbb{F}_p^n$ of bounded index such that A is Fourier-uniform with respect to almost all cosets of H. Arithmetic regularity lemmas, and their higher order analogues (see [16, 19]), are now important tools in arithmetic combinatorics. From a general perspective, one can view regularity lemmas as tools for decomposing mathematical objects into ingredients that are easier to study because they are either highly structured (e.g., a partition of a graph) or highly random (e.g., a regular pair in a graph).

Recently, a large body of work has developed around strengthened regularity lemmas for classes of graphs which forbid some particular bipartite configuration. This setting is fundamental in both combinatorics and model theory, although often for very different reasons. In combinatorics, forbidden configurations often lead to significant *quantitative* improvements in results about graphs, and several well-known open problems arise in this pursuit (e.g., the Erdős-Hajnal conjecture). In model theory, the focus is usually on infinite objects, and forbidden configurations are

Date: February 11, 2019.
2010 *MSC*. Primary: 03C45, 03C20, 20D60, 22C05; Secondary: 22E35.
The second author was supported by NSF grants DMS-1360702 and DMS-1665035. This work was finished during the Model Theory, Combinatorics and Valued Fields trimester program at Institut Henri Poincaré, at which the authors were participants. We thank IHP for their hospitality. We also thank Julia Wolf for many helpful conversations.

used to obtain *qualitative* results about definable sets in mathematical structures. Indeed, much of modern model theory developed from the study of mathematical structures in which every definable bipartite graph omits a finite "half-graph" as an induced subgraph (such structures are called *stable*).

In combinatorics and model theory, the practice of forbidding finite bipartite configurations is rigorously formulated using VC-dimension. By definition, the **VC-dimension** of a *bipartite* graph $(V, W; E)$ is the supremum of all $k \in \mathbb{Z}^+$ such that $(V, W; E)$ contains $([k], \mathcal{P}([k]); \in)$ as induced subgraph (where $[k] = \{1, \ldots, k\}$). We call $(V, W; E)$ k**-NIP** if it has VC-dimension at most $k - 1$.[1] While this definition is based on omitting one specific bipartite graph, an illuminating exercise is that if $(V, W; E)$ omits *some* finite bipartite graph (V', W', E') as induced subgraph, then $(V, W; E)$ is k-NIP for some $k \leq |V'| + \lceil \log_2 |W'| \rceil$. In other words, having finite VC-dimension is equivalent to omitting *some* finite bipartite configuration.

In [2], Alon, Fischer, and Newman proved a strengthened regularity lemma for finite graphs of bounded VC-dimension[2], in which the bound on the size of the partition is polynomial in the degree of irregularity. This is in contrast to Szemerédi's original work, where these bounds are necessarily tower-type [17]. Strengthened regularity lemmas have also been found for semi-algebraic hypergraphs (e.g., [14, 15]), as well as for various model theoretic settings inside NIP (see [4, 10, 11, 28]). The strongest result is in the setting of stable graphs,[3] where Malliaris and Shelah [28] prove the existence of regular partitions with polynomial bounds, no irregular pairs, and such that the uniform edge density in any regular pair is close to 0 or 1 (i.e., as a bipartite graph, each regular pair is almost empty or complete). This final condition on regular pairs is also obtained in the regularity lemmas for NIP graphs. In other words, for NIP graphs, the normally random ingredients of Szemerédi regularity are in fact also highly structured.

The goal of this article is to develop arithmetic regularity for arbitrary finite groups in the context of forbidden bipartite configurations, as quantified by VC-dimension. In analogy to the case of graphs, we show that by forbidding finite bipartite configurations, one obtains a strengthened version of arithmetic regularity in which all aspects are highly structured and the random ingredients have been removed. Our results qualitatively generalize and strengthen recent work of the third author and Wolf [43] and Alon, Fox, and Zhao [3] on arithmetic regularity lemmas for stable and NIP sets in finite abelian groups. Moreover, our proof methods deepen the connection between model theory and additive combinatorics, in that we use pseudofinite methods to extend combinatorial results for finite abelian groups to the nonabelian setting. This is in the same vein as Hrushovski's [21] celebrated work on approximate groups, and the subsequent structure theory proved by Breuillard, Green, and Tao [7]. We will use techniques and results from the literature on approximate groups in order to formulate arithmetic regularity in nonabelian groups using *Bohr sets*, which are fundamentally linked to abelian groups. Finally, our results show that NIP arithmetic regularity for finite sets in abelian groups coincides with a certain model theoretic phenomenon called "compact domination". This notion was first isolated by Hrushovski, Peterzil, and

[1] This terminology is from model theory, where a bipartite graph with infinite VC-dimension is said to have the *independence property*, and so NIP stands for "no independence property".

[2] The VC-dimension of a graph $(V; E)$ is that of the associated bipartite graph $(V, V; E)$.

[3] A bipartite graph $(V, W; E)$ is called k-*stable* if it omits $([k], [k]; \leq)$ as induced subgraph.

Pillay [22] in their proof of the so-called "Pillay conjectures" for groups definable in o-minimal theories, and later played an important role in the development of definably amenable groups definable in NIP theories [9, 22, **?**, 24].

Before stating the main results of this paper, we briefly recall our main result on stable arithmetic regularity from [13], as it provides a template for the "structure and regularity" statements will will obtain for NIP sets. In [43], the third author and Wolf consider *arithmetic regularity* for sets $A \subseteq \mathbb{F}_p^n$ such that the graph on $\mathbb{F}_p^n$ given by $x + y \in A$ is k-stable for some $k \geq 1$. They prove that such sets satisfy a strengthened version of Green's arithmetic regularity lemma above, in which there is an efficient bound on the index of H and there are no non-uniform cosets. They also show that any stable subset of $\mathbb{F}_p^n$ is approximately a union of cosets of a subgroup of small index, which is the arithmetic analogue of "0-1 density" in regular pairs. In [13], we generalized and strengthened the qualitative aspects of the stable arithmetic regularity lemma from [43] to the setting of arbitrary finite groups, but without explicit bounds (a quantitative analogue remains open).

Theorem 1.1 (Conant, Pillay, Terry [13])**.** *For any $k \geq 1$ and $\epsilon > 0$, there is $n = n(k, \epsilon)$ such that the following holds. Suppose G is a finite group and $A \subseteq G$ is k-stable. Then there is a normal subgroup $H \leq G$, of index at most n, satisfying the following properties.*

(i) (structure) *There is $D \subseteq G$, which is a union of cosets of H, such that*

$$|A \triangle D| < \epsilon|H|.$$

(ii) (regularity) *For any $g \in G$, either $|gH \cap A| < \epsilon|H|$ or $|gH \backslash A| < \epsilon|H|$.*

Moreover, H is in the Boolean algebra generated by $\{gAh : g, h \in G\}$.

Our first result on arithmetic regularity in the setting of bounded VC-dimension is for k-NIP subsets of finite groups with *uniformly bounded exponent*.

Theorem 3.3. *For any $k, r \geq 1$ and $\epsilon > 0$, there is $n = n(k, r, \epsilon)$ such that the following holds. Suppose G is a finite group of exponent r, and $A \subseteq G$ is k-NIP. Then there are*

∗ *a normal subgroup $H \leq G$ of index at most n, and*
∗ *a set $Z \subseteq G$, which is a union of cosets of H with $|Z| < \epsilon|G|$,*

satisfying the following properties.

(i) (structure) *There is $D \subseteq G$, which is a union of cosets of H, such that*

$$|(A \backslash Z) \triangle D| < \epsilon|H|.$$

(ii) (regularity) *For any $g \in G \backslash Z$, either $|gH \cap A| < \epsilon|H|$ or $|gH \backslash A| < \epsilon|H|$.*

Moreover, H is in the Boolean algebra generated by $\{gAh : g, h \in G\}$.

Thus the behavior of NIP sets in bounded exponent groups is almost identical to that of stable sets in arbitrary finite groups, where the only difference is in the the error set Z. Theorem 3.3 also qualitatively generalizes and strengthens a recent result of Alon, Fox, and Zhao [3] on the case of k-NIP subsets of finite *abelian* groups of uniformly bounded exponent.[4] A version of Theorem 3.3 with polynomial bounds, but slightly weaker qualitative ingredients, is conjectured in [3].

We then turn to k-NIP sets in arbitrary finite groups. Easy counterexamples show that one cannot expect a statement involving only subgroups, as in Theorem

[4]Once again, "qualitative" means that we do not obtain explicit bounds, in contrast to [3].

3.3. Such counterexamples live in groups which have very few subgroups (e.g. cyclic groups of prime order). In situations like this, one often works with *Bohr neighborhoods*, which are certain well-structured sets in groups. Bohr neighborhoods in cyclic groups were used in Bourgain's improvement of Roth's Theorem [6], and also form the basis of Green's arithmetic regularity lemma for finite abelian groups [18]. Results related to ours, with quantitative bounds but for abelian groups, were recently obtained by Sisask [38].

Given a group H, an integer $r \geq 0$, and a $\delta > 0$, we define a (δ, r)*-Bohr neighborhood in* H to be a set of the form $B^r_{\tau,\delta} := \{x \in H : d(\tau(x), 0) < \delta\}$, where $\tau \colon H \to \mathbb{T}^r$ is a homomorphism and d denotes the usual invariant metric on the r-dimensional torus $\mathbb{T}^r$. Our main structure and regularity result for NIP sets in finite groups is as follows.

Theorem 5.7. *For any* $k \geq 1$ *and* $\epsilon > 0$ *there is* $n = n(k, \epsilon)$ *such that the following holds. Suppose* G *is a finite group and* $A \subseteq G$ *is* k*-NIP. Then there are*

- *a normal subgroup* $H \leq G$ *of index at most* n*,*
- *a* (δ, r)*-Bohr neighborhood* B *in* H*, for some* $r \leq n$ *and* $\delta \geq \frac{1}{n}$*, and*
- *a subset* $Z \subseteq G$*, with* $|Z| < \epsilon|G|$*,*

satisfying the following properties.

(*i*) (structure) *There is* $D \subseteq G$*, which is a union of translates of* B*, such that*

$$|(A \triangle D)\backslash Z| < \epsilon|B|.$$

(*ii*) (regularity) *For any* $g \in G\backslash Z$*, either* $|gB \cap A| < \epsilon|B|$ *or* $|gB\backslash A| < \epsilon|B|$*.*

Moreover, H *and* Z *are in the Boolean algebra generated by* $\{gAh : g, h \in G\}$*, and if* G *is abelian then we may assume* $H = G$*.*

In order to prove Theorems 3.3 and 5.7, we will first prove companion theorems for these results involving definable sets in infinite pseudofinite groups (Theorems 3.2 and 5.5, respectively). We then prove the theorems about finite groups by taking ultraproducts of counterexamples in order to obtain infinite pseudofinite groups contradicting the companion theorems.

To prove these companion theorems, we work with a saturated pseudofinite group G, and an invariant NIP formula $\theta(x; \bar{y})$ (see Definitions 2.1 and 2.11). In [12], the first two authors proved "generic compact domination" for the quotient group $G/G^{00}_{\theta^r}$, where $G^{00}_{\theta^r}$ denotes the local type-definable connected component relative to $\theta(x; \bar{y})$. In this case, $G/G^{00}_{\theta^r}$ is a compact Hausdorff group, and generic compact domination roughly states that, given a θ-definable set $X \subseteq G$, the set of cosets of $G^{00}_{\theta^r}$, which intersect both X and $G\backslash X$ in "large" sets, has Haar measure 0 (see Fact 2.16). This is essentially a regularity statement for X with respect to the subgroup $G^{00}_{\theta^r}$, which is rather remarkable as generic compact domination originated in the global NIP setting (see [22]) related to conjectures of the second author on the Lie structure of groups definable in o-minimal theories [33].

The regularity provided by generic compact domination for θ-definable sets in G cannot be transferred directly to finite groups, as the statement depends entirely on type-definable data (such as $G^{00}_{\theta^r}$). Thus, much of the work in this article focuses on obtaining definable approximations to $G^{00}_{\theta^r}$ and the other objects involved in generic compact domination. We first investigate the situation when $G/G^{00}_{\theta^r}$ is a profinite group, in which case $G^{00}_{\theta^r}$ can be approximated by definable finite-index subgroups of G. Using this, we prove Theorem 3.2 (the pseudofinite companion to Theorem

3.3 above). The connection to Theorem 3.3 is that, if G is elementarily equivalent to an ultraproduct of groups of uniformly bounded exponent, then $G/G^{00}_{\theta^r}$ is a compact Hausdorff group of finite exponent, hence is profinite [25].

When $G/G^{00}_{\theta^r}$ is not profinite, there are not enough definable finite-index subgroups available to describe $G^{00}_{\theta^r}$, and it is for this reason that we turn to Bohr neighborhoods. This is somewhat surprising, as Bohr neighborhoods are fundamentally linked to commutative groups, and we do not make any assumptions of commutativity. However, by a result of the second author [34], since G is pseudofinite, the identity component in $G/G^{00}_{\theta^r}$ is commutative. It is at this point that we see the beautiful partnership between pseudofinite groups and NIP formulas. Specifically, we have generic compact domination of θ-definable sets by the commutative-by-profinite group $G/G^{00}_{\theta^r}$, which allows us to describe θ-definable sets in G using Bohr neighborhoods inside of definable finite-index subgroups. In order to obtain a statement involving only *definable* objects, we use approximate homomorphisms to formulate a notion of approximate Bohr neighborhoods. This leads to Theorem 5.5 (the pseudofinite companion of Theorem 5.7 above), and to a version of Theorem 5.7 involving approximate Bohr neighborhoods (see Lemma 5.6). We then apply a result of Alekseev, Glebskiĭ, and Gordon [1] on approximate homomorphisms in order to find actual Bohr neighborhoods inside of approximate Bohr neighborhoods and, ultimately, prove Theorem 5.7.

In Section 6, we prove similar results for *fsg* groups definable in *distal* NIP theories (see Theorems 6.6 and 6.7). This follows our theme, as such groups satisfy a strong form of compact domination (see Fact 6.4). In Section 7, we discuss compact p-adic analytic groups as one concrete example of the distal *fsg* setting.

2. Preliminaries

Let $\mathcal{L}$ be a first-order language expanding the language of groups, and let G be a sufficiently saturated $\mathcal{L}$-structure expanding a group. An $\mathcal{L}_G$*-formula* is a formula in the language $\mathcal{L}$ with parameters from G. Throughout the paper, we say that a set is *bounded* if its cardinality is strictly less than the saturation of G.[5]

We note some terminology for formulas, and refer the reader to [29] for an introduction to first-order languages and basic model theory.

Definition 2.1. Let $\theta(x;\bar{y})$ be an $\mathcal{L}_G$-formula.

(1) An **instance of** $\theta(x;\bar{y})$ is a formula of the form $\theta(x;\bar{b})$ for some $\bar{b}\in G^{\bar{y}}$.
(2) A **θ-formula** is a finite Boolean combination of instances $\theta(x;\bar{y})$ (we allow $x=x$ as a trivial θ-formula).
(3) Let $S_\theta(G)$ denote the space of **complete θ-types over** G (i.e. ultrafilters in the Boolean algebra of θ-formulas).
(4) Let $\theta^r(x;\bar{y},u)$ denote the formula $\theta(x\cdot u;\bar{y})$.
(5) $\theta(x;\bar{y})$ is **(left) invariant** if, for any $a\in G$ and $\bar{b}\in G^{\bar{y}}$, there is $\bar{c}\in G^{\bar{y}}$ such that $\theta(a\cdot x;\bar{b})$ is equivalent to $\theta(x;\bar{c})$.

The typical example of an invariant $\mathcal{L}_G$-formula is something of the form $\theta(x;y) := \phi(y^{-1}\cdot x)$, where $\phi(x)$ is an $\mathcal{L}_G$-formula in one variable.

Definition 2.2.

[5] This is not to be confused with later uses of the phrase "uniformly bounded" in the context of theorems about finite groups.

(1) A set $X \subseteq G$ is **definable** if $X = \phi(G)$ for some $\mathcal{L}_G$-formula $\phi(x)$. If $\phi(x)$ is a θ-formula, for some $\theta(x;\bar{y})$, then X is **θ-definable**.
(2) A set $X \subseteq G$ is **type-definable** if $X = \bigcap_{i\in I} \phi_i(G)$, where each $\phi_i(x)$ is an $\mathcal{L}_G$-formula and I is bounded. If each $\phi_i(x)$ is a θ-formula, for some (fixed) $\theta(x;\bar{y})$, then X is **θ-type-definable**.

2.1. **Compact quotients.** Throughout the paper, when we say that a topological space is *compact*, we mean compact and Hausdorff. Given a compact group K, we let K^0 denote the connected component of the identity in K. Recall that K is *profinite* if it is a projective limit of finite groups. If K is compact, then K is profinite if and only if K^0 is trivial.

Continuing with the fixed saturated group G from above, we recall some basic facts about quotients of G by type-definable subgroups.

Remark 2.3. Suppose $\Gamma \leq G$ is type-definable and normal of bounded index.
(1) G/Γ is a compact group under the *logic topology*, where $X \subseteq G/\Gamma$ is *closed* if and only if $\{a \in G : a\Gamma \in X\}$ is type-definable. See [33] for details.
(2) If $X \subseteq G$ is definable then $\{a\Gamma \in G/\Gamma : a\Gamma \subseteq X\}$ is an open subset of G/Γ. Indeed, by saturation of G and type-definability of Γ, the set $\{a \in G : a\Gamma \cap G\backslash X \neq \emptyset\}$ is type-definable.

Definition 2.4. Suppose $X \subseteq G$ is definable, Y is a compact space, and $f\colon X \to Y$ is a function. Then f is **definable** if $f^{-1}(C)$ is type-definable for any closed $C \subseteq Y$.

Remark 2.5. If $f\colon X \to Y$ is definable in the above sense then, by saturation of G, for any closed $C \subseteq Y$ and open $U \subseteq Y$, with $C \subseteq U$, there is $D \subseteq X$ definable such that $f^{-1}(C) \subseteq D \subseteq f^{-1}(U)$.

Given a type-definable bounded-index subgroup $\Gamma \leq G$, since G/Γ is a compact group, it can be analyzed using the structure theory involving (finite-dimensional) compact Lie groups. Since the topology on G/Γ is controlled by type-definable objects in G, this leads to an analysis of the type-definable structure of Γ involving compact Lie groups. The follow lemma collects the basic ingredients of this analysis. We say that a set $X \subseteq G$ is *θ^r-countably-definable* if it is an intersection of countably many θ^r-definable sets.

Lemma 2.6. *Let $\theta(x;\bar{y})$ be invariant and suppose Γ is a θ^r-countably-definable bounded-index normal subgroup of G. Then there is a sequence $(\Gamma_t)_{t=0}^\infty$ of θ^r-countably-definable bounded-index normal subgroups of G, and a sequence $(H_t)_{t=0}^\infty$ of θ^r-definable finite-index normal subgroups of G, with the following properties.*
(i) *For all $t \in \mathbb{N}$, $\Gamma_{t+1} \leq \Gamma_t$, $H_{t+1} \leq H_t$, and $\Gamma_t \leq H_t$.*
(ii) $\Gamma = \bigcap_{n=0}^\infty \Gamma_t$.
(iii) $G/\Gamma = \varprojlim G/\Gamma_t$ *and* $(G/\Gamma)^0 = \varprojlim H_t/\Gamma_t$.
(iv) *For all $t \in \mathbb{N}$, G/Γ_t is a compact Lie group, $H_t/\Gamma_t = (G/\Gamma_t)^0$ is a compact connected Lie group, and the projection map $G \to G/\Gamma_t$ is θ^r-definable.*

Proof. Since the logic topology on G/Γ is induced by θ^r-type-definability (see [12, Corollary 4.2]), we may assume $\mathcal{L}$ is countable, and thus G/Γ is separable (see [9, Remark 2.15], [36, Proposition 5.1]). Now G/Γ is a separable compact group, and so there is a sequence $(K_t)_{t=0}^\infty$ of closed normal subgroups of G/Γ such that $(G/\Gamma)/K_t$ is a Lie group, $K_{t+1} \leq K_t$ for all $t \in \mathbb{N}$, $\bigcap_{t=0}^\infty K_t = \{1\}$, and $G/\Gamma = \varprojlim (G/\Gamma)/K_t$ (see [40]). Let Γ_t be the pullback of K_t to G. Then Γ_t is a θ^r-type-definable bounded-index normal subgroup of G containing Γ. Using Remark 2.5, one sees that each Γ_t is θ^r-countably-definable. Since $\bigcap_{t=0}^\infty K_t = \{1\}$, we have $\bigcap_{t=0}^\infty \Gamma_t = \Gamma$. For any $t \in \mathbb{N}$, $K_{t+1} \leq K_t$ implies $\Gamma_{t+1} \leq \Gamma_t$. Now $(G/\Gamma)/K_t$ and G/Γ_t are isomorphic topological groups, and so we have $G/\Gamma = \varprojlim G/\Gamma_t$.

Fix $t \in \mathbb{N}$, and let H_t be the pullback to G of $(G/\Gamma_t)^0$. Since $(G/\Gamma_t)^0$ is a closed finite-index normal subgroup of G/Γ_t, and the projection from G to G/Γ_t is θ^r-definable, we have that H_t is a θ^r-type-definable finite-index normal subgroup of G. So H_t is θ^r-definable. Note also that $H_t/\Gamma_t = (G/\Gamma_t)^0$ and $(G/\Gamma)^0 = \varprojlim H_t/\Gamma_t$. In particular, H_t/Γ_t is a compact connected Lie group. Finally, the pullback of H_t/Γ_t to G/Γ_{t+1} is a clopen normal finite-index subgroup, and thus contains $(G/\Gamma_{t+1})^0 = H_{t+1}/\Gamma_{t+1}$. So $H_{t+1} \leq H_t$. □

Remark 2.7. Lemma 2.6 can be reformulated for θ^r-type-definable subgroups of G of bounded-index, using inverse systems indexed by directed sets.

Corollary 2.8. *Let $\theta(x;\bar{y})$ be invariant, and suppose Γ is a θ^r-countably-definable bounded-index normal subgroup of G. If G/Γ is profinite then Γ is a countable intersection of θ^r-definable finite-index subgroups of G.*

Proof. If G/Γ is profinite then, in Lemma 2.6, we have $H_t = \Gamma_t$ for all $t \in \mathbb{N}$. □

In the next section, we will focus on the case that G is **pseudofinite** which, as G is saturated, we take to mean that G is an elementary extension of an ultraproduct of finite $\mathcal{L}$-structures, each of which is a group. The following fact is the key ingredient which will eventually allow us to introduce Bohr sets into the setting of NIP subsets of possibly non-commutative groups.

Fact 2.9 (Pillay [34]). *Assume G is pseudofinite. If Γ is a type-definable bounded-index normal subgroup of G then $(G/\Gamma)^0$ is commutative.*[6]

Thus, in the case that G is pseudofinite and $\Gamma \leq G$ is type-definable of bounded index, $(G/\Gamma)^0$ is a projective limit of compact connected commutative Lie groups. Recall that compact connected commutative Lie groups are classified as precisely the finite-dimensional tori. So for $n \in \mathbb{N}$, let $\mathbb{T}^n$ denote the n-dimensional torus, i.e., $\mathbb{T}^n = \mathbb{R}/\mathbb{Z} \times \ldots \times \mathbb{R}/\mathbb{Z}$ (n times). Note $\mathbb{T}^0$ is the trivial group.

Corollary 2.10. *Assume G is pseudofinite. Let $\theta(x;\bar{y})$ be invariant, and suppose Γ is a θ^r-countably-definable bounded-index normal subgroup of G. Then, in Lemma 2.6, we have that for all $t \in \mathbb{N}$ there is $n_t \in \mathbb{N}$ such that $H_t/\Gamma_t \cong \mathbb{T}^{n_t}$.*

2.2. **NIP formulas in pseudofinite groups.** We continue to work with a saturated group G as in the previous section.

Definition 2.11. Let $\theta(x;\bar{y})$ be an $\mathcal{L}_G$-formula. Given $k \geq 1$, $\theta(x;\bar{y})$ is k-**NIP** if there do not exist sequences $(a_i)_{i\in[k]}$ in G and $(\bar{b}_I)_{I\subseteq[k]}$ in $G^{\bar{y}}$ such that $\theta(a_i, \bar{b}_I)$ holds if and only if $i \in I$. We say $\theta(x;\bar{y})$ is **NIP** if it is k-NIP for some $k \geq 1$.

Remark 2.12. Given $k \geq 1$, an $\mathcal{L}_G$-formula $\theta(x;\bar{y})$ is k-NIP if and only if the set system $\{\theta(G;\bar{b}) : \bar{b} \in G^{\bar{y}}\}$ on G has VC-dimension at most $k-1$.

Definition 2.13. A subset $A \subseteq G$ is **left generic** (resp. **right generic**) if G is a union of finitely many left (resp. right) translates of A. A formula $\phi(x)$ is **left generic** (resp. **right generic**) if $\phi(G)$ is left generic (resp. right generic).

[6] In [31], Nikolov, Schneider, and Thom generalize this to arbitrary compactifications of (abstract) pseudofinite groups.

Now we focus on the case that G is *pseudofinite* (as defined above). In this case, μ will always denote the pseudofinite counting measure on G, which is obtained from the ultralimit of normalized counting measures on finite $\mathcal{L}$-structures (and then lifted to G, which is an elementary extension of an ultraproduct of finite $\mathcal{L}$-structures). In several proofs, we will apply Łoś's Theorem to properties of μ, which requires an expanded language $\mathcal{L}^+$ containing a sort for the ordered interval $[0,1]$ with a distance function, and functions from the G-sort to $[0,1]$ giving the measures of $\mathcal{L}$-formulas. There are many accounts of this kind of formalism, and so we will omit further details and refer the reader to similar treatments in the literature, for example [12, Section 2.2] and [21, Section 2.6].

The next fact is a fundamental property of NIP formulas in pseudofinite groups, which forms the basis for much of the work in [12].

Fact 2.14 (Conant & Pillay [12])**.** *Assume G is pseudofinite. Suppose $\phi(x)$ is an $\mathcal{L}_G$-formula such that $\phi(y^{-1}\cdot x)$ is NIP. Then the following are equivalent:*

(i) *$\phi(x)$ is left generic;*
(ii) *$\phi(x)$ is right generic;*
(iii) *$\mu(\phi(x))>0$.*

A straightforward exercise is that, if $\theta(x;\bar{y})$ is invariant and NIP, and $\phi(x)$ is a θ^r-formula, then $\phi(y^{-1}\cdot x)$ is NIP. So, in this situation, we will just say that $\phi(x)$ is **generic** in the case it is left generic (equivalently, right generic). We also call a θ^r-type $p\in S_{\theta^r}(G)$ **generic** if every formula in p is generic.

Corollary 2.15. *Assume G is pseudofinite. If $\theta(x;\bar{y})$ is invariant and NIP then there are generic types in $S_{\theta^r}(G)$.*

Proof. The set of θ^r-formulas of μ-measure 1 is a filter, and so can be extended to a type in $S_{\theta^r}(G)$. This type must be generic by Fact 2.14. □

The next fact gives the "generic compact domination" statement from [12].

Fact 2.16 (Conant & Pillay [12])**.** *Assume G is pseudofinite. Suppose $\theta(x;\bar{y})$ is invariant and NIP. Then there is a normal subgroup $G^{00}_{\theta^r}$ of G satisfying the following properties.*

(a) *$G^{00}_{\theta^r}$ is θ^r-countably-definable of bounded index, and is the intersection of all θ^r-type-definable bounded-index subgroups of G.*
(b) *Suppose $X\subseteq G$ is θ^r-definable. Let $E_X\subseteq G/G^{00}_{\theta^r}$ be the set of $C\in G/G^{00}_{\theta^r}$ such that*
$$p\models C\cap X \quad \textit{and} \quad q\models C\cap (G\backslash X)$$
for some generic θ^r-types $p,q\in S_{\theta^r}(G)$. Then E_X has Haar measure 0 and so, for any $\epsilon>0$, there is a θ^r-definable set $Z\subseteq G$ such that $\mu(Z)<\epsilon$ and $\{a\in G: aG^{00}_{\theta^r}\subseteq E_X\}\subseteq Z$.

Remark 2.17. Except for Section 6, our use of NIP and finite VC-dimension is entirely concentrated in the application of Fact 2.16. So it is worth emphasizing that the proof of this fact heavily uses fundamental tools about set systems of finite VC-dimension, namely, the Sauer-Shelah Lemma, the VC-Theorem, and Matoušek's general version [30] of the (p,q)-theorem. See also [9] and [36, Chapter 6].

3. Structure and regularity: the profinite case

Let G be a saturated $\mathcal{L}$-structure expanding a group, as in Section 2. In this section, we prove a structure and regularity theorem for θ^r-definable sets in G, in the case that $\theta(x;\bar{y})$ is NIP, G is pseudofinite, and $G/G^{00}_{\theta^r}$ is profinite (see Theorem 3.2). As an application, we obtain a structure and regularity theorem for NIP sets in finite groups of uniformly bounded exponent (see Theorem 3.3).

First, we prove a key lemma. Specifically, given an invariant NIP $\mathcal{L}_G$-formula $\theta(x;\bar{y})$, we use generic compact domination for $G/G^{00}_{\theta^r}$ (Fact 2.16) to derive regularity for θ^r-definable sets in G, with respect to definable approximations of $G^{00}_{\theta^r}$.

Lemma 3.1. *Assume G is pseudofinite. Let $\theta(x;\bar{y})$ be invariant and NIP, and let $\{W_i : i \in \mathbb{N}\}$ be a collection of definable sets such that $G^{00}_{\theta^r} = \bigcap_{i=0}^{\infty} W_i$ and, for all $i \in \mathbb{N}$, $W_{i+1} \subseteq W_i$. Fix a θ^r-definable set $X \subseteq G$. Then, for any $\epsilon > 0$, there is a θ^r-definable set $Z \subseteq G$ and some $i \in \mathbb{N}$ such that*

(i) $\mu(Z) < \epsilon$, *and*
(ii) *for any $g \in G\backslash Z$, either $\mu(gW_i \cap X) = 0$ or $\mu(gW_i\backslash X) = 0$.*

Proof. To ease notation, let $\Gamma = G^{00}_{\theta^r}$ and let $K = G/G^{00}_{\theta^r}$. Let $\pi\colon G \to K$ be the canonical projection. Let $E_X \subseteq K$ be as in Fact 2.16, and fix $\epsilon > 0$.

Claim 1: For any $a \notin \pi^{-1}(E_X)$ there is some $i \in \mathbb{N}$ such that either $\mu(aW_i \cap X) = 0$ or $\mu(aW_i\backslash X) = 0$.

Proof: Since Γ is θ^r-type-definable, a routine compactness argument shows that we may assume each W_i is θ^r-definable. Suppose the claim fails, i.e. there is $a \notin \pi^{-1}(E_X)$ such that, for all $i \in \mathbb{N}$, $\mu(aW_i \cap X) > 0$ and $\mu(aW_i\backslash X) > 0$. Let $p_0 = \{aW_i \cap X : i \in \mathbb{N}\}$ and $q_0 = \{aW_i\backslash X : i \in \mathbb{N}\}$. Then p_0 and q_0 are partial types of generic θ^r-definable sets, and are closed under finite conjunctions. So there are generic $p, q \in S_{\theta^r}(G)$ such that $p_0 \subseteq p$ and $q_0 \subseteq q$. Note that $p \models a\Gamma \cap X$ and $q \models a\Gamma \cap (G\backslash X)$. Thus $a\Gamma \in E_X$, which is a contradiction. $\dashv_{\text{Claim 1}}$

By Fact 2.16, we may fix a θ^r-definable set $Z \subseteq G$ such that $\mu(Z) < \epsilon$ and $\pi^{-1}(E_X) \subseteq Z$. Let $U = \{C \in K : C \subseteq Z\}$. Then $E_X \subseteq U$, $\pi^{-1}(U) \subseteq Z$, and U is an open set in K by Remark 2.3(2). Toward a contradiction, suppose that for all $i \in \mathbb{N}$ there is some $a_i \notin \pi^{-1}(U)$ such that $\mu(a_iW_i \cap X) > 0$ and $\mu(a_iW_i\backslash X) > 0$. Then $(a_i\Gamma)_{i=0}^{\infty}$ is an infinite sequence in $K\backslash U$. Since U is open and K is compact, we may pass to a subsequence and assume that $(a_i\Gamma)_{i=0}^{\infty}$ converges to some $a\Gamma \in K\backslash U$.

Claim 2: For all $i \in \mathbb{N}$, $\mu(aW_i \cap X) > 0$ and $\mu(aW_i\backslash X) > 0$.

Proof: First, given $i \in \mathbb{N}$, let $U_i = \{C \in K : C \subseteq aW_i\}$. As above, each U_i is open in K. Moreover, for any $i \in \mathbb{N}$, since $\Gamma \subseteq W_i$, we have $a\Gamma \in U_i \subseteq aW_i/\Gamma$.

Next, since $W_{i+1} \subseteq W_i$ for all $i \in \mathbb{N}$, $\Gamma = \bigcap_{i=0}^{\infty} W_i$, and Γ is a group, it follows from compactness that, for all $i \in \mathbb{N}$, there is $n_i \in \mathbb{N}$ such that $W_{n_i}W_{n_i} \subseteq W_i$.

Now fix $i \in \mathbb{N}$. Since U_{n_i} is an open neighborhood of $a\Gamma$, there is $j \geq n_i$ such that $a_j\Gamma \in U_{n_i}$. In particular, $a_j \in \pi^{-1}(U_{n_i}) \subseteq aW_{n_i}$. Now we have $a_jW_j \subseteq a_jW_{n_i} \subseteq aW_{n_i}W_{n_i} \subseteq aW_i$. Since $\mu(a_jW_j \cap X) > 0$ and $\mu(a_jW_j\backslash X) > 0$, we have $\mu(aW_i \cap X) > 0$ and $\mu(aW_i\backslash X) > 0$. $\dashv_{\text{Claim 2}}$

Finally, since $a \notin \pi^{-1}(U) \supseteq \pi^{-1}(E_X)$, Claim 2 contradicts Claim 1. So there is some $i \in \mathbb{N}$ such that, for all $a \notin \pi^{-1}(U)$, we have $\mu(aW_i \cap X) = 0$ or $\mu(aW_i\backslash X) = 0$. Since $\pi^{-1}(U) \subseteq Z$, this finishes the proof. $\square$

Theorem 3.2. *Assume G is pseudofinite. Let $\theta(x;\bar{y})$ be invariant and NIP, and suppose $G/G^{00}_{\theta^r}$ is profinite. Fix a θ^r-definable set $X \subseteq G$ and some $\epsilon > 0$. Then there are*

∗ *a θ^r-definable finite-index normal subgroup $H \leq G$, and*
∗ *a set $Z \subseteq G$, which is a union of cosets of H with $\mu(Z) < \epsilon$,*

satisfying the following properties.

(i) (structure) *There is $D \subseteq G$, which is a union of cosets of H, such that*

$$\mu((X \backslash Z) \triangle D) = 0.$$

(ii) (regularity) *For any $g \in G \backslash Z$, either $\mu(gH \cap X) = 0$ or $\mu(gH \backslash X) = 0$.*

Proof. Let $X \subseteq G$ and $\epsilon > 0$ be fixed. By Corollary 2.8, there is a collection $\{H_t : t \in \mathbb{N}\}$ of θ^r-definable finite-index subgroups of G, such that $H_{t+1} \subseteq H_t$ for all $t \in \mathbb{N}$ and $G^{00}_{\theta^r} = \bigcap_{t=0}^{\infty} H_t$. By Lemma 3.1, there is a θ^r-definable $Z \subseteq G$ and some $t \in \mathbb{N}$ such that $\mu(Z) < \epsilon$ and $H := H_t$ satisfies (ii). Replace Z by $\{a \in G : aH \subseteq Z\}$. Then Z is still θ^r-definable, and we still have $\mu(Z) < \epsilon$ and condition (ii). For condition (i), take $D = \bigcup\{gH : gH \cap Z = \emptyset$ and $\mu(gH \cap X) > 0\}$. □

As an application, we now prove structure and regularity for NIP sets in finite groups of *uniformly bounded exponent*. In this case, we obtain the optimal situation where NIP sets are entirely controlled by finite-index subgroups, up to a small error set. This theorem is related to a recent result of Alon, Fox, and Zhao [3], which obtains a similar regularity lemma for *abelian* groups of bounded exponent, in which the index of the subgroup is polynomial in ϵ^{-1}. Our result is stronger in the sense that the abelian assumption is removed and the structural conclusions are improved, but also weaker in the sense that we do not obtain explicit bounds. This is analogous to the comparison of our stable arithmetic regularity lemma in [13] (Theorem 1.1) to the work of the third author and Wolf [43] on stable sets in $\mathbb{F}_p^n$.

Given an arbitrary group G and an integer $k \geq 1$, we say that a set $A \subseteq G$ is **k-NIP** if the collection of left translates $\{gA : g \in G\}$ has VC-dimension at most $k - 1$. Note that this is equivalent to saying that the invariant "formula" $\theta(x; y) := x \in y \cdot A$ is k-NIP in the sense of Definition 2.11.

Theorem 3.3. *For any $k, r \geq 1$ and $\epsilon > 0$, there is $n = n(k, r, \epsilon)$ such that the following holds. Suppose G is a finite group of exponent r, and $A \subseteq G$ is k-NIP. Then there are*

∗ *a normal subgroup $H \leq G$ of index at most n, and*
∗ *a set $Z \subseteq G$, which is a union of cosets of H with $|Z| < \epsilon|G|$,*

satisfying the following properties.

(i) (structure) *There is $D \subseteq G$, which is a union of cosets of H, such that*

$$|(A \backslash Z) \triangle D| < \epsilon|H|.$$

(ii) (regularity) *For any $g \in G \backslash Z$, either $|gH \cap A| < \epsilon|H|$ or $|gH \backslash A| < \epsilon|H|$.*

Moreover, H is in the Boolean algebra generated by $\{gAh : g, h \in G\}$.

Proof. Note that if we have condition (i), then condition (ii) follows immediately. So suppose condition (i) is false. Then we have fixed $k, r \geq 1$ and $\epsilon > 0$ such that, for all $i \in \mathbb{N}$, there is a finite group G_i of exponent r, which is a counterexample. Specifically, there is a k-NIP subset $A_i \subseteq G_i$ such that, if $H \leq G_i$ is normal with index at most i, and $Y, Z \subseteq G_i$ are unions of cosets of H with $|Z| < \epsilon|G_i|$, then $|(A_i \backslash Z) \triangle Y| > \epsilon|H|$.

Let $\mathcal{L}$ be the group language with a new predicate A, and consider (G_i, A_i) as a finite $\mathcal{L}$-structure. Let G be a saturated elementary extension of a nonprincipal ultraproduct of $(G_i, A_i)_{i\in\mathbb{N}}$. Let $\theta(x;y)$ be the formula $x \in y \cdot A$, and note that $\theta(x;y)$ is invariant and k-NIP (in G). Since G has exponent r, $G/G^{00}_{\theta^r}$ is a compact torsion group, and thus is profinite (see [25, Theorem 4.5]). By Theorem 3.2, there is a θ^r-definable finite-index normal subgroup $H \leq G$ and $D, Z \subseteq G$, which are unions of cosets of H, such that $\mu(Z) < \epsilon$ and $\mu((A\backslash Z) \triangle D) = 0$.

Let $n = [G : H]$, and fix θ^r-formulas $\phi(x;\bar{y})$, $\psi(x;\bar{z})$, and $\zeta(x;\bar{u})$ (without parameters) such that H, D, and Z are defined by instances of $\phi(x;\bar{y})$, $\psi(x;\bar{z})$, and $\zeta(x;\bar{u})$, respectively. Given $i \in \mathbb{N}$, let μ_i be the normalized counting measure on G_i. Let I be the set of $i \in \mathbb{N}$ such that, for some tuples $\bar{a}_i$, $\bar{b}_i$, and $\bar{c}_i$ from G_i,

(*i*) $\phi(x;\bar{a}_i)$ defines a normal subgroup H_i of G_i of index n,
(*ii*) $\psi(x;\bar{b}_i)$ and $\zeta(x;\bar{c}_i)$ define sets $D_i, Z_i \subseteq G_i$, respectively, which are each unions of cosets of H_i, and
(*iii*) $\mu_i(Z_i) < \epsilon$ and $\mu_i((A_i\backslash Z_i) \triangle D_i) < \frac{\epsilon}{n}$.

Then $I \in \mathcal{U}$ by elementarity and Łoś's Theorem, and so there is $i \in I$ such that $i \geq n$. This contradicts the choice of (G_i, A_i). □

The previous result is almost identical to Theorem 1.1 on stable subsets of arbitrary finite groups, except for the need for the error set Z. As in [13, Corollary 3.5], we can use this result to deduce a very strong *graph regularity* statement for the Cayley graphs determined by NIP subsets of finite groups of uniformly bounded exponent. Since we work with possibly nonabelian groups, it is more natural to consider bipartite graphs.

Given a bipartite graph $\Gamma = (V, W; E)$, subsets $X \subseteq V$ and $Y \subseteq W$, and vertices $v \in V$ and $w \in W$, define

$$\deg_\Gamma(v, Y) = |\{y \in Y : E(v, y)\}| \quad \text{and} \quad \deg_\Gamma(X, w) = |\{x \in X : E(x, w)\}|.$$

Given $\epsilon > 0$ and nonempty $X \subseteq V$ and $Y \subseteq W$, with $|X| = |Y|$, we say that the pair (X, Y) is *uniformly ϵ-good for* Γ if either:

(*i*) for any $x \in X$ and $y \in Y$, $\deg_\Gamma(x, Y) = \deg_\Gamma(X, y) \leq \epsilon|X|$, or
(*ii*) for any $x \in X$ and $y \in Y$, $\deg_\Gamma(x, Y) = \deg_\Gamma(X, y) \geq (1-\epsilon)|X|$.

It is straightforward to show that if the pair (X, Y) is uniformly ϵ^2-good then it is ϵ-regular in the usual sense, with edge density $\delta_\Gamma(X, Y)$ at most ϵ or at least $1-\epsilon$. In fact, it a stronger property holds: if $X_0 \subseteq X$ and $Y_0 \subseteq Y$ are nonempty and *either* $|X_0| \geq \epsilon|X|$ *or* $|Y_0| \geq \epsilon|Y|$, then $|\delta_\Gamma(X_0, Y_0) - \delta_\Gamma(X, Y)| \leq \epsilon$ (see [13]).

Now suppose G is a finite group and A is a subset of G. The *Cayley graph* $C_G(A)$ is a bipartite graph $(V, W; E)$ where $V = W = G$ and $E(x, y)$ holds if and only if $xy \in A$. Given $X \subseteq G$ and $g \in G$, note that $\deg_{C_G(A)}(g, X) = |A \cap gX|$, and $\deg_{C_G(A)}(X, g) = |A \cap Xg|$. We now observe that Theorem 3.3 implies a graph regularity statement for Cayley graphs of NIP subsets of finite groups of uniformly bounded exponent, in which the partition is given by cosets of a normal subgroup, almost all pairs are uniformly good, and even more pairs are regular. Given a group G, a normal subgroup $H \leq G$, and $C, D \in G/H$, let $C \cdot D$ denote the product of C and D in the quotient group G/H.

Corollary 3.4. *For any $k, r \geq 1$ and $\epsilon > 0$ there is $m = m(k, r, \epsilon)$ such that the following holds. Suppose G is a finite group of exponent r and $A \subseteq G$ is k-NIP. Then there is a normal subgroup H of index $n \leq m$, and set $\mathcal{I} \subseteq G/H$, with $|\mathcal{I}| \leq \epsilon^2 n$, satisfying the following properties.*

(*i*) *If* $\Sigma = \{(C,D) \in (G/H)^2 : C \cdot D \in \mathcal{I}\}$*, then* $|\Sigma| \leq \epsilon^2 n^2$ *and any* $(C,D) \notin \Sigma$ *is uniformly* ϵ^2*-good for* $C_G(A)$*.*

(*ii*) *If* $(C,D) \notin \mathcal{I} \times \mathcal{I}$ *then* (C,D) *is* ϵ*-regular for* $C_G(A)$*, with edge density at most* ϵ *or at least* $1-\epsilon$ *(note that* $|\mathcal{I} \times \mathcal{I}| \leq \epsilon^4 n^2$*).*

Proof. Fix $k, r \geq 1$ and $\epsilon > 0$ and let $m = n(k, r, \epsilon^2)$ be from Theorem 3.3. Fix a finite group G and k-NIP set $A \subseteq G$. By the theorem, there is a normal subgroup $H \leq G$, of index $n \leq m$, and a set $\mathcal{I} \subseteq G/H$, with $|\mathcal{I}| \leq \epsilon^2 n$ such that for any $C \notin \mathcal{I}$, either $|C \cap A| \leq \epsilon^2 |H|$ or $|C \cap A| \geq (1-\epsilon^2)|H|$. It follows that, with $\Sigma \subseteq (G/H)^2$ defined as in (i) above, any (C,D) not in Σ is uniformly ϵ^2-good for $C_G(A)$ (see also [13]). Note that $\Sigma = \bigcup_{C \in G/H} C^{-1} \cdot \mathcal{I}$, and so $|\Sigma| \leq \epsilon^2 n^2$.

Finally, suppose $(C,D) \notin \mathcal{I} \times \mathcal{I}$. Then, for $X \subseteq C$ and $Y \subseteq D$, with $|X| \geq \epsilon |H|$ *and* $|Y| \geq \epsilon |H|$, a straightforward calculation as in [13] shows that the edge density in $C_G(A)$ of the pair (X_0, Y_0) is either at most ϵ or at least $1-\epsilon$. □

Remark 3.5. The assumption of uniformly bounded exponent was used to obtain profinite quotients in ultraproducts, and so it is worth reviewing the actual content of the results above. Specifically, fix $k \geq 1$ and suppose $\mathcal{G}$ is a class of finite groups satisfying the following property. For any sequences $(G_i)_{i=0}^{\infty}$ and $(A_i)_{i=0}^{\infty}$, where $G_i \in \mathcal{G}$ and $A_i \subseteq G_i$ is k-NIP, and for any ultrafilter $\mathcal{U}$ on $\mathbb{N}$, if G is a sufficiently saturated extension of $\prod_{\mathcal{U}}(G_i, A_i)$, then $G/G^{00}_{\theta^r}$ is profinite, where $\theta(x;y) := x \in y \cdot A$. Then, for any $\epsilon > 0$ there is $n = n(k, \epsilon, \mathcal{G})$ such that any group $G \in \mathcal{G}$ and k-NIP set $A \subseteq G$ satisfy the conclusions of Theorem 3.3 using n. Indeed, Theorem 3.3 only uses that, for any $k, r \geq 1$ the class $\mathcal{G}_r$ of finite groups of exponent r satisfies the property above, for the rather heavy-handed reason that compact torsion groups are profinite.

Profinite quotients also arise when $A \subseteq G$ is *k-stable*, which means there do not exist $a_1, \ldots, a_k, b_1, \ldots, b_k \in G$ such that $a_i b_j \in A$ if and only if $i \leq j$ (note that this implies A is k-NIP). In fact, if G is pseudofinite and saturated, and $\theta(x;\bar{y})$ is a stable formula, then the group $G/G^{00}_{\theta^r}$ is actually *finite* (see [12, Corollary 3.17]). Therefore, in this case, the set E_X in Fact 2.16 is empty and so, if one replaces "NIP" with "stable" in Lemma 3.1 and Theorem 3.2, then the error set Z can be chosen to be empty. This yields Theorem 1.1 for stable subsets of groups.[7]

4. Bohr neighborhoods

In this section, we recall some basic definitions and facts concerning Bohr neighborhoods, and then define an approximate version of Bohr neighborhoods, which we will need for later arguments involving ultraproducts.

Given a group G, 1_G denotes the identity (if G is abelian we use 0_G). A *Bohr neighborhood* in a group G is a subset of G of the form $\pi^{-1}(U)$, where $\pi \colon G \to L$ is a homomorphism from G to a compact group L, with dense image, and $U \subseteq L$ is an identity neighborhood (see, e.g., [5]). Under this definition, $\{1_G\}$ is a Bohr neighborhood in any finite group G and so, in the setting of finite groups, one works with a more quantitative formulation (defined below). For our purposes, it

[7] It is worth noting that this explanation of Theorem 1.1 is not a faster proof than what is done in [13]. In particular, [12, Corollary 3.17] relies on the same results from [23] used in [13] to directly prove Theorem 1.1.

will suffice to consider the case that L is compact and metrizable.[8] Recall that the topology on a compact metrizable group L is induced by a bi-invariant metric, which we fix and denote d_L. When working with L^n under the product topology, we assume $d_{L^n}(\bar{x},\bar{y}) = \max_{1\leq i\leq n} d_L(x_i,y_i)$.

In the setting of finite groups, Bohr neighborhoods are usually defined in the case of abelian groups and characters to the torus (e.g. [6],[18]). In order to match these definitions more explicitly, we choose the usual invariant metric on $\mathbb{T}^1 = \mathbb{R}/\mathbb{Z}$, namely $d_{\mathbb{T}^1}(x,y) = \min\{|x-y|, 1-|x-y|\}$.

Definition 4.1. Let H be a group and L a compact metrizable group. Given some $\epsilon > 0$ and a homomorphism $\tau\colon H \to L$, define

$$B^L_{\tau,\epsilon} = \{x \in H : d_L(\tau(x), 1_L) < \epsilon\}.$$

A set $B \subseteq H$ is an (ϵ, L)**-Bohr neighborhood in** H if $B = B^L_{\tau,\epsilon}$ for some homomorphism $\tau\colon H \to L$.

If $L = \mathbb{T}^n$ for some $n \in \mathbb{N}$, then we use the notation $B^n_{\tau,\epsilon}$ for $B^{\mathbb{T}^n}_{\tau,\epsilon}$, and say $B \subseteq H$ is an (ϵ, n)**-Bohr neighborhood** in H if it is an $(\epsilon, \mathbb{T}^n)$-Bohr neighborhood in H.[9]

The next result gives a lower bound on the size of Bohr neighborhoods. The proof is a standard averaging argument (adapted from [42, Lemma 4.20], see also [18, Lemma 4.1]). We include the details for the sake of clarity and to observe that the method works when H is nonabelian and L is any compact metrizable group.

Proposition 4.2. *Let L be a compact metrizable group and, given $\delta > 0$, let ℓ_δ be the Haar measure of $\{t \in L : d_L(t, 1_L) < \delta\}$. For any finite group H and $\delta > 0$, if $B \subseteq H$ is a $(2\delta, L)$-Bohr neighborhood in H, then $|B| \geq \ell_\delta |H|$.*

Proof. Fix a finite group H, a homomorphism $\tau\colon H \to L$, and some $\delta > 0$. Given $x \in H$, let $f_x\colon L \to \{0,1\}$ be the characteristic function of $\{t \in L : d_L(t, \tau(x)) < \delta\}$. Then

$$\ell_\delta |H| = \sum_{x\in H} \int_L f_x \, d\eta = \int_L \sum_{x\in H} f_x \, d\eta,$$

where η is the (normalized) Haar measure on L. So there must be some $t \in L$ such that $\sum_{x\in H} f_x(t) \geq \ell_\delta |H|$, i.e. if $S = \{x \in H : d_L(\tau(x), t) < \delta\}$ then $|S| \geq \ell_\delta |H|$. Fix $a \in S$. For any $x \in S$, we have

$$d_L(\tau(xa^{-1}), 1_L) = d_L(\tau(x), \tau(a)) \leq d_L(\tau(x), t) + d_L(\tau(a), t) < 2\delta.$$

Therefore $Sa^{-1} \subseteq B^L_{\tau,2\delta}$, and so $|B^L_{\tau,2\delta}| \geq |Sa^{-1}| = |S| \geq \ell_\delta |H|$. □

Our ultimate goal is to transfer Bohr neighborhoods in a pseudofinite group G to (ϵ, n)-Bohr neighborhoods in finite groups. More generally, we wish to approximate arbitrary Bohr neighborhoods in expansions of groups by definable objects. This necessitates an approximate notion of Bohr neighborhood, which involves approximate homomorphisms of groups.

Definition 4.3. Let H be a group and L a compact metrizable group.

(1) Given $\delta > 0$, function $f\colon H \to L$ is a δ**-homomorphism** if $f(1_H) = 1_L$ and, for all $x, y \in H$, $d_L(f(xy), f(x)f(y)) < \delta$.

[8] In fact, we will ultimately only be concerned with compact connected Lie groups, and the only reason we consider groups other than $\mathbb{T}^n$ is for the work in Section 6.

[9] In this case n is sometimes referred to as the *dimension* of B, and ϵ is the *width* of B.

(2) Given $\epsilon, \delta > 0$, a set $Y \subseteq H$ is a **δ-approximate (ϵ, L)-Bohr set** if there is a δ-homomorphism $f\colon H \to L$ such that $Y = \{x \in H : d_L(f(x), 1_L) < \epsilon\}$.

Approximate homomorphisms have been studied extensively in the literature, with a large focus on the question of when an approximate homomorphism is "close" to an actual homomorphism. For our purposes, this is what is needed to replace approximate Bohr neighborhoods with actual Bohr neighborhoods. More precisely, we will start with a definable approximate Bohr neighborhood in G, and transfer this to find an approximate Bohr neighborhood in a finite group H. At this point, we will be working with an approximate homomorphism from H to a compact Lie group, which is a setting where one can always find a Bohr neighborhood inside of an approximate Bohr neighborhood, with only a negligible loss in size.

Fact 4.4 (Alekseev, Glebskiĭ, & Gordon [1, Theorem 5.13]). *Let L be a compact Lie group. Then there is an $\alpha_L > 0$ such that, for any $0 < \delta < \alpha_L$, if H is a compact group and $f\colon H \to L$ is a δ-homomorphism, then there is a homomorphism $\tau\colon H \to L$ such that $d_L(f(x), \tau(x)) < 2\delta$ for all $x \in H$.*

An easy consequence is that, in the setting of compact Lie groups, Bohr neighborhoods can be found inside of approximate Bohr neighborhoods.

Corollary 4.5. *Let L be a compact Lie group. Then there is an $\alpha_L > 0$ such that, if H is a compact group, $n \in \mathbb{N}$, and $0 < \delta < \alpha_L$, then every δ-approximate $(3\delta, L^n)$-Bohr neighborhood in H contains a (δ, L^n)-Bohr neighborhood in H.*

Proof. Fix $\alpha_L > 0$ from Fact 4.4. Suppose H is a compact group, and $Y \subseteq H$ is a δ-approximate $(3\delta, L^n)$-Bohr neighborhood in H, for some $n \in \mathbb{N}$ and $0 < \delta < \alpha_L$, witnessed by a δ-homomorphism $f\colon H \to L^n$. We may assume $n \geq 1$. For $1 \leq i \leq n$, let $f_i\colon H \to L$ be given by $f_i(x) = f(x)_i$. Then each f_i is a δ-homomorphism. Given $1 \leq i \leq n$, Fact 4.4 provides a homomorphism $\tau_i\colon H \to L$ such that $d_L(f_i(x), \tau_i(x)) < 2\delta$ for all $x \in H$. Let $\tau\colon H \to L^n$ be such that $\tau(x) = (\tau_1(x), \ldots, \tau_n(x))$. Then τ is a homomorphism and $d_{L^n}(f(x), \tau(x)) < 2\delta$ for all $x \in H$. Now we have $B^{L^n}_{\tau,\delta} \subseteq Y$ by the triangle inequality. □

Finally, we will need the following minor generalization of a standard and well-known exercise, namely, if G is an amenable group and $A \subseteq G$ has positive upper density then AA^{-1} is generic (or *syndetic*).

Proposition 4.6. *Suppose G is a group, $\mathcal{B}$ is a Boolean algebra of subsets of G, and ν is a left-invariant, finitely additive probability measure on $\mathcal{B}$. Suppose $A \in \mathcal{B}$ is such that $\nu(A) > 0$. Then, for any $Z \subseteq G$, there is a finite set $F \subseteq G \backslash Z$ such that $|F| \leq \frac{1}{\nu(A)}$ and $G \backslash Z \subseteq FAA^{-1}$.*

Proof. We say that $X \subseteq G$ *separates* A if $xA \cap yA = \emptyset$ for all distinct $x, y \in X$. By the assumptions on ν, if $X \subseteq G$ separates A then $|X| \leq \frac{1}{\nu(A)}$. Choose a finite set $F \subseteq G \backslash Z$ with maximal size among subsets of $G \backslash Z$ that separate A. Fix $x \in G \backslash Z$. Then there is $y \in F$ such that $xA \cap yA \neq \emptyset$, and so we may fix $z \in xA \cap yA$. Then $y^{-1}z \in A$ and $z^{-1}x \in A^{-1}$, which means $y^{-1}x \in AA^{-1}$, and so $x \in FAA^{-1}$. □

5. Structure and regularity: the general case

The next goal is a result analogous to Theorem 3.2, but without the assumption that $G/G^{00}_{\theta^r}$ is profinite. For this, we need to understand more about families $\mathcal{W}$ of definable sets such that $G^{00}_{\theta^r} = \bigcap \mathcal{W}$. The goal is to find properties of the sets in $\mathcal{W}$ which are both interesting algebraically, and are also sufficiently first-order so that they can be transferred down to to finite groups in arguments with ultraproducts. In particular, we will use Bohr neighborhoods.

We again let G denote a saturated expansion of a group (in the language $\mathcal{L}$), as in Section 2.1. The next result shows that type-definable bounded-index subgroups in G can be approximated by definable sets containing Bohr neighborhoods.

Proposition 5.1. *Let $\theta(x;\bar{y})$ be invariant, and suppose Γ is a θ^r-countably-definable bounded-index normal subgroup of G. Then there is a decreasing sequence $(W_i)_{i=0}^{\infty}$ of θ^r-definable subsets of G such that $\Gamma = \bigcap_{i=0}^{\infty} W_i$ and, for all $i \in \mathbb{N}$, there are*

* *a θ^r-definable finite-index normal subgroup $H_i \leq G$,*
* *a θ^r-definable homomorphism $\pi_i \colon H_i \to L_i$, where L_i is a compact connected Lie group, and*
* *a real number $\epsilon_i > 0$,*

such that $B^{L_i}_{\pi_i,\epsilon_i} \subseteq W_i \subseteq H_i$. Moreover:

(i) if G is pseudofinite then, for each $i \in \mathbb{N}$, we may assume $L_i = \mathbb{T}^{n_i}$ for some $n_i \in \mathbb{N}$;

(ii) if G/Γ is abelian then, for each $i \in \mathbb{N}$, we may assume $H_i = G$ and $L_i = \mathbb{T}^{n_i}$ for some $n_i \in \mathbb{N}$.

Proof. Let $(\Gamma_t)_{t=0}^{\infty}$ and $(H_t)_{t=0}^{\infty}$ be as in Lemma 2.6. For each $t \in \mathbb{N}$, let $L_t = H_t/\Gamma_t$ and let $\pi_t : H_t \to L_t$ be the projection map. Then π_t is θ^r-definable (as $G \to G/\Gamma_t$ is θ^r-definable).

Let $\mathcal{W}_t$ be a countable collection of θ^r-definable subsets of G such that $\Gamma_t = \bigcap \mathcal{W}_t$. Since $\Gamma_t \subseteq H_t$, and H_t is θ^r-definable, we may assume without loss of generality that $\Gamma_t \subseteq W \subseteq H_t$ for all $W \in \mathcal{W}_t$. Let $\mathcal{W} = \bigcup_{t\in\mathbb{N}} \mathcal{W}$ and let $\mathcal{W}^*$ be the collection of all finite intersections of elements of $\mathcal{W}$. By construction, and since $(H_t)_{t=0}^{\infty}$ and $(\Gamma_t)_{t=0}^{\infty}$ are decreasing sequences, we have that for all $W \in \mathcal{W}^*$, there is some $t \in \mathbb{N}$ such that $\Gamma_t \subseteq W \subseteq H_t$. Since $\mathcal{W}^*$ is countable and closed under finite intersections, we may assume $\mathcal{W}^* = \{W_i : i \in \mathbb{N}\}$ with $W_{i+1} \subseteq W_i$ for all $i \in \mathbb{N}$. Note that $\Gamma = \bigcap_{i=0}^{\infty} W_i$.

Finally, fix $i \in \mathbb{N}$ and let $t_i \in \mathbb{N}$ be such that $\Gamma_{t_i} \subseteq W_i \subseteq H_{t_i}$. The set $U = \{a\Gamma_{t_i} \in G/\Gamma_{t_i} : a\Gamma_{t_i} \subseteq W_i\}$ is an identity neighborhood in L_{t_i} (by Fact 2.16 and Remark 2.3(2)) and, by construction, $\pi^{-1}_{t_i}(U) \subseteq W_i$. So choose $\epsilon_i > 0$ such that U contains the open ball of radius ϵ_i around 1_{L_i}.

For the moreover statements, note first that if G is pseudofinite then, by Corollary 2.10, we may assume $L_t = \mathbb{T}^{n_t}$ for some $n_t \in \mathbb{N}$. On the other hand, if G/Γ is abelian, then G/Γ_t is a compact abelian Lie group, and thus isomorphically embeds in $\mathbb{T}^{n_t}$ for some $n_t \in \mathbb{N}$ (see, e.g., [8, Corollary 3.7]). So we use the same argument as above, but replace H_t and π_t with G and $G \to G/\Gamma_t \subseteq \mathbb{T}^{n_t}$. □

One drawback of the previous result is that the Bohr neighborhood $B^{L_i}_{\pi_i,\epsilon_i}$ is not necessarily definable (it is only co-type-definable). In order to work with definable objects, we will have to consider approximate Bohr neighborhoods.

Definition 5.2. Let $\theta(x;\bar{y})$ be a formula, and suppose $H \leq G$ is θ-definable. Fix a compact metrizable group L.

(1) Given $\delta, \epsilon > 0$, we say that $Y \subseteq H$ is a **θ-definable δ-approximate (ϵ, L)-Bohr neighborhood in** H if there is a θ-definable δ-homomorphism $f\colon H \to L$ such that $f(H)$ is finite and $Y = \{x \in H : d_L(f(x), 1_L) < \epsilon\}$.
(2) Given an integer $t \geq 1$, we say that a sequence $(Y_m)_{m=0}^{\infty}$ of subsets of H is a **(θ, t, L)-approximate Bohr chain in** H if $Y_{m+1} \subseteq Y_m$, for all $m \in \mathbb{N}$, and there is a decreasing sequence $(\delta_m)_{m=0}^{\infty}$ of positive real numbers converging to 0 such that, for all $m \in \mathbb{N}$, Y_m is a θ-definable δ_m-approximate $(\delta_m t, L)$-Bohr neighborhood in H.

The parameter t in the previous definition is introduced in order to control the "radius" ϵ and the "error" δ in a δ-approximate (ϵ, L)-Bohr neighborhood. Specifically, it is desirable to have ϵ be some constant multiple t of δ, and in the following results we will choose t arbitrarily. This will eventually be used to find actual Bohr neighborhoods inside of approximate Bohr neighborhoods, with $L = \mathbb{T}^n$ for some $n \in \mathbb{N}$, in which case setting $t = 3$ will suffice (via Corollary 4.5).

Proposition 5.3. *Suppose $H \leq G$ is θ-definable, L is a compact metrizable group, and $Y \subseteq H$ is a θ-definable δ-approximate (ϵ, L)-Bohr neighborhood in H, for some $\epsilon, \delta > 0$. Then Y is θ-definable.*

Proof. By definition, we have a θ-definable map $f\colon H \to L$ with $f(H)$ *finite*. So H is partitioned into θ-type-definable fibers $f^{-1}(\lambda)$ for $\lambda \in f(H)$. Since $f(H)$ is finite and H is θ-definable, the complement (in H) of each fiber is also θ-type-definable, and so all fibers are θ-definable by compactness. Now Y is a union of the fibers $f^{-1}(\lambda)$ over all $\lambda \in f(L)$ such that $d(\lambda, 1_L) < \epsilon$. So Y is θ-definable. $\square$

Lemma 5.4. *Fix a formula $\theta(x; \bar{y})$, and let H be a θ-definable subgroup of G. Suppose there is a θ-definable homomorphism $\pi\colon H \to L$ for some compact metrizable group L. Then, for any integer $t \geq 1$, there is a (θ, t, L)-approximate Bohr chain $(Y_m)_{m=0}^{\infty}$ in H such that* $\ker \pi = \bigcap_{m=0}^{\infty} Y_m$.

Proof. Let $d = d_L$. Given $\lambda \in L$ and $\epsilon > 0$, let $K(\lambda, \epsilon) \subseteq L$ and $U(\lambda, \epsilon) \subseteq L$ be the closed ball of radius ϵ around λ and the open ball of radius ϵ around λ, respectively.

Fix $m \geq 1$. Let $\epsilon = \frac{1}{m}$, and choose $\Lambda \subseteq L$ finite such that $L = \bigcup_{\lambda \in \Lambda} K(\lambda, \frac{\epsilon}{2})$ and $1_L \in \Lambda$. For any $\lambda \in \Lambda$, since π is θ-definable, there is a θ-definable set $D_\lambda \subseteq H$ such that $\pi^{-1}(K(\lambda, \frac{\epsilon}{2})) \subseteq D_\lambda \subseteq \pi^{-1}(U(\lambda, \epsilon))$ (see Remark 2.5). Enumerate $\Lambda = \{\lambda_1, \ldots, \lambda_k\}$, with $\lambda_1 = 1_L$, and, for each $1 \leq i \leq k$, let $D_i = D_{\lambda_i}$. Define inductively $E_1 = D_1$ and $E_{i+1} = D_{i+1} \backslash (E_1 \cup \ldots \cup E_i)$. Then $E_1, \ldots, E_k$ partition H into θ-definable sets. This determines a θ-definable function $f_m\colon H \to \Lambda$ such that $f_m(x) = \lambda_i$ if and only if $x \in E_i$. For any $x \in H$, we have $x \in D_{f_m(x)} \subseteq \pi^{-1}(U(f_m(x), \epsilon))$, and so $d(\pi(x), f_m(x)) < \epsilon$. Note that $f_m(1_H) = 1_L$ by definition. Also, given $x, y \in H$, we have

$$\begin{aligned} d(f_m(xy), f_m(x)f_m(y)) &\leq d(f_m(xy), \pi(xy)) + d(\pi(x)\pi(y), f_m(x)\pi(y)) \\ &\qquad + d(f_m(x)\pi(y), f_m(x)f_m(y)) \\ &= d(f_m(xy), \pi(xy)) + d(\pi(x), f_m(x)) + d(\pi(y), f_m(y)) \\ &< 3\epsilon. \end{aligned}$$

Altogether, $f_m\colon H \to L$ is a θ-definable 3ϵ-homomorphism.

Now fix an integer $t \geq 1$. For $m \in \mathbb{N}$, define

$$Y_m = \{x \in H : d(f_m(x), 1_L) < 3\epsilon t\}.$$

Note that $D_1 \subseteq Y_m$, and so $\ker\pi \subseteq \pi^{-1}(K(1_L, \frac{1}{2m})) \subseteq Y_m$. We now have a sequence $(Y_m)_{m=1}^{\infty}$ of θ-definable subsets of H, with $\ker\pi \subseteq Y_m$ for all $m \in \mathbb{N}$. Note also that if $\delta_m = \frac{3}{m}$, then Y_m is a δ_m-approximate $(\delta_m t, L)$-Bohr neighborhood in H, witnessed by $f_m : H \to L$. Moreover, for any $m \in \mathbb{N}$, if $x \in Y_m$ then

$$d(\pi(x), 1_L) \leq d(\pi(x), f_m(x)) + d(f_m(x), 1_L) < \tfrac{3t+1}{m}.$$

This implies $\ker\pi = \bigcap_{m=0}^{\infty} Y_m$. Finally, given $m \in \mathbb{N}$, we have $\pi^{-1}(K(1_L, \frac{1}{2m})) \subseteq Y_m \subseteq \pi^{-1}(U(1_L, \frac{3t+1}{m}))$. In particular, if $n \geq (6t+2)m$, then $Y_n \subseteq Y_m$. So, after thinning the sequence, we may assume $Y_{m+1} \subseteq Y_m$ for all $m \in \mathbb{N}$. □

We now combine the ingredients above to prove a structure and regularity theorem for θ^r-definable sets in G, in the case that G is pseudofinite and $\theta(x;\bar{y})$ is an arbitrary invariant NIP formula.

Theorem 5.5. *Assume G is pseudofinite. Let $\theta(x;\bar{y})$ be invariant and NIP. Fix a θ^r-definable set $X \subseteq G$ and some $\epsilon > 0$. Then there are*

- *a θ^r-definable finite-index normal subgroup $H \leq G$,*
- *a θ^r-definable homomorphism $\pi\colon H \to \mathbb{T}^n$, for some $n \in \mathbb{N}$, and*
- *a θ^r-definable set $Z \subseteq G$, with $\mu(Z) < \epsilon$,*

such that, for any integer $t \geq 1$, there is

- *a (θ^r, t, n)-approximate Bohr chain $(Y_m)_{m=0}^{\infty}$ in H, with $\ker\pi = \bigcap_{m=0}^{\infty} Y_m$,*

satisfying the following properties.

(*i*) (structure) *For any $m \in \mathbb{N}$, there is $D \subseteq G$, which is a finite union of translates of Y_m, such that*

$$\mu((X \triangle D)\backslash Z) = 0.$$

(*ii*) (regularity) *For any $m \in \mathbb{N}$ and any $g \in G\backslash Z$, either $\mu(gY_m \cap X) = 0$ or $\mu(gY_m\backslash X) = 0$.*

Moreover, if $G/G^{00}_{\theta^r}$ is abelian then we may assume $H = G$.

Proof. Let $X \subseteq G$ and $\epsilon > 0$ be fixed, and let $\{W_i : i \in \mathbb{N}\}$ be a collection of θ^r-definable sets in G satisfying the conditions of Proposition 5.1, with $\Gamma = G^{00}_{\theta^r}$. By Lemma 3.1, there is θ^r-definable $Z \subseteq G$ and some $i \in \mathbb{N}$ such that $\mu(Z) < \epsilon$ and, if $W := W_i$, then for all $g \in G\backslash Z$, we have $\mu(gW \cap X) = 0$ or $\mu(gW\backslash X) = 0$. Proposition 5.1 associates to W a θ^r-definable homomorphism $\pi\colon H \to \mathbb{T}^n$, where $H \leq G$ is θ^r-definable of finite-index. If $G/G^{00}_{\theta^r}$ is abelian then we may further assume $H = G$. Fix $t \geq 1$. By Lemma 5.4, there is a (θ, t, n)-approximate Bohr chain $(Y_m)_{m=0}^{\infty}$ in H such that $\ker\pi = \bigcap_{m=0}^{\infty} Y_m$. Since $\ker\pi$ is θ^r-type-definable and contained in the θ^r-definable set W, it follows from compactness that $Y_m \subseteq W$ for sufficiently large m. So for sufficiently large m we have that, for any $g \in G\backslash Z$, either $\mu(gY_m \cap X) = 0$ or $\mu(gY_m\backslash X) = 0$. So, after removing some finite initial segment of $(Y_m)_{m=0}^{\infty}$, we have condition (*ii*).

Toward proving condition (*i*), fix $m \in \mathbb{N}$. Since $\ker\pi$ is a group, we may use compactness (similar to as in the proof of Lemma 3.1), to find some $r \geq m$ such that $Y_r Y_r^{-1} \subseteq Y_m$. Since $\ker\pi$ has bounded index, Y_r is generic and so $\mu(Y_r) > 0$. By Proposition 4.6, there is a finite set $F \subseteq G\backslash Z$ such that $G\backslash Z \subseteq FY_m$. Let $I = \{g \in F : \mu(gY_m\backslash X) = 0\}$, and note that if $g \in F\backslash I$ then $\mu(gY_m \cap X) = 0$. Let

$D = \bigcup_{g\in I} gY_m$. Since $G\backslash Z \subseteq FY_m$, we have

$$X \triangle D \subseteq Z \cup \bigcup_{g\in I} gY_m\backslash X \cup \bigcup_{g\in F\backslash I} gY_m \cap X,$$

and so $\mu((X \triangle D)\backslash Z) = 0$. □

In analogy to how we obtained Theorem 3.3 from Theorem 3.2, we now use Theorem 5.5 to prove the main structure and regularity theorem for NIP sets in arbitrary finite groups. Roughly speaking, the result states that if A is a k-NIP set in a finite group G, then there is a normal subgroup $H \leq G$, and Bohr neighborhood B in H, such that almost all translates of B are almost contained in A or almost disjoint from A. Moreover, A is approximately a union of translates of B, and the index of H and complexity of B are bounded in terms of k and ϵ. We first prove a lemma, which gives a rather flexible version of the regularity aspect, and also contains the ultraproducts argument necessary to prove Theorem 5.7.

Lemma 5.6. *For any $k \geq 1$ and $\epsilon > 0$, and any function $\gamma\colon (\mathbb{Z}^+)^2 \times (0,1] \to \mathbb{R}^+$, there is $n = n(k,\epsilon,\gamma)$ such that the following holds. Suppose G is a finite group and $A \subseteq G$ is k-NIP. Then there are*

- ∗ *a normal subgroup $H \leq G$ of index $m \leq n$,*
- ∗ *a (δ, r)-Bohr neighborhood B in H and a δ-approximate $(3\delta, r)$-Bohr neighborhood Y in H, for some $r \leq n$ and $\delta \geq \frac{1}{n}$, and*
- ∗ *a set $Z \subseteq G$, with $|Z| < \epsilon|G|$,*

such that $B \subseteq Y \subseteq H$ and, for any $g \in G\backslash Z$, either $|gY \cap A| < \gamma(m,r,\delta)|B|$ or $|gY\backslash A| < \gamma(m,r,\delta)|B|$. Moreover, H, Y, and Z are in the Boolean algebra generated by $\{gAh : g,h \in G\}$, and if G is abelian then we may assume $H = G$.

Proof. Suppose not. Then we have $k \geq 1$, $\epsilon > 0$ and a fixed function $\gamma\colon (\mathbb{Z}^+)^2 \times (0,1] \to \mathbb{R}^+$ witnessing this. In particular, for any $n \in \mathbb{N}$, there is a finite group G_n and a k-NIP subset $A_n \subseteq G_n$ such that, for any $H, B, Y, Z \subseteq G_n$, if

- ∗ H is a normal subgroup of G_n of index $m \leq n$, and $H = G_n$ if G_n is abelian,
- ∗ B is a (δ, r)-Bohr neighborhood in H and Y is a δ-approximate $(3\delta, r)$-Bohr neighborhood in H, for some $r \leq n$ and $\delta \geq \frac{1}{n}$,
- ∗ $|Z| < \epsilon|G_n|$, and $B \subseteq Y \subseteq H$,

then there is $g \in G_n\backslash Z$ such that $|gY \cap A_n| \geq \gamma(m,r,\delta)|B|$ and $|gY\backslash A_n| \geq \gamma(m,r,\delta)|B|$.

Let $\mathcal{L}$ be the group language together with an extra predicate A, and consider each (G_n, A_n) as a finite $\mathcal{L}$-structure. Let G be a sufficiently saturated elementary extension of a nonprincipal ultraproduct of $(G_n, A_n)_{n\in\mathbb{N}}$. Let $\theta(x;y)$ be the formula $x \in y \cdot A$, and note that $\theta(x;y)$ is invariant and k-NIP (in G). Finally, let $\alpha := \alpha_{\mathbb{T}^1} > 0$ be as in Fact 4.4. By Theorem 5.5 (with $t = 3$), there are θ^r-definable $Y, Z \subseteq G$ and a θ^r-definable finite-index normal subgroup $H \leq G$ such that:

- ∗ if G is abelian then $H = G$,
- ∗ $\mu(Z) < \epsilon$,
- ∗ Y is a θ^r-definable δ-approximate $(3\delta, r)$-Bohr neighborhood in H for some $r \in \mathbb{N}$ and $0 < \delta < \alpha$, and
- ∗ for any $g \in G\backslash Z$, either $\mu(gY \cap A) = 0$ or $\mu(gY\backslash A) = 0$.

Let $m = [G : H]$, and set $\epsilon^* = \gamma(m,r,\delta)m^{-1}(\frac{\delta}{2})^r > 0$. Let $f\colon H \to \mathbb{T}^r$ be a θ^r-definable δ-homomorphism witnessing that Y is a θ^r-definable δ-approximate

$(3\delta, r)$-Bohr neighborhood in H. Let $\Lambda = f(H)$, and note that Λ is finite. Given $\lambda \in \Lambda$, let $F(\lambda) = f^{-1}(\lambda)$. Then each $F(\lambda)$ is θ^r-definable (as in Proposition 5.3).

Fix θ^r-formulas $\phi(x;\bar{y})$, $\psi(x;\bar{z})$, $\zeta(x;\bar{u})$, and $\xi_\lambda(x;\bar{v}_\lambda)$, for $\lambda \in \Lambda$, (without parameters) such that H, Y, Z, and $F(\lambda)$, for $\lambda \in \Lambda$, are defined by instances of $\psi(x;\bar{z})$, $\zeta(x;\bar{u})$, and $\xi_\lambda(x;\bar{v}_\lambda)$, respectively. Given $n \in \mathbb{N}$, let μ_n denote the normalized counting measure on G_n. Let d denote $d_{\mathbb{T}^r}$. Define $I \subseteq \mathbb{N}$ to be the set of $n \in \mathbb{N}$ such that, for some tuples $\bar{a}_n$, $\bar{b}_n$, $\bar{c}_n$, and $\bar{d}_{n,\lambda}$, for $\lambda \in \Lambda$,

- (i) $\phi(x;\bar{a}_n)$ defines a normal subgroup H_n of G_n of index m,
- (ii) $\zeta(x;\bar{c}_n)$ defines a subset Z_n with $\mu_n(Z_n) < \epsilon$,
- (iii) for each $\lambda \in \Lambda$, $\xi_\lambda(x;\bar{d}_{n,\lambda})$ defines a subset $F_n(\lambda)$ of H_n,
- (iv) if $f_n \colon H_n \to \Lambda$ is defined so that $f_n(x) = \lambda$ if and only if $x \in F_n(\lambda)$, then f_n is a δ-homomorphism,
- (v) $\psi(x;\bar{b}_n)$ defines the set $Y_n = \{x \in H_n : d(f_n(x), 0) < 3\delta\}$, and
- (vi) for all $g \in G_n \backslash Z_n$, either $\mu_n(gY_n \cap A_n) < \epsilon^*$ or $\mu_n(gY_n \backslash A_n) < \epsilon^*$.

We claim that, by Łoś's Theorem an elementarity, $I \in \mathcal{U}$. In other words, the claim is that conditions (i) through (vi) are first-order expressible (possibly using the expanded measure language discussed before Fact 2.14). This is clear for (i), (ii), (iii), (v), and (vi), and the only subtleties lie in (iv) and (v). In both cases, the crucial point is that Λ is finite, and so these conditions can be described by finite first-order sentences (using similar ideas as in Proposition 5.3). For instance, to express condition (iv), fix some $\lambda \in \Lambda$, and let $P_\lambda = \{(\lambda_1, \lambda_2) \in \Lambda^2 : d(\lambda, \lambda_1 + \lambda_2) < \delta\}$. Let σ_λ be a sentence expressing that, for any x, y, if $x \cdot y \in F(\lambda)$ then $x \in F(\lambda_1)$ and $y \in F(\lambda_2)$ for some $(\lambda_1, \lambda_2) \in P(\lambda)$. Then the conjunction of all such σ_λ, for all $\lambda \in \Lambda$, expresses precisely that f is a δ-homomorphism. Altogether, by Łoś's theorem and elementarity, the set of n for which f_n satisfies condition (iv) is in $\mathcal{U}$. The details for condition (v) are similar and left to the reader.

Since $I \in \mathcal{U}$, we may fix $n \in I$ such that $n \geq \max\{m, r, \delta^{-1}\}$. Since $0 < \delta < \alpha$ and Y_n is a δ-approximate $(3\delta, r)$-Bohr neighborhood in H_n, it follows from Corollary 4.5 that Y_n contains a (δ, r)-Bohr neighborhood B in H_n. So, by choice of (G_n, A_n), there must be $g \in G_n \backslash Z_n$ such that $\mu_n(gY_n \cap A_n) \geq \gamma(m, r, \delta)\mu_n(B)$ and $\mu_n(gY_n \backslash A_n) \geq \gamma(m, r, \delta)\mu_n(B)$. So, to obtain a contradiction, it suffices to show that $\epsilon^* \leq \gamma(m, r, \delta)\mu_n(B)$, i.e. (by choice of ϵ^*), show $m^{-1}(\frac{\delta}{2})^r \leq \mu_n(B)$. To see this, note that $|B| \geq (\frac{\delta}{2})^r |H_n| \geq (\frac{\delta}{2})^r m^{-1} |G_n|$, since $[G : H_n] = m$ and by Proposition 4.2 (applied with $L = \mathbb{T}^r$, and so $\ell_{\frac{\delta}{2}} = (\frac{\delta}{2})^r$). □

We now prove the main result for NIP subsets of arbitrary finite groups.

Theorem 5.7. *For any $k \geq 1$ and $\epsilon > 0$ there is $n = n(k, \epsilon)$ such that the following holds. Suppose G is a finite group and $A \subseteq G$ is k-NIP. Then there are*

- *a normal subgroup $H \leq G$ of index at most n,*
- *a (δ, r)-Bohr neighborhood B in H, for some $r \leq n$ and $\delta \geq \frac{1}{n}$, and*
- *a subset $Z \subseteq G$, with $|Z| < \epsilon |G|$,*

satisfying the following properties.

(i) (structure) *There is $D \subseteq G$, which is a union of translates of B, such that*

$$|(A \triangle D) \backslash Z| < \epsilon |B|.$$

(ii) (regularity) *For any $g \in G \backslash Z$, either $|gB \cap A| < \epsilon |B|$ or $|gB \backslash A| < \epsilon |B|$.*

Moreover, H and Z are in the Boolean algebra generated by $\{gAh : g, h \in G\}$, and if G is abelian then we may assume $H = G$.

Proof. Fix $k \geq 1$ and $\epsilon > 0$. Define $\gamma\colon (\mathbb{Z}^+)^2 \times (0,1] \to \mathbb{R}^+$ such that $\gamma(x,y,z) = \epsilon x^{-1}(\frac{z}{4})^y$. Let $n = n(k,\epsilon,\gamma)$ be given by Lemma 5.6. Fix a finite group G and a k-NIP subset $A \subseteq G$. By Lemma 5.6, there are

- a normal subgroup $H \leq G$ of index $m \leq n$,
- a subset $Y \subseteq H$,
- a (δ, r)-Bohr neighborhood B in H, for some $r \leq n$ and $\delta \geq \frac{1}{n}$, and
- a set $Z \subseteq G$, with $|Z| < \epsilon|G|$,

such that, $B \subseteq Y \subseteq G$ and, for all $g \in G\backslash Z$, either $|gY \cap A| < \gamma(m,r,\delta)|B|$ or $|gY \backslash A| < \gamma(m,r,\delta)|B|$. Moreover, if G is abelian then we may assume $H = G$. Since $B \subseteq Y$ and $\gamma(m,r,\delta) \leq \epsilon$, this immediately yields condition (ii).

Say $B = B^r_{\tau,\delta}$ for some homomorphism $\tau : H \to \mathbb{T}^r$. Let $B_0 = B^r_{\tau,\delta/2}$. By Proposition 4.2, and since $[G : H] = m$, we have

$$|B_0| \geq \left(\tfrac{\delta}{4}\right)^r |H| = m^{-1} \left(\tfrac{\delta}{4}\right)^r |G| = \epsilon^{-1}\gamma(m,\delta,r)|G|.$$

So, if ν denotes the normalized counting measure on G, then we have $\nu(B_0) \geq \epsilon^{-1}\gamma(m,\delta,r) > 0$. Note also that $B_0B_0^{-1} \subseteq B$ since, if $x, y \in B_0$ then

$$d(\tau(xy^{-1}), 0_{\mathbb{T}^r}) = d(\tau(x), \tau(y)) < d(\tau(x), 0_{\mathbb{T}^r}) + d(\tau(y), 0_{\mathbb{T}^r}) < \delta.$$

Altogether, by Proposition 4.6, we may fix $x_1, \ldots, x_t \in G\backslash Z$, such that $t \leq \epsilon(\gamma(m,\delta,r))^{-1}$ and $G\backslash Z \subseteq \bigcup_{i=1}^t x_iB$.

Let $I = \{i \in [t] : |x_iB\backslash A| < \gamma(m,r,\delta)|B|\}$, and note that if $i \notin I$ then $|x_iB \cap A| < \gamma(m,r,\delta)|B|$. Let $D = \bigcup_{i\in I} x_iB$. Since $G\backslash Z \subseteq \bigcup_{i=1}^t x_iB$, we have

$$A \triangle Y \subseteq Z \cup \bigcup_{i\in I} x_iB\backslash A \cup \bigcup_{i\notin I} x_iB \cap A, \text{ and so}$$

$$|(A \triangle Y)\backslash Z| \leq \sum_{i\in I} |x_iB\backslash A| + \sum_{i\notin I} |x_iB\backslash A| < t\gamma(m,r,\delta)|B| \leq \epsilon|B|.$$ □

Remark 5.8. Note that Theorem 3.2 can be derived a special case of Theorem 5.5, which then gives Theorem 3.3 as a special case of Theorem 5.7. Indeed, in the setting of Theorem 5.5, if $G/G^{00}_{\theta^r}$ is profinite then so is $H/\ker\pi \cong \mathbb{T}^n$, which implies $n = 0$ and so $H = \ker\pi$. Therefore, in each (θ^r, t, n)-approximate Bohr chain, we have $Y_m = H$ for all $m \in \mathbb{N}$. On the other hand, the proofs of Theorems 3.2 and 3.3 given above are much more direct, and entirely independent from the machinery involving Bohr neighborhoods.

6. Distal regularity and compact p-adic Lie groups

In this section, we briefly adapt the preceding results to the case of NIP *fsg* groups with smooth left-invariant measures (e.g. those definable in distal theories). In contrast to previous results, where we focused on a single NIP formula $\theta(x;\bar{y})$, here we will operate in the setting of an *fsg* group G definable in (a saturated model of) an NIP theory, and prove structure and regularity theorems for definable sets in G. The conclusions for suitable families of finite groups will be routine in light of the work above, but we will mention explicitly an application to the family of quotients of a compact p-adic analytic group H by its open normal subgroups.

For the rest of this section, T is a complete theory and we work in a saturated model M. Given a model $M_0 \preceq M$ and an M_0-definable set X in M, a *Keisler measure over* M_0 is a finitely additive probability measure defined on the Boolean algebra of M_0-definable subsets of X. In the case that $M = M_0$, we call such a measure a *global Keisler measure on* X. Recall that, for T NIP, a global Keisler measure μ on a definable set X is *generically stable* over a small model M_0 if it is both definable over M_0 and finitely satisfiable in M_0 (see [36, Section 7.5]).

In the previous sections, we focused on pseudofinite groups because, in addition to being the right setting for applications to finite groups, the local NIP group theory developed in [12] is for pseudofinite groups. However, as further discussed in Remark 6.8 below, the crucial property of the pseudofinite counting measure is that it is generically stable with respect to certain NIP formulas. So we recall the following notion (introduced in [24]).

Definition 6.1. (T is NIP.) A definable group G is ***fsg*** if it admits a generically stable left-invariant Keisler measure.

We should remark that this is not the original definition of *fsg*, but rather the right characterization for our purposes (see [36, Proposition 8.32]). As indicated above, we have the following example.

Example 6.2. If T is NIP and G is a definable pseudofinite group, then G is *fsg*, namely, the pseudofinite counting measure on G is generically stable. This follows directly from the VC-Theorem (see, e.g., [36, Example 7.32], [12, Section 2]).

The *fsg* property for a definable group has strong consequences, for instance such a group has a unique left-invariant Keisler measure, which is also the unique right-invariant Keisler measure. Moreover, genericity and positive measure coincide (as in Fact 2.14). Such groups also satisfy a generic compact domination statement similar to Fact 2.16 (see [24, 37]), which can be further strengthened in the case that the unique left-invariant measure is *smooth*.

Definition 6.3. Let M_0 be a small model and fix an M_0-definable set X in M. A Keisler measure on X over M_0 is **smooth** if it has a unique extension to a global Keisler measure on X. A global Keisler measure μ on X is **smooth over** M_0 if $\mu|_{M_0}$ is smooth (in which case μ is the unique global extension of $\mu|_{M_0}$).

Suppose T is NIP. If a definable group G admits a left-invariant smooth Keisler measure μ, then μ is generically stable (see [36, Proposition 7.10, Theorem 7.29]), and so G is *fsg*. Recall also that if G is a definable group then there is a normal type-definable subgroup G^{00} of G, which is the smallest type-definable bounded-index subgroup of G. In [36, Section 8.4], Simon proves the following consequence of smoothness, which strengthens generic compact domination to outright compact domination, and extends compact domination for *definably compact groups in o-minimal theories* (after one knows that o-minimal theories are distal, see below). The reader should compare this statement to Fact 2.16.

Fact 6.4. *Assume T is NIP and G is a definable fsg group. Let μ be the unique left-invariant Keisler measure on G, and assume μ is smooth. Fix a definable set $X \subseteq G$, and let $F_X \subseteq G/G^{00}$ be the set of $C \in G/G^{00}$ such that*

$$C \cap X \neq \emptyset \quad \textit{and} \quad C \cap (G \backslash X) \neq \emptyset.$$

Then F_X has Haar measure 0, and so for any $\epsilon > 0$, there is a definable set $Z \subseteq G$ such that $\mu(Z) < \epsilon$ and $\{a \in G : aG^{00} \subseteq F_X\} \subseteq Z$.

A place to find smooth measures is in the setting of *distal theories*, which were introduced by Simon in [35] to capture the notion of an NIP theory with "no stable part". For the purposes of this paper we will take the following characterization from [35] as a *definition* of distality.

Definition 6.5. T is **distal** if it is NIP and every global generically stable Keisler measure is smooth.

So, in particular, if T is distal and G is a definable *fsg* group, then the unique left-invariant Keisler measure on G is smooth.

We now prove an analogue of Theorem 5.5 for an *fsg* group G, definable in an NIP theory, such that the unique left-invariant measure on G is smooth. Following this, we prove an analogue of Theorem 3.2 for the case when G/G^{00} is profinite (which will hold in our application to compact p-adic analytic groups). In these results, the assumptions are stronger in the sense that the whole theory is assumed to be NIP. Moreover, in the conclusions we have definability of the data, but no claims about definability in a certain Boolean fragment (see Remark 6.8). On the other hand, we have outright inclusion or disjointness, rather than up to ϵ, which yields stronger structure and regularity statements.

Theorem 6.6. *Assume T is NIP. Let G be a definable fsg group, and let μ be the unique left-invariant Keisler measure on G. Suppose μ is smooth (e.g. if T is distal). Fix a definable set $X \subseteq G$ and some $\epsilon > 0$. Then there are*

* *a definable finite-index normal subgroup H of G,*
* *a compact connected Lie group L,*
* *a definable homomorphism $\pi \colon H \to L$, and*
* *a definable set $Z \subseteq G$, with $\mu(Z) < \epsilon$,*

such that, for any integer $t \geq 1$, there is

* *a (t, L)-approximate Bohr chain $(Y_m)_{m=0}^{\infty}$ in H, where is each Y_m definable and $ker(\pi) = \bigcap_{m=0}^{\infty} Y_m$,*

satisfying the following properties.

(*i*) (structure) *For any $m \in \mathbb{N}$, there is $D \subseteq G$, which is a finite union of translates of Y_m, such that $D \subseteq X \subseteq D \cup Z$.*

(*ii*) (regularity) *For any $m \in \mathbb{N}$ and $g \in G \backslash Z$, either $gY_m \cap X = \emptyset$ or $gY_m \subseteq X$.*

Moreover, if G is pseudofinite then we may assume $L = \mathbb{T}^n$ for some $n \in \mathbb{N}$, and if G/G^{00} is abelian then we may assume $H = G$.

Proof. The proof is similar to that of Theorem 5.5, and so we provide a sketch. First, we may restrict to a countable language in which G and X are definable and μ is still smooth (see Lemma 7.8 and the remarks after Proposition 8.38 in [36]).

To prove (*i*) and (*ii*), one first finds a definable normal finite-index subgroup $H \leq G$, a compact connected Lie group L, a definable homomorphism $\pi \colon H \to L$, and definable sets $W \subseteq H$ and $Z \subseteq G$, such that $\ker \pi \subseteq W$ and, for any $g \in G \backslash Z$, either $gW \cap X = \emptyset$ or $gW \subseteq X$. This requires "non-local" applications of Proposition 5.1 and, in turn, Lemma 2.6. Here "non-local" means that we work with all definable sets, rather than around a fixed formula $\theta(x; \bar{y})$. The Lie structure associated to G/G^{00}, described as in Lemma 2.6, remains valid as we assume T is countable. In place of Lemma 3.1, one uses the fact that, for any decreasing sequence $(W_i)_{i=0}^{\infty}$ of definable sets, with $\bigcap_{i=0}^{\infty} W_i = G^{00}$, there are $Z \subseteq G$ definable and some $i \in \mathbb{N}$ such that $\mu(Z) < \epsilon$, and for any $g \in G \backslash Z$, either $gW_i \cap X = \emptyset$ or $gW_i \subseteq X$ (this is immediate from compactness and Fact 6.4). Finally, we obtain the definable Bohr chains in W using a non-local analogue of Lemma 5.4, which yields condition (ii). For condition (i), one mimics the end of the proof of Theorem 5.5 to find $D \subseteq G$, which is a finite union of translates of Y_m, such that $D \subseteq X \subseteq D \cup Z$ (simply replace each instance of $\mu(X \cap gY_m) = 0$ with $X \cap gY_m = \emptyset$, and each instance of $\mu(X \backslash gY_m) = 0$ with $gY_m \subseteq X$). □

Next, we give a strengthened version of the previous theorem, under the additional assumption that G/G^{00} is profinite.

Theorem 6.7. *Assume T is NIP. Let G be a definable fsg group, and let μ be the unique left-invariant Keisler measure on G. Suppose μ is smooth (e.g. if T is distal), and G/G^{00} is profinite. Fix a definable set $X \subseteq G$ and some $\epsilon > 0$. Then there are*

∗ *a definable finite-index normal subgroup H of G, and*
∗ *a set $Z \subseteq G$, which is a union of cosets of H with $\mu(Z) < \epsilon$,*

satisfying the following properties.

(i) (structure) *$X \backslash Z$ is a union of cosets of H.*
(ii) (regularity) *For any $g \in G \backslash Z$, either $gH \cap X = \emptyset$ or $gH \subseteq X$.*

Proof. This can be argued as a special case of Theorem 6.6, since if G/G^{00} is profinite then we may take the Lie group L to be trivial, and so $Y_m = H$ for any $t \geq 1$ and $m \in \mathbb{N}$. Alternatively, as in the proof of Theorem 3.2, one can deduce condition (ii) directly from compactness and Fact 6.4. In this case, we again assume T is countable (without loss of generality), in order to write G^{00} as an intersection of a countable decreasing sequence of definable subgroups. For condition (i), set $D = \bigcup \{g \in G : gH \subseteq X \text{ and } gH \cap Z = \emptyset\}$. Then D is a union of cosets of H, and $D \subseteq X \subseteq D \cup Z$. Since Z is a union of cosets of H, it follows from the definition of D that $D \cap Z = \emptyset$. So $X \backslash Z = D$. □

Remark 6.8. Results similar to those above can be shown just for *fsg* groups definable in NIP theories, and without the smoothness assumption. In this case, the results would be nearly identical to Theorems 5.5 and 3.2. Indeed, *fsg* groups in NIP theories satisfy generic compact domination just as in Fact 2.16, but with $G^{00}_{\theta^r}$ replaced by G^{00}, see [24, 37]. However, as discussed in [12, Remark 1.2], since the unique left-invariant measure in an NIP *fsg* group G is generically stable, one could fix an invariant formula $\theta(x;\bar{y})$ and construct $G^{00}_{\theta^r}$ as in Fact 2.16. This yields structure and regularity theorems as before, but with additional information about the definability of the data. Precisely:

Assume T is NIP. Suppose G is definable and *fsg*, and let μ be the unique left-invariant Keisler measure on G. Fix an invariant formula $\theta(x;\bar{y})$, where x is in the sort for G. Then, for any θ^r-definable $X \subseteq G$ and any $\epsilon > 0$, we have the conclusion of Theorem 5.5, except with $\mathbb{T}^n$ replaced by some compact connected Lie group L. If $G/G^{00}_{\theta^r}$ is profinite, then the conclusion of Theorem 3.2 holds exactly as stated.

It would be interesting to pursue notions of smoothness for local measures, or "local distality" for formulas, and recover local versions of Theorems 6.6 and 6.7.

For instance, one might consider an NIP formula $\phi(x,\bar{y})$ such that every generically stable global Keisler measure on the Boolean algebra of ϕ-formulas is smooth.

7. Compact p-adic analytic groups

We give an application of Theorem 6.7 to compact p-adic analytic groups. See Theorems 7.2 and 7.4 below. In fact, as we point out below, this can also be seen as a fairly direct application of [32, Proposition 2.8], and an extension of certain results in that paper.

We assume some familiarity with the p-adic field $\mathbb{Q}_p$ and p-adic model theory. The topology on $\mathbb{Q}_p$ is given by the valuation where open neighborhoods of a point a are defined by $v(x-a) \geq n$ for $n \in \mathbb{Z}$. The topology on $\mathbb{Q}_p^n$ is the product topology. A *p-adic analytic function* is a function f, from some open $V \subseteq \mathbb{Q}_p^n$ to $\mathbb{Q}_p$, such that for every $a \in V$, there is an open neighborhood of $a \in V$ in which f is given by a convergent power series. We obtain the notions of a p-adic analytic mainfold and a p-adic analytic (or Lie) group.

We let $\mathbb{Q}_p^{\mathrm{an}}$ denote the expansion of the field $(\mathbb{Q}_p,+,\cdot)$ by symbols for all convergent power series in $\mathbb{Z}_p[[X_1,\ldots,X_n]]$ for all n. Then any compact p-adic analytic manifold or group is seen to be naturally definable in the structure $\mathbb{Q}_p^{\mathrm{an}}$. (We conflate definable and interpretable at this point.) It is well known that $\mathrm{Th}(\mathbb{Q}_p^{\mathrm{an}})$ is distal, and that distality passes to T^{eq} (see Exercise 9.12 of [36]).

Let us fix a compact p-adic analytic group K (so definable in $\mathbb{Q}_p^{\mathrm{an}}$). The open subgroups of K are definable (in $\mathbb{Q}_p^{\mathrm{an}}$) and of course have finite index. (In fact it is pointed out in [27] that the family of open normal subgroups of K is even *uniformly* definable in $\mathbb{Q}_p^{\mathrm{an}}$, although we will not need this additional information.) If M is a saturated elementary extension of $\mathbb{Q}_p^{\mathrm{an}}$ then $K(M)$ denotes the group definable in the structure M by the same formula as the one defining K in $\mathbb{Q}_p$.

Fact 7.1.

(*i*) *$K(M)$ is an fsg group.*
(*ii*) *The unique left invariant Keisler measure on $K(M)$ is smooth,*
(*iii*) *$K(M)^{00} = K(M)^0$, and so $K(M)/K(M)^{00}$ is profinite. Moreover, $K(M)^{00}$ is precisely the group of "infinitesimals" of $K(M)$, and the standard part map $\mathrm{st}\colon K(M) \to K$ is an isomorphism of topological groups.*

Proof. Part (*i*) is Corollary 2.3(*iv*) of [32]. Part (*ii*) follows from distality of $\mathrm{Th}(\mathbb{Q}_p^{\mathrm{an}})$. Part (*iii*) follows from Section 6 of [22]; see also [32]. □

From Theorem 6.7 and its proof, as well as restricting to sets X defined over the standard model, we obtain the following:

Theorem 7.2. *Let K be a compact p-adic analytic group (definable, as mentioned above, in the structure $\mathbb{Q}_p^{\mathrm{an}}$). Let μ be its unique normalized Haar measure. Let $X \subseteq K$ be definable in $\mathbb{Q}_p^{\mathrm{an}}$, and let $\epsilon > 0$. Then there are*

* *an open (so finite-index) normal subgroup H of K, and*
* *a set $Z \subseteq K$, which is a union of cosets of H with $\mu(Z) < \epsilon$,*

satisfying the following properties.

(*i*) (structure) *$X \setminus Z$ is a union of cosets of H.*
(*ii*) (regularity) *For any $g \in K \backslash Z$, either $gH \cap X = \emptyset$, or $gH \subseteq X$.*

Remark 7.3.

(1) Proposition 2.8 of [32] states that $K(M)$ is compactly dominated via the map $K(M) \to K(M)/K(M)^{00}$. So we could also have deduced Theorem 1.2 from this result, together with the standard methods.
(2) We have received suggestions that Theorem 7.2 (and so 7.4 below) can be obtained directly from the cell decomposition results of Denef and others, at least when K is $\mathbb{Z}_p^n$ (some n).

We obtain easily from Theorem 7.2 the following statement about the family of quotients of K by open normal subgroups.

Theorem 7.4. *Let K be as above. Let $(G_i)_{i\in I}$ be the family of finite groups obtained as quotients of K by open normal subgroups. Let $A \subseteq K$ be definable in $\mathbb{Q}_p^{\text{an}}$ and for $i \in I$, let $A_i \subseteq G_i$ be the image of A under the quotient map. Then, for any $\epsilon > 0$, there is some $n \geq 1$ such that for any $i \in I$, there are*

* *a normal subgroup $H_i \leq G_i$ of index at most n, and*
* *a set $Z_i \subseteq G_i$, which is a union of cosets of H_i with $|Z_i| < \epsilon|G_i|$,*

satisfying the following properties.

(*i*) (structure) *$A_i\backslash Z_i$ is a union of cosets of Z_i.*
(*ii*) (regularity) *For any $g \in G_i\backslash Z_i$, either $gH_i \cap A_i = \emptyset$ or $gH_i \subseteq A_i$.*

References

[1] M. A. Alekseev, L. Yu. Glebskiĭ, and E. I. Gordon, *On approximations of groups, group actions and Hopf algebras*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) **256** (1999), no. Teor. Predst. Din. Sist. Komb. i Algoritm. Metody. 3, 224–262, 268. MR 1708567

[2] Noga Alon, Eldar Fischer, and Ilan Newman, *Efficient testing of bipartite graphs for forbidden induced subgraphs*, SIAM J. Comput. **37** (2007), no. 3, 959–976.

[3] Noga Alon, Jacob Fox, and Yufei Zhao, *Efficient arithmetic regularity and removal lemmas for induced bipartite patterns*, arXiv:1801.04675, 2018.

[4] Saugata Basu, *Combinatorial complexity in o-minimal geometry*, Proceedings of the London Mathematical Society **100** (2009), no. 2, 405–428.

[5] Vitaly Bergelson, Hillel Furstenberg, and Benjamin Weiss, *Piecewise-Bohr sets of integers and combinatorial number theory*, Topics in discrete mathematics, Algorithms Combin., vol. 26, Springer, Berlin, 2006, pp. 13–37. MR 2249261

[6] J. Bourgain, *On triples in arithmetic progression*, Geom. Funct. Anal. **9** (1999), no. 5, 968–984. MR 1726234

[7] Emmanuel Breuillard, Ben Green, and Terence Tao, *The structure of approximate groups*, Publ. Math. Inst. Hautes Études Sci. **116** (2012), 115–221. MR 3090256

[8] Theodor Bröcker and Tammo tom Dieck, *Representations of compact Lie groups*, Graduate Texts in Mathematics, vol. 98, Springer-Verlag, New York, 1985. MR 781344

[9] Artem Chernikov and Pierre Simon, *Definably amenable NIP groups*, arXiv:1502.04365, 2015, to appear in J. Amer. Math. Soc.

[10] Artem Chernikov and Sergei Starchenko, *Regularity lemma for distal structures*, arXiv:1507.01482, 2015, J. Eur. Math. Soc., accepted.

[11] ———, *Definable regularity lemmas for NIP hypergraphs*, arXiv:1607.07701, 2016.

[12] Gabriel Conant and Anand Pillay, *Pseudofinite groups and VC-dimension*, arXiv:1802.03361, 2018.

[13] Gabriel Conant, Anand Pillay, and Caroline Terry, *A group version of stable regularity*, arXiv:1710.06309, 2017.

[14] Jacob Fox, Mikhail Gromov, Vincent Lafforgue, Assaf Naor, and János Pach, *Overlap properties of geometric expanders*, J. Reine Angew. Math. **671** (2012), 49–83. MR 2983197

[15] Jacob Fox, János Pach, and Andrew Suk, *A polynomial regularity lemma for semialgebraic hypergraphs and its applications in geometry and property testing*, SIAM J. Comput. **45** (2016), no. 6, 2199–2223. MR 3585030

[16] Timothy Gowers and Julia Wolf, *Linear forms and higher-degree uniformity for functions on* $\mathbb{F}_p^n$, Geom. Funct. Anal. **21** (2011), no. 1, 36–69.
[17] Timothy Gowers, *Lower bounds of tower type for Szemerédi's uniformity lemma*, Geom. Funct. Anal. **7** (1997), no. 2, 322–337.
[18] B. Green, *A Szemerédi-type regularity lemma in abelian groups, with applications*, Geom. Funct. Anal. **15** (2005), no. 2, 340–376. MR 2153903
[19] Ben Green and Terence Tao, *An arithmetic regularity lemma, an associated counting lemma, and applications*, An Irregular Mind (2010), 261–334.
[20] Ben Green and Tom Sanders, *Fourier uniformity on subspaces*, arXiv:1607.07701 (2015).
[21] Ehud Hrushovski, *Stable group theory and approximate subgroups*, J. Amer. Math. Soc. **25** (2012), no. 1, 189–243. MR 2833482
[22] Ehud Hrushovski, Ya'acov Peterzil, and Anand Pillay, *Groups, measures, and the NIP*, J. Amer. Math. Soc. **21** (2008), no. 2, 563–596. MR 2373360 (2008k:03078)
[23] Ehud Hrushovski and Anand Pillay, *Groups definable in local fields and pseudo-finite fields*, Israel J. Math. **85** (1994), no. 1-3, 203–262. MR 1264346
[24] Ehud Hrushovski, Anand Pillay, and Pierre Simon, *Generically stable and smooth measures in NIP theories*, Trans. Amer. Math. Soc. **365** (2013), no. 5, 2341–2366. MR 3020101
[25] Richard Iltis, *Some algebraic structure in the dual of a compact group*, Canad. J. Math. **20** (1968), 1499–1510. MR 0232892
[26] László Lovász and Balázs Szegedy, *Regularity partitions and the topology of graphons*, An irregular mind, Bolyai Soc. Math. Stud., vol. 21, János Bolyai Math. Soc., Budapest, 2010, pp. 415–446. MR 2815610
[27] Dugald Macpherson and Katrin Tent, *Profinite groups with NIP theory and p-adic analytic groups*, Bull. Lond. Math. Soc. **48** (2016), no. 6, 1037–1049. MR 3608949
[28] M. Malliaris and S. Shelah, *Regularity lemmas for stable graphs*, Trans. Amer. Math. Soc. **366** (2014), no. 3, 1551–1585. MR 3145742
[29] David Marker, *Model theory*, Graduate Texts in Mathematics, vol. 217, Springer-Verlag, New York, 2002. MR 1924282 (2003e:03060)
[30] Jiří Matoušek, *Bounded VC-dimension implies a fractional Helly theorem*, Discrete Comput. Geom. **31** (2004), no. 2, 251–255. MR 2060639
[31] Nikolay Nikolov, Jakob Schneider, and Andreas Thom, *Some remarks on finitarily approximable groups*, arXiv:1703.06092, 2017.
[32] A. Onshuus and A. Pillay, *Definable groups and compact p-adic Lie groups*, J. Lond. Math. Soc. (2) **78** (2008), no. 1, 233–247. MR 2427062
[33] Anand Pillay, *Type-definability, compact Lie groups, and o-minimality*, J. Math. Log. **4** (2004), no. 2, 147–162. MR 2114965
[34] ———, *Remarks on compactifications of pseudofinite groups*, Fund. Math. **236** (2017), no. 2, 193–200. MR 3591278
[35] Pierre Simon, *Distal and non-distal NIP theories*, Ann. Pure Appl. Logic **164** (2013), no. 3, 294–318. MR 3001548
[36] Pierre Simon, *A guide to NIP theories*, Lecture Notes in Logic, vol. 44, Association for Symbolic Logic, Chicago, IL; Cambridge Scientific Publishers, Cambridge, 2015. MR 3560428
[37] Pierre Simon, *VC-sets and generic compact domination*, Israel J. Math. **218** (2017), no. 1, 27–41. MR 3625123
[38] Olof Sisask, *Convolutions of sets with bounded VC-dimension are uniformly continuous*, arXiv:1802.02836, 2018.
[39] Endre Szemerédi, *Regular partitions of graphs*, Problèmes combinatoires et théorie des graphes (Colloq. Internat. CNRS, Univ. Orsay, Orsay, 1976), Colloq. Internat. CNRS, vol. 260, CNRS, Paris, 1978, pp. 399–401. MR 540024
[40] Terence Tao, *Hilbert's fifth problem and related topics*, Graduate Studies in Mathematics, vol. 153, American Mathematical Society, Providence, RI, 2014. MR 3237440
[41] ———, *Expanding polynomials over finite fields of large characteristic, and a regularity lemma for definable sets*, Contrib. Discrete Math. **10** (2015), no. 1, 22–98. MR 3386249
[42] Terence Tao and Van Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2006. MR 2289012
[43] C. Terry and J. Wolf, *Stable arithmetic regularity in the finite-field model*, arXiv:1710.02021, 2017.

Department of Mathematics, University of Notre Dame, Notre Dame, IN, 46656, USA
E-mail address: gconant@nd.edu

Department of Mathematics, University of Notre Dame, Notre Dame, IN, 46656, USA
E-mail address: apillay@nd.edu

Department of Mathematics, University of Maryland, College Park, MD 20742, USA
E-mail address: cterry@umd.edu