Canad. J. Math. Vol. 73 (1), 2021 pp. 63–107
http://dx.doi.org/10.4153/S0008414X1900049X
© Canadian Mathematical Society 2019

The Erdős–Moser Sum-free Set Problem

Tom Sanders

Abstract. We show that there is an absolute $c > 0$ such that if A is a finite set of integers, then there is a set $S \subset A$ of size at least $\log^{1+c}|A|$ such that the restricted sumset $\{s + s' : s, s' \in S \text{ and } s \neq s'\}$ is disjoint from A. (The logarithm here is to base 3.)

1 Introduction

In this paper we are interested in some problems in additive combinatorics. The standard introduction to this area is the book [TV06] by Tao and Vu, and we have tried to give references to this book where possible.

Given finite sets A and S in an abelian group, we write $A+S$ for the sumset $\{a+s : a \in A \text{ and } s \in S\}$ and $A\widehat{+}S$ for the restricted sumset $\{a + s : a \in A, s \in S \text{ and } a \neq s\}$. Erdős [Erd65, p. 187] describes joint work with Moser in which they investigate the following question: if A is a finite set of integers, then what is the size of the largest set $S \subset A$ such that $(S\widehat{+}S) \cap A = \varnothing$? They consider the restricted sumset to make the problem non-trivial: if A is a set of consecutive powers of 2 and $S \subset A$ has $(S+S)\cap A = \varnothing$, then S has size at most 1.

To discuss the problem we need a definition: for a finite set A of integers, define

$$M(A) := \max\{|S| : S \subset A \text{ and } (S\widehat{+}S) \cap A = \varnothing\}.$$

We are interested in lower bounds on $M(A)$ that are uniform in the size of A. We record the history relevant to our interests below. [TV06, §6.2.1] contains more details, and a full survey can be found in [TV17].

In [Erd65] a simple example is given to show (that for any natural number N there is a set A of size N such) that $M(A) \leq \frac{1}{3}|A| + O(1)$. This was improved, first by Selfridge [Erd65, p. 187], then by Choi [Erd65, p. 190], and then more substantially by Choi [Cho71, (2)] where it is shown that $M(A) \leq |A|^{2/5+o(1)}$. The $o(1)$-term was refined by Baltz, Schoen, and Srivastav in [BSS00, Corollary 3][1] before Ruzsa adapted a classical construction of Behrend [Beh46] in [Ruz05, Theorem] to show the following theorem.

Theorem 1.1 (Ruzsa) *Given a natural number, there is a set A of that size such that*

$$M(A) = \exp(O(\sqrt{\log|A|})).$$

Received by the editors May 17, 2018; revised July 31, 2019.
Published online on Cambridge Core September 23, 2019.
AMS subject classification: **11P70**.
Keywords: sum-free, summing, Freiman's theorem, additive energy, hereditarily energetic.
[1]Equivalently [BSS99, Corollary 1].

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

In the other direction, Erdős and Moser showed that $M(A) \to \infty$ as $|A| \to \infty$, and Klarner showed that $M(A) = \Omega(\log |A|)$ (both results are mentioned in [Erd65, p187] though the proofs, or at least Klarner's, seem to have been lost [Cho71, †, p. 630]). Ruzsa showed that $M(A) > 2\log_3 |A| - 1$ in [Ruz05, Theorem] by a greedy algorithm, and then Sudakov, Szemerédi, and Vu made an important breakthrough in [SSV05], giving the first super-logarithmic lower bound on $M(A)$ in [SSV05, Theorem 1.1]. Their argument was improved by Dousse [Dou13, §4] and then Shao [Sha15, Corollary 1.3], who showed that $M(A) = (\log\log |A|)^{\frac{1}{2}-o(1)} \log |A|$. We shall show the following theorem.

Theorem 1.2 *For every finite set of integers A, we have*

$$M(A) = \log^{1+\Omega(1)} |A|.$$

The result as recorded in the abstract follows immediately from this.[2]

Our argument has two parts: the first makes use of specific properties of the integers and is dealt with in Section 3. The second part does not (at least it works in any abelian group with no 2-torsion), and is covered in the remainder of the paper from Section 4 onwards. Indeed, because the integers do not play a significant role, we are able to give a model argument in Section 4, and we hope the reader familiar with the area will be able to understand the main ideas of our proof from Sections 3 and 4 alone.

Our approach falls within the general strategy proposed by [SSV05], so we begin in the next section with an overview of it, which also serves to explain how the improvements of Dousse and Shao arise.

2 Overview of the Sudakov–Szemerédi–Vu Strategy

Given sets A and X in an abelian group, we say that A is (k, X)-summing[3] if for any set $S \subset A$ with $|S| \geq k$, we have $(S \dot{+} S) \cap X \neq \emptyset$. (We shall always take $k \geq 2$.)

The following proposition is the focus of the Sudakov–Szemerédi–Vu strategy.

Proposition 2.1 *Suppose that $A \subset X \subset \mathbb{Z}$ have $|X| \leq (1+\eta)|A|$ and A is (k, X)-summing for some $k \in \mathbb{N}$. Then either $\eta = k^{-O(1)}$ or $|A| \leq F(k)$ for some universal (monotonically increasing) function $F\colon \mathbb{N} \to \mathbb{N}$.*

[2] For $|A| \in \{1, 2, 3\}$, it is trivial, since $\log_3 |A| \leq 1$ and any $S \subset A$ of size 1 has $S \dot{+} S = \emptyset$. On the other hand Theorem 1.2 immediately gives the result in the abstract for $|A| \geq C$ for some absolute $C > 0$. Finally, for $4 \leq |A| < C$ [Ruz05, Theorem] shows that

$$M(A) > 2\log_3 |A| - 1 > \Big(1 + \frac{\log 4 - \log 3}{\log 4}\Big) \log_3 |A| > \exp\Big(\frac{1}{6}\Big) \log_3 |A|$$
$$> \log_3^{1+\frac{1}{6\log\log_3 C}} |A|.$$

The result is proved. (Note that if $A = \{-1, 0, 1\}$, then $M(A) = 1$, and so the best lower bound on $M(A)$ over all size 3 sets is 1, which is the reason for taking logarithms to the base 3.)

[3] Our terminology is not standard. In [SSV05] the authors say a set S is sum-free with respect to X if $(S \dot{+} S) \cap X = \emptyset$. In [TV16] the authors say that a set S is summing in X for the same thing.

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

[SSV05, Theorem 1.2] says we can take

$$F(k) = \exp\big(\exp\big(\exp(\exp(\exp(O(k))))\big)\big).$$

Applying Proposition 2.1 with $X = A$ gives us that $M(A) \to \infty$ as $|A| \to \infty$, but the point of the proposition is the extra flexibility afforded by being able to take X to be a little larger than A. This means that it can be bootstrapped to give [SSV05, Theorem 1.1] (as is done in [SSV05, §2], and as we will do in Section 3), and in general, given F, one gets

$$M(A) = \Omega\Big(\frac{F^{-1}(|A|)}{\log F^{-1}(|A|)} \log |A|\Big). \tag{2.1}$$

Many tools in additive combinatorics do not distinguish between different abelian groups, and so to make use of them, one needs to be in a situation where the conclusion does not depend on the underlying group. In this case, if we try to replace the integers in Proposition 2.1 with a general abelian group we run into the problem that A might be a subgroup. In some sense this is the only obstacle, as shown by Tao and Vu.

Theorem 2.2 ([TV16, Theorem 1.2]) *Suppose that G is an abelian group and $A \subset G$ is finite and (k, A)-summing. Then there is an integer $m \le k$ and subgroups $H_1, \dots, H_m \le G$ such that $|A \setminus (H_1 \cup \cdots \cup H_m)| = O_k(1)$ and $|A \cap H_i| = \Omega_k(|H_i|)$ for $1 \le i \le m$.*

The order property of the integers (not available in general abelian groups) is used in the derivation of (2.1) from Proposition 2.1, and the integers are also used essentially in Lemma 2.3 (and hence Lemma 2.5), but all other uses in this discussion section are for convenience.

Sudakov, Szemerédi, and Vu capture a property of the integers that eliminates the subgroup examples of Theorem 2.2 in the next lemma. For this we require a definition. Given a set A in an abelian group G, the additive energy of A is defined[4] to be

$$E(A) := \|1_A * 1_{-A}\|_{\ell_2(G)}^2 = \sum_z \Big(\sum_y 1_A(y) 1_{-A}(z-y)\Big)^2.$$

Lemma 2.3 *Suppose that $X \subset \mathbb{Z}$ has $E(X) \ge \eta |X|^3$. Then there is a set $X' \subset X$ such that $|X'| \ge \eta^{O(1)}|X|$ and $(2 \cdot X') \cap X = \varnothing$.*

The proof of this is [SSV05, Lemma 5.2] coupled with the Balog–Szemerédi–Gowers Theorem,[5] but we only need the statement for our discussion so do not record the details. The rough idea (which we shall use to prove Lemma 3.4) is to partition $\mathbb{Z}$ into sets $T_i := \{x \in \mathbb{Z} : 2^i \mid x \text{ and } 2^{i+1} \nmid x\}$ for $0 \le i \le \infty$. There cannot be a lot of additive quadruples with each element in a different T_i; we then essentially take the largest index i where the intersection with X is not too small.

[4] See [TV06, Definition 2.8] for a discussion.

[5] In its usual form, which corresponds to [TV06, Theorem 2.31((i) ⇒ (iv))] and then [TV06, Exercise 2.3.15]).

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

It turns out that a short application of Turán's theorem from graph theory (essentially [SSV05, Lemma 3.1]) shows that if A is (k,X)-summing, then A has large additive energy. One can think of this as saying a large part of A is highly structured.

Lemma 2.4 *Suppose that G is an abelian group and $A, X \subset G$ are such that $|X| \leq K|A|$ and A is (k,X)-summing. Then either $|A| = k^{O(1)}$ or $E(A) = (kK)^{-O(1)}|A|^3$.*

The above lemma is a consequence of Lemma 3.2 proved later, though the method of [SSV05, Lemma 3.1] gives better constants for the $O(1)$-terms; again we only need the above form for the discussion.

In fact Lemma 3.2 is importantly stronger than Lemma 2.4 and tells us that if A is (k,X)-summing, then either A is small or else every subset of A that is not small has large energy. We formalise this in Section 3 in the notion of "hereditarily energetic". With this additional fact Lemmas 2.3 and 2.4 can be combined to give the following.

Lemma 2.5 *Suppose that $A \subset X \subset \mathbb{Z}$ are such that $|X| \leq (1+\eta)|A|$ and A is (k,X)-summing. Then either $|A| = k^{O(1)}$; or $\eta \geq k^{-O(1)}$; or there is a set $A' \subset A$ with $|A'| \geq k^{-O(1)}|A|$ such that $(2\cdot A') \cap X = \varnothing$ and $E(A') \geq k^{-O(1)}|A'|^3$.*

Again we omit the details, as we only need the statement for discussion. The key point is that if we are not in the first two outcomes, then after applying Lemma 2.4 to get that A has large additive energy, we apply Lemma 2.3 to X (which inherits large additive energy from A), and since η is small enough, the set X' in Lemma 2.3 is large enough that it necessarily has large intersection with A.

Although it was more unusual at the time of [SSV05], it is now commonplace to apply the Balog–Szemerédi–Gowers–Freiman machinery in this sort of situation. This tells us that if $A \subset \mathbb{Z}$ and $E(A) \geq \eta|A|^3$, then there is an arithmetic progression P such that

$$|P| \geq |A|^{\eta^{o(1)}} \quad \text{and} \quad |A \cap P| \geq \eta^{O(1)}|P|. \tag{2.2}$$

This result with the $o(1)$-term replaced by $O(1)$ follows from [TV06, Theorem 2.29] combined with [TV06, Theorem 5.32]; the stronger bounds require the replacement of [TV06, Theorem 2.29] by the improved estimates of Schoen [Sch11].

For our purposes, arithmetic progressions are the same as intervals, and the final ingredient we need is the following proposition.

Proposition 2.6 *Suppose that $A \subset \{1,\dots,N\}$ has size αN and there is no proper k-tuple $(a_1,\dots,a_k) \in A^k$ with $a_i + a_j \in 2\cdot A$ for all $i<j$. Then $N \leq F'(\alpha,k)$ for some universal function $F' : (0,1]\times\mathbb{N} \to \mathbb{N}$ (decreasing in the first coordinate and increasing in the second).*

To see why this is enough, apply the proposition to the output A' of Lemma 2.5 after applying (2.2). If $S \subset A'$ has $s+s' \in 2\cdot A'$ for all $s \neq s' \in S$, then since $(2\cdot A') \cap X = \varnothing$, we have that $(S\widehat{+}S) \cap X = \varnothing$, and so A' is not (k,X)-summing. It follows that we can take

$$F(k) = F'(k^{-O(1)},k)^{k^{-o(1)}}$$

in Proposition 2.1.

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Sudakov, Szemerédi, and Vu [SSV05, Corollary 3.3] proved that

$$F'(\alpha, k) \leq \exp\left(\exp\left(\alpha^{-\exp(\exp(O(k)))}\right)\right)$$

by using Gowers' bounds for Szemerédi's Theorem. Dousse noted that the system being counted in Proposition 2.6 has complexity 1 (in the sense of [GT10, Definition 1.5]), so one can study the configurations in Proposition 2.6 using Fourier analysis rather than the higher order analogues of Gowers. This is cheaper, and she proved [Dou13, Corollary 3.4] that

$$F'(\alpha, k) \leq \exp\left(\exp\left(\alpha^{-O(k^2)}\right)\right).$$

Finally Shao implemented Dousse's Fourier argument in Bohr sets (following Bourgain [Bou99] for three-term progressions, which is the special case $k = 2$ of Proposition 2.6; see [TV06, Theorem 10.29] for an exposition) to show [Sha15, Theorem 1.3] that

$$F'(\alpha, k) \leq \exp\left(\alpha^{-O(k^2)}\right).$$

(In fact, this does not quite do justice to Shao's work; Dousse used a weaker version of the Balog–Szemerédi–Gowers–Freiman machinery that arises from applying a version of Freiman's theorem providing a progression containing the whole of the set A, rather than a progression inside $2A - 2A$. Shao noted that Ruzsa's Embedding Lemma [TV06, Lemma 5.26] suffices and gives better bounds, although those improvements do not impact the level of our discussion above.)

It is natural to ask what sort of bounds we can expect on F'. The worst example of bad α-dependence in F' comes from Behrend's construction [Beh46] (see [Elk10, GW10] for the state of the art), which tells us that $F'(\alpha, 2) = \exp(\Omega(\log^2 \alpha^{-1}))$. On the other hand, if we choose $A \subset \{1, \dots, N\}$ by selecting elements independently with probability α, then[6] there is a choice of $\alpha = \Omega(1)$ such that any $S \subset A$ with $S \widehat{+} S \subset 2 \cdot A$ has $|S| = O(\log N)$, from which it follows that $F'(\Omega(1), k) = \exp(\Omega(k))$. By monotonicity of F' in each of its variables we conclude that

$$F'(\alpha, k) = \exp(\Omega(k + \log^2 \alpha^{-1})), \tag{2.3}$$

and it is a natural question to ask if one can do better.

[6] Any $S \subset \{1, \dots, N\}$ has $\mathbb{P}(S \widehat{+} S \subset 2 \cdot A) \leq \alpha^{|S \widehat{+} S|}$. Moreover, if S has size k then $|S \widehat{+} S| \geq 2k - 3$, as can be seen by writing $S = \{s_1 < \cdots < s_k\}$ and noting that $s_1 + s_2, \dots, s_1 + s_k, s_k + s_2, \dots, s_k + s_{k-1}$ are distinct elements. By embedding $\{1, \dots, N\}$ in $\mathbb{Z}/2N\mathbb{Z}$ and applying [Gre05a, Proposition 23] it follows that

$$\begin{aligned}
&\mathbb{E}|\{S \subset A : |S| = k \text{ and } (S \widehat{+} S) \subset 2 \cdot A\}| \\
&\quad \leq \mathbb{E}|\{S \subset \{1, \dots, N\} : |S| = k \text{ and } (S \widehat{+} S) \subset 2 \cdot A\}| \\
&\quad \leq \sum_{m=2k-3}^{\binom{k}{2}} \alpha^m |\{S \subset \{1, \dots, N\} : |S| = k \text{ and } |S \widehat{+} S| = m\}| \\
&\quad \leq \sum_{m=2k-3}^{\binom{k}{2}} \alpha^m (2N)^{1+O(\frac{m}{k})} O\Big(\frac{m}{k}\Big)^{O(k)}.
\end{aligned}$$

If $k \geq C \log N$ for some absolute $C > 0$ sufficiently large, then each term in this last sum is $\alpha^m \exp(O(k \log m k^{-1})) = \exp(m(O(1) - \log \alpha^{-1}))$ from which it follows that we can choose $\alpha = \Omega(1)$ such that the expectation is strictly less than $\frac{1}{2}$. On the other hand, for N sufficiently large $\mathbb{P}(|A| \geq \frac{1}{2}\alpha N) > \frac{1}{2}$, and it follows that there is some set A with the claimed property.

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Although we may not have given the best combination of examples above, as far as we know, it might be that $F'(\alpha, k) = \exp(O(k\alpha^{-o(1)}))$. This would imply a considerable strengthening of Roth's theorem[7] and also that $F(k) \leq \exp(k^{1+o(1)})$, which in turn would give $M(A) \geq \log^{2-o(1)} |A|$ improving Theorem 1.2. Any bound of the form $F'(\alpha, k) = \exp((k\alpha^{-1})^{O(1)})$ already leads to $F(k) = \exp(k^{O(1)})$ and a different proof of Theorem 1.2. That being said, it is not completely clear how to get a singly rather than doubly exponential dependence on k in F', since Fourier methods seem to rely on regularising a set that is exponentially small in k; perhaps a first step of showing $F'(\alpha, k) \leq \exp(\alpha^{-O(k)})$ is within reach of those methods.

We do not prove Theorem 1.2 by proving better bounds for F'. Our advantage comes from one weakness of the above: all we are looking for is sums from A that are in A^ϵ, in principle a much less demanding condition (at least if A is thin) than that in Proposition 2.6 where we ask that they are in $2 \cdot A$. We explain this further in Section 4.

The aim of the remainder of the paper is to prove the following quantitative version of Proposition 2.1.

Proposition 2.7 *Suppose that $A \subset X \subset \mathbb{Z}$, $|X| \leq (1+\eta)|A|$, and A is (k, X)-summing for some $k \in \mathbb{N}$. Then either $\eta = k^{-O(1)}$, or $|A| \leq \exp(k^{C+o(1)})$ for some absolute $C > 0$.*

A value of C could be calculated from our work, but there is considerable scope for optimising it. We have tried to indicate some places this might be possible, but it also seems likely that some of the convenient decoupling of the argument (for example the introduction of hereditarily-energetic sets in the next section) might be lost in a very careful optimisation.

It seems quite possible that $C = 1$ is achievable, though it is not clear that one can expect to do better without a new idea, and in light of (2.3), one cannot do better by trying to improve F' in Proposition 2.6. Given this, it seems to us that the natural next question is whether or not $M(A) = \log^{2+\Omega(1)} |A|$.

3 Overview of Our Argument

In this section, we decouple our arguments into those that require order and divisibility properties of the integers and those that in some sense do not. The latter are captured by Proposition 3.5 below; everything else is covered in the present section.

We begin by using the order structure on the integers to prove Theorem 1.2 using Proposition 2.7. The argument is included for completeness; it is essentially[8] the same as the bootstrapping of [SSV05, Theorem 1.2] to get [SSV05, Theorem 1.1] in [SSV05, §2].

Proof of Theorem 1.2 Let $k, m \in \mathbb{N}$ be parameters to be optimised; let $l \in \mathbb{N}$ be such that $l = k^{O(1)}$, and if $\eta \geq l^{-1}$, then the first conclusion of Proposition 2.7 does not hold; and let $D = \exp(k^{O(1)})$ be a natural number such that if $|A| \geq D$, then the second conclusion does not hold.

[7] The current best bounds there are due to Bloom [Blo16] and imply that $F'(\alpha, 2) \leq \exp(\alpha^{-1-o(1)})$.

[8] Our argument is slightly sloppier, which we can afford, because of the strength of Proposition 2.7.

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

We construct sets iteratively as follows: let $m \in \mathbb{N}_0$ be such that $2D(2l(mk+1))^m \le |A|$, and for $0 \le i \le m$ let Z_i be the $2D(2l(mk+1))^i$ largest elements of A so that $|Z_i| \ge 2l(mk+1)|Z_{i-1}|$. For $0 \le r < m$, we will define sets $S_0, \dots, S_r$ of size k such that $(S_i \widehat{+} S_i) \cap Z_i = \varnothing$ and an auxiliary sequence of A_is with $S_i \subset A_i$ and

$$A_0 := Z_0 \text{ and } A_{i+1} := Z_{i+1} \setminus \Big(Z_i - \Big(\{0\} \cup \bigcup_{j \le i} S_j \Big) \Big) \text{ for } 0 \le i \le r.$$

For clarity we note that $\setminus$ denotes relative complement of sets here and $-$ denotes the difference of two sets. If $i < m$, then

$$\begin{aligned} |Z_{i+1}| &\le |A_{i+1}| + |Z_i| \Big(1 + \sum_{j \le i} |S_j| \Big) \\ &\le |A_{i+1}| + \frac{1}{2l(mk+1)} |Z_{i+1}|(1+mk) = |A_{i+1}| + \frac{1}{2l} |Z_{i+1}|. \end{aligned}$$

Rearranging, we have that $|Z_{i+1}| \le (1+l^{-1})|A_{i+1}|$ and by design $A_{i+1} \subset Z_{i+1}$. Moreover, $|A_{i+1}| \ge \frac{1}{2}|Z_{i+1}| \ge D$, and so it follows from Proposition 2.7 that A_{i+1} is not (k, Z_{i+1})-summing; equivalently, there is a set $S_{i+1} \subset A_{i+1}$ with $(S_{i+1} \widehat{+} S_{i+1}) \cap Z_{i+1} = \varnothing$, as required.

For $0 \le r < m$, consider the set $S := \bigcup_{i=0}^{r} S_i$. Then

$$\begin{aligned} (S \widehat{+} S) \cap A &= \Big(\Big(\bigcup_{0 \le i < j \le r} S_i + S_j \Big) \cup \Big(\bigcup_{i=0}^{r} S_i \widehat{+} S_i \Big) \Big) \cap A \\ &= \Big(\bigcup_{0 \le i < j \le r} (S_i + S_j) \cap A \Big) \cup \Big(\bigcup_{i=0}^{r} (S_i \widehat{+} S_i) \cap A \Big). \end{aligned}$$

Suppose that $0 \le i \le r$. Since $S_i \subset Z_i$ (of the largest $2D(2l(mk+1))^i$ elements of A) and when two positive integers are added their size increases, we see that $(S_i \widehat{+} S_i) \cap A = (S_i \widehat{+} S_i) \cap Z_i = \varnothing$.

Then suppose that $0 \le i < j \le r$. By similar reasoning $(S_i + S_j) \cap A = (S_i + S_j) \cap Z_{j-1}$ and $S_j \cap (Z_{j-1} - S_i) = \varnothing$ by design, so $(S_i + S_j) \cap A = \varnothing$.

Combining these we see that $(S \widehat{+} S) \cap A = \varnothing$. Since the sets $(A_i)_i$ are disjoint, so are the S_is, so $|S| \ge mk$, and our task is to maximise this subject to $2D(2l(mk+1))^m \le |A|$. We can certainly take $k = \log^{\Omega(1)} |A|$ and $2D \le \sqrt{|A|}$ and $m = \Omega(\log|A| / \log\log|A|)$ such that $(2l(mk+1))^m \le \sqrt{|A|}$, from which the result follows. ∎

It will be useful to have some notation for the Fourier transform. This is developed in [TV06, Chapter 4], but we will use different conventions. Suppose that G is an abelian group. Given $f, g \in \ell_1(G)$, we write $f * g$ for the convolution of f and g, defined point-wise by

$$f * g(x) := \sum_z f(z) g(x - z) \text{ for all } x \in G.$$

We write $\widehat{G}$ for the compact abelian group of characters on G and define the Fourier transform of $f \in \ell_1(G)$ to be

$$\widehat{f} \colon \widehat{G} \longrightarrow \mathbb{C}; \quad \gamma \longmapsto \sum_{z \in G} f(z) \overline{\gamma(z)}.$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

The group $\widehat{G}$ is endowed with a Haar probability measure in such a way that we have Plancherel's theorem (see [Rud90, Theorem 1.6.1]):

$$\sum_x f(x)^2 = \int |\widehat{f}(\gamma)|^2 d\gamma \quad \text{for all } f.$$

Sets with small doubling (see [TV06, §2.2]) and large additive energy [TV06, Theorem 2.31] are staples of additive combinatorics. We need an intermediate concept. We say that $A \subset G$ is ν-hereditarily energetic if

$$E(S) \geq \nu\sigma|S|^3 \quad \text{for all } S \subset A \text{ with } |S| \geq \sigma|A|.$$

If A has $E(A) \geq \nu|A|^3$ and $S \subset A$ is chosen independently at random with probability σ, then typically

$$|S| \approx \sigma|A| \text{ and } E(S) \gtrsim \nu\sigma^4|A|^3 \approx \nu\sigma|S|^3;$$

the definition says we never do worse than this. We take the name from the related concept of hereditarily non-uniform sets defined in [Gre02, §2], and the notion is implicit in numerous papers.

There are many examples including the α-spectrum [TV06, Definition 4.33], which we shall not discuss,[9] and symmetry sets, which we shall. Recall, following [TV06, Definition 2.32], that if $\nu \in (0,1]$ and $X \subset G$, then the symmetry set of X at threshold ν is the set

$$\mathrm{Sym}_\nu(X) := \{x \in G : 1_X * 1_{-X}(x) > \nu|X|\}.$$

Lemma 3.1 (Basic facts about hereditarily energetic sets)

(i) (Sets with small doubling) *Suppose that* $|A+A| \leq K|A|$. *Then* A *is* K^{-1}*-hereditarily energetic.*

(ii) (Monotonicity) *Suppose that* A *is* ν*-hereditarily energetic and* $A' \subset A$ *has size* $\epsilon|A|$. *Then* A' *is* $\epsilon\nu$*-hereditarily energetic.*

(iii) (Unions) *Suppose that* A *and* A' *are* ν*-hereditarily energetic. Then* $A \cup A'$ *is* $\Omega(\nu)$*-hereditarily energetic.*

(iv) (Symmetry sets) *Suppose that* $E(A) \geq \nu|A|^3$. *Then* $\mathrm{Sym}_{\frac{1}{2}\nu}(A)$ *is* $\Omega(\nu^4)$*-hereditarily energetic.*

Proof For (i) suppose $S \subset A$ has $|S| \geq \sigma|A|$. Then

$$\|1_S * 1_{-S}\|_{\ell_2(G)}^2 \geq \frac{|S|^4}{|S+S|} \geq \frac{|S|^4}{|A+A|} \geq K^{-1}\sigma|S|^3,$$

and we have the claim; (ii) is trivial. For (iii), suppose that $S \subset A \cup A'$ has size $\sigma|A \cup A'|$. Then without loss of generality, $|S \cap A| \geq \frac{1}{2}\sigma|A \cup A'|$. It follows that

$$E(S) \geq E(S \cap A) \geq \nu\frac{1}{|A|}|S \cap A|^4 \geq \frac{\nu}{8}\sigma|S|^3,$$

as claimed.

[9] If a set has large additive energy then the large spectrum is large (for a suitable threshold) by Parseval's theorem. The case $k = 2$ of [Shk08, Theorem 5], then essentially shows that the large spectrum is hereditarily energetic in a similar way to Lemma 3.1(iv).

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Finally, to prove (iv), note that since $E(A) \geq \nu|A|^3$, we have $|\operatorname{Sym}_{\frac{1}{2}\nu}(A)| \geq \frac{1}{2}\nu|A|$. Suppose that $S \subset \operatorname{Sym}_{\frac{1}{2}\nu}(A)$. Then

$$E(S)|A|^5 \geq \|\widehat{1_S}\|_{L_4(\widehat{G})}^4 \||\widehat{1_A}|^2\|_{L_{\frac{4}{3}}(\widehat{G})}^4 \geq \left|\langle \widehat{1_S}, |\widehat{1_A}|^2\rangle_{L_2(\widehat{G})}\right|^4$$
$$= |\langle 1_S, 1_A * 1_{-A}\rangle_{\ell_2(G)}|^4 \geq \left(\frac{1}{2}\nu|S||A|\right)^4.$$

The final result follows. ∎

In view of (i) and (iii) above, the union of two sets with small doubling is hereditarily energetic, but of course the union need not have small doubling so the notion is strictly weaker than having small doubling. At the other end of the range, if A is ν-hereditarily energetic, then it has large additive energy but the converse need not be true, as can be seen by taking any set of large additive energy and adjoining a dissociated set of the same size.

We formulate the version of Lemma 2.4 we need as follows.

Lemma 3.2 *Suppose that G is an abelian group and $A, X \subset G$ are such that $|X| \leq K|A|$ and A is (k, X)-summing. Then A is $(kK)^{-O(1)}$-hereditarily energetic.*

Proof Suppose that $|A| \geq k^4$, and note that if $x \in A^k$, then there are two indices $i \neq j$ such that either $x_i = x_j$ or $x_i + x_j \in X$. It follows that if $S \subset A$, we have

$$\sum_{x \in S^k} \sum_{i<j} 1_X(x_i + x_j) \geq |S|^k - \binom{k}{2}|S|^{k-1}.$$

Defining σ by $|S| = \sigma|A|$, we conclude that either $|S| \leq k^2$, in which case $E(S) \geq |S|^2 \geq \sigma|S|^3$; or $|S| > k^2$ in which case

$$\binom{k}{2}|S|^{k-2}\langle 1_S * 1_S, 1_X\rangle_{\ell_2(G)} = \sum_{x \in S^k} \sum_{i<j} 1_X(x_i + x_j) \geq \frac{1}{2}|S|^k,$$

but then,

$$\frac{1}{k^2}|S|^2 \leq \langle 1_S * 1_S, 1_X\rangle_{\ell_2(G)} = \int \widehat{1_S}(\gamma)^2\overline{\widehat{1_X}(\gamma)}d\gamma \leq E(S)^{\frac{1}{2}}|X|^{\frac{1}{2}},$$

and so

$$E(S) \geq \frac{1}{k^4} \cdot \frac{|S|}{|X|} \cdot |S|^3 \geq \frac{1}{Kk^4}\sigma|S|^3.$$

The result is proved. ∎

We need a little notation. We use the language of 2-adic valuations, but this is only for brevity. Direct discussion of divisibility properties rather, say, than the use of the ultra-metric property, is very easy.

If $z \in \mathbb{Z}$, we write $|z|_2$ for the 2-adic valuation of z, that is $|z|_2 = 2^{-i}$ where $2^i \mid z$ and $2^{i+1} \nmid z$, with the convention that $|0|_2 = 0$. For $A \subset \mathbb{Z}$, we write

$$A_i := \{z \in A : |z|_2 = 2^{-i}\} \quad \text{for all } i \in \mathbb{N}_0 \cup \{\infty\}$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

(with the convention that $2^{-\infty}=0$). Thus, $\{A_i : i \in \mathbb{N}_0 \cup \{\infty\}\}$ is a set of disjoint sets whose union is A. For $\epsilon \in (0,1]$, we write

$$I_\epsilon(A) := \{i \in \mathbb{N}_0 \cup \{\infty\} : |A_i| > \epsilon|A|\} \quad \text{and} \quad A_\epsilon := \bigcup \{A_i : i \in I_\epsilon(A)\};$$

note that $|I_\epsilon(A)| < \epsilon^{-1}$. This method of decomposing A is used in the proof of [SSV05, Lemma 5.2], and the following lemma is very much in the spirit of that result.

Lemma 3.3 *Suppose that $A \subset \mathbb{Z}$ is ν-hereditarily energetic, and $\epsilon \in (0,1]$ is a parameter. Then $|A \setminus A_{\epsilon^2\nu}| = O(\epsilon|A|)$.*

Proof Suppose that $(a_1,a_2,a_3,a_4) \in A^4$ are such that $a_1+a_2=a_3+a_4$. Let i be such that $|a_i|_2$ is maximal, and note that $|a_i| = |a_j + a_k - a_l|$ for any j,k,l with $\{i,j,k,l\} = \{1,2,3,4\}$. Since the 2-adic valuation induces an ultra-metric $|a_i|_2 \leq \max\{|a_j|_2, |a_k|_2, |a_l|_2\} \leq |a_i|_2$, so there is some $j \in \{1,2,3,4\}$ with $j \neq i$ such that $|a_j|_2 = |a_i|_2$. Writing $A^- := A \setminus A_{\epsilon^2\nu}$, it follows that

$$\begin{aligned} E(A^-) \leq \sum_{i \notin I_{\epsilon^2\nu}(A)} \big(& \langle 1_{A_i} * 1_{-A_i}, 1_{A^-} * 1_{-A^-}\rangle_{\ell_2(\mathbb{Z})} \\ & + \langle 1_{A_i} * 1_{-A^-}, 1_{A_i} * 1_{-A^-}\rangle_{\ell_2(\mathbb{Z})} \\ & + \langle 1_{A_i} * 1_{-A^-}, 1_{A^-} * 1_{-A_i}\rangle_{\ell_2(\mathbb{Z})} + \langle 1_{A^-} * 1_{-A_i}, 1_{A_i} * 1_{-A^-}\rangle_{\ell_2(\mathbb{Z})} \\ & + \langle 1_{A^-} * 1_{-A_i}, 1_{A^-} * 1_{-A_i}\rangle_{\ell_2(\mathbb{Z})} + \langle 1_{A^-} * 1_{-A^-}, 1_{A_i} * 1_{-A_i}\rangle_{\ell_2(\mathbb{Z})} \big) \\ \leq 6 \sum_{i \notin I_{\epsilon^2\nu}(A)} & |A_i|^2|A^-| \leq 6\epsilon^2\nu|A||A^-| \sum_{i \notin I_{\epsilon^2\nu}(A)} |A_j| = 6\epsilon^2\nu|A||A^-|^2. \end{aligned}$$

On the other hand, the left-hand side is at least $\nu|A^-|^4/|A|$, since A is ν-hereditarily energetic. The result follows on rearranging. ∎

We shall use the above to get the following lemma.

Lemma 3.4 *Suppose that $A, S \subset \mathbb{Z}$ are both ν-hereditarily energetic; $|S| \leq K|A|$, and $\kappa \in (0,1]$ and $r \in \mathbb{N}$ are parameters. Then either $r \leq (\kappa^{-1}\nu^{-1}K)^{O(1)}$ or there is some $1 \leq j \leq r$ such that $|(2^j \cdot S) \cap A| < \kappa|A|$.*

Proof Let $\epsilon = \Omega(\kappa K^{-1})$ be such that the $O(\epsilon)$-term in Lemma 3.3 is at most $\frac{\kappa}{2(1+K)}$. It follows that

$$|A \setminus A_{\epsilon^2\nu}| \leq \frac{\kappa}{2}|A| \quad \text{and} \quad |S \setminus S_{\epsilon^2\nu}| \leq \frac{\kappa}{2(1+K)}|S| < \frac{\kappa}{2}|A|.$$

Suppose that $|(2^j \cdot S) \cap A| \geq \kappa|A|$ for all $1 \leq j \leq r$. Then by the triangle inequality,

$$|(2^j \cdot S_{\epsilon^2\nu}) \cap A_{\epsilon^2\nu}| > \kappa|A| - \frac{\kappa}{2}|A| - \frac{\kappa}{2}|A| > 0 \quad \text{for all } i \leq j \leq r.$$

Let $t \in (2^j \cdot S_{\epsilon^2\nu}) \cap A_{\epsilon^2\nu}$; then there are odd numbers a and s such that $t = 2^i a$ and $t = 2^{j+i'}s$ with $i \in I_{\epsilon^2\nu}(A)$ and $i' \in I_{\epsilon^2\nu}(S)$. It follows that $j = i - i'$. However, $|I_{\epsilon^2\nu}(A)| < \epsilon^{-2}\nu^{-1}$ and similarly $|I_{\epsilon^2\nu}(S)| < \epsilon^{-2}\nu^{-1}$, and so $r \leq \epsilon^{-4}\nu^{-2}$, and the result follows. ∎

The main ingredient replacing Proposition 2.6 and allied arguments is the following proposition.

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Proposition 3.5 *Suppose that G has no 2-torsion, $A, X \subset G$, $|X \setminus A| \leq \eta|A|$, A is (k, X)-summing, and $r \in \mathbb{N}$ is a parameter. Then either $\eta^{-1} = (kr)^{O(1)}$, $|A| \leq \exp((kr)^{O(1)})$, or there is a $k^{-O(1)}$-hereditarily energetic set S with $|S| \geq k^{-O(1)}|A|$ and $|(2^j \cdot S) \cap A| \geq k^{-O(1)}|S|$ for all $1 \leq j \leq r$.*

There are a couple of remarks worth making. First, the proof we give actually shows the stronger conclusion that S has doubling $k^{O(1)}$. We do not need this, but it may be worth noting and is probably helpful in understanding the structure of the proof. The notion of hereditarily energetic is necessary for dealing with (k, Z)-summing sets (see Lemma 3.2), which need not have small doubling.

Secondly, it is not surprising that we ask for G to have no 2-torsion in view of the conclusion that no analogue of Proposition 3.5 can be true for $G = (\mathbb{Z}/2\mathbb{Z})^n$, since the final conclusion collapses to $|A| \leq k^{O(1)}$, since whatever S is, $2 \cdot S = \{0_G\}$ in this case. Thus, if $A = X = G$, then A is $(2, X)$-summing (and so (k, X)-summing), and we can take $\eta = 0$, but there is no bound on $|A|$ in terms of k.

We first prove Proposition 3.5 in a model setting in Section 4 to illustrate our arguments, before moving on to the general result.

With this in hand, we are ready for the main result. It is worth giving a word of explanation. A number of our arguments involve careful dependences between various parameters, and we say things like 'let ϵ_0 be such that the first conclusion of Theorem X does not hold'. When we say this the conclusions of Theorem X will begin with a number of inequalities between parameters, and we will want to choose things so that those inequalities do not hold, leading to the more substantial conclusion(s) of the theorem. The proof below while short will give a flavour; the later arguments are more involved.

Proof of Proposition 2.7 By Lemma 3.2, A is $\nu_0 = k^{-O(1)}$-hereditarily energetic (assuming, as we may, that $\eta \leq 1$). Let $\nu_1 = k^{-O(1)}$ such that the set S in Proposition 3.5 is always ν_1-hereditarily energetic; let $\sigma_0 = k^{-O(1)}$ be such that $|S| \geq \sigma_0|A|$, and let $\tau_0 = k^{-O(1)}$ be such that $|(2^j \cdot S) \cap A| \geq \tau_0|S|$. Finally, let $r = k^{O(1)}$ be such that the first conclusion of Lemma 3.4 does not hold when applied to sets that are $\min\{\nu_0, \nu_1\}$-hereditarily energetic, whose size ratio is at most σ_0^{-1}, and with parameters $\sigma_0\tau_0$ and r.

Apply Proposition 3.5 with parameter r. Then either $\eta^{-1} = k^{O(1)}$, or $|A| \leq \exp(k^{O(1)})$, or else there is a ν_1-hereditarily energetic set S with $|S| \geq \sigma_0|A|$ and $|(2^j \cdot S) \cap A| \geq \tau_0|S| \geq \tau_0\sigma_0|A|$ for all $1 \leq j \leq r$. By Lemma 3.4, whose first conclusion does not hold by design, we get a contradiction, and the result is proved. ∎

4 A Model for Proposition 3.5

The finite field model is useful for illustrating arguments in arithmetic combinatorics without a lot of the technical difficulties involved in general abelian groups. One of the earliest introductions to the model is in [Gre05b], and a summary of more recent developments can be found in [Wol15].

Throughout this section, V denotes a vector space over a finite field $\mathbb{F}_p$ where p is an odd prime. Our aim is to prove the following model version of Proposition 3.5.

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Proposition 4.1 *Suppose that $A, X \subset V$, $|X \setminus A| \leq \eta |A|$ for some $\eta \in (0,1]$, A is (k,X)-summing for some $k \geq 2$, and $r \in \mathbb{N}$ is a parameter. Then either $\eta^{-1} \leq (rk)^{O(1)}$, $|A| \leq p^{(kr)^{O(1)}}$, or there is a $k^{-O(1)}$-hereditarily energetic set S with $|S| \geq k^{-O(1)}|A|$ and $|(2^j \cdot S) \cap A| \geq k^{-O(1)}|S|$ for all $1 \leq j \leq r$.*

There are three qualifying remarks to make. First, there is (necessarily) no analogue of Lemma 3.4 in V, and without such it is difficult to argue that the conclusion of Proposition 4.1 is terribly significant.

Secondly, the fact that the bound on $|A|$ is dependent on p may look odd to those not familiar with this sort of model, since there is no equivalent in Proposition 3.5. This is standard for the model and arises because, in this setting, we use genuine subspaces rather than "approximate" subgroups. The size of genuine subspaces is a power of p, and replacing an 'approximate' subgroup with a genuine subspace usually involves shrinking by a factor that is a power of p.

Typically in the model we think of p as fixed, although this could be seen to be in conflict with the fact that the r-dependence is important. Indeed, for any natural number $j \in (p, r]$, there will be some $j' \in \{1, \dots, p\}$ such that $2^j \equiv 2^{j'} \pmod{p}$; it follows that one might as well take $r \leq p$. We shall not make this simplification, as it is an artefact of the model.

Finally, it may be that a better result could be proved using the new polynomial techniques of Croot, Lev, and Pach [CLP17]. We have not used their method, because at present there is no known way to convert it to give arguments in the setting we ultimately need (the integers).

We begin with a sketch in which we aim to convey the structure of the argument. The more detailed work afterwards is to explain where the bounds come from.

Before going into the sketch we recall the two sets of hypotheses in Proposition 4.1 and mention where they arise: we have that A is (k,X)-summing which is used in Step II and Step V; we also (more or less) have that $|X \setminus A| \leq \eta |A|$ for η small in terms of k, which is used in Step V.

Given a non-empty set S, we write m_S for the uniform probability measure supported on S.

(Step I) We study A with respect to an average of uniform probability measures on translates of (possibly different) subspaces. Formally, a weighted cover of S by subspaces[10] is a pair (z, Z) where z is a V-valued random variable, Z is a finite-subspace-of-V-valued random variable, and $\mathbb{E} m_{z+Z} = m_S$. Since we can take V to be finite, we can assume these random variables take finitely many values so that there are no analytic issues to worry about.

We construct new weighted covers by specifying conditional joint distributions: given a weighted cover (z, Z) of S, and a weighted cover $(w^{(x,U)}, W^{(x,U)})$ of $x + U$ for each $x \in V$ and $U \leq V$ (finite), we can define a new weighted cover (z', Z') of S by specifying that

$$\mathbb{P}(z' = x' \text{ and } Z' = U' | z, Z) = \mathbb{P}(w^{(z,Z)} = x' \text{ and } W^{(z,Z)} = U').$$

[10] We tend to drop the "by subspaces" part.

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Indeed, we have

$$\mathbb{E}m_{z'+Z'} = \mathbb{E}\big(\mathbb{E}(m_{z'+Z'}|z,Z)\big) = \mathbb{E}\big(\mathbb{E}(m_{w^{(z,Z)}+W^{(z,Z)}}|z,Z)\big) = \mathbb{E}m_{z+Z} = m_S.$$

When we need to refer to the underlying sample space, we call it Ω, and ω will always denote an element of Ω.

Given a weighted cover (z,Z) of S, we also need weighted covers of $2^j \cdot S$ and conversely. In the model setting, this is particularly easy: if (z,Z) is a weighted cover of S, then $(2^j z, Z)$ is a weighted cover of $2^j \cdot S$ (since 2^j is just a scalar, and so $2^j \cdot (z+Z) = 2^j z + Z$) whenever $j \in \mathbb{N}_0$. We do not get this as cheaply in the non-model setting.

(Step II) First we use the Balog–Szemerédi–Gowers–Freiman machinery to find a set S having large intersection with A, and a subspace U of size not too much smaller than A, such that $m_S = m_S * m_U$. One can also require that S has small doubling, and so it is hereditarily energetic. Considering (S, m_S) as a probability space and writing $z : S \to V$ for the natural inclusion, and Z the constant function taking the value U, we have that $m_S = \mathbb{E}m_{z+Z}$, and so (z,Z) is a weighted cover of S. This is Lemma 4.2.

(Step III) Suppose that (z,Z) is a weighted cover and A is not highly uniform on $z(\omega)+Z(\omega)$; then a Fourier argument tells us that there is a large subspace $Z' \leq Z(\omega)$ such that $m_{z(\omega)+Z(\omega)} = \mathbb{E}_{z' \in z(\omega)+Z(\omega)} m_{z'+Z'}$ and

$$\mathbb{E}_{z' \in z(\omega)+Z(\omega)} m_{z'+Z'}(A)^2 > m_{z(\omega)+Z(\omega)}(A)^2,$$

where the size of the difference between left and right increases with the level of non-uniformity.

If there is a small (but not too small) proportion of $\omega \in \Omega$ with A not highly uniform on $z(\omega)+Z(\omega)$, then the above can be used to produce a new weighted cover (z', Z') of S, where

$$\mathbb{E}m_{z'+Z'}(A)^2 > \mathbb{E}m_{z+Z}(A)^2.$$

Again, the size of the difference is dependent on the notion of small above and the level of non-uniformity. As a result, this can be set in an iteration until we end up with a weighted cover where none of the subspaces are too small and where A is highly uniform on $z(\omega)+Z(\omega)$ for almost all $\omega \in \Omega$. This is Lemma 4.3. In the end we shall need A to be highly uniform on $2^j z(\omega) + Z(\omega)$ for almost all $\omega \in \Omega$ and all $0 \leq j \leq r$. This can be done at a cost of iterating r times as often. This (combined with Step II) is Corollary 4.4.

(Step IV) By averaging (using the fact that (z,Z) is a weighted cover of S) most of the mass of A in S corresponds to ωs where $m_{z(\omega)+Z(\omega)}(A)$ is not too small. By Step III, A is highly uniform on most of these ωs and *if* $m_{2z(\omega)+Z(\omega)}(X^c)$ is very large (meaning close to 1), we can use this uniformity, the size of A on $z(\omega)+Z(\omega)$, and the pigeonhole principle to show that there are many $z \in (A \cap (z(\omega)+Z(\omega)))^k$ such that $z_i + z_j \in A^c$ whenever $i < j$. Crucially, counting with the pigeonhole principle requires much less uniformity than counting zs for which $z_i + z_j \in 2 \cdot A$ as in Proposition 2.6. This counting is Lemma 4.5.

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

(Step V) Assuming that A is large enough that most $z \in (A \cap (z(\omega) + Z(\omega)))^k$ have $z_i \neq z_j$ for $i \neq j$, then since A is (k, X)-summing, it follows from Step IV that $m_{2z(\omega)+Z(\omega)}(X^c)$ is not very large for ωs supporting most of the mass of A, equivalently, $m_{2z(\omega)+Z(\omega)}(X)$ is not too small. Since X is only slightly bigger than A (this is the condition on η), on average, this means that $m_{2z(\omega)+Z(\omega)}(A)$ is not too small. Now, we arranged the same uniformity properties in (Step III) for the weighted cover $(2z, Z)$ of $2 \cdot S$. It follows from this that $m_{4z(\omega)+Z(\omega)}(X)$ is not too small for a large mass of points, and this process can be iterated to get Proposition 4.1.

We now turn to the details.

Lemma 4.2 *Suppose that $A, X \subset V$; $|X| \leq K|A|$, and A is (k, X)-summing. Then there is a $(kK)^{-O(1)}$-hereditarily energetic set S and a weighted cover (z, Z) of S such that*

$$|A \cap S| \geq (kK)^{-O(1)}|A|, \quad |S| \leq (kK)^{O(1)}|A| \text{ and } \min_{\omega} |Z(\omega)| \geq p^{-(kK)^{O(1)}}|A|.$$

Proof Apply Lemma 3.2 to get that A has $E(A) = \Omega((kK)^{-O(1)}|A|^3)$. The Balog–Szemerédi–Gowers Theorem[11] then gives us a subset $A' \subset A$ with $|A'| = \Omega((kK)^{-O(1)}|A|)$ such that $|A' + A'| = (kK)^{O(1)}|A'|$. By Chang's theorem for r-torsion groups [TV06, Corollary 5.29] (applied to A' despite only needing the large energy hypothesis), we get a subspace $U \leq V$ of size $p^{-(kK)^{O(1)}}|A'|$ such that $U \subset 2A' - 2A'$. Plünnecke's Inequality [TV06, Corollary 6.27] then tells us that $|U + A'| \leq |3A' - 2A'| \leq (kK)^{O(1)}|A'|$. The claimed result follows by letting $S = U + A'$ (which has $|S + S| \leq (kK)^{O(1)}|S|$ and so is $(kK)^{-O(1)}$-hereditarily energetic by Lemma 3.1(i)), where z is the random variable taking values uniformly from S, and Z is the constant random variable taking the value U. ∎

If we were interested in getting a good constant for the $O(1)$-terms in Proposition 4.1 (and hence the $\Omega(1)$-term in Theorem 1.2), then improvements could easily be made here. Unusually, with Freiman's theorem, one is most interested in the size of the intersection of the set on the subspace, and not so much with the size of the subspace. We give the argument we do because there are easy references to results in the literature.

We now turn to the Fourier argument in Step III. This is a routine 'energy increment' argument.

Lemma 4.3 *Suppose that $A, S \subset V$; (z, Z) is a weighted cover of S and $\delta \in (0, 1]$ is a parameter. Then either*

(i) $\mathbb{P}(\|1_{A\cap(z+Z)} * (1_A dm_{z+Z}) - m_{z+Z}(A)^2\|_{L_2(m_{2z+Z})} \leq \delta) \geq 1 - \delta$,
(ii) *or else there is a weighted cover (z', Z') of S with*

$$\min |Z'| \geq p^{-O(\delta^{-2})} \min |Z| \quad \text{and} \quad \mathbb{E}m_{z'+Z'}(A)^2 \geq \mathbb{E}m_{z+Z}(A)^2 + \Omega(\delta^3).$$

[11] In its usual form, which corresponds to [TV06, Theorem 2.31((i) ⇒ (iv))] and then [TV06, Exercise 2.3.15]).

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Proof Suppose that $z = x$ and $Z = U$ are such that

$$\|1_{A\cap(x+U)} * (1_A dm_{x+U}) - m_{x+U}(A)^2\|_{L_2(m_{2x+U})} > \delta. \tag{4.1}$$

Write $A' := (A - x) \cap U$ and work inside the space U considered as endowed with Haar probability measure m_U. That means that while the definition of $\widehat{U}$ is the same as in Section 3, it is convenient (for this proof) to take different normalisations for convolution and the Fourier transform: we define convolution to be

$$f * g(y) := \int f(x)g(y-x)dm_U(x) \quad \text{for all } y \in U \text{ for all } f, g \in L_1(m_U),$$

and the Fourier transform by

$$\widehat{f}(\gamma) := \int f(x)\overline{\gamma(x)}dm_U(x) \quad \text{for all } \gamma \in \widehat{U}.$$

These conventions are in line with those in [TV06, Definition 4.7] and [TV06, Definition 4.6], respectively, and the key identities are summarised in [TV06, (4.1),(4.2), and (4.8)].

Set

$$\Gamma := \left\{\gamma \in \widehat{U} : |\widehat{1_{A'}}(\gamma)| \geq \frac{1}{2}\delta\right\} \quad \text{and} \quad U' := \bigcap_{\gamma\in\Gamma} \ker \gamma.$$

By Parseval's theorem for A' in U, we have

$$\left(\frac{1}{2}\delta\right)^2 |\Gamma| \leq \sum_{\gamma\in\widehat{U}} |\widehat{1_{A'}}(\gamma)|^2 = \int 1_{A'}(x)^2 dm_U(x) \leq 1,$$

so $|\Gamma| = O(\delta^{-2})$ and $|U'| \geq p^{-O(\delta^{-2})}|U|$. On the other hand,

$$\begin{aligned}
&\sum_{\gamma\in\Gamma\setminus\{0_{\widehat{U}}\}} |\widehat{1_{A'}}(\gamma)|^2 + \frac{1}{2}\delta^2 \\
&\qquad \geq \sum_{\gamma\in\Gamma\setminus\{0_{\widehat{U}}\}} |\widehat{1_{A'}}(\gamma)|^4 + \left(\frac{1}{4}\delta^2\right) \sum_{\gamma\notin\Gamma\cup\{0_{\widehat{U}}\}} |\widehat{1_{A'}}(\gamma)|^2 \\
&\qquad \geq \sum_{\gamma\neq 0_{\widehat{U}}} |\widehat{1_{A'}}(\gamma)|^4 = \|1_{A'} * 1_{A'} - m_U(A')^2\|^2_{L_2(m_U)} > \delta^2,
\end{aligned}$$

whence

$$\begin{aligned}
\|1_{A\cap(x+U)} * m_{U'}\|^2_{L_2(m_{x+U})} &= \|1_{A'} * m_{U'}\|^2_{L_2(m_U)} \\
&= \sum_{\gamma\in\widehat{U}} |\widehat{1_{A'}}(\gamma)|^2 |\widehat{m_{U'}}(\gamma)|^2 \\
&\geq \sum_{\gamma\in\Gamma} |\widehat{1_{A'}}(\gamma)|^2 > m_U(A')^2 + \frac{1}{2}\delta^2.
\end{aligned}$$

Let (z', Z') be defined conditional on $z = x$ and $Z = U$, as follows: if (4.1) holds, then let Z' be U' with (conditional) probability 1 and choose z' uniformly from $x + U$; if (4.1) does not hold, then let Z' be U with (conditional) probability 1 and $z' = x$ with

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

(conditional) probability 1. It then follows that (z', Z') is a weighted cover of S and

$$\mathbb{E}m_{z'+Z'}(A)^2 \geq \mathbb{E}m_{z+Z}(A)^2 + \mathbb{P}\Big(\|1_{A\cap(z+Z)} * (1_A dm_{z+Z}) - m_{z+Z}(A)^2\|_{L_2(m_{2z+Z})} > \delta\Big)\cdot\frac{1}{2}\delta^2.$$

It follows that if we are not in the first case of the lemma, then we must be in the second. ∎

If we were interested in optimising our arguments, then it might be more effective to use an L_p-version of the above in the style of Croot and Sisask (compare, for example, [CS10, Proposition 3.3] with [CS10, Proposition 3.1]).

The above lemma leads to a weighted cover (z, Z) where Z is not necessarily constant. This is necessary for us to ensure good bounds; if we wanted Z to be constant, then examples such as those in [Gre05c, Theorem 10.2] show we would have to have tower-type dependencies. This phenomenon is discussed after [Gre05b, Proposition 5.8] as an important part of Shkredov's argument from [Shk06] (see also [Shk04]).

Corollary 4.4 *Suppose that $A, X \subset V$; $|X| \leq K|A|$; A is (k, X)-summing, and $r \in \mathbb{N}$ and $\delta \in (0,1]$ are parameters. Then there is a $(kK)^{-O(1)}$-hereditarily energetic set S and a weighted cover (z, Z) such that*

$$|A\cap S| \geq (kK)^{-O(1)}|A|, \quad |S| \leq (kK)^{O(1)}|A|,$$
$$\min_\omega |Z(\omega)| \geq p^{-(\delta^{-1}rkK)^{O(1)}}|A|,$$

and for all $0 \leq i \leq r-1$, we have

$$\mathbb{P}\Big(\|1_{A\cap(2^i z+Z)} * (1_A dm_{2^i z+Z}) - m_{2^i z+Z}(A)^2\|_{L_2(m_{2^{i+1}z+Z})} \leq \delta\Big) \geq 1-\delta.$$

Proof We first apply Lemma 4.2 to get a $(kK)^{-O(1)}$-hereditarily energetic set S and a weighted cover (z_0, Z_0) of S such that

$$|A\cap S| \geq (kK)^{-O(1)}|A|, \quad |S| \leq (kK)^{O(1)}|A|,$$
$$\min_\omega\{|Z_0(\omega)|\} \geq p^{-(kK)^{O(1)}}|A|.$$

Suppose that we have a weighted cover (z_i, Z_i) of S and there is some $0 \leq j < r$ such that

$$\mathbb{P}\big(\|1_{A\cap(2^j z_i+Z_i)} * (1_A dm_{2^j z_i+Z_i}) - m_{2^j z_i+Z_i}(A)^2\|_{L_2(m_{2^{j+1}z_i+Z_i})} \leq \delta\big) < 1-\delta.$$

Then apply Lemma 4.3 to the weighted cover $(2^j z_i, Z_i)$ of $2^j\cdot S$. We get a new weighted cover (z', Z_{i+1}) of $2^j\cdot S$ (and we put $z_{i+1} = 2^{-j}z'$, so (z_{i+1}, Z_{i+1}) is a weighted cover of S) such that

$$\min_\omega |Z_{i+1}(\omega)| \geq p^{-O(\delta^{-2})}\min_\omega |Z_i(\omega)|,$$
$$\mathbb{E}m_{2^j z_{i+1}+Z_{i+1}}(A)^2 \geq \mathbb{E}m_{2^j z_i+Z_i}(A)^2 + \Omega(\delta^3).$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Since for any weighted cover (z, Z) of S, we have

$$\sum_{j=0}^{r-1} \mathbb{E} m_{2^j z+Z}(A)^2 \le r,$$

we see that we can be in this situation at most $O(\delta^{-3} r)$ times. The iteration terminates with the desired outcome. ∎

The second key ingredient as far as bounds are concerned comes in the next lemma covering Step IV. The level of uniformity necessary to count in the case we are interested in is polynomial in k, whereas in *e.g.*, [Sha15, Corollary 4.2], it is exponential in k.

Lemma 4.5 *Suppose that $x \in V$, $U \le V$, $A, X \subset V$, $\alpha := m_{x+U}(A)$, $m_{2x+U}(X) \le \epsilon$,*

$$\| 1_{A\cap(x+U)} * (1_A dm_{x+U}) - \alpha^2 \|_{L_2(m_{2x+U})} \le \epsilon\alpha^2,$$

and $k \in \mathbb{N}$ is a parameter. Then either $k = \Omega(\epsilon^{-\frac{1}{2}})$ or

$$\int \Big(\prod_{1\le i<j\le k} 1_{(2x+U)\setminus X}(z_i + z_j) \Big) \prod_{i=1}^{k} 1_A(z_i) dm_{x+U}(z_i) = \Omega(\alpha^k). \tag{4.2}$$

Proof First note that

$$\begin{aligned} &\int 1_X(z_1+z_2) 1_A(z_1) 1_A(z_2) dm_{x+U}(z_1) dm_{x+U}(z_2) \\ &\quad = \alpha^2 - \langle 1_{A\cap(x+U)} * (1_A dm_{x+U}), 1_{(2x+U)\setminus X} \rangle_{L_2(m_{2x+U})} \\ &\quad \le \alpha^2 - \alpha^2 m_{2\cdot x+U}((2x+U)\setminus X) + \epsilon\alpha^2 \le 2\epsilon\alpha^2. \end{aligned}$$

Now write Q for the integral on the left of (4.2). Apply Bonferroni's inequality[12] to the events $\{z \in (A\cap(x+U))^k : z_i + z_j \in X\}$,

$$\begin{aligned} Q &\ge \int \Big(\prod_{1\le i<j\le k} 1_{2x+U}(z_i+z_j) \Big) \prod_{i=1}^{k} 1_A(z_i) dm_{x+U}(z_i) \\ &\quad - \sum_{1\le i'<j'\le k} \int 1_X(z_{i'}+z_{j'}) \prod_{\substack{1\le i<j\le k \\ (i,j)\ne(i',j')}} 1_{2\cdot x+U}(z_i+z_j) \prod_{i=1}^{k} 1_A(z_i) dm_{x+U}(z_i) \\ &= \alpha^k - \sum_{1\le i'<j'\le k} \int 1_X(z_{i'}+z_{j'}) \prod_{i=1}^{k} 1_A(z_i) dm_{x+U}(z_i) \ge \alpha^k - 2\epsilon \binom{k}{2} \alpha^k. \end{aligned}$$

The lemma follows. ∎

It seems likely that a more careful argument using Turán's theorem (in a form like [SSV05, Lemma 3.1]) or the Lovász Local Lemma [TV06, Corollary 1.2.6] could be used to let us take $k = \Omega(\epsilon^{-1})$. Again, such improvements would impact the $O(1)$-term in Proposition 4.1 and, ultimately, the $\Omega(1)$-term in Theorem 1.2, but this is not the concern of this paper.

For us the crucial aspect of the above is that the conclusion $k = \Omega(\epsilon^{-\frac{1}{2}})$ does not depend on the density α. If it were allowed to depend on α, then we would not need

[12] In the form $\mathbb{P}(\bigcap_i E_i) \ge 1 - \sum_i \mathbb{P}(E_i^c)$.

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

the uniformity argument in Lemma 4.3. The reason that it is not is that it would lead to a lower bound on the intersections $(2^j \cdot S) \cap A$ which decreases with j. This in turn is not enough for our application in the non-model setting.

Finally we have the tools to complete the argument for Step V.

Proof of Proposition 4.1 Take $\epsilon_0 = \Omega(k^{-2})$ such that the first conclusion of Lemma 4.5 does not happen and let $\nu_0 = k^{-O(1)}$ be such that $m_S(A) \geq \nu_0$ always holds in the conclusion of Corollary 4.4 when $K \leq 2$.

Apply Corollary 4.4 with $K = 2$ and $\delta = 2^{-3}r^{-1}\epsilon_0 \min\{\nu_0^2, \epsilon_0^2\}$ to get a $k^{-O(1)}$-hereditarily energetic set S and a weighted cover (z, Z) (supported on the probability space $(\Omega, \mathbb{P})$) such that

$$|A \cap S| \geq k^{-O(1)}|A|, \quad |S| \leq k^{O(1)}|A|, \tag{4.3}$$
$$\min_{\omega} |Z(\omega)| \geq p^{-(rk)^{O(1)}}|A|,$$

and, for $0 \leq s < r$, writing

$$E_s := \left\{\omega \in \Omega : \|1_{A\cap(2^s z+Z)} * (1_A d m_{2^s z+Z}) - m_{2^s z+Z}(A)^2\|_{L_2(m_{2^{s+1}z+Z})} \leq \delta\right\},$$

we have $\mathbb{P}(E_s^c) < \delta$. For $1 \leq s \leq r$, write

$$L_s := \left\{\omega \in \Omega : m_{2^s z+Z}(X \setminus A) \leq \frac{1}{2}\epsilon_0\right\}.$$

Then

$$\mathbb{P}(L_s^c) < 2\epsilon_0^{-1}\mathbb{E}m_{2^s z+Z}(X \setminus A) = 2\epsilon_0^{-1} m_{2^s \cdot S}(X \setminus A) \leq k^{O(1)}\frac{|X \setminus A|}{|A|} = \eta k^{O(1)}.$$

Either $\eta^{-1} = (rk)^{O(1)}$ or else $\mathbb{P}(L_s^c) \leq \frac{1}{8r}\nu_0$ for all $1 \leq s \leq r$; we can assume the latter. Write

$$B := \left\{\omega \in \Omega : m_{z+Z}(A) > \frac{1}{2}\nu_0\right\} \quad \text{and} \quad \Omega' := B \cap \left(\bigcap_{i=1}^{r} L_i\right) \cap \left(\bigcap_{i=0}^{r-1} E_i\right),$$

and note that

$$\begin{aligned}\mathbb{E}m_{z+Z}(A)1_{\Omega'} &\geq \mathbb{E}m_{z+Z}(A) - \mathbb{E}m_{z+Z}(A)\left(1_{B^c} + \sum_{s=1}^{r} 1_{L_s^c} + \sum_{s=0}^{r-1} 1_{E_s^c}\right)\\ &\geq \nu_0 - \frac{1}{2}\nu_0 - r \cdot \frac{1}{8r}\nu_0 - r\delta \geq \frac{1}{8}\nu_0.\end{aligned}$$

Claim *Either $|A| \leq p^{(kr)^{O(1)}}$ or else for all $\omega \in \Omega'$ we have*

$$m_{2^s z(\omega)+Z(\omega)}(A) > \min\left\{\frac{1}{2}\nu_0, \frac{1}{2}\epsilon_0\right\} \quad \textit{for all } 0 \leq s \leq r.$$

Proof We proceed by induction. When $s = 0$, we have $\omega \in B$, and so the hypothesis holds. Suppose that it holds for $0 \leq s < r$. Then since $\omega \in E_s$,

$$\left\|1_{A\cap(2^s z(\omega)+Z(\omega))} * (1_A d m_{2^s z(\omega)+Z(\omega)}) - m_{2^s z(\omega)+Z(\omega)}(A)^2\right\|_{L_2(m_{2^{s+1}z(\omega)+Z(\omega)})} \leq \epsilon_0 m_{2^s z(\omega)+Z(\omega)}(A)^2.$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

If $m_{2^{s+1}z(\omega)+Z(\omega)}(X) \le \epsilon_0$, then it follows from Lemma 4.5 that

$$\int \Big(\prod_{1\le i<j\le k} 1_{(2^{s+1}z(\omega)+Z(\omega))\setminus X}(z_i+z_j)\Big) \prod_{i=1}^{k} 1_A(z_i)\, dm_{2^s z(\omega)+Z(\omega)}(z_i) = \Omega\Big(\Big(\min\Big\{\frac{1}{2}\nu_0, \frac{1}{2}\epsilon_0\Big\}\Big)^k\Big).$$

Since A is (k,X)-summing, we know that the first product on the left is 0 on $z \in A^k$ unless there is $1 \le i < j \le k$ with $z_i = z_j$. It follows that the integral is at most

$$\binom{k}{2} m_{2^s z(\omega)+Z(\omega)}(A)^{k-1} \cdot \frac{1}{|2^s z(\omega)+Z(\omega)|}.$$

The upper bound on $|A|$ follows from the lower bound on $|Z(\omega)|$ in (4.3). Thus, we can assume $m_{2^{s+1}z(\omega)+Z(\omega)}(X) > \epsilon_0$, and since $\omega \in L_{s+1}$, we have

$$\begin{aligned} m_{2^{s+1}z(\omega)+Z(\omega)}(A) &\ge m_{2^{s+1}z(\omega)+Z(\omega)}(X) - m_{2^{s+1}z(\omega)+Z(\omega)}(X\setminus A) > \frac{1}{2}\epsilon_0 \\ &\ge \min\Big\{\frac{1}{2}\nu_0, \frac{1}{2}\epsilon_0\Big\}. \end{aligned}$$

The claim is proved. ∎

We conclude that for all $1 \le s \le r$, we have

$$\begin{aligned} m_{2^s\cdot S}(A) = \mathbb{E} m_{2^s z+Z}(A) &\ge \mathbb{E} m_{2^s z+Z}(A) 1_{\Omega'} \ge \min\Big\{\frac{1}{2}\nu_0, \frac{1}{2}\epsilon_0\Big\}\mathbb{E}1_{\Omega'} \\ &\ge \min\Big\{\frac{1}{2}\nu_0, \frac{1}{2}\epsilon_0\Big\}\mathbb{E} m_{z+Z}(A)1_{\Omega'} = k^{-O(1)}, \end{aligned}$$

as required. ∎

5 Translating the Model

The remainder of the paper deals with converting the model argument of Section 4 to give a proof of Proposition 3.5; throughout, G denotes an abelian group. This is analogous to converting Meshulam's proof of the Roth–Meshulam Theorem [TV06, Proposition 10.12] to Bourgain's proof of Roth's Theorem [TV06, Theorem 10.29].

The reader already familiar with this sort of translation can move to Section 6. The key definitions are τ-closed pairs, defined before Lemma 5.1; covering numbers – $\mathcal{C}$ and $\mathcal{C}^\flat$, defined in (5.1) and before Lemma 5.3, respectively; and $\mathcal{S}(G)$, systems, and dimension defined after the proof of Lemma 5.3. The main variation in the definitions we have chosen here is in defining $\mathcal{C}^\flat$, which is set up the way it is so that dimension is sub-additive with respect to intersection rather than sub-additive up to a multiplicative constant.

The basic idea is to replace subspaces by pairs of sets that are "almost" closed, and we start with a definition for this purpose. We write $\mathcal{N}(G)$ for the set of finite symmetric neighbourhoods of the identity, that is, sets $S \subset G$ with $S = -S$ and $0_G \in S$. (We use topological language here for motivation; the systems we define later can be

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

thought of as bases for particular topologies.) Given $Z, W \in \mathcal{N}(G)$ we say that (Z, W) is τ-closed if there are sets $Z^-, Z^+ \in \mathcal{N}(G)$ such that

$$Z^- + W \subset Z, Z + W \subset Z^+ \quad \text{and} \quad |Z^+| \leq (1+\tau)|Z^-|.$$

Given a measure μ on G and $x \in G$, we write $\tau_x(\mu)$ for the measure induced by

$$C(G) \longrightarrow C(G); \quad f \longmapsto \Big(y \longmapsto \int f(y-x)d\mu(y)\Big).$$

As in Section 4, if S is a finite non-empty subset of G, we write m_S for the uniform probability measure supported on S.

Lemma 5.1 (Basic properties of τ-closed pairs) *Suppose that (Z, W) is a τ-closed. Then*

(i) $\|\tau_w(m_Z) - m_Z\| \leq \tau$ *for all* $w \in W$;
(ii) *if* $W' \subset W$, *then* (Z, W') *is* τ*-closed;*
(iii) *if* G *has no* 2*-torsion, then* $(2^m \cdot Z, 2^m \cdot W)$ *is* τ*-closed.*

Proof The first property is immediate by the triangle inequality; the second is immediate, and the third equally so, once we recall that $x \mapsto 2^m x$ is an injection in a group with no 2-torsion. ∎

These sorts of pairs behave enough like subspaces that we can prove counting-type results (as we will in Section 9), but not so much that they are amenable to iteration without greater losses than we can afford to bear. One usually deals with this by recording auxiliary data in the form of (the-data-necessary-to-generate) Bohr sets. We discuss this in detail before Lemma 8.2.

The fact that we are using Freiman's theorem suggests that we actually need Bohr sets inside coset progressions (defined before Lemma 7.1), and the notion of a Bourgain system was formulated in [GS08, Definition 4.1] to deal with exactly this situation. This gives a common framework for Bohr sets and generalised arithmetic progressions, and it turns out much of the technology Bourgain developed for Bohr sets in [Bou99] extends. We shall use a very similar definition.

It turns out that we only need to use Freiman's theorem once, so rather than proceeding with the more general systems below, we could simply pass to a long arithmetic progression within the generalised arithmetic progression (in the style of [TV06, Exercise 3.2.5]) and consider Bohr sets inside that. This does not seem to afford any great simplification.

We now turn to the basic definitions. For $X, Y \subset G$, we write

$$\mathcal{C}(X;Y) := \min\{|T| : X \subset T + Y\}. \tag{5.1}$$

We call these numbers *covering numbers*. Here and throughout, we take the usual conventions concerning ∞.

Lemma 5.2 (Basic properties of covering numbers) *Suppose that G and H are abelian groups.*

(i) (Chaining) *For all $X, Y, Z \subset G$, we have*

$$\mathcal{C}(X;Z) \leq \mathcal{C}(X;Y)\mathcal{C}(Y;Z).$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

(ii) (Homomorphisms) *For all $X, Y \subset G$ and homomorphisms $\phi\colon G \to H$, we have*

$$\mathcal{C}\big(\phi(X);\phi(Y)\big) \le \mathcal{C}(X;Y).$$

(iii) (Covering and size) *For all $X, Y \subset G$, we have*

$$|X| \le \mathcal{C}(X;Y)|Y|.$$

(iv) (Ruzsa's covering lemma) *For all $X, Y \subset G$, we have*

$$\mathcal{C}(X;Y-Y) \le \min\Big\{\frac{|X+Y|}{|Y|}, \frac{|X-Y|}{|Y|}\Big\}.$$

Proof The first three facts are immediate from the definition. The last is just Ruzsa's covering lemma [TV06, Lemma 2.14], which can be proved by letting $T \subset X$ be a maximal subset such that if $(t+Y)\cap(t'+Y) \ne \varnothing$ and $t, t' \in T$, then $t = t'$. This gives the first bound in the minimum; applying the first with Y replaced by $-Y$ and noticing that $(-Y)-(-Y) = Y-Y$ and $|-Y| = |Y|$ gives the second. ∎

Covering numbers do not behave well with respect to intersections, and so we define[13]

$$\mathcal{C}^\flat(X;Y) := \min_{\substack{H\in\mathbf{Ab}\\ \phi\in\mathrm{Hom}(\langle X\rangle,H)}} \big\{\mathcal{C}(W;Z) : Z, W \subset H, \; X \subset \phi^{-1}(W), \phi^{-1}(Z-Z) \subset Y\big\}.$$

Here, $\langle X\rangle$ denotes the group generated by X and is there to make the definition independent of the particular ambient group in which X and Y live.

These numbers are close to covering numbers but also respect intersections.

Lemma 5.3 (Basic properties of $\mathcal{C}^\flat$) *Suppose that G and K are abelian groups.*

(i) (Order) *For all $X' \subset X$ and $Y \subset Y'$, we have*

$$\mathcal{C}^\flat(X';Y') \le \mathcal{C}^\flat(X;Y).$$

(ii) (Homomorphisms) *For all $X, Y \subset G$ and $\psi\colon K \to G$ a homomorphism, we have*

$$\mathcal{C}^\flat\big(\psi^{-1}(X);\psi^{-1}(Y)\big) \le \mathcal{C}^\flat(X;Y).$$

(iii) (Equivalence) *Whenever $X, Y \subset G$, we have*

$$\mathcal{C}^\flat(X;Y-Y) \le \mathcal{C}(X;Y) \le \mathcal{C}^\flat(X;Y).$$

(iv) (Meets) *Whenever $X, X', Y, Y' \subset G$ we have*

$$\mathcal{C}^\flat(X\cap X';Y\cap Y') \le \mathcal{C}^\flat(X;Y)\mathcal{C}^\flat(X';Y').$$

Proof Suppose for all parts that $H \in \mathbf{Ab}$, $\phi \in \mathrm{Hom}(\langle X\rangle, H)$, and $Z, W \subset H$ are such that $X \subset \phi^{-1}(W)$ and $\phi^{-1}(Z-Z) \subset Y$, and $\mathcal{C}^\flat(X;Y) = \mathcal{C}(W;Z)$.

For (i), note that $\psi := \phi|_{\langle X'\rangle} \in \mathrm{Hom}(\langle X'\rangle, H)$ and $X' \subset X \cap \langle X'\rangle \subset \psi^{-1}(W)$, and $\psi^{-1}(Z-Z) \subset Y \subset Y'$. It follows that $\mathcal{C}^\flat(X';Y') \le \mathcal{C}(W;Z) = \mathcal{C}^\flat(X;Y)$, as claimed.

[13]Here **Ab** denotes the category of abelian groups. One could equally say that H is an abelian group and ϕ is a homomorphism from G to H.

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

For (ii), note $\pi := (\phi \circ \psi)|_{\langle \psi^{-1}(X)\rangle} \in \mathrm{Hom}(\langle \psi^{-1}(X)\rangle, H)$ and $\psi^{-1}(X) \subset \pi^{-1}(W)$ and $\pi^{-1}(Z-Z) \subset \psi^{-1}(Y)$. Again, it follows that $\mathcal{C}^\flat(X';Y') \leq \mathcal{C}(W;Z) = \mathcal{C}^\flat(X;Y)$ as claimed.

For (iii), we get the left-hand inequality by considering the canonical embedding $\phi\colon \langle X\rangle \to G \in \mathrm{Hom}(\langle X\rangle, G)$, and noting that $\phi^{-1}(Y-Y) \subset Y-Y$ and $X \subset \phi^{-1}(X)$.

For the right-hand inequality let $S \subset H$ be such that $W \subset S+Z$ and $|S| = \mathcal{C}(W;Z)$, and $T \subset \langle X\rangle$ be minimal such that if $s \in S$ has $(s+Z) \cap \phi(\langle X\rangle) \neq \varnothing$, then there is some $t \in T$ such that $\phi(t) \in s+Z$. It follows that $|T| \leq |S|$, and if $x \in X$, then $x \in \phi^{-1}(W)$, and so $\phi(x) \in W$ and there is some $s \in S$ with $\phi(x) \in s+Z$. By the definition of T there is some $t \in T$ such that $\phi(t) \in s+Z$, whence $\phi(x-t) \in Z-Z$ and so $x \in T+\phi^{-1}(Z-Z) \subset T+Y$. We conclude that $\mathcal{C}(X;Y) \leq |T| \leq |S|$. The claimed inequality follows.

Finally, for (iv), suppose additionally that $H' \in \mathbf{Ab}$, $\psi \in \mathrm{Hom}(\langle X'\rangle, H')$, and $Z', W' \subset H'$ are such that $X' \subset \psi^{-1}(W')$ and $\psi^{-1}(Z'-Z') \subset Y'$, and $\mathcal{C}^\flat(X';Y') = \mathcal{C}(W';Z')$.

Then $H \times H'$ is an abelian group, $W \times W', Z \times Z' \subset H \times H'$, $\pi\colon \langle X \cap X'\rangle \to H \times H', x \mapsto (\phi(x), \psi(x))$ is a homomorphism, and

$$\pi^{-1}(W \times W') = \langle X \cap X'\rangle \cap (\phi^{-1}(W) \cap \psi^{-1}(W')) \supset X \cap X',$$

and

$$\begin{aligned}\pi^{-1}((Z \times Z') - (Z \times Z')) &= \pi^{-1}((Z-Z) \times (Z'-Z')) \\ &\subset \phi^{-1}(Z-Z) \cap \psi^{-1}(Z'-Z') \subset Y \cap Y'.\end{aligned}$$

On the other hand,

$$\mathcal{C}(W \times W'; Z \times Z') \leq \mathcal{C}(W;Z)\mathcal{C}(W';Z') = \mathcal{C}^\flat(X;Y)\mathcal{C}^\flat(X';Y').$$

The result follows. ∎

We use a slight variant of [GS08, Definition 4.1], defining a system on G to be a vector $B = (B_i)_{i\in\mathbb{N}_0}$ of symmetric neighbourhoods of the identity such that $B_{i+1} + B_{i+1} \subset B_i$ for all $i \in \mathbb{N}_0$. We define its dimension to be

$$\dim B := \sup_{i\in\mathbb{N}_0} \log_2 \mathcal{C}^\flat(B_i; B_{i+1}),$$

and write $\mathcal{S}(G)$ for the set of systems on G. For $S \subset G$, we also write $\mathcal{C}^\flat(S;B)$ for $\mathcal{C}^\flat(S;B_0)$. This is how we record the "size" of B relative to some reference set S. (As an aside we remark that while we have found it convenient to use this notion of 'size', the more conventional $|B_0|$ works better in some ways; while it does not deal so well with intersections, it would allow us to dispense with the second part of Lemma 5.5(iii).)

There are many examples of systems; coset progressions naturally define systems, as we will show in Lemma 7.1, as do Bohr sets (which we cover in Lemma 8.2) and subgroups, which we deal with now.

Lemma 5.4 (Subgroup systems) *Suppose that $H \leq G$. Then the $\mathbb{N}_0$-indexed vector B taking the constant value H is a 0-dimensional system.*

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Proof B is easily seen to be a system. Moreover, $\mathcal{C}(H;H) \le \mathcal{C}^\flat(H;H) = \mathcal{C}^\flat(H;H-H) \le \mathcal{C}(H;H)$ by Lemma 5.3(iii) and $\mathcal{C}(H;H) = 1$, so $\dim B = 0$, as claimed. ∎

There are three ways of creating new systems that will be useful to us. Given $B, B' \in \mathcal{S}(G)$ and $m \in \mathbb{N}_0$, we make the following definitions.

- The intersection of B and B' is the vector $B \wedge B' := (B_i \cap B'_i)_{i\in\mathbb{N}_0}$. It is easy to check that $(\mathcal{S}(G), \wedge)$ forms a meet semi-lattice. (Meaning that $B \wedge B'$ is indeed a system, that $B \wedge (B' \wedge B'') = (B \wedge B') \wedge B''$, that $B \wedge B' = B' \wedge B$, and that $B \wedge B = B$.)
- The 2^{-m}-dilate of B is the vector $2^{-m}B := (B_{i+m})_{i\in\mathbb{N}_0}$. It is easy to check that this is an action of the (additive) monoid $\mathbb{N}_0$ on $\mathcal{S}(G)$. (Meaning that $2^{-m}B$ is indeed a system, that $1B = B$, and that $2^{-m}(2^{-m'}B) = 2^{-(m+m')}B$.)
- The 2^m-multiple of B is the vector $2^m \cdot B := (2^m \cdot B_i)_{i\in\mathbb{N}_0}$. Again, it is easy to check that this is an action of $\mathbb{N}_0$ on $\mathcal{S}(G)$.

The meet semi-lattice structure induces a partial order on $\mathcal{S}(G)$, and $B' \le B$ if and only if $B'_i \subset B_i$ for all $i \in \mathbb{N}_0$.

Dilates distribute over intersection, meaning that $2^{-m}(B \wedge B') = (2^{-m}B) \wedge (2^{-m}B')$, but in general for multiples, we only have $2^m \cdot (B \wedge B') \le (2^m \cdot B) \wedge (2^m \cdot B')$; if G has no 2-torsion then we do have equality.

Finally, it is worth noting that multiples and dilates do not interact terribly well; in particular, we will need to consider systems of the form $2^m \cdot (2^{-m'}B)$ and this does not in general simplify.

Lemma 5.5 *Suppose that $B, B' \in \mathcal{S}(G)$, $S \subset G$, and $m \in \mathbb{N}_0$.*

(i) (Intersections)

$$\dim B \wedge B' \le \dim B + \dim B' \text{ and } \mathcal{C}^\flat(S; B \wedge B') \le \mathcal{C}^\flat(S;B)\mathcal{C}^\flat(S;B').$$

(ii) (Dilations)

$$\dim 2^{-m}B \le \dim B \text{ and } \mathcal{C}^\flat(S; 2^{-m}B) \le \mathcal{C}^\flat(S;B)2^{(m+1)\dim B}.$$

(iii) (Multiples)

$$\dim 2^m \cdot B \le \dim B,$$

and if G has no 2-torsion, then

$$\mathcal{C}^\flat(S; 2^m \cdot B) \le \mathcal{C}^\flat(S;B) \exp\big(O(m \dim B)\big).$$

Proof The first part follows immediately from Lemma 5.3(iv) and the definitions.

For (ii), the dimension inequality is immediate from the definition. For the second estimate we have the following chain of inequalities, justified afterwards:

$$\begin{aligned}\mathcal{C}^\flat(S; 2^{-m}B) = \mathcal{C}^\flat(S;B_m) &\le \mathcal{C}^\flat(X; B_{m+1} - B_{m+1}) \\ &\le \mathcal{C}(X;B_{m+1}) \le \mathcal{C}(X;B_0) \prod_{i=0}^{m} \mathcal{C}(B_i;B_{i+1}) \\ &\le \mathcal{C}^\flat(X;B_0)2^{(m+1)\dim B}.\end{aligned}$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Since B is a system, $B_{m+1} - B_{m+1} \subset B_m$, and so Lemma 5.3(i) gives the first inequality. The second follows from the first inequality in Lemma 5.3(iii). The third follows by Lemma 5.2(i), and then the fourth by the second inequality in Lemma 5.3(iii).

For the dimension bound in (iii), note that the isomorphism $\psi\colon 2^m \cdot G \to G, x \mapsto 2^{-m}x$ has $\psi^{-1}(B_i) = (2^m \cdot B)_i$, and the bound follows from Lemma 5.3(ii) and the definition of dimension.

For the second part of (iii), suppose that G has no 2-torsion. We have the same chain of inequalities as above. Again, they are justified afterwards:

$$\begin{aligned}\mathcal{C}^\flat(S; 2^m \cdot B) &= \mathcal{C}^\flat(S; 2^m \cdot B_0)\\ &\le \mathcal{C}^\flat(S; 2^m \cdot B_1 - 2^m \cdot B_1)\\ &\le \mathcal{C}(S; 2^m \cdot B_1) \le \mathcal{C}(S; B_0)\mathcal{C}(B_0; B_1)\prod_{t=0}^{m-1}\mathcal{C}(2^t \cdot B_1; 2^{t+1} \cdot B_1).\end{aligned}$$

Then $2^m \cdot B$ is a system so $2^m \cdot B_1 - 2^m \cdot B_1 \subset 2^m \cdot B_0$, and then Lemma 5.3(i) gives the first inequality. The second follows from the first inequality in Lemma 5.3(iii), and then the third by Lemma 5.2(i).

By Lemma 5.2 (ii), the fact that $2 \cdot B$ is a system so $2 \cdot B_2 - 2 \cdot B_2 \subset 2 \cdot B_1$, and Ruzsa's covering lemma (Lemma 5.2(iv)) we have

$$\mathcal{C}(2^t \cdot B_1; 2^{t+1} \cdot B_1) \le \mathcal{C}(B_1; 2 \cdot B_1) \le \mathcal{C}(B_1; 2 \cdot B_2 - 2 \cdot B_2) \le \frac{|B_1 + 2 \cdot B_2|}{|2 \cdot B_2|}.$$

Since

$$B_1 + 2 \cdot B_2 \subset B_1 + B_2 + B_2 \subset B_1 + B_1 \subset B_0,$$

we have by Lemma 5.2 (iii) and (i) that

$$|B_1 + 2 \cdot B_2| \le |B_0| \le \mathcal{C}(B_0; B_1)\mathcal{C}(B_1; B_2)|B_2|.$$

Combining all this gives

$$\mathcal{C}^\flat(S; 2^m \cdot B) \le \mathcal{C}(S; B_0)\mathcal{C}(B_0; B_1)\prod_{t=0}^{m-1}\Big(\mathcal{C}(B_0; B_1)\mathcal{C}(B_1; B_2)\frac{|B_2|}{|2 \cdot B_2|}\Big).$$

Since G has no 2-torsion, $|2 \cdot B_2| = |B_2|$,and so the second part of Lemma 5.3 (iii) can then be used to give the result. ∎

The requirement that G has no 2-torsion is clearly necessary for the second part of (iii) above; the proof is essentially the same as the proof of [Buk08, Theorem 15] with Ruzsa's triangle inequality [TV06, Lemma 2.6] replaced by Lemma 5.2(i).

At the start of the section we introduced the notion of τ-closed pair, and we can use the pigeonhole principle to produce a plentiful supply of such pairs from systems. Although we do not need it, a stronger result [TV06, Lemma 4.24] is available.

Lemma 5.6 *Suppose that $Z \in \mathcal{N}(G)$, $B \in \mathcal{S}(G)$, $|Z + B_0| \le K|Z|$, and $\tau \in (0, 1]$ is a parameter. Then there is a set $Z_0 \in \mathcal{N}(G)$ with $Z \subset Z_0 \subset Z + B_0$ and a natural number $m = \log_2 \log 2K + \log_2 \tau^{-1} + O(1)$ such that (Z_0, B_m) is τ-closed.*

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Proof Let $m \in \mathbb{N}_0$ be a parameter to be optimised shortly;

$$\prod_{j=0}^{2^{m-1}-1} \frac{|Z+(2j+2)B_m|}{|Z+2jB_m|} \le \frac{|Z+2^mB_m|}{|Z|} \le \frac{|Z+B_0|}{|Z|} \le K.$$

By the pigeonhole principle, there is some $0 \le j \le 2^{m-1}-1$ such that

$$\frac{|Z+(2j+2)B_m|}{|Z+2jB_m|} \le K^{\frac{1}{2^{m-1}}},$$

so we can take $m = \log_2 \log_2 K + \log_2 \tau^{-1} + O(1)$ such that the right-hand side is at most $1+\tau$. In that case, let $Z_0 := Z+(2j+1)B_m$, $Z_0^- := Z+2jB_m$ and $Z_0^+ := Z_0+(2j+2)B_m$, which are all symmetric neighbourhoods of the identity and have $Z_0^- + B_m = Z_0$ and $Z_0 + B_m = Z_0^+$. Moreover, the choice of j ensures that $|Z_0^+| \le (1+\tau)|Z_0^-|$, and the result is proved. ∎

In particular, for low-dimensional systems, we have the following lemma.

Lemma 5.7 *Suppose that $B \in \mathcal{S}(G)$ is d-dimensional, and $\tau \in (0,1]$ is a parameter. Then there is a set $Z_0 \in \mathcal{N}(G)$ with $B_1 \subset Z_0 \subset B_0$ and a natural number $m = \log_2 d + \log_2 \tau^{-1} + O(1)$ such that (Z_0, B_m) is τ-closed.*

Proof By Lemma 5.2(iii) and the second inequality in Lemma 5.3(iii), we have

$$|B_1 + B_1| \le |B_0| \le \mathcal{C}(B_0; B_1)|B_1| \le 2^d|B_1|.$$

Apply Lemma 5.6 to $Z := B_1$ and the system $2^{-1}B$ to get the result. ∎

6 Proof of Proposition 3.5

It is convenient to use the language of probability theory, but all our probability measures will have finite support, so there is no analysis involved. (This can be easily checked, as the only places where we produce new probability spaces are in Lemmas 6.1 and 8.5.) We say that a probability space $(\Omega', \mathbb{P}')$ is an extension of a probability space $(\Omega, \mathbb{P})$ if there is a map $\phi\colon \Omega' \to \Omega$ such that $\mathbb{P}'(\phi^{-1}(A)) = \mathbb{P}(A)$ for all (measurable) A in Ω. Note that if $(\Omega'', \mathbb{P}'')$ is an extension of $(\Omega', \mathbb{P}')$ and $(\Omega', \mathbb{P}')$ is an extension of $(\Omega, \mathbb{P})$, then $(\Omega'', \mathbb{P}'')$ is an extension of $(\Omega, \mathbb{P})$.

Given a random variable X on Ω, and an extension $(\Omega', \mathbb{P}')$ of $(\Omega, \mathbb{P})$, then for convenience, we also write X for the pull-back of X on Ω'.

We follow the plan in Section 4; the analogue of Step II and Lemma 4.2 is following lemma proved in Section 7.

Lemma 6.1 *Suppose that A is (k, X)-summing, $|X| \le K|A|$, and $\tau \in \left(0, \frac{1}{2}\right]$ is a parameter. Then there is a $(kK)^{-O(1)}$-hereditarily energetic set S with $|A \cap S| \ge (kK)^{-O(1)}|S|$ and $|S| \ge k^{-O(1)}|A|$, a probability space $(\Omega, \mathbb{P})$ supporting a G-valued random variable z, an $\mathcal{N}(G)$-valued random variable Z, and an $\mathcal{S}(G)$-valued random variable T such that*

$$\|\mathbb{E}m_{z+Z} - m_S\| \le \tau,$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

and for all $\omega \in \Omega$,

$$\dim T(\omega) \le (kK)^{O(1)}, \quad \mathcal{C}^\flat(S; T(\omega)) \le \exp\left((kK \log \tau^{-1})^{O(1)}\right)$$

and $(Z(\omega), T(\omega)_0)$ *is* τ*-closed.*

The most technically demanding part is the following analogue of Step III and Corollary 4.4, which we prove in Section 8.

Corollary 6.2 *Suppose that* G *has no 2-torsion,* $A, S \subset G$, $(\Omega, \mathbb{P})$ *is a probability space supporting a* G*-valued random variable* z*, an* $\mathcal{N}(G)$*-valued random variable* Z *and an* $\mathcal{S}(G)$*-valued random variable* T *such that for all* $\omega \in \Omega$,

$$\dim T(\omega) \le d \quad \textit{and} \quad \mathcal{C}^\flat(S; T(\omega)) \le D$$

and $(Z(\omega), T(\omega)_0)$ *is* τ*-closed. Let* $\delta \in (0,1]$ *and* $r \in \mathbb{N}$ *be parameters.*

Then either $\tau^{-1} \le (\delta^{-1} r)^{O(1)}$ *or there is a probability space* $(\Omega', \mathbb{P}')$ *extending* $(\Omega, \mathbb{P})$*, supporting a* G*-valued random variable* z'*, and a* $\mathcal{N}(G)$*-valued random variable* Z' *with*

$$\|\mathbb{E}' m_{z'+Z'} - \mathbb{E} m_{z+Z}\| \le \delta$$

and a further extension $(\Omega'', \mathbb{P}'')$ *of* $(\Omega', \mathbb{P}')$ *supporting a* G*-valued random variable* z'' *and* $\mathcal{N}(G)$*-valued random variables* $Z_1, \dots, Z_k$ *such that*

(i)
$$\|\mathbb{E}'' m_{2^s z''+Z_i} - \mathbb{E}' m_{2^s \cdot (z'+Z')}\| \le \delta$$
for all $1 \le i \le k$ *and* $0 \le s \le r$*;*

(ii) (U_1-uniformity)
$$\mathbb{E}'' |m_{2^s z''+Z_i}(A) - m_{2^s \cdot (z'+Z')}(A)|^2 \le \delta$$
for all $1 \le i \le k$ *and* $0 \le s \le r$*;*

(iii) (U_2-uniformity)
$$\mathbb{E}'' \left\| 1_{A \cap (2^s z''+Z_i)} * (1_A d m_{2^s z''+Z_j}) - m_{2^s \cdot (z'+Z')}(A)^2 \right\|^2_{L_2(m_{2^{s+1} z''+Z_i})} \le \delta$$
for all $1 \le i < j \le k$ *and* $0 \le s \le r$*;*

and for all $\omega'' \in \Omega''$, $(Z_i(\omega''), Z_{i+1}(\omega''))$ *is* δ*-closed for all* $1 \le i < k$*, and* $\mathcal{C}^\flat(S; Z_k(\omega'')) \le D \exp((dkr\delta^{-1})^{O(1)})$.

Finally, we have the analogue of Step IV and Lemma 4.5 which is not terribly different from the model case and which is proved in Section 9.

Lemma 6.3 *Suppose that* $A, X \subset G$, $z_0 \in G$, (Z_i, Z_{i+1}) *is* τ*-closed for all* $1 \le i < k$*, and*

(i) $|m_{Z_i}(A - z_0) - \alpha| \le \tau$ *for* $1 \le i \le k$*;*
(ii) $m_{Z_i}(X - 2z_0) \le \epsilon$ *for all* $1 \le i < k$*;*
(iii) $\|1_{(A-z_0) \cap Z_i} * (1_{A-z_0} d m_{Z_j}) - \alpha^2\|^2_{L_2(m_{Z_i})} \le \delta$ *for all* $1 \le i < j \le k$.

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Then either $\delta^{-1} = O(k^2\alpha^{-4})$, $\tau^{-1} = O(k\alpha^{-1})$, $\epsilon^{-1} = O(k^2)$, *or*

$$\int \Big(\prod_{i<j} 1_{(Z_i+2z_0)\setminus X}(z_i+z_j+2z_0) 1_{Z_i}(z_i+z_j) \Big) \times \prod_{i=1}^{k} 1_A(z_i+z_0)\, dm_{Z_i}(z_i) = \Omega(\alpha^k).$$

With these tools we can proceed with the main proof. (We begin by recalling the statement for convenience.)

Proposition (Proposition 3.5) *Suppose that G has no 2-torsion; $A, X \subset G$; $|X \setminus A| \leq \eta|A|$; A is (k, X)-summing; and $r \in \mathbb{N}$ is a parameter. Then either $\eta^{-1} = (kr)^{O(1)}$, or $|A| \leq \exp((kr)^{O(1)})$, or there is a $k^{-O(1)}$-hereditarily energetic set S with $|S| \geq k^{-O(1)}|A|$ and $|(2^j \cdot S) \cap A| \geq k^{-O(1)}|S|$ for all $1 \leq j \leq r$.*

Proof We begin by making a number of choices for parameters.

- Let $\epsilon_0 = \Omega(k^{-2})$ be such that the third conclusion of Lemma 6.3 does not follow in any application of that lemma with parameters ϵ_0 and k.
- Let $\nu_0 = k^{-O(1)}$ be such that $m_S(A) \geq \nu_0$ and $|S| \geq \nu_0|A|$ in the conclusion of any application of Lemma 6.1 with $K = 2$.
- Let $\delta_1 = k^{-O(1)}$ be such that the first conclusion in Lemma 6.3 does not follow in any application of that lemma with parameters $\alpha \geq \min\{\frac{1}{2}\nu_0, \frac{1}{2}\epsilon_0\}$ and k.
- Let $\frac{1}{4}\epsilon_0 \geq \tau_1 = k^{-O(1)}$ be such that the second conclusion of Lemma 6.3 does not follow in any application of that lemma with parameters $\alpha \geq \min\{\frac{1}{2}\nu_0, \frac{1}{2}\epsilon_0\}$ and k.
- Let $\delta_0 = (kr)^{-O(1)}$ be such that

$$\big(1 + 4\epsilon_0^{-1}rk + (r+1)k\tau_1^{-2} + (r+1)k^2\delta_1^{-1}\big)\delta_0 \leq \frac{\nu_0^2}{8}.$$

- Finally, let $\frac{\nu_0^2}{8(4rk\epsilon_0^{-1}+1)} \geq \tau_0 = (kr)^{-O(1)}$ be such that the first conclusion of Corollary 6.2 does not follow in any application of that corollary with parameters δ_0 and r.

Apply Lemma 6.1 with $K = 2$ and parameter τ_0 to get a $k^{-O(1)}$-hereditarily energetic set S with $|A \cap S| \geq \nu_0|S|$ and $|S| \geq k^{-O(1)}|A|$ and a probability space $(\Omega, \mathbb{P})$ supporting random variables z, Z, and T such that

$$\|\mathbb{E} m_{z+Z} - m_S\| \leq \tau_0 \tag{6.1}$$

and for all $\omega \in \Omega$, $(Z(\omega), T(\omega)_0)$ is τ_0-closed, $\dim T(\omega) \leq k^{O(1)}$, and $\mathbb{C}^\flat(S; T(\omega)_0) \leq \exp((kr)^{O(1)})$. Apply Corollary 6.2 with parameters δ_0 and r, so that (in light of the choice of τ_0) we get an extension $(\Omega', \mathbb{P}')$ supporting random variables z', Z', such that

$$\|\mathbb{E}' m_{z'+Z'} - \mathbb{E} m_{z+Z}\| \leq \delta_0$$

and a further extension $(\Omega'', \mathbb{P}'')$ supporting z'', $Z_1, \ldots, Z_k$ such that

$$\|\mathbb{E}'' m_{2^s z''+Z_i} - \mathbb{E}' m_{2^s\cdot(z'+Z')}\| \leq \delta_0 \tag{6.2}$$

for all $1 \leq i \leq k$ and $0 \leq s \leq r$,

$$\mathbb{E}''|m_{2^s z''+Z_i}(A) - m_{2^s\cdot(z'+Z')}(A)|^2 \leq \delta_0 \tag{6.3}$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

for all $1 \leq i \leq k$ and $0 \leq s \leq r$, and

$$\mathbb{E}''\big\|1_{A\cap(2^s z''+Z_i)} * \big(1_A d m_{2^s z''+Z_j}\big) - m_{2^s\cdot(z'+Z')}(A)^2\big\|^2_{L_2(m_{2^{s+1}z''+Z_i})} \leq \delta_0 \tag{6.4}$$

for all $1 \leq i < j \leq k$ and $0 \leq s \leq r$, and for all $\omega'' \in \Omega''$ we have

$$\big(Z_i(\omega''), Z_{i+1}(\omega'')\big) \text{ is } \delta_0\text{-closed for all } 1 \leq i < k$$

and

$$\mathcal{C}^\flat\big(S; Z_k(\omega'')\big) \leq \exp\big((kr)^{O(1)}\big)\mathcal{C}^\flat\big(S; T(\omega'')\big) = \exp((kr)^{O(1)}).$$

For each $0 \leq s < r$ and $1 \leq i < j \leq k$, write

$$E_{s,i,j} := \Big\{\omega'' \in \Omega'' : \big\|1_{A\cap(2^s z''(\omega'')+Z_i(\omega''))} * \big(1_A d m_{2^s z''(\omega'')+Z_j(\omega'')}\big) - m_{2^s\cdot(z'(\omega'')+Z'(\omega''))}(A)^2\big\|^2_{L_2\left(m_{2^{s+1}z''(\omega'')+Z_i(\omega'')}\right)} \leq \delta_1\Big\},$$

so that by averaging and (6.4), we have $\mathbb{P}''(E^c_{s,i,j}) \leq \delta_1^{-1}\delta_0$. For $0 \leq s \leq r$ and $1 \leq i \leq k$, write

$$F_{s,i} := \Big\{\omega'' \in \Omega'' : |m_{2^s z''(\omega'')+Z_i(\omega'')}(A) - m_{2^s\cdot(z'(\omega'')+Z'(\omega''))}(A)| \leq \tau_1\Big\},$$

so that by averaging and (6.3) we have $\mathbb{P}''(F^c_{s,i}) \leq \tau_1^{-2}\delta_0$. For each $1 \leq s \leq r$ and $1 \leq i \leq k$, write

$$L_{s,i} := \Big\{\omega'' \in \Omega'' : m_{2^s z''(\omega'')+Z_i(\omega'')}(X \setminus A) \leq \frac{1}{4}\epsilon_0\Big\}.$$

Then by (6.2) and (6.1), we have

$$\begin{aligned}\mathbb{P}''(L^c_{s,i}) &\leq 4\epsilon_0^{-1}\mathbb{E}'' m_{2^s z''(\omega'')+Z_i(\omega'')}(X \setminus A)\\ &\leq 4\epsilon_0^{-1}\big(\delta_0 + \mathbb{E}' m_{2^s\cdot(z'+Z')}(X \setminus A)\big)\\ &\leq 4\epsilon_0^{-1}\big(\delta_0 + \tau_0 + m_S(X \setminus A)\big) \leq 4\epsilon_0^{-1}(\delta_0 + \tau_0 + \nu_0^{-1}\eta).\end{aligned}$$

Finally, let

$$B := \Big\{\omega'' \in \Omega'' : m_{z'(\omega'')+Z'(\omega'')}(A) \geq \frac{1}{2}\nu_0\Big\},$$

$$R := B \cap \Big(\bigcap_{s=1}^{r}\bigcap_{i=1}^{k} L_{s,i}\Big) \cap \Big(\bigcap_{s=0}^{r}\bigcap_{i=1}^{k} F_{s,i}\Big) \cap \Big(\bigcap_{s=0}^{r-1}\bigcap_{1\leq i<j\leq k} E_{s,i,j}\Big).$$

Then we have

$$\begin{aligned}&\mathbb{E}''1_R m_{z'+Z'}(A)\\ &\quad\geq \mathbb{E}'' m_{z'+Z'}(A)\\ &\qquad - \sum_{s=1}^{r}\sum_{i=1}^{k}\mathbb{P}''(L^c_{s,i}) - \sum_{s=0}^{r}\sum_{i=1}^{k}\mathbb{P}''(F^c_{s,i}) - \sum_{s=0}^{r-1}\sum_{1\leq i<j\leq k}\mathbb{P}''(E^c_{s,i,j})\\ &\quad\geq \big(m_S(A) - \|\mathbb{E}'' m_{z'+Z'} - \mathbb{E}m_{z+Z}\| - \|\mathbb{E}m_{z+Z} - m_S\|\big)\\ &\qquad - rk\cdot 4\epsilon_0^{-1}(\delta_0 + \tau_0 + \nu_0^{-1}\eta) - (r+1)k\tau_1^{-2}\delta_0 - (r+1)k^2\delta_1^{-1}\delta_0\\ &\quad\geq \nu_0 - \delta_0 - \tau_0\\ &\qquad - rk\cdot 4\epsilon_0^{-1}(\delta_0 + \tau_0 + \nu_0^{-1}\eta) - (r+1)k\tau_1^{-2}\delta_0 - (r+1)k^2\delta_1^{-1}\delta_0.\end{aligned} \tag{6.5}$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

This is at least $\frac{1}{2}\nu_0$ by choice of δ_0 and τ_0 provided $\eta \le \frac{\nu_0^2\epsilon_0}{16rk}$ (if not, then we are in the first conclusion of the proposition).

Claim *Either* $|A| \le \exp((kr)^{O(1)})$ *or else for all* $\omega'' \in E$, *we have*[14]

$$m_{2^s\cdot(z'(\omega'')+Z'(\omega''))}(A) \ge \min\left\{\frac{1}{2}\nu_0, \frac{1}{2}\epsilon_0\right\} \quad \textit{for all } 0 \le s \le r.$$

Proof We proceed by induction on s. When $s = 0$, we have $\omega'' \in B$ and so the hypothesis holds. Suppose that it holds for some $s < r$. Then since $\omega'' \in E_{s,i,j}$ for all $1 \le i < j \le k$, we have

$$\left\|1_{A\cap(2^sz''(\omega'')+Z_i(\omega''))} * \left(1_A dm_{2^sz''(\omega'')+Z_j(\omega'')}\right) - m_{2^s\cdot(z'(\omega'')+Z'(\omega''))}(A)^2\right\|^2_{L_2\left(m_{2^{s+1}z''(\omega'')+Z_i(\omega'')}\right)} \le \delta_1,$$

and since $\omega \in F_{s,i}$ for all $1 \le i \le k$, we have

$$|m_{2^sz''(\omega'')+Z_i(\omega'')}(A) - m_{2^s\cdot(z'(\omega'')+Z'(\omega''))}(A)| \le \tau_1.$$

Apply Lemma 6.3 with parameters ϵ_0, δ_1, τ_1, and $\alpha = m_{2^s\cdot(z'(\omega'')+Z'(\omega''))}(A)$ to the set A with $z_0 = 2^sz''(\omega'')$ and the δ_0-closed pairs $(Z_i(\omega''), Z_{i+1}(\omega''))$ (for $1 \le i < k$) valid, since $\delta_0 \le \tau_1$.

The inductive hypothesis tells us that $\alpha \ge \min\left\{\frac{1}{2}\nu_0, \frac{1}{2}\epsilon_0\right\}$, so in view of the choice of δ_1 and τ_1, either there is some $1 \le i < k$ such that $m_{2^{s+1}z''(\omega'')+Z_i(\omega'')}(X) > \epsilon_0$ or

$$\begin{aligned}\int \Bigg(\prod_{i<j} 1_{(Z_i+2^{s+1}z''(\omega''))\setminus X}(z_i+z_j+2^sz''(\omega''))1_{Z_i(\omega'')}(z_i+z_j)\Bigg) \\ \times \prod_{i=1}^k 1_A(z_i+2^sz''(\omega''))dm_{Z_i(\omega'')}(z_i) \\ = \Omega\left(\left(\min\left\{\frac{1}{2}\nu_0, \frac{1}{2}\epsilon_0\right\}\right)^k\right).\end{aligned}$$

Since A is (k, X)-summing, we know that the first product on the left is 0 on $z \in A^k$ unless there is $1 \le i < j \le k$ with $z_i = z_j$. It follows that the integral is at most

$$\begin{aligned}\binom{k}{2} \cdot \max\left\{\frac{1}{|Z_i(\omega'')|} : 1 \le i \le k\right\} \\ \le k^2 \max\left\{\frac{\mathcal{C}(S; Z_i(\omega''))}{|S|} : 1 \le i \le k\right\} \\ \le k^2\frac{1}{\nu_0|A|}\mathcal{C}^\flat(S; Z_k(\omega'')) \le \exp((rk)^{O(1)})|A|^{-1},\end{aligned}$$

where the first inequality follows from Lemma 5.2(iii); and the second follows from Lemma 5.3(i) and (iii) using the nesting $Z_k(\omega'') \subset Z_i(\omega'')$.

The upper bound on $|A|$ follows.

[14]It may help to note that apart from the $s = 0$ case, the lower bound we establish is always $\frac{1}{2}\epsilon_0$.

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Thus, we can assume $m_{2^{s+1}z''(\omega'')+Z_i(\omega'')}(X) > \epsilon_0$, and since $\omega \in L_{s+1,i}$, we have

$$m_{2^{s+1}z''(\omega'')+Z_i(\omega'')}(A) \geq m_{2^{s+1}z''(\omega'')+Z_i(\omega'')}(X) - m_{2^{s+1}z''(\omega'')+Z_i(\omega'')}(X \setminus A) > \frac{3}{4}\epsilon_0.$$

Finally, the claim is proved once we note that $\omega'' \in F_{s+1,i}$ and that $\tau_1 \leq \frac{1}{4}\epsilon_0$. ∎

We conclude that for all $1 \leq s \leq r$, we have

$$\begin{aligned} m_{2^s\cdot S}(A) &\geq \mathbb{E}'' m_{2^s\cdot(z'+Z')}(A) - \|m_{2^s\cdot S} - \mathbb{E}'' m_{2^s\cdot(z'+Z')}\| \\ &\geq \mathbb{E}'' 1_R m_{2^s\cdot(z'+Z')}(A) - \delta_0 - \tau_0 \\ &\geq \mathbb{E}'' 1_R m_{z'+Z'}(A) m_{2^s\cdot(z'+Z')}(A) - \delta_0 - \tau_0 \\ &\geq \min\left\{\frac{1}{2}\nu_0, \frac{1}{2}\epsilon_0\right\}\cdot\frac{1}{2}\nu_0 - \delta_0 - \tau_0 \end{aligned}$$

from the lower bound on the left of (6.5) noted immediately afterwards. The result follows, since $\delta_0, \tau_0 \leq \frac{1}{16}\min\{\nu_0^2, \nu_0\epsilon_0\}$. ∎

7 Finding Structure

In this section, we prove the following lemma.

Lemma (Lemma 6.1) *Suppose that A is (k, X)-summing, $|X| \leq K|A|$, and $\tau \in \left(0, \frac{1}{2}\right]$ is a parameter. Then there is a $(kK)^{-O(1)}$-hereditarily energetic set S with $|A \cap S| \geq (kK)^{-O(1)}|S|$ and $|S| \geq k^{-O(1)}|A|$, a probability space $(\Omega, \mathbb{P})$ supporting a G-valued random variable z, an $\mathcal{N}(G)$-valued random variable Z and an $\mathcal{S}(G)$-valued random variable T such that*

$$\|\mathbb{E}m_{z+Z} - m_S\| \leq \tau,$$

and for all $\omega \in \Omega$,

$$\dim T(\omega) \leq (kK)^{O(1)}, \quad \mathcal{C}^\flat(S; T(\omega)) \leq \exp\left((kK\log\tau^{-1})^{O(1)}\right),$$

and $(Z(\omega), T(\omega)_0)$ is τ-closed.

Recall from [TV06, Definition 0.2] that P is a generalised arithmetic progression of dimension d if there are elements $x_0, \dots, x_d$ and natural numbers $N_1, \dots, N_d \in \mathbb{N}$ such that

$$P = \left\{x_0 + t_1x_1 + \cdots + t_dx_d : 0 \leq t_j \leq N_j \text{ for all } 1 \leq j \leq d\right\}. \tag{7.1}$$

A coset progression of dimension d is a set of the form $P + H$ where P is a generalised arithmetic progression of dimension d and $H \leq G$. This corresponds to [TV06, Definition 4.20] where dimension is given the name rank instead.

Coset progressions are easily related to systems as follows.

Lemma 7.1 *Suppose that M is a d-dimensional coset progression. Then there is an $O(d)$-dimensional system B with $B_0 \subset M - M$ and $M \subset x_0 + B_0$ for some x_0.*

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Proof Write $M = P + H$ where P is as in (7.1) and put

$$B_i := \left\{t_1x_1 + \cdots + t_dx_d : t_j \in \mathbb{Z} \text{ and } |t_j| \le 2^{-i}N_j \text{ for all } 1 \le j \le d\right\} + H$$

so that $B_0 \subset M - M$ and $M \subset x_0 + B_0$. By the triangle inequality, $B_i + B_i \subset B_{i-1}$, and each set is a symmetric neighbourhood of 0_G, so $B = (B_i)_{i\in\mathbb{N}_0}$ is a system. Moreover,

$$B_i \subset B_{i+1} + \left\{\sigma_1\lceil 2^{-(i+1)}N_1\rceil x_1 + \cdots + \sigma_d\lceil 2^{-(i+1)}N_d\rceil x_d : \sigma_j \in \{-1,0,1\} \text{ for all } 1 \le j \le d\right\},$$

so $\mathcal{C}(B_i; B_{i+1}) \le 3^d$. It follows (*cf.* the proof of Lemma 5.5(ii)) that

$$\begin{aligned}\mathcal{C}^\flat(B_i; B_{i+1}) &\le \mathcal{C}^\flat(B_i; B_{i+2} - B_{i+2})\\ &\le \mathcal{C}(B_i; B_{i+2}) \le \mathcal{C}(B_i; B_{i+1})\mathcal{C}(B_{i+1}; B_{i+2}) \le \exp\big(O(d)\big),\end{aligned}$$

and we see that B is $O(d)$-dimensional. ∎

Proof of Lemma 6.1 By Lemma 3.2 we see that $E(A) \ge (kK)^{-O(1)}|A|^3$ and so by the Balog–Szemerédi–Gowers Theorem[15], there is a set $A' \subset A$ with $|A'+A'| \le (kK)^{O(1)}|A'|$ and $|A'| = (kK)^{-O(1)}|A|$. Apply the Ruzsa–Chang theorem [TV06, Theorem 5.46] to see that $2A' - 2A'$ contains a coset progression M of dimension $(kK)^{O(1)}$ such that $|M| \ge \exp(-(kK)^{O(1)})|A|$. By Lemma 7.1 there is a system B such that $\dim B = (kK)^{O(1)}$, $B_0 \subset M - M \subset 4A' - 4A'$, and $|B_0| \ge |M| \ge \exp(-(kK)^{O(1)})|A|$.

Since $|A' - A' + B_2| \le |A' - A' + B_0| \le |5A' - 5A'| \le (kK)^{O(1)}|A' - A'|$ by Plünnecke's inequality [TV06, Corollary 6.27], we can apply Lemma 5.6 to the system $2^{-2}B$ and set $A' - A' \in \mathcal{N}(G)$ to get $m = \log_2\log_2 kK + \log_2\tau^{-1} + O(1)$ and a set $S \in \mathcal{N}(G)$ with $A' - A' \subset S \subset A' - A' + B_2$ so that (S, B_{m+2}) is τ-closed. By Lemma 5.1(i), it follows that

$$\|\tau_t(m_S) - m_S\| \le \tau \quad \text{for all } t \in B_{m+2}.$$

Apply Lemma 5.7 to the system $2^{-(m+2)}B$ (which has dimension $(kK)^{O(1)}$ by Lemma 5.5 (ii)) to get a natural number $m' = O(\log_2 kK) + \log_2\tau^{-1}$ and a set Z with $B_{m+3} \subset Z_* \subset B_{m+2}$ such that $(Z_*, B_{m'+m+2})$ is τ-closed.

Let z be the random variable taking values in S uniformly; let Z be the constant random variable taking the value Z_*; and let T be the constant random variable taking the value $2^{-(m'+m+2)}B$. Then

$$\|\mathbb{E}m_{z+Z} - m_S\| = \|m_S * m_{Z_*} - m_S\| \le \int \|\tau_t(m_S) - m_S\|dm_{Z_*}(t) \le \tau.$$

Since $S \subset A' - A' + B_2$, we have $|S| \le (kK)^{O(1)}|A|$; since $S \supset A' - A'$, S intersects a translate of A in a set of size at least $|A'| = (kK)^{-O(1)}|A| = (kK)^{-O(1)}|S|$. By translating A if necessary (which results in translations of z and S too), we can assume that this translate is at 0_G, so that $|A \cap S| \ge (kK)^{-O(1)}|S|$. Since $S \subset A' - A' + B_2 \subset 5A' - 5A'$, we see that $|S + S| \le |10A' - 10A'| \le (kK)^{O(1)}|S|$ by Plünnecke's inequality [TV06, Corollary 6.27]. It follows from Lemma 3.1(i) that S is $(kK)^{-O(1)}$-hereditarily energetic.

[15] In its usual form, which corresponds to [TV06, Theorem 2.31((i) ⇒ (iv))] and then [TV06, Exercise 2.3.15]).

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Finally, we let T be the random variable taking the constant value $2^{-(m+m'+2)}B$. Then $(Z(\omega), T(\omega)_0)$ is τ-closed by design; Lemma 5.5(ii) tells us that

$$\dim T(\omega) = \dim 2^{-(m+m'+2)}B \leq \dim B = (kK)^{O(1)},$$

and also

$$\begin{aligned}\mathcal{C}^\flat\big(S; T(\omega)\big) &\leq \exp\big(O((m+m'+2))(kK)^{O(1)}\big)\mathcal{C}^\flat(S;B)\\ &= \exp\big((kK\log\tau^{-1})^{O(1)}\big)\mathcal{C}^\flat(S;B).\end{aligned}$$

Now, by Lemma 5.2(iii) and (i) and the second inequality in Lemma 5.3(iii), we have

$$\begin{aligned}\frac{|B_0|}{|B_2|} &\leq \mathcal{C}(B_0;B_2) \leq \mathcal{C}(B_0;B_1)\mathcal{C}(B_1;B_2) \leq \mathcal{C}^\flat(B_0;B_1)\mathcal{C}^\flat(B_1;B_2)\\ &\leq \exp((kK)^{O(1)}),\end{aligned}$$

so $|B_2| \geq \exp(-(kK)^{O(1)})|A|$. By Lemma 5.3(i) and (iii) and then Ruzsa's covering lemma (Lemma 5.2(iv)), we have

$$\begin{aligned}\mathcal{C}^\flat(S;B) = \mathcal{C}^\flat(S;B_0) &\leq \mathcal{C}^\flat(S;(B_2-B_2)-(B_2-B_2))\\ &\leq \mathcal{C}(S;B_2-B_2) \leq \frac{|S+B_2|}{|B_2|} \leq \frac{|A'-A'+B_2+B_2|}{|B_2|}.\end{aligned}$$

By Plünnecke's inequality [TV06, Corollary 6.27] and the lower estimate on $|B_2|$, we have $\mathcal{C}^\flat(S;B) \leq \exp((kK)^{O(1)})$, and the result is proved. ∎

8 Uniformity

The aim of this section is to prove Corollary 6.2 We will do this through the Fourier transform which we introduced in Section 3. The reader unfamiliar with its use in this context may wish to consult [TV06, Chapter 4] or the book [Rud90].

We begin with some definitions. The Fourier transform of a measure $\mu \in M(G)$ (where $M(G)$ denotes the set of measures on G with finite support) is

$$\widehat{\mu}(\gamma) := \int \overline{\gamma(x)}d\mu(x) \quad \text{for all } \gamma \in \widehat{G},$$

and given $f \in C(G)$, we define

$$\mu * f(x) = f * \mu(x) = \int f d\tau_x(\mu) \quad \text{for all } x \in G.$$

Given $\mu \in M(G)$ and a parameter $\epsilon > 0$, we write (*c.f.* [TV06, Definition 4.33])

$$\mathrm{Spec}_\epsilon(\mu) := \{\gamma \in \widehat{G} : |\widehat{\mu}(\gamma)| > \epsilon\}. \tag{8.1}$$

Motivated by [Rud90, Theorem 1.2.6], for $W \subset G$ and $\epsilon > 0$ a parameter, we write

$$\mathrm{Ann}(W,\epsilon) := \{\gamma \in \widehat{G} : |\gamma(x) - 1| < \epsilon \text{ for all } x \in W\}.$$

The spectrum and approximate annihilators of η-closed pairs are closely related by the next result, which is [GK09, Lemma 3.6].

Lemma 8.1 *Suppose that (Z,W) is η-closed, and $\kappa \in (0,1]$ is a parameter. Then* $\mathrm{Spec}_\kappa(m_Z) \subset \mathrm{Ann}(W, \eta\kappa^{-1})$.

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Proof For $x \in W$, simply note that

$$\begin{aligned}\kappa|\gamma(x)-1| &< |\gamma(x)\widehat{m_Z}(\gamma)-\widehat{m_Z}(\gamma)|\\ &= \Big|\int \overline{\gamma(y)}d(\tau_x(m_Z)-m_Z)(y)\Big| \le \|\tau_x(m_Z)-m_Z\| \le \eta. \qquad \blacksquare\end{aligned}$$

Dually to approximate annihilators, we have Bohr sets, which provide us with a ready supply of low dimensional systems.

Lemma 8.2 (Bohr sets) *Suppose that $\gamma \in \widehat{G}$ and $\rho > 0$. Then there is a system B with* $\dim B = O(1)$, $\mathcal{C}^\flat(G; B_0) = O(\rho^{-O(1)})$, *and*

$$|\gamma(x)-1| < \rho \quad \textit{for all } x \in B_0.$$

Proof Let $A_i := \{z \in S^1 : |1-z| < 2^{2-i}\}$ so that $A_{i+1} - A_{i+1} \subset A_i$ by the triangle inequality. Let $T_i := \{z \in S^1 : |1-z| \in \{0, 2^{-i}, 2\cdot 2^{-i}, 3\cdot 2^{-i}, 4\cdot 2^{-i}\}\}$ so that $|T_i| \le 9$ and $A_i \subset T_i + A_{i+2}$, whence $\mathcal{C}(A_i; A_{i+2}) = O(1)$. Put $B_i' = \gamma^{-1}(A_i)$ for all $i \in \mathbb{N}_0$. By the definition of $\mathcal{C}^\flat$, we have

$$\mathcal{C}^\flat(B_i'; B_{i+1}') \le \mathcal{C}(A_i; A_{i+2}) = O(1),$$

so B' is $O(1)$-dimensional. On the other hand, there is $m = \log_2 \rho^{-1} + O(1)$ such that $B_m' \subset \{x : |\gamma(x)-1| < \rho\}$; setting $B := 2^{-m}B$ gives the result by Lemma 5.5(ii). $\blacksquare$

The Parseval bound is the standard way to bound the size of spectrum (*cf.* [TV06, (4.38)]). In our approximate setting, we do not have perfectly orthogonal characters necessary for Parseval's theorem, but fortunately we do have a notion of "almost orthogonal". This was first exploited to achieve a result of the following type by Green and Tao; see [GT08, Corollary 8.6].

Lemma 8.3 (The Parseval bound) *Suppose that (Z, W) is τ-closed, $f \in L_2(m_Z)$ has $\|f\|_{L_2(m_Z)} \le 1$, and $\epsilon > 0$ and $\delta \in (0, \frac{1}{2}]$ are parameters. Then there is a system B with* $\dim B = O(\epsilon^{-2})$ *and* $\mathcal{C}^\flat(G; B_0) \le \delta^{-O(\epsilon^{-2})}$ *and*

$$\mathrm{Spec}_\epsilon(f dm_Z) \subset \mathrm{Ann}(B_0 \cap W, 4\epsilon^{-2}\tau + \delta).$$

Proof Let $\Lambda \subset \mathrm{Spec}_\epsilon(f dm_Z)$ be a maximal subset such that if $(\lambda + \mathrm{Ann}(W, 2\epsilon^{-2}\tau)) \cap (\lambda' + \mathrm{Ann}(W, 2\epsilon^{-2}\tau)) \ne \varnothing$ for some $\lambda, \lambda' \in \Lambda$, then $\lambda = \lambda'$. By maximality, we have

$$\mathrm{Spec}_\epsilon(f dm_Z) \subset \Lambda + \mathrm{Ann}(W, 2\epsilon^{-2}\tau) - \mathrm{Ann}(W, 2\epsilon^{-2}\tau) \subset \Lambda + \mathrm{Ann}(W, 4\epsilon^{-2}\tau),$$

where the last inclusion is by the triangle inequality.

On the other hand, writing σ_λ for the sign of $\widehat{f dm_Z}(\lambda)$, we can use linearity and the Cauchy–Schwarz inequality to see that

$$\begin{aligned}|\Lambda|\epsilon \le \sum_{\lambda\in\Lambda} |\widehat{f dm_Z}(\lambda)| &= \Big\langle f, \sum_{\lambda\in\Lambda} \sigma_\lambda \lambda \Big\rangle_{L_2(m_Z)}\\ &\le \|f\|_{L_2(m_Z)} \Big(\sum_{\lambda,\lambda'\in\Lambda} |\langle \lambda, \lambda'\rangle_{L_2(m_Z)}| \Big)^{\frac{1}{2}}.\end{aligned}$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

By Lemma 8.1, we have $\mathrm{Spec}_{\frac{1}{2}\epsilon^2}(m_Z) \subset \mathrm{Ann}(W, 2\epsilon^{-2}\tau)$, so

$$\sum_{\lambda,\lambda'\in\Lambda} |\langle \lambda, \lambda'\rangle_{L_2(m_Z)}| \leq |\Lambda| + \frac{1}{2}\epsilon^2|\Lambda|^2.$$

Rearranging and using the fact that $\|f\|_{L_2(m_Z)} \leq 1$ (by hypothesis), we see that $|\Lambda| \leq 2\epsilon^{-2}$. For each $\lambda \in \Lambda$ let $B^{(\lambda)}$ be the system given by Lemma 8.2; let $B := \bigcap_{\lambda\in\Lambda} B^{(\lambda)}$. Then by Lemma 5.5(i) we see that $\dim B = O(\epsilon^{-2})$, and by Lemma 5.3(iii) we have $\mathcal{C}^\flat(G; B_0) \leq \delta^{-O(\epsilon^{-2})}$. By definition, $\Lambda \subset \mathrm{Ann}(B_0, \delta)$; the result follows. ∎

We are now in a position to prove that large local U_2-norm gives rise to a density increment. As a small word of caution, we remark that normalising constants below may not be exactly as expected; this is already apparent from the definition of the spectrum in (8.1), where we look at characters satisfying $|\widehat{\mu}(\gamma)| > \epsilon$ but *not* $|\widehat{\mu}(\gamma)| > \epsilon\|\mu\|$. The latter is essentially the definition in [TV06, Definition 4.33].

Lemma 8.4 *Suppose that $A \subset G$, (Z_0, Z_1) and (Z_1, Z_2) are τ-closed, $|m_{Z_0}(A) - \alpha| < \tau$ and $|m_{Z_1}(A) - \alpha| < \tau$, and $\epsilon, \delta > 0$ are parameters. Then there is a Bohr system B with $\dim B = O(\epsilon^{-2})$ and $\mathcal{C}^\flat(G; B_0) \leq \delta^{-O(\epsilon^{-2})}$ such that*

$$\|1_A * m_{Z'} - \alpha\|^2_{L_2(m_{Z_0})} \geq \|1_{A\cap Z_0} * (1_A dm_{Z_1}) - \alpha^2\|^2_{L_2(m_{Z_0})} - O(\epsilon^2 + \epsilon^{-2}\tau + \delta)$$

for all $Z' \subset Z_2 \cap B_0$.

Proof First note that

$$1_{Z_0} * (1_A dm_{Z_1})(z) = m_{Z_1}(A) \quad \text{for all } z \in Z_0^-,$$

whence

$$\begin{aligned}\|1_{Z_0} * (1_A dm_{Z_1}) - \alpha\|_{L_1(m_{Z_0})} &\leq |m_{Z_1}(A) - \alpha| + \max\{\alpha, 1-\alpha\} m_{Z_0}(Z_0 \setminus Z_0^-)\\ &= O(\tau).\end{aligned}$$

The inequality $\|f\|_{L_2}^2 \leq \|g\|_{L_2}^2 + (\|f\|_{L_\infty} + \|g\|_{L_\infty})\|f - g\|_{L_1}$ and Plancherel's theorem then give

$$\begin{aligned}&|Z_0|\|1_{A\cap Z_0} * (1_A dm_{Z_1}) - \alpha^2\|^2_{L_2(m_{Z_0})}\\ &\quad\leq |Z_0|\|(1_{A\cap Z_0} - \alpha 1_{Z_0}) * (1_A dm_{Z_1})\|^2_{L_2(m_{Z_0})}\\ &\qquad + O\big(|Z_0|\|1_{Z_0} * (1_A dm_{Z_1}) - \alpha\|_{L_1(m_{Z_0})}\big)\\ &\quad= |Z_0|\|(1_{A\cap Z_0} - \alpha 1_{Z_0}) * (1_A dm_{Z_1})\|^2_{L_2(m_{Z_0})} + O(\tau|Z_0|)\\ &\quad\leq \|(1_{A\cap Z_0} - \alpha 1_{Z_0}) * (1_A dm_{Z_1})\|^2_{\ell_2(G)} + O(\tau|Z_0|)\\ &\quad= \int |(1_{A\cap Z_0} - \alpha 1_{Z_0})\hat{}(\gamma)|^2 |(1_A dm_{Z_1})\hat{}(\gamma)|^2 d\gamma + O(\tau|Z_0|).\end{aligned}$$

On the other hand

$$\int |(1_{A\cap Z_0} - \alpha 1_{Z_0})\hat{}(\gamma)|^2 d\gamma = \|1_{A\cap Z_0} - \alpha 1_{Z_0}\|^2_{\ell_2(G)} \leq (1 + O(\tau))|Z_0|,$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

and so

$$|Z_0|\|1_{A\cap Z_0} * (1_A dm_{Z_1}) - \alpha^2\|^2_{L_2(m_{Z_0})} \leq \int_{\mathrm{Spec}_\epsilon(1_A dm_{Z_1})} |(1_{A\cap Z_0} - \alpha 1_{Z_0})^\wedge(\gamma)|^2 d\gamma + \epsilon^2|Z_0| + O(\tau|Z_0|).$$

Apply Lemma 8.3 (we can certainly assume $\delta \leq \frac{1}{2}$, or else there is nothing to prove) to the τ-closed pair (Z_1, Z_2) and 1_A (which has $\|1_A\|_{L_2(m_{Z_1})} \leq 1$) to get a system B with $\dim B = O(\epsilon^{-2})$ and $\mathcal{C}^\flat(G; B_0) \leq \delta^{-O(\epsilon^{-2})}$ such that

$$\mathrm{Spec}_\epsilon(1_A dm_{Z_1}) \subset \mathrm{Ann}(B_0 \cap Z_2, 4\epsilon^{-2}\tau + \delta).$$

It follows that if $Z' \subset B_0 \cap Z_2$, then

$$|\widehat{m_{Z'}}(\gamma) - 1| = O(\epsilon^{-2}\tau + \delta) \quad \text{for all } \gamma \in \mathrm{Spec}_\epsilon(1_A dm_{Z_1}),$$

and hence by the triangle inequality,

$$\begin{aligned}\int |(1_{A\cap Z_0} - \alpha 1_{Z_0})^\wedge(\gamma)|^2 |\widehat{m_{Z'}}(\gamma)|^2 d\gamma &\geq \int_{\mathrm{Spec}_\epsilon(1_A dm_{Z_1})} |(1_{A\cap Z_0} - \alpha 1_{Z_0})^\wedge(\gamma)|^2 d\gamma - O\big((\epsilon^{-2}\tau + \delta)|Z_0|\big) \\ &\geq |Z_0|\|1_{A\cap Z_0} * (1_A dm_{Z_1}) - \alpha^2\|^2_{L_2(m_{Z_0})} - O\big((\epsilon^2 + \epsilon^{-2}\tau + \delta)|Z_0|\big).\end{aligned}$$

The result follows on dividing by $|Z_0|$ and applying Parseval's theorem again to see that

$$\begin{aligned}\int |(1_{A\cap Z_0} - \alpha 1_{Z_0})^\wedge(\gamma)|^2 |\widehat{m_{Z'}}(\gamma)|^2 d\gamma &= \|(1_{A\cap Z_0} - \alpha 1_{Z_0}) * m_{Z'}\|^2_{\ell_2(G)} \\ &= \|(1_{A\cap Z_0} - \alpha 1_{Z_0}) * m_{Z'}\|^2_{\ell_2(Z_0^-)} + O(|Z_0^+ \setminus Z_0^-|) \\ &= \|1_A * m_{Z'} - \alpha\|^2_{\ell_2(Z_0^-)} + O(|Z_0^+ \setminus Z_0^-|) \\ &\leq \|1_A * m_{Z'} - \alpha\|^2_{L_2(m_{Z_0})}|Z_0| + O(\tau|Z_0|),\end{aligned}$$

since $Z' \subset Z_2 \subset Z_1$, and (Z_0, Z_1) is τ-closed. ∎

We now have all the Fourier tools we need. The next lemma captures some facts about the non-model analogue of weighted covers (from Step I in Section 4) in a useful package for the main iteration lemma in our argument.

Lemma 8.5 *Suppose that G has no 2-torsion, $A \subset G$, $(\Omega, \mathbb{P})$ is a probability space, z is a G-valued random variable, and Z is an $\mathcal{N}(G)$-valued random variable. Then there is an extension $(\Omega', \mathbb{P}')$ of $(\Omega, \mathbb{P})$ supporting a G-valued random variable z', such that for any $\mathcal{N}(G)$-valued random variable W on Ω we have the following.*

- *For all $0 \leq s \leq r$,*

$$\mathbb{E}\|1_A * m_{2^r\cdot W} - m_{2^s\cdot(z+Z)}(A)\|^2_{L_2(m_{2^s\cdot(z+Z)})} = \mathbb{E}'|m_{2^s z' + 2^r\cdot W}(A) - m_{2^s\cdot(z+Z)}(A)|^2 \tag{8.2}$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

and

$$\|\mathbb{E}'m_{z'+W} - \mathbb{E}m_{z+Z}\| \leq \tau \text{ if } (Z(\omega), W(\omega)) \text{ is } \tau\text{-closed for all } \omega \in \Omega. \tag{8.3}$$

- *For all* $0 \leq s_0 \leq r$, *we have*

$$\sum_{s=0}^{r} \mathbb{E}'m_{2^s\cdot(z'+2^{r-s_0}\cdot W)}(A)^2 \geq \sum_{s=0}^{r} \mathbb{E}m_{2^s\cdot(z+Z)}(A)^2 + \mathbb{E}'|m_{2^{s_0}z'+2^r\cdot W}(A) - m_{2^{s_0}\cdot(z+Z)}(A)|^2 - O(\tau' r) \tag{8.4}$$

if $(Z(\omega), 2^r W(\omega))$ *is* τ'*-closed for all* $\omega \in \Omega$.

Proof Let $\Omega' := \{(\omega, z_*) : \omega \in \Omega, z_* \in Z(\omega)\}$ and

$$\mathbb{P}'(\{(\omega, z_*)\}) = \frac{1}{|Z(\omega)|} \cdot \mathbb{P}(\{\omega\}) \quad \text{for all } (\omega, z_*) \in \Omega'.$$

The space $(\Omega', \mathbb{P}')$ is an extension of $(\Omega, \mathbb{P})$ via the canonical projection $\Omega' \to \Omega$. Let $z'(\omega, z_*) := z(\omega) + z_*$.

The first part is immediate once the definition has been unpacked (using the fact that G has no 2-torsion so that $|2^s \cdot (z(\omega) + Z(\omega))| = |Z(\omega)|$. For the second part, use Lemma 5.1(i) to see that

$$\begin{aligned}
\|\mathbb{E}'m_{z'+W} - \mathbb{E}m_{z+Z}\| &= \left\|\mathbb{E}\int m_{z(\omega)+z_*+W(\omega)}dm_{Z(\omega)}(z_*) - \mathbb{E}m_{z(\omega)+Z(\omega)}\right\| \\
&= \left\|\mathbb{E}\int \left(\tau_w(m_{z(\omega)+Z(\omega)}) - m_{z(\omega)+Z(\omega)}\right)dm_{W(\omega)}(w)\right\| \\
&\leq \mathbb{E}\int \left\|\tau_w(m_{z(\omega)+Z(\omega)}) - m_{z(\omega)+Z(\omega)}\right\|dm_{W(\omega)}(w) \\
&= \mathbb{E}\int \left\|\tau_w(m_{Z(\omega)}) - m_{Z(\omega)}\right\|dm_{W(\omega)}(w) \leq \tau.
\end{aligned}$$

For the third part, note that for $0 \leq s \leq r$, we have

$$\begin{aligned}
&\mathbb{E}'|m_{2^s z'+2^{r+s-s_0}\cdot W}(A) - m_{2^s\cdot(z+Z)}(A)|^2 \\
&= \mathbb{E}\int |m_{2^s\cdot(z+z_*+2^{r-s_0}\cdot W)}(A) - m_{2^s\cdot(z+Z)}(A)|^2 dm_Z(z_*) \\
&= \mathbb{E}\Bigg(\int m_{2^s\cdot(z+z_*+2^{r-s_0}\cdot W)}(A)^2 dm_Z(z_*) \\
&\qquad - 2m_{2^s\cdot(z+Z)}(A)\int m_{2^s\cdot(z+z_*+2^{r-s_0}\cdot W)}(A)dm_Z(z_*) \\
&\qquad + m_{2^s\cdot(z+Z)}(A)^2\Bigg) \\
&= \mathbb{E}'m_{2^s\cdot(z'+2^{r-s_0}\cdot W)}(A)^2 - \mathbb{E}'m_{2^s\cdot(z+Z)}(A)^2 + O(\tau').
\end{aligned} \tag{8.5}$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

The last equality follows, since

$$\begin{aligned}\int & m_{2^s\cdot(z(\omega)+z_*+2^{r-s_0}\cdot W(\omega))}(A)\,dm_{Z(\omega)}(z_*)\\ &= \int m_{2^s\cdot(z(\omega)+Z(\omega)+2^{r-s_0}w)}(A)\,dm_{W(\omega)}(w)\\ &= \int \tau_{2^{r-s_0}w}(m_{z(\omega)+Z(\omega)})(2^{-s}\cdot(A\cap 2^s\cdot G))\,dm_{W(\omega)}(w)\\ &= m_{z(\omega)+Z(\omega)}(2^{-s}\cdot(A\cap 2^s\cdot G))+O(\tau') = m_{2^s\cdot(z(\omega)+Z(\omega))}(A)+O(\tau'),\end{aligned}$$

in view of the τ'-closure of $(Z(\omega), 2^{r-s_0}\cdot W(\omega))$ (inherited since $2^{r-s_0}\cdot W(\omega)\subset 2^r W(\omega)$) and Lemma 5.1 (i) and the fact that G has no 2-torsion.

Sum (8.5) over s, and for $s\neq s_0$ use the lower bound of 0 for the left-hand side, valid since it is an average over a square. The result follows. ∎

With these results in hand we turn to the main technical ingredient of the whole argument, which we shall iterate to get Corollary 6.2.

Lemma 8.6 *Suppose that G has no 2-torsion, $A, S\subset G$, and $(\Omega,\mathbb{P})$ is a probability space supporting a G-valued random variable z, an $\mathcal{N}(G)$-valued random variable Z, and an $\mathcal{S}(G)$-valued random variable T such that for all $\omega\in\Omega$, we have*

$$\big(Z(\omega), T(\omega)_0\big) \text{ is } \tau\text{-closed, } \dim T(\omega)\le d, \text{ and } \mathcal{C}^\flat(S; T(\omega))\le D.$$

Let $\delta,\nu\in(0,\frac12]$ and $r\in\mathbb{N}$ be parameters. Then either $\tau^{-1}\le(\delta^{-1}r)^{O(1)}$, or

(i) *there is an extension $(\Omega'',\mathbb{P}'')$ of $(\Omega,\mathbb{P})$ supporting a G-valued random variable z'', an $\mathcal{N}(G)$-valued random variable Z'', and a $\mathcal{S}(G)$-valued random variable T'' with*

$$\|\mathbb{E}''m_{z''+Z''}-\mathbb{E}m_{z+Z}\|\le\tau$$

and

$$\Big(\sum_{s=0}^{r}\mathbb{E}''m_{2^s\cdot(z''+Z'')}(A)^2\Big)\ge\Big(\sum_{s=0}^{r}\mathbb{E}m_{2^s\cdot(z+Z)}(A)^2\Big)+\delta^{O(1)},\tag{8.6}$$

such that for all $\omega''\in\Omega''$, $(Z''(\omega''), T''(\omega'')_0)$ is ν-closed, and

$$\dim T''(\omega'')\le d+\delta^{-O(1)},\quad \mathcal{C}^\flat\big(S; T''(\omega'')\big)\le D\exp\big((dkr\delta^{-1}\log\nu^{-1})^{O(1)}\big);$$

(ii) *or there is an extension $(\Omega',\mathbb{P}')$ of $(\Omega,\mathbb{P})$, supporting a G-valued random variable z' and $\mathcal{N}(G)$-valued random variables $Z_1,\dots,Z_k$ such that*

(a)
$$\|\mathbb{E}'m_{2^sz'+Z_i}-\mathbb{E}m_{2^s\cdot(z+Z)}\|\le\tau$$
for all $1\le i\le k$ and $0\le s\le r$;

(b) (U_1-uniformity)
$$\mathbb{E}'|m_{2^sz'+Z_i}(A)-m_{2^s\cdot(z+Z)}(A)|^2\le\delta$$
for all $1\le i\le k$ and $0\le s\le r$;

(c) (U_2-uniformity)
$$\mathbb{E}'\big\|1_{A\cap(2^sz'+Z_i)}*(1_A\,dm_{2^sz'+Z_j})-m_{2^s\cdot(z+Z)}(A)^2\big\|^2_{L_2(m_{2^{s+1}z'+Z_i})}\le\delta$$
for all $1\le i<j\le k$ and $0\le s\le r$;

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

and for all $\omega' \in \Omega'$, $(Z_i(\omega'), Z_{i+1}(\omega'))$ *is* δ*-closed for all* $1 \le i < k$ *and* $\mathcal{C}^\flat(S; Z_k(\omega')) \le D\exp((dkr\delta^{-1}\log\nu^{-1})^{O(1)})$.

Proof Let $\delta_0, \delta_1, \delta_2, \delta_3$ be related constants to be optimised later. (They will all be of the shape $\delta^{O(1)}$.)

For each $\omega \in \Omega$ we shall create sets $Z_i'(\omega)$ and natural numbers m_i iteratively. Let $m_0 := r$, and suppose that m_i has been defined for some $0 \le i < k$. Apply Lemma 5.7 to the system $2^{-m_i}T(\omega)$ to get a set $T(\omega)_{m_i+1} \subset Z_{i+1}'(\omega) \subset T(\omega)_{m_i}$ and a natural number $m_{i+1} = m_i + \log_2 dr\delta_0^{-1}\nu^{-1} + O(1)$ such that $(Z_{i+1}'(\omega), T(\omega)_{m_{i+1}})$ is $r^{-1}\delta_0\nu$-closed. Although the m_is may depend on ω, by construction we have the universal bound $m_i = O(r + k\log dr\delta_0^{-1}\nu^{-1})$ for $0 \le i \le k$, and we have that

$$(Z_i'(\omega), Z_{i+1}'(\omega)) \text{ is } r^{-1}\delta_0\nu\text{-closed for all } 1 \le i < k \text{ and } \omega \in \Omega, \tag{8.7}$$

and (since $2^r Z_i'(\omega) \subset 2^r T(\omega)_{m_0} = 2^r T(\omega)_r \subset T(\omega)_0$)

$$(Z(\omega), 2^r Z_i'(\omega)) \text{ is } \tau\text{-closed for all } 1 \le i \le k \text{ and } \omega \in \Omega. \tag{8.8}$$

Apply Lemma 8.5 to z, Z to get an extension $(\Omega', \mathbb{P}')$ of $(\Omega, \mathbb{P})$. Suppose that there is some $1 \le i_0 \le k$ and $0 \le s_0 \le r$ such that

$$\mathbb{E}'|m_{2^{s_0}z'+2^r\cdot Z_{i_0}'}(A) - m_{2^{s_0}\cdot(z+Z)}(A)|^2 \ge \delta_1. \tag{8.9}$$

Put $Z'(\omega) := 2^{r-s_0}\cdot Z_{i_0}'(\omega)$ and use conclusion (8.4) of Lemma 8.5 (with $W = Z_{i_0}'$ applicable in view of (8.8)) to see that

$$\Big(\sum_{s=0}^{r}\mathbb{E}'m_{2^s\cdot(z'+Z')}(A)^2\Big) \ge \Big(\sum_{s=0}^{r}\mathbb{E}m_{2^s\cdot(z+Z)}(A)^2\Big) + \delta_1 - O(\tau r).$$

It follows that either $\tau^{-1} = O(r\delta_1^{-1})$ (and we are done) or else

$$\Big(\sum_{s=0}^{r}\mathbb{E}'m_{2^s\cdot(z'+Z')}(A)^2\Big) \ge \Big(\sum_{s=0}^{r}\mathbb{E}m_{2^s\cdot(z+Z)}(A)^2\Big) + \Omega(\delta_1).$$

Let $T'(\omega) := 2^{r-s_0}\cdot(2^{-m_{i_0}}T(\omega))$. By Lemma 5.1(iii), the fact that $(2^{-m_{i_0}}T(\omega))_0 = T(\omega)_{m_{i_0}}$, and the construction of $Z_{i_0}'(\omega)$ we have that $(Z'(\omega), T'(\omega)_0)$ is δ_0-closed. Moreover, by Lemma 5.5 (iii) and then Lemma 5.5 (ii)

$$\begin{aligned}\mathcal{C}^\flat(S; T'(\omega)_0) &= \mathcal{C}^\flat(S; 2^{r-s_0}\cdot T(\omega)_{m_{i_0}})\\ &\le \exp(O(rd))\mathcal{C}^\flat(S; T(\omega)_{m_{i_0}})\\ &\le \exp(O(rd + m_{i_0}))\mathcal{C}^\flat(S; T(\omega)_0)\\ &\le D\exp\big((dkr\delta_0^{-1}\log\nu^{-1})^{O(1)}\big).\end{aligned} \tag{8.10}$$

Finally, use conclusion (8.3) of Lemma 8.5 with $W = 2^{r-s_0}\cdot Z_{i_0}'$ (applicable in view of (8.8)) and the fact that

$$\begin{aligned}2^{r-s_0}\cdot Z_{i_0}'(\omega) &\subset 2^{r-s_0}Z_{i_0}'(\omega) \subset 2^r Z_{i_0}'(\omega)\\ &\subset 2^r T(\omega)_{m_{i_0-1}} \subset 2^r T(\omega)_{m_0} = 2^r T(\omega)_r \subset T(\omega)_0\end{aligned}$$

to see that

$$\|\mathbb{E}'m_{z'+Z'} - \mathbb{E}m_{z+Z}\| \le \tau.$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

Set $(\Omega'', \mathbb{P}'') := (\Omega', \mathbb{P}')$, $Z'' := Z'$, $z'' := z'$, and $T'' := T'$, and we are in case (i) of the lemma.

In view of this, we assume that there is no i_0 or s_0 (in the given ranges) such that (8.9) holds. Put $Z_i := 2^r \cdot Z_i'$ for all $1 \le i \le k$. Then for $1 \le i \le k$ and $0 \le s \le r$, we have

$$\begin{aligned}\left\|\mathbb{E}' m_{2^s z' + Z_i} - \mathbb{E} m_{2^s \cdot (z+Z)}\right\| &= \left\|\mathbb{E}' m_{2^s z' + 2^r \cdot Z_i'} - \mathbb{E} m_{2^s \cdot (z+Z)}\right\| \\ &= \left\|\mathbb{E}' m_{z' + 2^{r-s} \cdot Z_i'} - \mathbb{E} m_{z+Z}\right\| \le \tau,\end{aligned}$$

by (8.3) with $W = 2^{r-s} \cdot Z_i'$, since $2^{r-s} \cdot Z_i'(\omega) \subset 2^r Z_i'(\omega) \subset 2^r T(\omega)_{m_0} \subset T(\omega)_0$; we have established the conclusion (a).

Now, by (8.7) and Lemma 5.1(iii) we have that

$$\big(Z_i(\omega), Z_{i+1}(\omega)\big) \text{ is } \delta_0\text{-closed for all } 1 \le i < k. \tag{8.11}$$

Since $Z_k'(\omega) \supset T(\omega)_{m_k+1}$, Lemma 5.3(i) and almost exactly the same argument as in (8.10) shows that

$$\mathcal{C}^\flat\big(S; Z_k'(\omega)\big) \le D \exp\big((dkr\delta_0^{-1} \log \nu^{-1})^{O(1)}\big).$$

Since (8.9) does not hold for any $1 \le i \le k$ or $0 \le s \le r$, we have that

$$\mathbb{E}' |m_{2^s z' + Z_i}(A) - m_{2^s \cdot (z+Z)}(A)|^2 < \delta_1 \quad \text{for all } 1 \le i \le k, 0 \le s \le r.$$

Conclusion (b) follows.

Suppose that for all $1 \le i < j \le k$ and $0 \le s \le r$, we have

$$\mathbb{E}' \|1_{A \cap (2^s z' + Z_i)} * (1_A d m_{2^s z' + Z_j}) - m_{2^s \cdot (z+Z)}(A)\|^2_{L_2(m_{2^{s+1} z' + Z_i})} < \delta_2.$$

Then conclusion (c) follows, and we are in case (ii) of the lemma. Thus, we assume so that there are not elements $1 \le i_1 < j_1 \le k$ and $0 \le s_1 \le r$ such that

$$\mathbb{E}' \|1_{A \cap (2^{s_1} z' + Z_{i_1})} * (1_A d m_{2^{s_1} z' + Z_{j_1}}) - m_{2^{s_1} \cdot (z+Z)}(A)\|^2_{L_2(m_{2^{s_1+1} z' + Z_{i_1}})} \ge \delta_2. \tag{8.12}$$

For each $\omega \in \Omega$ put $T'(\omega) := 2^r \cdot (2^{-m_k} T(\omega))$ so that Lemma 5.1(iii) tells us that $(Z_k(\omega), T'(\omega)_0)$ is δ_0-closed.

For each $\omega' \in \Omega'$, apply Lemma 8.4 with $\alpha = m_{2^{s_1} \cdot (z(\omega') + Z(\omega'))}(A)$ and both free parameters equal to δ_3 (for reasons that will become clear) to $(Z_{i_1}(\omega'), Z_{j_1}(\omega'))$ and $(Z_{j_1}(\omega'), T'(\omega')_0)$ (which are both δ_0-closed, the former by (8.11)), and the set $A - 2^{s_1} z'(\omega')$.

Out of the lemma, we get a Bohr system $B''(\omega')$ with $\dim B''(\omega') \le \delta_3^{-O(1)}$ and $\mathcal{C}^\flat(G; B''(\omega')_0) \le \exp(\delta_3^{-O(1)})$ such that for any $W''(\omega') \subset T'(\omega')_0 \cap B''(\omega')_0$, we

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

have

$$\begin{aligned}
&\big\|1_{(A-2^{s_1}z'(\omega'))} * m_{W''(\omega')} - m_{2^{s_1}\cdot(z(\omega')+Z(\omega'))}(A)\big\|^2_{L_2(m_{Z_{i_1}(\omega')})}\\
&\quad\geq \big\|1_{(A-2^{s_1}z'(\omega'))\cap Z_{i_1}(\omega')} * (1_{A-2^{s_1}z'(\omega')}dm_{Z_{j_1}(\omega')})\\
&\qquad\qquad - m_{2^{s_1}\cdot(z(\omega')+Z(\omega'))}(A)^2\big\|^2_{L_2(m_{Z_{i_1}(\omega')})}\\
&\quad - O\big(\delta_3 + \delta_3^{-2}\delta_0 + \delta_3^{-2}|m_{2^{s_1}z'(\omega')+Z_{i_1}(\omega')}(A) - m_{2^{s_1}\cdot(z(\omega')+Z(\omega'))}(A)|\big)\\
&\quad - O\big(\delta_3^{-2}|m_{2^{s_1}z'(\omega')+Z_{j_1}(\omega')}(A) - m_{2^{s_1}\cdot(z(\omega')+Z(\omega'))}(A)|\big)\\
&\quad= \big\|1_{A\cap(2^{s_1}z'(\omega')+Z_{i_1}(\omega'))} * (1_A dm_{2^{s_1}z'(\omega')+Z_{j_1}(\omega')})\\
&\qquad - m_{2^{s_1}\cdot(z(\omega')+Z(\omega'))}(A)^2\big\|^2_{L_2(m_{2^{s_1+1}z'(\omega')+Z_{i_1}(\omega')})}\\
&\qquad - O\big(\delta_3 + \delta_3^{-2}\delta_0 + \delta_3^{-2}|m_{2^{s_1}z'(\omega')+Z_{i_1}(\omega')}(A) - m_{2^{s_1}\cdot(z(\omega')+Z(\omega'))}(A)|\big)\\
&\qquad - O\big(\delta_3^{-2}|m_{2^{s_1}z'(\omega')+Z_{j_1}(\omega')}(A) - m_{2^{s_1}\cdot(z(\omega')+Z(\omega'))}(A)|\big).
\end{aligned}$$

Apply Lemma 5.7 to the system $2^{-2r}(T'(\omega')\wedge B''(\omega'))$ (which has dimension $d+\delta_3^{-O(1)}$ by (Lemma 5.5(i) and (ii)) to get a set $Z''(\omega')\subset (T'(\omega')\wedge B''(\omega'))_{2r}$ and a natural number $m'=O(\log_2 d\delta_3^{-1}r\delta_0^{-1})$ such that $(Z''(\omega'),(T'(\omega')\wedge B''(\omega'))_{m'+2r})$ is $r^{-1}\delta_0$-closed. Put $T''(\omega'):=2^r\cdot(2^{-(m'+r)}(T'(\omega')\wedge B''(\omega')))$, which has dimension $d+\delta_3^{-O(1)}$, and

$$\mathcal{C}^\flat\big(S;T''(\omega')_0\big)\leq D\exp\big((dkr\delta_3^{-1}\delta_0^{-1}\log\nu^{-1})^{O(1)}\big).$$

Taking $W''(\omega'):=2^r\cdot Z''(\omega')$ (which is contained in $T'(\omega')_0\cap B''(\omega')_0$ by design), and averaging against $\mathbb{P}'$, we have

$$\begin{aligned}
&\mathbb{E}'\|1_{A-2^{s_1}z'} * m_{2^r\cdot Z''} - m_{2^{s_1}\cdot(z+Z)}(A)\|^2_{L_2(m_{Z_{i_1}})}\\
&\quad\geq \delta_2 - O\big(\delta_3+\delta_3^{-2}\delta_0+\delta_3^{-2}\mathbb{E}'|m_{2^{s_1}z'+Z_{i_1}}(A)-m_{2^{s_1}\cdot(z+Z)}(A)|\big)\\
&\qquad - O\big(\delta_3^{-2}\mathbb{E}'|m_{2^{s_1}z'+Z_{j_1}}(A)-m_{2^{s_1}\cdot(z+Z)}(A)|\big)\\
&\quad\geq \delta_2 - O\big(\delta_3+\delta_3^{-2}\delta_0+\delta_3^{-2}\big(\mathbb{E}|m_{2^{s_1}z'+Z_{i_1}}(A)-m_{2^{s_1}\cdot(z+Z)}(A)|^2\big)^{\frac12}\big)\\
&\qquad - O\big(\delta_3^{-2}\big(\mathbb{E}'|m_{2^{s_1}z'+Z_{j_1}}(A)-m_{2^{s_1}\cdot(z+Z)}(A)|^2\big)^{\frac12}\big)\\
&\quad\geq \delta_2 - O(\delta_3+\delta_3^{-2}\delta_0+\delta_3^{-2}\delta_1^{\frac12})
\end{aligned}$$

by (8.12), the Cauchy–Schwarz inequality, and the fact that (8.9) does not hold for $i_0=i_1$ and $s_0=s_1$, or $i_0=j_1$ and $s_0=s_1$.

Again, since (8.9) does not hold for $s_0=s_1$ and $i_0=i_1$, we have from the Cauchy–Schwarz inequality that

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

$$\begin{aligned}
&\mathbb{E}'\|1_A * m_{2^r\cdot Z''} - m_{2^{s_1}z'+Z_{i_1}}(A)\|^2_{L_2(m_{2^{s_1}z'+Z_{i_1}})}\\
&\quad\geq \mathbb{E}'\|1_A * m_{2^r\cdot Z''} - m_{2^{s_1}\cdot(z+Z)}(A)\|^2_{L_2(m_{2^{s_1}z'+Z_{i_1}})}\\
&\qquad - O(\mathbb{E}'|m_{2^{s_1}\cdot(z+Z)}(A) - m_{2^{s_1}z'+Z_{i_1}}(A)|)\\
&\quad\geq \mathbb{E}'\|1_{A-2^{s_1}z'} * m_{2^r\cdot Z''} - m_{2^{s_1}\cdot(z+Z)}(A)\|^2_{L_2(m_{Z_{i_1}})} - O(\delta^{-\frac{1}{2}}).
\end{aligned}$$

Combining all this tells us that

$$\mathbb{E}'\|1_A * m_{2^r\cdot Z''} - m_{2^{s_1}z'+Z_{i_1}}(A)\|^2_{L_2(m_{2^{s_1}z'+Z_{i_1}})} \geq \delta_2 - O(\delta_3 + \delta_3^{-2}\delta_0 + \delta_3^{-2}\delta_1^{\frac{1}{2}}).$$

Recall that $Z_{i_1} = 2^r\cdot Z'_{i_1}$ and apply Lemma 8.5 to $(\Omega',\mathbb{P}')$, z' and $2^{r-s_1}\cdot Z'_{i_1}$ to get $(\Omega'',\mathbb{P}'')$ and z''. Equation (8.2) combined with the above tells us that

$$\mathbb{E}''|m_{2^{s_1}z''+2^r\cdot Z''}(A) - m_{2^{s_1}\cdot(z'+2^{r-s_1}\cdot Z'_{i_1})}(A)|^2 \geq \delta_2 - O(\delta_3 + \delta_3^{-2}\delta_0 + \delta_3^{-2}\delta_1^{\frac{1}{2}}).$$

Since $(2^{r-s_1}\cdot Z'_{i_1}(\omega'), 2^rZ''(\omega'))$ is $r^{-1}\delta_0$-closed (since $2^{r-s_1}\cdot(2^rZ''(\omega'))\subset T'(\omega')_0$ by design), (8.4) for $s_0 = s_1$ tells us that

$$\begin{aligned}
&\sum_{s=0}^{r}\mathbb{E}''m_{2^s\cdot(z''+2^{r-s_1}\cdot Z'')}(A)^2 \geq\\
&\qquad\qquad \sum_{s=0}^{r}\mathbb{E}'m_{2^s\cdot(z'+2^{r-s_1}\cdot Z'_{i_1})}(A)^2 + \delta_2 - O(\delta_3 + \delta_3^{-2}\delta_0 + \delta_3^{-2}\delta_1^{\frac{1}{2}}).
\end{aligned}$$

On the other hand, from (8.4) (when we applied Lemma 8.5 to get $(\Omega',\mathbb{P}')$) and the fact that $(Z(\omega), 2^rZ'_{i_1}(\omega))$ is τ-closed (8.8), we also have

$$\sum_{s=0}^{r}\mathbb{E}'m_{2^s\cdot(z'+2^{r-s_1}\cdot Z'_{i_1})}(A)^2 \geq \sum_{s=0}^{r}\mathbb{E}m_{2^s\cdot(z+Z)}(A)^2 - O(\tau r).$$

Hence (either $\tau^{-1} = O(\delta_2^{-1}r^{-1})$), or else

$$\begin{aligned}
&\sum_{s=0}^{r}\mathbb{E}''m_{2^s\cdot(z''+2^{r-s_1}\cdot Z'')}(A)^2 \geq\\
&\qquad\qquad \sum_{s=0}^{r}\mathbb{E}m_{2^s\cdot(z+Z)}(A)^2 + \delta_2 - O(\delta_3 + \delta_3^{-2}\delta_0 + \delta_3^{-2}\delta_1^{\frac{1}{2}}).
\end{aligned}$$

Taking $\delta_2 = \delta$, $\delta_3 = c\delta$, $\delta_1 = c\delta^6$, and $\delta_0 = c\delta^3$ for some sufficiently small c gives the result, and we are in case (i) of the lemma. ∎

Corollary (Corollary 6.2) *Suppose that G has no 2-torsion; $A, S\subset G$; $(\Omega,\mathbb{P})$ is a probability space supporting a G-valued random variable z, an $\mathcal{N}(G)$-valued random variable Z, and an $\mathcal{S}(G)$-valued random variable T such that for all $\omega\in\Omega$,*

$$\dim T(\omega)\leq d \quad \textit{and} \quad \mathbb{C}^\flat(S;T(\omega))\leq D$$

and $(Z(\omega), T(\omega)_0)$ is τ-closed. Let $\delta\in(0,1]$ and $r\in\mathbb{N}$ be parameters. Then either $\tau^{-1}\leq(\delta^{-1}r)^{O(1)}$, or there is a probability space $(\Omega',\mathbb{P}')$ extending $(\Omega,\mathbb{P})$, supporting a G-valued random variable z' and an $\mathcal{N}(G)$-valued random variable Z' with

$$\|\mathbb{E}'m_{z'+Z'} - \mathbb{E}m_{z+Z}\|\leq\delta,$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

and a further extension $(\Omega'', \mathbb{P}'')$ *of* $(\Omega', \mathbb{P}')$*, supporting a* G*-valued random variable* z'' *and* $\mathcal{N}(G)$*-valued random variables* $Z_1, \dots, Z_k$ *such that*

(i)
$$\|\mathbb{E}'' m_{2^s z''+Z_i} - \mathbb{E}' m_{2^s\cdot(z'+Z')}\| \le \delta$$
for all $1 \le i \le k$ *and* $0 \le s \le r$*;*

(ii) (U_1-uniformity)
$$\mathbb{E}''|m_{2^s z''+Z_i}(A) - m_{2^s\cdot(z'+Z')}(A)|^2 \le \delta$$
for all $1 \le i \le k$ *and* $0 \le s \le r$*;*

(iii) (U_2-uniformity)
$$\mathbb{E}''\|1_{A\cap(2^s z''+Z_i)} * (1_A d m_{2^s z''+Z_j}) - m_{2^s\cdot(z'+Z')}(A)^2\|^2_{L_2(m_{2^{s+1}z''+Z_i})} \le \delta$$
for all $1 \le i < j \le k$ *and* $0 \le s \le r$*,*

and for all $\omega'' \in \Omega''$*,* $(Z_i(\omega''), Z_{i+1}(\omega''))$ *is* δ*-closed for all* $1 \le i < k$*, and* $\mathcal{C}^\flat(S; Z_k(\omega'')) \le D\exp((dkr\delta^{-1})^{O(1)})$.

Proof Let $\tau_0^{-1} = (r\delta^{-1})^{O(1)}$ be the function in the first conclusion of Lemma 8.6 applied with parameters r and δ. If $\tau^{-1} \le 2\delta^{-1}$, then terminate with $\tau^{-1} = (r\delta^{-1})^{O(1)}$, so assume not. Let $\delta_0 = \delta^{O(1)}$ be the lower bound in (8.6) (when that lemma is applied with parameter δ), and let $\nu_0^{-1} = (r\delta^{-1})^{O(1)}$ be such that $\nu_0 < \tau_0$ and $\nu_0(r+1)\lceil\delta_0^{-1}\rceil + \tau \le \delta$, which is possible, since $\tau \le \frac{1}{2}\delta$.

We proceed inductively to define $\Omega^{(i)}, \mathbb{P}^{(i)}$ such that $(\Omega^{(i)}, \mathbb{P}^{(i)})$ is a probability space, $z^{(i)}$ is a G-valued random variable, $Z^{(i)}$ is an $\mathcal{N}(G)$-valued random variable, and $T^{(i)}$ is an $\mathcal{S}(G)$-valued random variable such that for all $\omega^{(i)} \in \Omega^{(i)}$, we have

$$\big(Z^{(i)}(\omega^{(i)}), T^{(i)}(\omega^{(i)})\big) \text{ is } \tau\text{-closed if } i = 0, \text{ and } \nu_0\text{-closed if } i > 0,$$
$$\dim T^{(i)}(\omega^{(i)}) \le d + i\delta^{-O(1)},$$
$$\mathcal{C}^\flat(S; T^{(i)}(\omega^{(i)})) \le D\exp(i(dkr\delta^{-1}\log\nu_0^{-1})^{O(1)}),$$

$$\|\mathbb{E}^{(i)} m_{z^{(i)}+Z^{(i)}} - \mathbb{E} m_{z+Z}\| \le \begin{cases} 0 & \text{if } i = 0 \\ \tau + (i-1)\nu_0 & \text{if } i > 0, \end{cases} \tag{8.13}$$

and

$$\sum_{s=0}^{r} \mathbb{E}^{(i)} m_{2^s\cdot(z^{(i)}+Z^{(i)})}(A)^2 \ge i\delta_0. \tag{8.14}$$

We initialise with $\Omega^{(0)} := \Omega$, $\mathbb{P}^{(0)} := \mathbb{P}$, $z^{(0)} := z$, $Z^{(0)} := Z$, and $T^{(0)} := T$, which satisfies the above requirements trivially. At stage i apply Lemma 8.6 to the space $(\Omega^{(i)}, \mathbb{P}^{(i)})$; random variables $z^{(i)}$, $Z^{(i)}$, and $T^{(i)}$; parameter ν_0 in place of ν; ν_0 or τ in place of τ (as named in Lemma 8.6) depending on whether $i = 0$ or $i > 0$; and δ and r as given.

Since $\nu_0^{-1} > \tau_0^{-1}$, we are not in the first case of the lemma. (And if $i = 0$, we can assume we are not in the first case or else we are in the τ large conclusion of the corollary.) We shall terminate if we are in case (ii) of the lemma, so assume not. It follows we are in case (i) of the lemma, which gives us an extension $(\Omega^{(i+1)}, \mathbb{P}^{(i+1)})$ of

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

$(\Omega^{(i)}, \mathbb{P}^{(i)})$ and random variables $z^{(i+1)}$, $Z^{(i+1)}$, and $T^{(i+1)}$ with (8.13) being a result of the triangle inequality.

In view of (8.14), this iteration cannot proceed for more than $(r+1)\lceil \delta_0^{-1} \rceil$ steps, at which point we are in case (ii) of Lemma 8.6. The conclusion follows from the triangle inequality again. ∎

9 Counting

In this section, we prove the following lemma, which is the analogue of the model Lemma 4.5. The key feature is that the bound on ϵ in the third of the four possible conclusions only depends on k. If we were prepared to admit α-dependence, then the uniformity of hypothesis (iii) would not be necessary.

Lemma (Lemma 6.3) *Suppose that $A, X \subset G$; $z_0 \in G$; (Z_i, Z_{i+1}) is τ-closed for all $1 \le i < k$, and*

(i) $|m_{Z_i}(A - z_0) - \alpha| \le \tau$ *for* $1 \le i \le k$;
(ii) $m_{Z_i}(X - 2z_0) \le \epsilon$ *for all* $1 \le i < k$;
(iii)

$$\left\| 1_{(A-z_0)\cap Z_i} * (1_{A-z_0} dm_{Z_j}) - \alpha^2 \right\|^2_{L_2(m_{Z_i})} \le \delta \text{ for all } 1 \le i < j \le k.$$

Then either $\delta^{-1} = O(k^2\alpha^{-4})$, $\tau^{-1} = O(k\alpha^{-1})$, $\epsilon^{-1} = O(k^2)$, or

$$\int \Big(\prod_{i<j} 1_{(Z_i+2z_0)\setminus X}(z_i + z_j + 2z_0) 1_{Z_i}(z_i + z_j) \Big) \prod_{i=1}^{k} 1_A(z_i + z_0) dm_{Z_i}(z_i) = \Omega(\alpha^k). \tag{9.1}$$

Proof Replacing A by $A + z_0$ and X by $X + 2z_0$, we can assume that $z_0 = 0_G$. Recall the inequality

$$\prod_{1\le i'<j'\le k} (1 - x_{i'j'}) \ge 1 - \sum_{1\le i'<j'\le k} x_{i'j'}$$

whenever $0 \le x_{i'j'} \le 1$ for all $1 \le i' < j' \le k$; this is what we call the pigeonhole principle. Write I for the integral in (9.1). Then using the stated pigeonhole principle and integrating, we have

$$I \ge \int \Big(\prod_{1\le i<j\le k} 1_{Z_i}(z_i + z_j) \Big) \prod_{i=1}^{k} 1_A(z_i) dm_{Z_i}(z_i) - \sum_{1\le i'<j'\le k} \int 1_{X\cap Z_{i'}}(z_{i'} + z_{j'}) \Big(\prod_{\substack{1\le i<j\le k \\ (i,j)\ne(i',j')}} 1_{Z_i}(z_i + z_j) \Big) \prod_{i=1}^{k} 1_A(z_i) dm_{Z_i}(z_i).$$

Since the Z_is are nested, for fixed $1 \le i < k$, we have

$$\prod_{i<j\le k} 1_{Z_i}(z_i + z_j) \ge 1_{Z_i^-}(z_i) \quad \text{for all } z_{i+1} \in Z_{i+1}, \dots, z_k \in Z_k.$$

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

From this and hypothesis (i), we conclude that

$$\begin{aligned} I &\geq \prod_{i=1}^{k} m_{Z_i}(A \cap Z_i^-) - \sum_{1\leq i'<j'\leq k} \int 1_{X\cap Z_{i'}}(z_{i'}+z_{j'}) \prod_{i=1}^{k} 1_A(z_i) dm_{Z_i}(z_i) \\ &\geq (\alpha-2\tau)^k - (\alpha+\tau)^{k-2} \sum_{1\leq i'<j'\leq k} \int 1_{X\cap Z_{i'}} \\ &\quad \times (z_{i'}+z_{j'}) 1_A(z_{i'}) dm_{Z_{i'}}(z_{i'}) 1_A(z_{j'}) dm_{Z_{j'}}(z_{j'}) \\ &= (\alpha-2\tau)^k - (\alpha+\tau)^{k-2} \sum_{1\leq i'<j'\leq k} \langle 1_X, 1_{A\cap Z_{i'}} * (1_A dm_{Z_{j'}}) \rangle_{L_2(m_{Z_{i'}})}. \end{aligned}$$

Now for fixed $1 \leq i' < j' \leq k$, the Cauchy–Schwarz inequality and hypothesis (iii) tell us that

$$\left| \langle 1_X, 1_{A\cap Z_{i'}} * (1_A dm_{Z_{j'}}) \rangle_{L_2(m_{Z_{i'}})} - \langle 1_X, \alpha^2 \rangle_{L_2(m_{Z_{i'}})} \right| \leq \delta^{\frac{1}{2}} m_{Z_{i'}}(X)^{\frac{1}{2}},$$

and so by (ii), we see that

$$\begin{aligned} &\langle 1_X, 1_{A\cap Z_{i'}} * (1_A dm_{Z_{j'}}) \rangle_{L_2(m_{Z_{i'}})} \leq \\ &\qquad \langle 1_X, \alpha^2 \rangle_{L_2(m_{Z_{i'}})} + (\delta m_{Z_{i'}}(X))^{\frac{1}{2}} < \epsilon\alpha^2 + (\epsilon\delta)^{\frac{1}{2}}. \end{aligned}$$

Combining all this, we get that

$$I \geq (\alpha-2\tau)^k - \binom{k}{2} (\alpha+\tau)^{k-2} (\epsilon\alpha^2 + (\epsilon\delta)^{\frac{1}{2}}).$$

It follows that either $\delta^{-1} = O(k^2\alpha^{-4})$, $\tau^{-1} = O(k\alpha^{-1})$, or $\epsilon^{-1} = O(k^2)$, and if none of these holds, then $I = \Omega(\alpha^k)$ as claimed. The result is proved. ∎

Acknowledgments The author should like to thank the three referees both for their comments, which very much improved the paper, and also their care in reading the paper, which, given its technical nature, was no small ask.

References

[Beh46] F. A. Behrend, *On sets of integers which contain no three terms in arithmetical progression*. Proc. Nat. Acad. Sci. USA 32(1946), 331–332. https://doi.org/10.1073/pnas.32.12.331

[Blo16] T. F. Bloom, *A quantitative improvement for Roth's theorem on arithmetic progressions*. J. Lond. Math. Soc. (2) 93(2016), 643–663. https://doi.org/10.1112/jlms/jdw010

[Bou99] J. Bourgain, *On triples in arithmetic progression*. Geom. Funct. Anal. 9(1999), 968–984. https://doi.org/10.1007/s000390050105

[BSS99] A. Baltz, T. Schoen, and A. Srivastav, *Probabilistic construction of small strongly sum-free sets via large Sidon sets*. In: *Randomization, approximation, and combinatorial optimization. Algorithms and techniques, Berlin, Heidelberg, 1999*. Springer, Berlin Heidelberg, 1999, pp. 138–143. https://doi.org/10.1007/978-3-540-48413-4_15.

[BSS00] A. Baltz, T. Schoen, and A. Srivastav, *Probabilistic construction of small strongly sum-free sets via large Sidon sets*. Colloq. Math. 86(2000), 171–176. https://doi.org/10.4064/cm-86-2-171-176

[Buk08] B. Bukh, *Sums of dilates*. Combin. Probab. Comput. 17(2008), 627–639. https://doi.org/10.1017/S096354830800919X

[Cho71] S. L. G. Choi, *On a combinatorial problem in number theory*. Proc. London Math. Soc. (3) 23(1971), 629–642. https://doi.org/10.1112/plms/s3-23.4.629

[CLP17] E. S. Croot, V. F. Lev, and P. P. Pach, *Progression-free sets in $\mathbb{Z}_4^n$ are exponentially small*. Ann. of Math. (2) 185(2017), 331–337. https://doi.org/10.4007/annals.2017.185.1.7

[CS10] E. S. Croot and O. Sisask, *A probabilistic technique for finding almost-periods of convolutions*. Geom. Funct. Anal. 20(2010), 1367–1396. https://doi.org/10.1007/s00039-010-0101-8

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press

[Dou13] J. Dousse, *On a generalisation of Roth's theorem for arithmetic progressions and applications to sum-free subsets*. Math. Proc. Cambridge Philos. Soc. 155(2013), 331–341. https://doi.org/10.1017/S0305004113000327

[Elk10] M. Elkin, *An improved construction of progression-free sets*. In: *Symposium on Discrete Algorithms*. 2010, pp. 886–905. https://doi.org/10.1007/s11856-011-0061-1.

[Erd65] P. Erdős, *Extremal problems in number theory*. In: *Proc. Sympos. Pure Math., Vol. VIII*. Amer. Math. Soc., Providence, RI, 1965, pp. 181–189. https://pdfs.semanticscholar.org/6754/29cbb9a130028ce8af5380a0330bb9733d9b.pdf.

[Gre02] B. J. Green, *Arithmetic progressions in sumsets*. Geom. Funct. Anal. 2(2002), 584–597. https://doi.org/10.1007/s00039-002-8258-4

[Gre05a] B. J. Green, *Counting sets with small sumset, and the clique number of random Cayley graphs*. Combinatorica 25(2005), 307–326. https://doi.org/10.1007/s00493-005-0018-2

[Gre05b] B. J. Green, *Finite field models in additive combinatorics*. In: *Surveys in combinatorics 2005*. London Math. Soc. Lecture Note Ser., 327, Cambridge University Press, Cambridge, 2005, pp. 1–27. https://doi.org/10.1017/CBO9780511734885.002.

[Gre05c] B. J. Green, *A Szemerédi-type regularity lemma in abelian groups, with applications*. Geom. Funct. Anal. 15(2005), 340–376. https://doi.org/10.1007/s00039-005-0509-8

[GK09] B. J. Green and S. V. Konyagin, *On the Littlewood problem modulo a prime*. Canad. J. Math. 61(2009), 141–164. https://doi.org/10.4153/CJM-2009-007-4

[GS08] B. J. Green and T. Sanders, *A quantitative version of the idempotent theorem in harmonic analysis*. Ann. of Math. (2) 168(2008), 1025–1054. https://doi.org/10.4007/annals.2008.168.1025

[GT08] B. J. Green and T. C. Tao, *An inverse theorem for the Gowers $U^3(G)$ norm*. Proc. Edinb. Math. Soc. (2) 51(2008), 73–153. https://doi.org/10.1017/S0013091505000325

[GT10] B. J. Green and T. C. Tao, *Linear equations in primes*. Ann. of Math. (2) 171(2010), 1753–1850. https://doi.org/10.4007/annals.2010.171.1753

[GW10] B. J. Green and J. Wolf, *A note on Elkin's improvement of Behrend's construction*. In: *Additive number theory: Festschrift in honor of the sixtieth birthday of Melvyn B. Nathanson*, First ed., Springer-Verlag, 2010, pp. 141–144. https://doi.org/10.1007/978-0-387-68361-4_9.

[Rud90] W. Rudin, *Fourier analysis on groups*. Wiley Classics Library. John Wiley & Sons Inc., New York, 1990. https://doi.org/10.1002/9781118165621

[Ruz05] I. Z. Ruzsa, *Sum-avoiding subsets*. Ramanujan J. 9(2005), 77–82. https://doi.org/10.1007/s11139-005-0826-4

[Sch11] T. Schoen, *Near optimal bounds in Freiman's theorem*. Duke Math. J. 158(2011), 1–12. https://doi.org/10.1215/00127094-1276283

[Sha15] X. Shao, *Finding linear patterns of complexity one*. Int. Math. Res. Not. 2015 2311–2327. https://doi.org/10.1093/imrn/rnu004

[Shk04] I. D. Shkredov, *On one problem of Gowers*. 2004. arxiv:math/0405406.

[Shk06] I. D. Shkredov, *On a problem of Gowers*. Izv. Ross. Akad. Nauk Ser. Mat. 70(2006), 179–221. https://doi.org/10.1070/IM2006v070n02ABEH002316

[Shk08] I. D. Shkredov, *On sets of large trigonometric sums*. Izv. Ross. Akad. Nauk Ser. Mat. 72(2008), 161–182. https://doi.org/10.1070/IM2008v072n01ABEH002396

[SSV05] B. Sudakov, E. Szemerédi, and V. H. Vu, *On a question of Erdős and Moser*. Duke Math. J. 129(2005), 129–155. https://doi.org/10.1215/S0012-7094-04-12915-X

[TV06] T. C. Tao and V. H. Vu, *Additive combinatorics*. Cambridge Studies in Advanced Mathematics, 105, Cambridge University Press, Cambridge, 2006. https://doi.org/10.1017/CBO9780511755149

[TV16] T. C. Tao and V. H. Vu, *Sum-avoiding sets in groups*. Discrete Anal. 31(2016), Paper No. 15. https://doi.org/10.19086/da.887.

[TV17] T. C. Tao and V. H. Vu, *Sum-free sets in groups: a survey*. J. Comb. 8(2017), 541–552. https://doi.org/10.4310/JOC.2017.v8.n3.a7

[Wol15] J. Wolf, *Finite field models in arithmetic combinatorics—ten years on*. Finite Fields Appl. 32(2015), 233–274. https://doi.org/10.1016/j.ffa.2014.11.003

Mathematical Institute, University of Oxford, Radcliffe Observatory Quarter, Woodstock Road, Oxford OX2 6GG, United Kingdom
e-mail : tom.sanders@maths.ox.ac.uk

https://doi.org/10.4153/S0008414X1900049X Published online by Cambridge University Press