



A model-theoretic note on the Freiman–Ruzsa theorem

Amador Martin-Pizarro¹ · Daniel Palacín¹ · Julia Wolf²

Accepted: 2 June 2021 / Published online: 19 June 2021
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2021

Abstract

A non-quantitative version of the Freiman–Ruzsa theorem is obtained for finite stable sets with small tripling in arbitrary groups, as well as for (finite) weakly normal subsets in abelian groups.

Keywords Model theory · Local stability · Additive combinatorics · Freiman–Ruzsa

Mathematics Subject Classification 03C13 · 03C45 · 11B30

Introduction

A finite subset A of a group G is said to have *doubling* K if the product set $A \cdot A = \{a \cdot b \mid a, b \in A\}$ has size at most $K|A|$. Archetypal examples of sets with small doubling (where K is constant as the size of the group G , and the set A , tend to infinity) are cosets of subgroups. Theorems of Freiman–Ruzsa type assert that sets with small doubling are “not too far” from being subgroups in a suitable sense. Specifically, Freiman’s original theorem [6] asserts that a finite subset of the integers with small doubling is efficiently contained in a generalized arithmetic progression. A proof of

The first two authors conducted research partially supported by MTM2017-86777-P as well as by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - Project Number 2100310201 and 2100310301, part of the ANR-DFG program GeoMod.

✉ Daniel Palacín
palacin@math.uni-freiburg.de
Amador Martin-Pizarro
pizarro@math.uni-freiburg.de
Julia Wolf
julia.wolf@dpmms.cam.ac.uk

- ¹ Abteilung für Mathematische Logik, Mathematisches Institut, Albert-Ludwig-Universität Freiburg, Ernst-Zermelo-Straße 1, 79104 Freiburg, Germany
- ² Department of Pure Mathematics and Mathematical Statistics, Centre for Mathematical Sciences, Wilberforce Road, Cambridge CB3 0WB, UK

an analogous statement for arbitrary abelian groups was given by Green and Ruzsa [7], based on Ruzsa's proof of Freiman's theorem [16]. A version of the result for abelian groups of bounded exponent with a particularly pleasing proof was given by Ruzsa in [17]. His result asserts that if A is a finite subset of an abelian group $(G, +)$ of exponent r such that $|A + A| \leq K|A|$, then A is contained in a subgroup H of G of size at most $K^2 r^{K^4} |A|$. For $G = \mathbb{F}_p^n$ with p a fixed prime, the exponent can be improved to $2K - 1$ (see [5] and references therein). By considering the union of a subspace and K arbitrary linearly independent elements, it is not difficult to see that any bound on the size of a subgroup containing A must be exponential in K .

However, this example is still highly structured in the sense that a large part of the set has the structure of a subgroup, which suggests a natural reformulation of the problem. The *Polynomial Freiman–Ruzsa Conjecture*, which remains one of the central open problems in additive combinatorics, asserts that a subset A of doubling K in \mathbb{F}_2^∞ can be covered by $C_1(K)$ many cosets of some subspace of size $C_2(K)|A|$, where both $C_1(K)$ and $C_2(K)$ are polynomials in K ; or equivalently, that there are constants $C_3(K)$ and $C_4(K)$, each polynomial in K , such that for some coset $v + H$ of a subspace H of size $C_3(K)|A|$, we have that $A \cap (v + H)$ has size at least $|A|/C_4(K)$. For the best bounds known to date see [18,20].

The above formulation of the Freiman–Ruzsa theorem resonates with a classical setting in model theory, namely weakly normal groups. Weakly normal groups, also known as 1-based stable groups, are groups for which every definable set is a boolean combination of instances of weakly normal formulae (see Sect. 2). In a weakly normal (stable) group, every definable subset is a boolean combination of cosets of definable subgroups [9]. Furthermore, every type over a model is the generic type of a coset of a (type-)definable subgroup: the subgroup is its model-theoretic stabiliser. Roughly speaking, a large proportion of a given definable set intersects a coset of a definable group, so they are *commensurable*.

For non-abelian groups, the suitable notion of doubling is *tripling* K , that is, the cardinality of $A \cdot A \cdot A$ is bounded by $K|A|$. Indeed, sets of small tripling have small doubling, but the converse need not hold. In this context, phenomena of Freiman–Ruzsa type are present in recent work of Hrushovski [11, Corollary 4.18], who showed that a set of small tripling in a (possibly infinite) group of bounded exponent is commensurable with a subgroup, inspired by classical results and techniques from stability theory in a non-standard setting.

Motivated by Hrushovski's work, in this note we adapt the local approach to stability of Hrushovski and Pillay in [10, Theorem 4.1] in order to obtain a non-quantitative version of the Freiman–Ruzsa theorem for arbitrary (possibly infinite) groups under the assumption of stability. We say that a subset A of G is *r-stable* if there are no elements $a_1, \dots, a_r, b_1, \dots, b_r$ in G such that $b_j \cdot a_i$ belongs to A if and only if $i \leq j$. In particular, we prove the following result.

Theorem A *Given real numbers $K \geq 1$ and $\epsilon > 0$ and a natural number $r \geq 2$, there exists a natural number $n = n(K, \epsilon, r)$ such that for any (possibly infinite) group G and any finite r -stable subset $A \subseteq G$ with tripling K , there is a subgroup $H \subseteq A \cdot A^{-1}$ of G with $A \subseteq C \cdot H$ for some $C \subseteq A$ of size at most n . Moreover, there exists $C' \subseteq C$*

such that

$$|A\Delta(C' \cdot H)| \leq \epsilon|H|.$$

In particular, it follows from the Plünnecke–Ruzsa inequalities that

$$|A\Delta(C' \cdot H)| \leq \epsilon K^2|A|.$$

In the case when G is abelian, it suffices to assume that A has doubling K . Furthermore, we shall prove that, when G is abelian, the subgroup H can be taken to be a boolean combination (of complexity only depending on K , ϵ and r) of translates of A .

In particular, on choosing $\epsilon = 1$, Theorem A implies that there is some natural number $n_0 = n(K, 1, r)$ such that any finite r -stable subset A of tripling K is contained in n_0 translates of a subgroup $H \subseteq A \cdot A^{-1}$. It follows that $|A \cap g \cdot H| \geq |A|/n_0$ for some subgroup $H \leq G$ and some $g \in A$, which is a qualitative result in the spirit of Freiman–Ruzsa. As in the previous paragraph, when G is abelian, the complexity of such a subgroup H as a boolean combination of translates of A can be bounded solely in terms of K and r .

The above result dovetails with a suite of arithmetic regularity lemmas under the additional assumption of stability that have been obtained recently by Terry and the third author [25,26], as well as by Conant, Pillay and Terry [2]. However, without the assumption of small doubling/tripling, the bound on the symmetric difference is at best $\epsilon|H|$. Furthermore, the group H so obtained in [2,25,26] has finite index in G so its size comparable to $|G|$, but not necessarily to $|A|$. The Theorem A is also reminiscent of work of Sisask [21, Theorem 5.4], who combined the assumption of small doubling with that of bounded VC-dimension in vector spaces over finite fields. Finally, we remark that closely related results were obtained by Conant [3, Corollary 1.4] for groups of bounded exponent. In a previous version of this article, Theorem A was stated with the upper bound

$$|A\Delta(C' \cdot H)| \leq \epsilon|A|,$$

which was subsequently improved by Conant [4, Theorem 1.6] to the current upper bound. Conant’s methods do not use the full power of stability but instead work in the more general setting of finite VC-dimension. We later noticed that our (non-standard) techniques already implied the finer bound in terms of H .

We also explore the interaction between model theory and recent work in additive combinatorics in a second direction. In [8] Green and Sanders showed that subsets of a locally compact abelian group G which are elements of the Fourier algebra $\mathcal{A}(G)$ belong to the coset ring $\mathcal{W}(G)$. They also gave an upper bound for the boolean complexity of the representation as elements in $\mathcal{W}(G)$ of such sets in terms of their Wiener norm. More recently, Sanders [19] showed that smallness of this norm implies stability, hence $\mathcal{A}(G) = \mathcal{W}(G) \subseteq \mathcal{S}(G)$, where $\mathcal{S}(G)$ denotes the ring of stable subsets of G . He further observed that when G is not finite, it is possible for the latter inclusion to be strict.

In this paper we shall consider the ring $\mathcal{WN}(G)$ of subsets of G generated by all instances of weakly normal formulae, defined in Sect. 2. It is not difficult to see that $\mathcal{WN}(G)$ is contained in yet not identical to the stability ring $\mathcal{S}(G)$. Actually, we have the chain of inclusions

$$\mathcal{W}(G) \subseteq \mathcal{WN}(G) \subseteq \mathcal{S}(G).$$

In fact, we shall show that $\mathcal{WN}(G)$ is equal to $\mathcal{W}(G)$ for abelian G (see Proposition 4.3). This is a local reformulation of the celebrated result by Hrushovski and Pillay [9, Theorem 4.1]. We further deduce a result of Freiman–Ruzsa type for finite subsets in $\mathcal{WN}(G)$.

Mimicking the definition of Sanders in [19], we say that a subset A of G has an (r, k, l) -weakly normal representation if

$$A = \bigcup_{j=1}^k B_j \cap \bigcup_{i=1}^l G \setminus C_r,$$

where all the relations $B_1(x + y), \dots, B_k(x + y), C_1(x + y), \dots, C_l(x + y)$ are r -weakly normal.

Theorem B *Given natural numbers r, k , and l , there are natural numbers $n = n(r, k, l)$ and $m = m(r, k, l)$ such that for any abelian group G and any subset $A \subseteq G$ with an (r, k, l) -weakly normal representation, there are subgroups H_1, \dots, H_n of G , each contained in $A - A$, with*

$$A \subseteq \bigcup_{i=1}^n g_i + H_i,$$

for some g_1, \dots, g_n in A . Furthermore, each H_i is a boolean combinations of complexity at most m of translates of A .

In particular, if A is finite, we have that $|A \cap (g + H)| \geq |A|/n$ for some g in A and some subgroup $H \leq G$ contained in $A - A$.

In contrast to Theorem A, the subset A above need not have small doubling or tripling. Indeed, there is no correlation for finite sets between having a weakly normal representation and small doubling: consider the group $G = \mathbb{F}_p^2$ and let A be the subset $(\mathbb{F}_p \times \{0\} \cup \{0\} \times \mathbb{F}_p)$, which has a $(2, 2, 0)$ -weakly normal representation. However, the quantity

$$\frac{|A + A|}{|A|} = \frac{p^2}{2p - 1}$$

is not uniformly bounded for large p .

Throughout this paper, we will assume a certain familiarity with model theory. We refer the reader to [24] for an excellent introduction to the subject. Basic notions

related to local stability and weak normality are recalled and developed in Sects. 1 and 2, respectively. Section 3 is devoted to a discussion of Keisler measures on a certain boolean algebra arising from sets of small tripling, and the associated measure-theoretic stabilizers. The proofs of our main results are given in Sect. 4.

1 Local stability

We work inside a sufficiently saturated model \mathbb{U} of a complete theory T with infinite models in a language \mathcal{L} .

Recall that a formula $\varphi(x, y)$ is r -stable with respect to the partition of the variables into the tuples x and y if there are no tuples $a_1, \dots, a_r, b_1, \dots, b_r$ such that $\varphi(a_i, b_j)$ holds if and only if $i \leq j$.

A formula is stable if it is r -stable for some r . Stable formulae are closed under boolean combinations (see [25] for a finitary version of this fact). A set X is φ -definable over a subset A of parameters if it is definable by a boolean combination of instances $\varphi(x, a)$ with a in A . By a φ -type over a subset A we mean a maximal finitely consistent collection of instances of the form $\varphi(x, a)$ or $\neg\varphi(x, a')$ for a, a' in A .

The space of φ -types $S_\varphi(\mathbb{U})$ is a compact Hausdorff 0-dimensional topological space, with basic clopen sets of the form

$$[X] = \{p \in S_\varphi(\mathbb{U}) \mid p \cup \{X(x)\} \text{ is finitely consistent}\},$$

where $X(x)$ is φ -definable. Given a stable formula $\varphi(x, y)$ and a partial \mathcal{L} -type $\pi(x)$, the collection

$$X_\pi = \{q(x) \in S_\varphi(\mathbb{U}) \mid q(x) \cup \pi(x) \text{ is finitely consistent}\}$$

is a closed, hence compact, subset of $S_\varphi(\mathbb{U})$ with integer-valued *Cantor-Bendixson rank* $\text{CB}_\varphi(\pi)$. Thus, any element $q(x)$ in X_π can be isolated from all other types of rank at least $\text{CB}_\varphi(q)$ by the neighbourhood $[\chi]$ of some formula $\chi(x)$. Furthermore, the space X_π contains only finitely many elements of maximal rank. The number of such elements is the φ -multiplicity of π , see [1, Chapter 6].

If $\varphi(x, y)$ is stable, then every φ -type $p(x)$ over a small submodel M is *definable*, that is, there is a formula $\theta(y)$ with parameters over M such that

$$\varphi(x, m) \in p \iff \theta(m),$$

for all m in M . Furthermore, the definable set $\theta(y)$ above is unique and can be defined by a positive boolean combination of instances $\varphi(a, y)$ with parameters in M (cf. [10, Lemma 5.4]). We refer to this definable set as the φ -definition $(d_p\varphi)(y)$ of p . Given a superset $B \supseteq M$ of \mathbb{U} , there is a unique φ -type over B extending p which is again definable over M , namely

$$\{\varphi(x, b) \mid (d_p\varphi)(b)\} \cup \{\neg\varphi(x, b') \mid \neg(d_p\varphi)(b')\}.$$

We refer to this type as the *non-forking extension* $p|_B(x)$ of $p(x)$ to B . The *global non-forking extension* of p is the φ -type $p|_{\mathbb{U}}$. In fact, the unique global non-forking extension of $p(x)$ is the only element in X_p of rank $\text{CB}_\varphi(p)$, so p has φ -multiplicity 1 (cf. [1, Proposition 6.13 & Corollary 6.15]).

Henceforth, we will assume that the underlying structure \mathbb{U} carries a definable group structure (G, \cdot) without parameters. In order to analyse the structure of an arbitrary stable subset A of G , it suffices expand the language by a distinguished unary predicate, whose realisations are exactly the elements in A . Thus, we may assume that the formula $\varphi(x, y) = A(y \cdot x)$ is stable, for some fixed definable subset A of G .

Note that $\varphi(x, y)$ is *equivariant* (see [10, Definition 5.13]), that is, every left-translate of an instance of φ is again an instance of φ . Given a stable equivariant formula $\varphi(x, y)$, there is a distinguished subgroup of G which is φ -definable, relative to G , as first observed in [10]. The following fact can be found in [2, Theorem 2.3].

Fact 1.1 Given a stable equivariant formula $\varphi(x, y)$ and a definable group G over a model M , there is a subgroup G_φ^0 of finite index in G which is φ -definable over M (relative to G) such that for any coset C of G_φ^0 and any φ -definable subset X , either $X \cap C$ or $C \setminus X$ is generic, in the sense that finitely many translates cover G .

The Cantor-Bendixson rank of a union is the maximum of the ranks of the sets in the union, so every generic φ -definable subset of G has maximal Cantor-Bendixson rank $\text{CB}_\varphi(G(x))$. On the other hand, if X is a φ -definable subset of G of maximal Cantor-Bendixson rank $\text{CB}_\varphi(G(x))$, it must be generic: indeed, since G_φ^0 has finite index in G , there must be a coset C of G_φ^0 such that $X \cap C$ has rank $\text{CB}_\varphi(G(x))$. We need only show that $C \cap X$ is generic. Otherwise, the set $C \setminus X$ is generic, so finitely many translates will cover G . Each such translate has maximal rank, yet every coset of G_φ^0 contains a unique φ -type of maximal rank.

Given a φ -type $p(x)$ over a submodel M , we define its *stabilizer* to be the subgroup

$$\text{Stab}_\varphi(p) = \{g \in G \mid \forall u((d_p\varphi)(u) \leftrightarrow (d_p\varphi)(u \cdot g))\}.$$

The stabilizer is clearly a definable subgroup of G with parameters from M . The following elementary remark shows that the stabilizer is φ -definable whenever G is abelian.

Remark 1.2 If $(G, +)$ is abelian, then the subgroup $\text{Stab}_\varphi(p)$ is φ -definable over M .

Proof Let $q(x)$ be the unique global non-forking extension of $p(x)$. Choose a φ -formula

$$\chi(x) = \bigvee_{j \in J} \bigwedge_{i \in I} \varphi(x, b_{ij}) \wedge \neg\varphi(x, c_{ij})$$

such that q lies in the neighborhood $[\chi]$, with $\chi(x)$ of Cantor-Bendixson rank $\text{CB}_\varphi(p)$ and φ -multiplicity 1.

Now, an element g in G belongs to $\text{Stab}_\varphi(p)$ if and only if the φ -type $g + q$ equals q , that is, if and only if $\chi(x) - g$ belongs to q , or equivalently, if and only if

$$\bigvee_{j \in J} \bigwedge_{i \in I} (d_p \varphi)(b_{ij} + g) \wedge \neg (d_p \varphi)(c_{ij} + g).$$

Recall that $(d_p \varphi)(y)$ is a positive boolean combination of instances $\varphi(a, y)$. Since G is abelian, the formula $\varphi(x, y + z)$ is equivalent to $\varphi(x + y, z)$, so the above condition on g is equivalent to a boolean combination $\psi(z, a')$ of instances of $\varphi(a', z)$, for some choice of parameters a' in G . In particular, the formula

$$\exists u \forall z (\text{Stab}_\varphi(p)(z) \leftrightarrow \psi(z, u))$$

holds in \mathbb{U} . Since M is an elementary substructure of \mathbb{U} , there are some parameters m in M such that $\text{Stab}_\varphi(p)(M)$ equals $\psi(M, m)$, and thus the φ -formula $\psi(z, m)$ defines the subgroup $\text{Stab}_\varphi(p)$ in G . \square

Note that if G is abelian, then $\varphi(x, y) = \varphi(y, x)$. Given global φ -types $p(x)$ and $q(y)$ in $S_\varphi(\mathbb{U})$, Harrington’s lemma [1, Lemma 6.8] yields that

$$q(y) \in [(d_p \varphi)(y)] \Leftrightarrow p(x) \in [(d_q \varphi)(x)].$$

A standard argument yields the following result, whose short proof we include for completeness.

Remark 1.3 If $(G, +)$ is abelian, then given a φ -type p over M

$$\text{CB}_\varphi(\text{Stab}_\varphi(p)) \leq \text{CB}_\varphi(p).$$

Proof Let q be a global type in $[\text{Stab}_\varphi(p)]$ of maximal rank, and choose a realization b of $q \upharpoonright_M$. Note that q is a non-forking extension of $q \upharpoonright_M$, since $\text{Stab}_\varphi(p)$ is definable over the model M . Let a realize the non-forking extension $p \upharpoonright_{M \cup \{b\}}$, which is definable over M by the formula $(d_p \varphi)(y)$. Thus, the element $a + b$ realizes p , since $-b$ belongs to $\text{Stab}_\varphi(p)$.

Let us first show that b realizes the non-forking extension $q \upharpoonright_{M \cup \{a\}}$ of $q \upharpoonright_M$. It suffices to see that $\varphi(a, b)$ holds if and only if $(d_q \varphi)(a)$. Now,

$$\begin{aligned} (d_q \varphi)(a) &\iff p \upharpoonright_{\mathbb{U}}(x) \in [(d_q \varphi)(x)] \stackrel{\text{Harrington}}{\iff} q(y) \in [(d_p \varphi)(y)] \\ &\iff (d_p \varphi)(b) \iff \varphi(x, b) \in p \upharpoonright_{M \cup \{b\}} \iff \varphi(a, b) \text{ holds.} \end{aligned}$$

As the formula φ is equivariant, addition by an element preserves the rank of formulae, so

$$\begin{aligned} \text{CB}_\varphi(\text{Stab}_\varphi(p)) &= \text{CB}_\varphi(q \upharpoonright_M) = \text{CB}_\varphi(q \upharpoonright_{M \cup \{a\}}) = \text{CB}_\varphi(b/M \cup \{a\}) \\ &= \text{CB}_\varphi(a + b/M \cup \{a\}) \leq \text{CB}_\varphi(a + b/M) = \text{CB}_\varphi(p), \end{aligned}$$

as desired. \square

2 Weak normality

Given a natural number k , a formula $\psi(x, y)$ is k -weakly normal if, whenever the instances $\psi(x, b_1), \dots, \psi(x, b_k)$ are pairwise distinct, the intersection $\bigcap_{i=1}^k \psi(x, b_i)$ is empty [9]. A formula is weakly normal if it is k -weakly normal for some natural number k . The conjunction of weakly normal formulae is again weakly normal. However, neither the negation nor the disjunction of two weakly normal formulae need necessarily be weakly normal.

It is easy to see that a k -weakly normal formula is k -stable. If not, there is a sequence $(a_i, b_i)_{1 \leq i \leq k}$ witnessing the failure of stability. Since a_j belongs to $\psi(x, b_j)$ but not to $\psi(x, b_i)$ for $i < j$, the instances are pairwise distinct. However, the element a_1 belongs to their common intersection, so $\psi(x, y)$ is not k -weakly normal. Furthermore, a formula $\psi(x, y)$ is 2-stable precisely if it is 2-weakly normal. Indeed, if $\psi(x, y)$ is not 2-weakly normal, we can find two distinct instances $\psi(x, b_1)$ and $\psi(x, b_2)$ with non-empty intersection. We may assume that there is some a_2 in $\psi(x, b_2) \setminus \psi(x, b_1)$. As the intersection $\psi(x, b_1) \cap \psi(x, b_2)$ is non-empty, choose a_1 in $\psi(x, b_1) \cap \psi(x, b_2)$ and note that

$$\psi(a_i, b_j) \Leftrightarrow 1 \leq i \leq j \leq 2,$$

so $\psi(x, y)$ is not 2-stable.

Formulae which are 2-stable are very special. For example, in the setting of a group G with a fixed definable subset A , the formula $\varphi(x, y) = A(y \cdot x)$ is 2-stable if and only if A is either empty or a coset of a subgroup of G . Recall that a subset A of an abelian group G is Sidon if, whenever the 4-tuple (a_1, a_2, a_3, a_4) of elements of A satisfies $a_1 - a_2 = a_3 - a_4$, then $a_1 = a_2$ (and hence $a_3 = a_4$) or $a_1 = a_3$ (and thus $a_2 = a_4$). Sidon subsets of the integers, such as $2^{\mathbb{N}}$ or $3^{\mathbb{N}}$, are 3-stable, but need not lie in the coset ring $\mathcal{W}(\mathbb{Z})$ [19].

Remark 2.1 In general, stability need not imply weak normality. For a Sidon set A of cardinality at least k , the formula $A(x + y)$ cannot be k -weakly normal. Choose k distinct elements a_1, \dots, a_k in A and consider the collection of sets $(-a_j + A)_{1 \leq j \leq k}$. The element 0 belongs to their common intersection, yet they are pairwise distinct sets.

Since a definable set is defined over a submodel N if and only if it only has finitely many distinct automorphic copies over N (see for example [1, Proposition 1.11]), we deduce the following easy observation concerning sets defined by an instance of a weakly normal formula.

Remark 2.2 Let X be a definable set given by an instance of a weakly normal formula. Then the set X is definable over any submodel containing a realization of X .

A remarkable property of every weakly normal formula ψ is that the definition $(d_p \psi)$ of every local type p over an arbitrary set of parameters is explicit, in contrast to a general stable formula (cf. [24, Theorem 8.3.1]): indeed, given a k -weakly normal

formula $\psi(x, y)$, an instance $\psi(x, a)$ belongs to the ψ -type $p = \text{tp}_\psi(c/A)$ if and only if it contains the set

$$\mathcal{X}_{p,\psi} = \bigcap_{\psi(x,a') \in p} \psi(x, a').$$

This set is definable since it is the intersection of at most $k - 1$ instances in p (notice that $\mathcal{X}_{p,\psi}$ is the empty set if and only if the collection of positive instances $\psi(x, a')$ in p is empty). It suffices to set

$$(d_p\psi)(y) = \forall x (\mathcal{X}_{p,\psi}(x) \rightarrow \psi(x, y)) \text{ when } \mathcal{X}_{p,\psi} \neq \emptyset,$$

and

$$(d_p\psi)(y) = (y \neq y) \text{ otherwise.}$$

Remark 2.3 Assume that the formula $\chi(x, y)$ is a boolean combination of weakly normal formulae. Then every χ -type p is definable over any submodel containing a realization of p .

Note that we do not require that the submodel contains the parameter set of p , which is assumed to be small with respect to the saturation of \mathbb{U} .

Proof If the formula $\chi(x, y)$ is a boolean combination of the weakly normal formulae $\psi_1(x, y), \dots, \psi_r(x, y)$, then the χ -type $p = \text{tp}_\chi(c/A)$ is determined by the collection of types $q_1 = \text{tp}_{\psi_1}(c/A), \dots, q_r = \text{tp}_{\psi_r}(c/A)$. Hence, the χ -definition $(d_p\chi)$ is determined by the definable sets $\{(d_{q_i}\psi_i)\}_{1 \leq i \leq r}$. Each $(d_{q_i}\psi_i)$ is determined by the corresponding definable set \mathcal{X}_{q_i,ψ_i} , as in the previous discussion, which is definable over any submodel containing c , by Remark 2.2. \square

A well-known result of Hrushovski and Pillay [9, Lemma 4.2] shows that, in a theory where all formulae are boolean combinations of weakly normal ones, types are generic in cosets of their stabilizers. In particular, groups definable in such theories are virtually abelian, that is, abelian-by-finite. We will provide a local version of their results for abelian groups, following closely [15, Lemma 2.6 & Remark 2.7].

Lemma 2.4 *Let $(G, +)$ be abelian and assume that the formula $\varphi(x, y) = A(x + y)$ is a boolean combination of weakly normal formulae. Given a φ -type p over a model M , there exists some element m in M such that $p|_{\mathbb{U}}$ lies in the neighborhood $[m + \text{Stab}_\varphi(p)]$, that is, the type p implies the φ -formula over M defining the coset $m + \text{Stab}_\varphi(p)$.*

Furthermore, the proof of the above result yields that $\text{CB}_\varphi(\text{Stab}_\varphi(p)) = \text{CB}_\varphi(p)$, but this fact will not be used in the sequel.

Proof Let a be a realization of $p(x)$. Since G is abelian, every coset of $H = \text{Stab}_\varphi(p)$ is φ -definable (since $\varphi(x, y)$ is equivariant). We want to show that the coset $H + a$ is φ -definable over M . It suffices to show that it is definable over M , for M is an elementary substructure. Since the element a lies in $H + a$, if this coset is definable

over M , then the type p must imply the corresponding φ -formula over M defining it. Remark 1.3 yields then the equality of ranks.

To prove that $H + a$ is definable over M , we need only show that $H + a$ is definable over a submodel $N \succeq M$ such that $\text{tp}(a/N)$ is an heir of $\text{tp}(a/M)$: indeed, suppose that $H + a$ is definable over such N , so there are an \mathcal{L}_M -formula $\theta(x, z)$ and some tuple n in N such that the formula $\theta(x, n)$ defines $H + a$. Note that this coset is definable over $M \cup \{a\}$. In particular, the formula

$$\forall x((H + u)(x) \leftrightarrow \theta(x, n))$$

belongs to $\text{tp}(a/N)$. Thus, we find a tuple m in M such that $\theta(x, m)$ defines $H + a$, by inheritance of $\text{tp}(a/N)$ over M .

Now choose a φ -type $q(x)$ over M of maximal Cantor-Bendixson rank $\text{CB}_\varphi(G)$. The extension $q|_{M \cup \{a\}}$ is definable over M , so it is finitely satisfiable over M [14, Lemma I.2.16]. Thus we can find some element c such that the type $\text{tp}(c/M \cup \{a\})$ is finitely satisfiable over M and extends $q|_{M \cup \{a\}}$. By a dual argument, we can find a submodel $N \succeq M$ containing c such that $\text{tp}(a/N)$ is an heir of $\text{tp}(a/M)$.

Claim The element a realizes $p|_{M \cup \{c-a\}}$.

Proof of Claim We need only show that $\varphi(a, c - a)$ holds if and only if $(d_p \varphi)(c - a)$ holds. Observe first that

$$\begin{aligned} \text{CB}_\varphi(G) &= \text{CB}_\varphi(q) = \text{CB}_\varphi(q|_{M \cup \{a\}}) = \text{CB}_\varphi(\text{tp}_\varphi(c/M \cup \{a\})) \\ &= \text{CB}_\varphi(\text{tp}_\varphi(c - a/M \cup \{a\})) \leq \text{CB}_\varphi(\text{tp}_\varphi(c - a/M)) \leq \text{CB}_\varphi(G). \end{aligned}$$

Hence, equality holds everywhere, so $r(y) = \text{tp}_\varphi(c - a/M \cup \{a\})$ is definable over M and has maximal rank $\text{CB}_\varphi(G)$.

Now, the formula $\varphi(a, c - a)$ holds if and only if $\varphi(a, y)$ belongs to $r(y)$, that is, if and only if the element a realizes $(d_r \varphi)(x)$, which is definable over M . Hence, the formula $\varphi(a, c - a)$ holds if and only if $p|_{\mathbb{U}}$ lies in $[(d_r \varphi)]$, which is equivalent to $r|_{\mathbb{U}}$ being contained in $[(d_p \varphi)]$, by Harrington's lemma. Since $[(d_p \varphi)]$ is definable over M and the element $c - a$ realizes r , the latter is equivalent to $(d_p \varphi)(c - a)$, as desired.

□ Claim

Since $c = a + (c - a)$, the element c realizes the complete φ -type

$$\begin{aligned} p|_{M \cup \{c-a\}} + (c - a) &= \{\theta(x) \varphi\text{-formula over } M \cup \{c - a\} \mid \theta(x - (c - a)) \\ &\in p|_{M \cup \{c-a\}}\}, \end{aligned}$$

which is again a complete φ -type over $M \cup \{c - a\}$. In particular, the global φ -type $p|_{\mathbb{U}} + (c - a)$ is a non-forking extension of $p|_{M \cup \{c-a\}} + (c - a)$. By Remark 2.3, both types are definable over N (which contains c).

Let us now show that the coset $H + a$ is definable over N . It suffices to show that every automorphism σ fixing N pointwise fixes the coset (setwise). Since $p|_{\mathbb{U}} + (c - a)$

is definable over N , the automorphism σ fixes $p|_{\mathbb{U}} + (c - a)$, so

$$p|_{\mathbb{U}} + (c - a) = \sigma(p|_{\mathbb{U}} + (c - a)) = p|_{\mathbb{U}} + (c - \sigma(a)),$$

since $p|_{\mathbb{U}}$ is definable over M . Thus, we have $p|_{\mathbb{U}} - a = p|_{\mathbb{U}} - \sigma(a)$, that is,

$$p|_{\mathbb{U}} = p|_{\mathbb{U}} + (a - \sigma(a)),$$

and hence $a - \sigma(a)$ lies in $H = \text{Stab}_\varphi(p)$, as desired. \square

In analogy to the classical result for weakly normal theories, we conclude that φ -definable sets are boolean combination of cosets of φ -definable groups whenever $\varphi(x, y) = A(x + y)$ is a boolean combination of weakly normal formulae.

Corollary 2.5 *In an abelian group G (written additively), assume that the formula $\varphi(x, y) = A(x + y)$ is a boolean combination of weakly normal formulae. Every φ -definable set is a boolean combination of cosets of φ -definable subgroups.*

Note in particular that a coset of a φ -definable subgroup is a boolean combination of translates of A .

Proof By a straightforward application of [24, Lemma 3.1.1], it suffices to show that whenever two φ -types p_1 and p_2 over a submodel M imply the same M -definable cosets of φ -definable subgroups (over M), then p_1 and p_2 are the same.

Let a_1 realize p_1 and choose a realization a_2 of $p_2|_{M \cup \{a_1\}}$. Set $H_1 = \text{Stab}_\varphi(p_1)$ and $H_2 = \text{Stab}_\varphi(p_2)$. By Lemma 2.4, both cosets $a_1 + H_1$ and $a_2 + H_2$ are M -definable. By assumption, since p_1 clearly implies the formula defining $a_1 + H_1$, every realization of p_2 lies in $a_1 + H_1$, and similarly for p_1 . In particular, the element $a_1 - a_2$ lies in $H_1 \cap H_2$. The rank computation

$$\begin{aligned} \text{CB}_\varphi(H_2) &\leq \text{CB}_\varphi(p_2) = \text{CB}_\varphi(p_2|_{M \cup \{a_1\}}) = \text{CB}_\varphi(\text{tp}_\varphi(a_2/M \cup \{a_1\})) \\ &= \text{CB}_\varphi(\text{tp}_\varphi(a_2 - a_1/M \cup \{a_1\})) \leq \text{CB}_\varphi(\text{tp}_\varphi(a_2 - a_1/M)) \leq \text{CB}_\varphi(H_2) \end{aligned}$$

yields that $a_2 - a_1$ realizes the non-forking extension of $q = \text{tp}_\varphi(a_2 - a_1/M)$ over $M \cup \{a_1\}$. As in the proof of Remark 1.3, Harrington’s Lemma implies that a_1 realizes $p_1|_{M \cup \{a_2 - a_1\}}$. Since $a_2 - a_1$ lies in $H_1 = \text{Stab}_\varphi(p_1)$, we have that $a_2 = a_1 + (a_2 - a_1)$ realizes p_1 . Thus the types p_1 and p_2 are equal, as desired. \square

3 Ideals and measures

A *Keisler measure* μ is a finitely additive probability measure on some boolean algebra of definable subsets of the ambient model [12]. Archetypal examples are measures μ_p , with two possible values 0 and 1, given by global φ -types p , that is, for every φ -definable set X ,

$$\mu_p(X) = 1 \Leftrightarrow p \in [X].$$

Given a Keisler measure μ , the collection of sets of measure zero forms an *ideal*, that is, it is closed under subsets and finite unions. A partial type is said to be *wide* (with respect to μ) if it contains no definable set of measure zero. In particular, since the collection of measure-0 sets forms an ideal, every wide partial type $\pi(x)$ over a parameter set A can be completed to a wide complete type over any arbitrary subset B containing A : indeed, the collection of formulae

$$\pi(x) \cup \{\neg\varphi(x) \mid \varphi(x) \in \mathcal{L}_B \text{ with } \mu(\varphi(x)) = 0\}$$

is clearly finitely consistent, and any completion of this partial type is wide. Note that we do not require that every formula in the completion has measure.

The measure μ is said to be *definable over the submodel M* (see [22, Definition 3.19]) if for every \mathcal{L} -formula $\varphi(x, y)$ and every $\epsilon > 0$, there is a partition of $\mathbb{U}^{|y|}$ into \mathcal{L}_M -formulae $\rho_1(y), \dots, \rho_m(y)$ such that for all pairs (b, b') realizing $\rho_i(y) \wedge \rho_i(z)$, we have that

$$|\mu(\varphi(x, b)) - \mu(\varphi(x, b'))| < \epsilon.$$

In particular, the set of tuples b with $\mu(\varphi(x, b)) = 0$ is type-definable over M and the map

$$\begin{aligned} S_y(M) &\rightarrow [0, 1] \\ \text{tp}(b/M) &\mapsto \mu(\varphi(x, b)) \end{aligned}$$

is continuous, so the value $\mu(\varphi(x, b))$ only depends on $\text{tp}(b/M)$. Note that a global φ -type p is definable over M if and only if the corresponding measure μ_p is.

Every Keisler measure admits an expansion of the original language \mathcal{L} in which it becomes definable (cf. [11, Section 2.6]). In this case, a formula of positive measure does not fork over \emptyset , see [11, Lemma 2.9 & Example 2.12].

In the presence of an ambient group G , we will consider the following notion of an *acceptable* set, which was introduced as a *near-subgroup* in [11, Definition 3.9].

Definition 3.1 A definable subset A of G is *acceptable* if there exists a Keisler measure μ on a boolean algebra of definable subsets of $(A \cup A^{-1} \cup \{\text{id}_G\})^3$ such that $\mu(A) > 0$, the set $A \cup A^{-1} \cup \{\text{id}_G\}$ has measure 1 and $\mu(Y) = \mu(X)$ for all definable measurable subsets X and Y of $(A \cup A^{-1} \cup \{\text{id}_G\})^3$ whenever Y is a translate of X .

Example 3.2 Let G be an abelian group, or more generally, an amenable group, equipped with a finitely additive probability measure μ . Every subset of positive measure becomes an acceptable subset of G witnessed by the restriction of μ with respect to a suitable boolean algebra of $(A \cup A^{-1} \cup \{\text{id}_G\})^3$. As above, we can expand the language of groups to a suitable language \mathcal{L} in such a way that both A and the measure μ are definable.

Example 3.3 Consider a finite non-empty subset A of a (possibly infinite group) G with tripling K , that is, with $|A \cdot A \cdot A| \leq K|A|$. Then $B = A \cup A^{-1} \cup \{\text{id}_G\}$ has size at least

$|A|$ and most $2|A| + 1$, and it follows from Ruzsa calculus that $|B \cdot B \cdot B| \leq 14K^3|A|$. Given a subset $X \subseteq B \cdot B \cdot B$, set

$$\mu(X) = \frac{|X|}{|B \cdot B \cdot B|}.$$

We have thus obtained a finitely additive measure μ such that $\mu(A) \geq (14K^3)^{-1}$. Hence, the set A is acceptable (with respect to the measure μ).

Furthermore, if G is abelian, it follows from the Plünnecke–Ruzsa inequality [23, Corollary 6.29] that we need only assume that A has small doubling.

Given an acceptable subset A of G with respect to the finitely additive definable measure μ and a complete type $p(x)$ over a submodel M containing the formula $A(x)$, define its (*measure-theoretic*) stabilizer $\text{Stab}(p)$ to be the group generated by the set

$$\text{st}(p) = \{g \in G : g \cdot p(x) \cup p(x) \text{ is wide}\}.$$

Note that $\text{st}(p)$ contains the identity element of G , whenever the type p is wide. The set $\text{st}(p)$, and hence $\text{Stab}(p)$, is invariant under automorphisms of \mathbb{U} fixing M pointwise.

Inspired by the corresponding results in geometric stability theory, Hrushovski proved in [11, Theorem 3.5] that the measure-theoretic stabilizer is type-definable and equals the set $\text{st}(p) \cdot \text{st}(p)$, whenever p is a wide type containing an acceptable subset $A(x)$ with respect to some (definable) measure μ . Furthermore, the stabilizer is a normal subgroup of the group $\langle A \rangle$ generated by A and has of bounded index in $\langle A \rangle$, that is, the number of cosets of $\text{Stab}(p)$ in $\langle A \rangle$ is bounded by the cardinality of saturation of the ambient model \mathbb{U} . For our purposes, we need a much weaker statement, namely, that $\text{st}(p)$ contains some wide complete type, for which we will now give a simple proof.

Lemma 3.4 *Let A be an acceptable subset of G with respect to the measure μ , which we assume to be definable over a submodel M . Given a wide type $p(x)$ over M containing the formula $A(x)$, there exists a wide type $q(x)$ over M whose realizations belong to the M -invariant set $\text{st}(p)$.*

Proof Consider a sequence $(a_i)_{i \in \mathbb{N}}$ of realizations of p such that $\text{tp}(a_i/M \cup \{a_j\}_{j < i})$ is wide. By a standard Ramsey argument, we may assume that the sequence is indiscernible over M . Set $q = \text{tp}(a_2^{-1} \cdot a_1/M)$. Since $\text{st}(p)$ is M -invariant, it suffices to show that the realization $a_2^{-1} \cdot a_1$ of q belongs to $\text{st}(p)$. Otherwise, the partial type $a_2^{-1} \cdot a_1 \cdot p(x) \cup p(x)$ is not wide and neither is $a_1 \cdot p(x) \cup a_2 \cdot p(x)$ by translation-invariance of the measure (for A is acceptable and a_2 is an element of A). Thus, we can find some M -definable set $X(x)$ in $p(x)$ contained in $A(x)$ such that

$$\mu(a_j \cdot X \cap a_i \cdot X) = 0 \text{ for } i \neq j.$$

As the definable set X is wide, there exists some natural number k such that $\frac{1}{k} < \mu(X) = \mu(a_j \cdot X)$ for all j in \mathbb{N} , so

$$\mu\left(\bigcup_{i=1}^k a_i \cdot X\right) = \sum_{i=1}^k \mu(a_i \cdot X) = k \cdot \mu(X) > 1,$$

which contradicts the assumption that μ is a probability measure. □

A standard application of Ruzsa’s covering argument (cf. [23, Lemma 2.14]) yields the following auxiliary result, which resonates with Lemma 2.4.

Lemma 3.5 *Let A be an acceptable subset of G with respect to the measure μ , which we assume to be definable over a submodel M . Given a wide M -definable subgroup $H \subseteq A \cdot A^{-1}$ of G , there is some finite subset C of $A(M)$ such that A is contained in $C \cdot H$.*

Proof Note that $a \cdot H \subseteq (A \cup A^{-1})^3$ is wide for every a in A . Hence, choose a maximal subset C of A (in \mathbb{U}) such that $(c \cdot H) \cap (c' \cdot H) = \emptyset$ for every $c \neq c'$ in C . In particular, given any a in A , there is some c in C such that $(a \cdot H) \cap (c \cdot H)$ is non-empty. Thus, the element a lies in $c \cdot H$ and so $A \subseteq C \cdot H$.

For each c in C , we have that $\mu(c \cdot H) = \mu(H) > \frac{1}{k}$ for some k in \mathbb{N} . As in the proof of Lemma 3.4, we deduce that C is finite. Since both A and H are definable over the model M and A is contained in $C \cdot H$, we can take C to be a subset of $A(M)$, as desired. □

Proposition 3.6 *Let A be an acceptable subset of G with respect to the measure μ , which we assume to be definable over a submodel M , and assume that the formula $\varphi(x, y) = A(y \cdot x)$ is stable. Then there exists some M -definable subgroup H of G contained in $A \cdot A^{-1}$ such that $A \subseteq C \cdot H$ for some finite subset C of $A(M)$.*

Furthermore, if G is abelian, then H can be taken to be a boolean combination of translates of A .

Proof Since the definable set $A(x)$ is wide, we may extend it to a wide complete \mathcal{L} -type $p(x)$ over M . The φ -type $q = p \upharpoonright_\varphi$ contains the instance $\varphi(x, \text{id}_G) = A(x)$, so the M -definable group $H = \text{Stab}_\varphi(q)$ is clearly contained in $A \cdot A^{-1}$.

In order to conclude the result by Lemma 3.5 (together with Remark 1.2 when G is abelian), it suffices to show that H is wide. By Lemma 3.4, the set

$$\text{st}(p) = \{g \in G(\mathbb{U}) : g \cdot p(x) \cup p(x) \text{ is wide}\}$$

contains a wide type over M . So we need only show that $\text{st}(p) \subseteq H$. Let $g \in \text{st}(p)$ and denote by $q_{|\mathbb{U}}(x)$ in $S_\varphi(\mathbb{U})$ the global non-forking extension of $q = p \upharpoonright_\varphi$. Since formulae of positive measure do not fork over \emptyset , we have that the partial type $g \cdot q(x) \cup q(x)$ does not fork over M , hence it is a restriction of $q_{|\mathbb{U}}(x)$. In particular, the global φ -type $g^{-1} \cdot q_{|\mathbb{U}}(x)$ is a non-forking extension of $q(x)$. By uniqueness of the global non-forking extension, we conclude that

$$g^{-1} \cdot q_{|\mathbb{U}}(x) = q_{|\mathbb{U}}(x),$$

so g^{-1} , and thus g , lies in H as desired. □

4 Main results

We are now in a position to prove our main results. Let us begin by recalling Theorem A in the introduction for the reader’s convenience.

Theorem 4.1 *Given real numbers $K \geq 1$ and $\epsilon > 0$ and a natural number $r \geq 2$, there exists a natural number $n = n(K, \epsilon, r)$ such that for any (possibly infinite) group G and any finite r -stable subset $A \subseteq G$ with tripling K , there is a subgroup $H \subseteq A \cdot A^{-1}$ of G with $A \subseteq C \cdot H$ for some $C \subseteq A$ of size at most n . Moreover, there exists $C' \subseteq C$ such that*

$$|A \Delta (C' \cdot H)| < \epsilon |H|.$$

In particular, we conclude that

$$|A \Delta (C' \cdot H)| < (\epsilon K^2) |A|,$$

by the Plünnecke–Ruzsa inequality.

Remark 4.2 It follows from Proposition 3.6 that when G is abelian, the subgroup H can be taken to be a boolean combination (whose complexity only depends on K, ϵ and r) of translates of A .

Proof The proof proceeds by contradiction. Assuming that the statement does not hold, there are fixed K, ϵ and r such that for each n in \mathbb{N} , we find a finite r -stable subset A_n of a group G_n with tripling K such that there are no subgroup $H \subseteq A_n \cdot A_n^{-1}$ and a finite subset C of A_n of size n with A_n contained in $C \cdot H$ and $|A_n \Delta (C' \cdot H)| < \epsilon |H|$ for some $C' \subseteq C$.

Following the approach of [11, Section 2.6] (see also [2, Proof of Theorem 1.3] and [13, Section 2.3]), we consider a suitable expansion \mathcal{L} of the language of groups and regard each group G_n as an \mathcal{L} -structure M_n . Choose a non-principal ultrafilter \mathcal{U} on \mathbb{N} and consider the ultraproduct $M = \prod_{\mathcal{U}} M_n$. The language \mathcal{L} is chosen in such a way that the sets $A = \prod_{\mathcal{U}} A_n$ and $B = \prod_{\mathcal{U}} B_n$ are \mathcal{L} -definable in the group $G = \prod_{\mathcal{U}} G_n$, where $B_n = A_n \cup A_n^{-1} \cup \{\text{id}_{G_n}\}$. Furthermore, the counting measure

$$\mu_n(X_n) = \frac{|X_n|}{|B_n \cdot B_n \cdot B_n|},$$

induces a definable Keisler measure μ , namely the standard part of $\lim_{\mathcal{U}} \mu_n$, on the boolean algebra of \mathcal{L} -definable subsets of $B \cdot B \cdot B$, such that the set A is acceptable. Now choose a sufficiently saturated elementary extension \mathbb{U} of the model M and note that μ is definable over M . Also, every collection of subgroups of M_n induces an M -definable group in $G(\mathbb{U})$, and vice versa.

Note that the formula $\varphi(x, y) = A(y \cdot x)$ is r -stable, by Łoś's theorem. Moreover, by construction the set A is acceptable. Hence, by Proposition 3.6, there is an M -definable subgroup H contained in $A \cdot A^{-1}$ such that $A \subseteq C \cdot H$, for some finite set C in $A(M)$. By Fact 1.1, after possibly increasing the size of C , we may assume that H is H_φ^0 .

Let C' be the collection of coset representatives c in C such that $A \cap (c \cdot H)$ is wide. Since μ is finitely additive and $A \subseteq C \cdot H$, we have that

$$\mu(A) = \sum_{c \in C} \mu(A \cap (c \cdot H)) = \sum_{c \in C'} \mu(A \cap (c \cdot H)) = \mu(A \cap (C' \cdot H)),$$

so

$$\mu(A \setminus (C' \cdot H)) = \mu(A) - \mu(A \cap (C' \cdot H)) = 0.$$

Thus, in order to compute $\mu(A \Delta (C' \cdot H))$, we need only consider

$$\mu((C' \cdot H) \setminus A) = \sum_{c \in C'} \mu((c \cdot H) \setminus A).$$

For c in C' , note that the definable set $c^{-1}((c \cdot H) \setminus A)$ equals $H \setminus (c^{-1} \cdot A)$, which is contained in $A \cdot A^{-1}$, and hence has a (real-valued) measure. Since $\mu(H) > 0$, the measure μ normalised by $\mu(H)$ induces a left-invariant Keisler measure on the definable subsets of H of the form $X \cap H$, where X is φ -definable over M . By [2, Theorem 2.3 (vi)], such a measure is unique and furthermore wide sets are exactly the generic sets (cf. Fact 1.1). By construction of H_φ^0 , we conclude that $\mu(H \setminus c^{-1} \cdot A) = 0$ for c in C' , so

$$\mu(A \Delta (C' \cdot H)) = \mu(A \setminus (C' \cdot H)) + \mu((C' \cdot H) \setminus A) = 0 < \frac{\epsilon}{2} \cdot \mu(H).$$

However, for $n \geq |C|$ sufficiently large, we conclude by Łoś's theorem that the corresponding trace A_n is contained in $C(A_n) \cdot H(G_n)$ and

$$|A_n \Delta (C'(A_n) \cdot H(G_n))| < \epsilon |H(G_n)|.$$

This yields the desired contradiction. \square

We now turn to the proof of Theorem B in the introduction. First, we note that a straightforward compactness argument yields (non-quantitative) bounds on the complexity of the representation of a weakly normal subset as a boolean combination of suitably chosen subgroups. The fact that the subgroups in Proposition 4.3 below can be expressed as bounded boolean combinations of translates of A goes beyond [19, Theorem 1.4].

Proposition 4.3 *Given natural numbers $r, k,$ and $l,$ there are natural numbers $n = n(r, k, l), m = m(r, k, l)$ and $t = t(r, k, l)$ such that for any abelian group G and any subset $A \subseteq G$ with an (r, k, l) -weakly normal representation, the set A is a boolean combination of complexity at most t of cosets of subgroups H_1, \dots, H_n of $G.$ Moreover, each subgroup H_i is a boolean combination of complexity at most m of translates of $A.$*

Proof Otherwise, as in the proof of Theorem 4.1, assume that, for fixed positive integers r, k and $l,$ there are n, m and t in $\mathbb{N},$ and an abelian group $G_{n,m,t}$ and a subset $A_{n,m,t}$ such that $A_{n,m,t}$ admits an (r, k, l) -weakly normal representation, but it is not a boolean combination of complexity at most t of cosets of n subgroups of $G,$ each of which is a boolean combination of complexity at most m of translates of the set $A_{n,m,t}.$

Set $G_n = G_{n,n,n}$ and $A_n = A_{n,n,n}.$ Since each A_n has an (r, k, l) -weakly normal representation, there are r -weakly normal subsets $B_{n,1}, \dots, B_{n,k}$ and $C_{n,1}, \dots, C_{n,l}$ such that

$$A_n = \bigcup_{j=1}^k B_{n,j} \cap \bigcup_{i=1}^l G_n \setminus C_{n,i}.$$

We consider an expansion \mathcal{L}' of \mathcal{L} with $k + l$ new predicates and regard each group G_n as an \mathcal{L}' -structure $M_n,$ where the predicates are interpreted as the sets $B_{n,1}, \dots, B_{n,k}, C_{n,1}, \dots, C_{n,l}.$ Choose a non-principal ultrafilter \mathcal{U} on $\mathbb{N},$ and consider the ultraproduct $G = \prod_{\mathcal{U}} G_n.$ The definable set $A = \prod_{\mathcal{U}} A_n$ admits an (r, k, l) -weakly normal representation, by Łoś’s theorem, thus the definable set A is a boolean combination of r -weakly normal formulae.

By Corollary 2.5, the definable set A is a boolean combination of complexity at most t_0 of n_0 definable subgroups, each of which is itself a boolean combination of complexity at most m_0 of translates of $A.$

Łoś’s theorem gives the desired contradiction, by choosing $n \geq n_0 + m_0 + t_0$ sufficiently large. □

Theorem 4.4 *Given natural numbers $r, k,$ and $l,$ there are natural numbers $n = n(r, k, l)$ and $m = m(r, k, l)$ such that for any abelian group G and any subset $A \subseteq G$ with an (r, k, l) -weakly normal representation, there are subgroups H_1, \dots, H_n of $G,$ each contained in $A - A,$ with*

$$A \subseteq \bigcup_{i=1}^n g_i + H_i,$$

for some g_1, \dots, g_n in $A.$ Furthermore, each H_i is a boolean combination of complexity at most m of translates of $A.$

Proof As in the proof of Proposition 4.3, if the statement does not hold, there are fixed integers r, k and l such that for each n and m in $\mathbb{N},$ we find an abelian group $G_{n,m}$ and

a subset $A_{n,m}$ such that $A_{n,m}$ admits an (r, k, l) -weakly normal representation, yet

$$A_{n,m} \not\subseteq \bigcup_{i=1}^n g_i + H_i,$$

for any g_1, \dots, g_n in $A_{n,m}$, for all subgroups H_1, \dots, H_n of $G_{n,m}$, each contained in $A_{n,m} - A_{n,m}$, which are boolean combinations of complexity at most m of translates of $A_{n,m}$.

Set $G_n = G_{n,n}$ and $A_n = A_{n,n}$. Choose a non-principal ultrafilter \mathcal{U} on \mathbb{N} , and consider the ultraproduct $G = \prod_{\mathcal{U}} G_n$. Choose a sufficiently saturated elementary extension \mathbb{U} of the model $M = \prod_{\mathcal{U}} M_n$. As observed before, the definable set $A = \prod_{\mathcal{U}} A_n$ admits an (r, k, l) -weakly normal representation, by Łoś’s theorem, so the formula $\varphi(x, y) = A(x + y)$ is a boolean combination of r -weakly normal formulae.

Let \mathcal{F} be the family of cosets $m + H$, where m belongs to $A(M)$ and the definable subgroup $H \subseteq A - A$ over M is given by a finite boolean combination of translates of A . By Łoś’s theorem, no finite collection of \mathcal{F} covers A : otherwise, considering the traces of these subgroups in the corresponding $G_{n,n}$ for sufficiently large n , we would have that $A_{n,n}$ is covered by a finite union of translates of these subgroups, which have bounded complexity as boolean combinations of translates of $A_{n,n}$.

By compactness, we obtain a complete φ -type $p(x)$ over M containing the formula $A(x) = \varphi(x, \text{id}_G)$ such that the partial type

$$p(x) \cup \{\neg(m + H)(x)\}_{m+H \in \mathcal{F}},$$

is consistent. By Remark 1.2, the subgroup $\text{Stab}_\varphi(p)$ is φ -definable, so it is a boolean combination of translates of A . Clearly $\text{Stab}_\varphi(p) \subseteq A - A$, by construction. Lemma 2.4 yields that the type p contains the definable set $(A \cap (m + \text{Stab}_\varphi(p)))(x)$, for some m in M . In particular, the M -definable set $(A \cap (m + \text{Stab}_\varphi(p)))(x)$ is non-empty, so there is a realisation in M , since M is an elementary substructure. Replacing the element m , we may assume that it lies in $A(M)$. Hence, the coset $m + \text{Stab}_\varphi(p)$ belongs to the family \mathcal{F} , which contradicts our choice of p and hence implies the result. \square

As an immediate consequence we obtain the result for finite sets stated as Theorem B in the introduction. An analogous result holds whenever G carries a finitely additive probability measure μ on the boolean algebra of translates of A with $\mu(A) > 0$.

Acknowledgements The authors wish to thank Gabriel Conant and Caroline Terry for many helpful conversations and useful comments on a previous version of this note.

References

1. Casanovas, E.: Simple theories and hyperimaginaries. Lecture Notes in Logic 39, Association Symb. IL; Cambridge University Press, Cambridge, Logic, Chicago (2011)
2. Conant, G., Pillay, A., Terry, C.: A group version of stable regularity. *Math. Proc. Camb. Philos. Soc.* **168**(2), 405–413 (2020)
3. Conant, G.: On finite sets of small tripling or small alternation in arbitrary groups. *Comb. Probab. Comput.* **29**, 807–829 (2020)

4. Conant, G.: Quantitative structure of stable sets in arbitrary finite groups. Proc. Am. Math. Soc. (2021). <https://doi.org/10.1090/proc/15479>
5. Even-Zohar, C., Lovett, S.: The Freiman–Ruzsa theorem over finite fields. J. Combin. Theory Ser. A **125**, 333–341 (2014)
6. Freĭman, G. A.: *Foundations of a structural theory of set addition (Translated from the Russian)*, Transl. Math. Monographs, **37**, AMS, Providence, RI (1973)
7. Green, B., Ruzsa, I.Z.: Freĭman’s theorem in an arbitrary abelian group. J. Lond. Math. Soc. **75**, 163–175 (2007)
8. Green, B., Sanders, T.: A quantitative version of the idempotent theorem in harmonic analysis, Ann. Math. (2) **168**, 1025–1054 (2008)
9. Hrushovski, E., Pillay, A.: Weakly normal groups, in Logic Colloquium 85, Stud. Logic Found. Math. **122**, North-Holland, Amsterdam, 233–244 (1987)
10. Hrushovski, E., Pillay, A.: Groups definable in local fields and pseudo-finite fields. Israel J. Math. **85**, 203–262 (1994)
11. Hrushovski, E.: Stable group theory and approximate subgroups. J. Am. Math. Soc. **25**, 189–243 (2012)
12. Keisler, H.J.: Measures and forking. Ann. Pure Appl. Logic **34**, 119–169 (1987)
13. Palacín, D.: On compactifications and product-free sets, to appear in. J. Lond. Math. Soc. **101**(1), 156–174 (2020)
14. Pillay, A.: Geometric stability theory, Oxford Logic Guides 32. The Clarendon Press, Oxford University Press, Oxford Science Pub, New York (1996)
15. Pillay, A.: Model-theoretic consequences of a theorem of Campana and Fujiki. Fund. Math. **174**, 187–192 (2002)
16. Ruzsa, I.Z.: Generalized arithmetical progressions and sumsets. Acta Math. Hungar. **65**, 379–388 (1994)
17. Ruzsa, I.Z.: An analog of Freĭman’s theorem in groups. Astérisque **258**, 323–326 (1999)
18. Sanders, T.: On the Bogolyubov–Ruzsa lemma. Anal. PDE **5**, 627–655 (2012)
19. Sanders, T.: The coset and stability rings. Online J. Anal. Comb. **15**, 1 (2020)
20. Schoen, T.: Near optimal bounds in Freĭman’s theorem. Duke Math. J. **158**, 1–12 (2011)
21. Sisask, O.: Convolutions of sets with bounded VC-dimension are uniformly continuous, Discrete Anal. 2021:1
22. Starchenko, S.: *NIP, Keisler measures and combinatorics*, in Séminaire Bourbaki, Astérisque **390**, 303–334 (2017)
23. Tao, T., Vu, V.: Additive Combinatorics, Cambridge Studies in Advanced Mathematics 105. Cambridge University Press, Cambridge (2006)
24. Tent, K., Ziegler, M.: A course in model theory. Lecture Notes in Logic 40, Association Symb. CA; Cambridge University Press, Cambridge, Logic, La Jolla (2012)
25. Terry, C., Wolf, J.: Quantitative structure of stable sets in finite abelian groups. Trans. Am. Math. Soc. **373**, 3885–3903 (2020)
26. Terry, C., Wolf, J.: Stable arithmetic regularity lemma in the finite-field model. Bull. London Math. Soc. **51**, 70–88 (2019)