OPTIMAL INVERSE LITTLEWOOD-OFFORD THEOREMS

HOI NGUYEN AND VAN VU

ABSTRACT. Let $\eta_i, i = 1, \dots, n$ be iid Bernoulli random variables, taking values ± 1 with probability $\frac{1}{2}$. Given a multiset V of n integers $v_1, \dots, v_n$, we define the *concentration probability* as

$$\rho(V) := \sup_x \mathbf{P}(v_1\eta_1 + \dots v_n\eta_n = x).$$

A classical result of Littlewood-Offord and Erdős from the 1940s asserts that if the v_i are non-zero, then $\rho(V)$ is $O(n^{-1/2})$. Since then, many researchers obtained improved bounds by assuming various extra restrictions on V.

About 5 years ago, motivated by problems concerning random matrices, Tao and Vu introduced the Inverse Littlewood-Offord problem. In the inverse problem, one would like to give a characterization of the set V, given that $\rho(V)$ is relatively large.

In this paper, we introduce a new method to attack the inverse problem. As an application, we strengthen a previous result of Tao and Vu, obtaining an optimal characterization for V. This immediately implies several classical theorems, such as those of Sárközy-Szemerédi and Halász.

The method also applies in the continuous setting and leads to a simple proof for the β-net theorem of Tao and Vu, which plays a key role in their recent studies of random matrices.

All results extend to the general case when V is a subset of an abelian torsion-free group and η_i are independent variables satisfying some weak conditions.

1. INTRODUCTION

1.1. **The Forward Littlewood-Offord problem.** Let $\eta_i, i = 1, \dots, n$ be iid Bernoulli random variables, taking values ± 1 with probability $\frac{1}{2}$. Given a multiset V of n integers $v_1, \dots, v_n$, we define the random walk S with steps in V to be the random variable $S := \sum_{i=1}^n v_i\eta_i$. The *concentration probability* is defined to be

$$\rho(V) := \sup_x \mathbf{P}(S = x).$$

Motivated by their study of random polynomials, in the 1940s Littlewood and Offord [13] raised the question of bounding $\rho(V)$. (We call this the *forward* Littlewood-Offord problem, in contrast with the *inverse* Littlewood-Offord problem discussed in the next section.) They showed that $\rho(V) = O(n^{-1/2} \log n)$. Shortly after Littlewood-Offord paper, Erdős [3] gave a beautiful combinatorial proof of the refinement

1991 *Mathematics Subject Classification.* 11B25.

Both authors are supported by research grants DMS-0901216 and AFOSAR-FA-9550-09-1-0167.

$$\rho(V) \leq \frac{\binom{n}{n/2}}{2^n} = O(n^{-1/2}). \quad (1)$$

Erdős' result is sharp, as demonstrated by $V = \{1, \dots, 1\}$.

Notation. Here and later asymptotic notations such as O, Ω, Θ, etc, are used under the assumption that $n \to \infty$. Notation such as $O_C(.)$ emphasizes that the hidden constant in O depends on C. If $a = \Omega(b)$, we write $b \ll a$ or $a \gg b$. All logarithm has natural base, if not specified otherwise.

The Littlewood-Offord and Erdős results are classic in combinatorics and have generated an impressive wave of research, in particular from the early 1960s to the late 1980s.

One direction of research was to generalize Erdős' result to other groups. For example, in 1966 and 1970, Kleitman extended Erdős' result to complex numbers and normed vectors, respectively. Several results in this direction can be found in [6, 10].

Another direction was motivated by the observation that (1) can be improved significantly under additional assumptions on V. The first such result was discovered by Erdős and Moser [4], who showed that if v_i are distinct, then $\rho(V) = O(n^{-3/2} \log n)$. They conjectured that the logarithmic term is not necessary and this was confirmed by Sárközy and Szemerédi [18].

Theorem 1.2. *Let V be a set of n different integers, then*

$$\rho(V) = O(n^{-3/2}).$$

In [8], Halász proved very general theorems that imply Theorem 1.2 and many others. One of his results can be formulated as follows.

Theorem 1.3. *Let l be a fixed integer and R_l be the number of solutions of the equation $v_{i_1} + \dots + v_{i_l} = v_{j_1} + \dots + v_{j_l}$. Then*

$$\rho(V) = O(n^{-2l-\frac{1}{2}} R_l).$$

It is easy to see, by setting $l = 1$, that Theorem 1.3 implies Theorem 1.2.

Another famous result in this area is that of Stanley [19], which, solving a conjecture of Erdős and Moser, shows when $\rho(V)$ attains its maximum under the assumption that the v_i are different.

Theorem 1.4. *Let n be odd and $V_0 := \{-\lfloor n/2 \rfloor, \dots, \lfloor n/2 \rfloor\}$. Then*

$$\rho(V) \leq \rho(V_0).$$

A similar result holds for the case n is even [19]. Stanley's proof of Theorem 1.4 used sophisticated machineries from algebraic geometry, in particular the hard-Lepschetz theorem.

Few years later, a more elementary proof was given by Proctor [15]. This proof is also of algebraic nature, involving the representation of the Lie algebra $sl(2, \mathbf{C})$. As far as we know, there is no purely combinatorial proof.

It is natural to ask for the actual value of $\rho(V_0)$. From Theorem 1.2, one would guess (under the assumption that the elements of V are different) that

$$\rho(V_0) = (C_0 + o(1))n^{-3/2}$$

for some constant $C_0 > 0$. However, the algebraic proofs does not give the value of C_0. In fact, it is not obvious that $\lim_{n\to\infty} n^{3/2}\rho(V_0)$ exists.

Assuming that C_0 exists for a moment, one would next wonder if V_0 is a stable maximizer. In other words, if some other set V_0' has $\rho(V_0')$ close to $C_0 n^{-3/2}$, then should V_0' (possibly after a normalization) "close" to V_0 ? (Notice that ρ is invariant under dilation so a normalization would be necessary.)

1.5. **The inverse Littlewood-Offord problem.** Motivated by inverse theorems from additive combinatorics (see [30, Chapter 5]) and a variant for random sums in [27, Theorem 5.2], Tao and the second author [25] brought a different view to the problem. Instead of trying to improve the bound further by imposing new assumptions as done in the forward problems, they tried to provide the full picture by finding the underlying reason for the concentration probability to be large (say, polynomial in n).

Notice that the (multi)-set V has 2^n subsums, and $\rho(V) \geq n^{-C}$ mean that at least $\frac{2^n}{n^C}$ among these take the same value. This suggests that the set should have a very strong additive structure. In order to determine this structure, we first discuss a few examples of V where $\rho(V)$ is large. For a set A, we denote by lA the set $\{a_1 + \cdots + a_l | a_i \in A\}$.

Example 1.6. Let $I = [-N, N]$ and $v_1, \dots, v_n$ be elements of I. Since $S \in nI$, by the pigeon hole principle, $\rho(V) \geq \frac{1}{|nI|} = \Omega(\frac{1}{nN})$. In fact, a short consideration yields a better bound. Notice that with probability at least .99, we have $S \in 10\sqrt{n}I$, thus again by the pigeonhole principle, we have $\rho(V) = \Omega(\frac{1}{\sqrt{n}N})$. If we set $N = n^{C-1/2}$ for some constant $C \geq 1/2$, then

$$\rho(V) = \Omega(\frac{1}{n^C}). \tag{2}$$

The next, and more general, construction comes from additive combinatorics. A very important concept in this area is that of *generalized arithmetic progressions* (GAPs). A set Q is a *GAP of rank* r if it can be expressed as in the form

$$Q = \{a_0 + x_1 a_1 + \cdots + x_r a_r | M_i \leq x_i \leq M_i' \text{ for all } 1 \leq i \leq r\}$$

for some $a_0, \dots, a_r, M_1, \dots, M_r, M_1', \dots, M_r'$.

It is convenient to think of Q as the image of an integer box $B := \{(x_1, \dots, x_r) \in \mathbf{Z}^r | M_i \leq m_i \leq M_i'\}$ under the linear map

$$\Phi : (x_1, \dots, x_r) \mapsto a_0 + x_1 a_1 + \cdots + x_r a_r.$$

The numbers a_i are the *generators* of P, the numbers M_i', M_i are the *dimensions* of P, and $\mathrm{Vol}(Q) := |B|$ is the *volume* of B. We say that Q is *proper* if this map is one to one, or equivalently if $|Q| = \mathrm{Vol}(Q)$. For non-proper GAPs, we of course have $|Q| < \mathrm{Vol}(Q)$. If $-M_i = M_i'$ for all $i \geq 1$ and $a_0 = 0$, we say that Q is *symmetric*.

Example 1.7. Let Q be a proper symmetric GAP of rank r and volume N. Let $v_1, \dots, v_n$ be (not necessarily distinct) elements of P. The random variable $S = \sum_{i=1}^n v_i \eta_i$ takes values in the GAP nP. Since $|nP| \leq \mathrm{Vol}(nB) = n^r N$, the pigeonhole principle implies that $\rho(V) \geq \Omega(\frac{1}{n^r N})$. In fact, using the same idea as in the previous example, one can improve the bound to $\Omega(\frac{1}{n^{r/2} N})$. If we set $N = n^{C-r/2}$ for some constant $C \geq r/2$, then

$$\rho(V) = \Omega(\frac{1}{n^C}). \tag{3}$$

The above examples show that if the elements of V belong to a proper GAP with small rank and small cardinality then $\rho(V)$ is large. A few years ago, Tao and the second author [25] showed that this is essentially the only reason:

Theorem 1.8 (Weak inverse theorem)**.** [25] *Let $C, \epsilon > 0$ be arbitrary constants. There are constants r and C' depending on C and ϵ such that the following holds. Assume that $V = \{v_1, \dots, v_n\}$ is a multiset of integers satisfying $\rho(V) \geq n^{-C}$. Then there is a proper symmetric GAP Q of rank at most r and volume at most $n^{C'}$ which contains all but at most $n^{1-\epsilon}$ elements of V (counting multiplicity).*

Remark 1.9. The presence of the small set of exceptional elements is not completely avoidable. For instance, one can add $o(\log n)$ completely arbitrary elements to V and only decrease $\rho(V)$ by a factor of $n^{-o(1)}$ at worst. Nonetheless we expect the number of such elements to be less than what is given by the results here.

The reason we call Theorem 1.8 *weak* is that the dependence between the parameters is not optimal. In particular, they are far from reflecting the relations in (2) and (8). In a later paper [23], Tao and the second author refined the approach to obtain the following stronger result.

Theorem 1.10 (Strong inverse theorem)**.** [23] *Let C and $1 > \varepsilon$ be positive constants. Assume that*

$$\rho(V) \geq n^{-C}.$$

Then there exists a proper symmetric GAP Q of rank $r = O_{C,\varepsilon}(1)$ which contains all but $O_r(n^{1-\varepsilon})$ elements of V (counting multiplicity), where

$$|Q| = O_{C,\varepsilon}(n^{C-\frac{r}{2}+\varepsilon}).$$

The bound on $|Q|$ matches Example 1.7, up to the n^ϵ term. However, this error term seems to be the limit of the approach. The proofs of Theorem 1.8 and 1.10 rely on a replacement argument and various lemmas about random walks and GAPs.

Let us now consider another application of Theorem 1.10. Notice that Theorem 1.10 enables us to make very precise counting arguments. Assume that we would like to count the number of (multi)-sets V of integers with $\max |v_i| \leq N = n^{O(1)}$ such that $\rho(V) \geq \rho := n^{-C}$.

Fix $d \geq 1$, fix [1] a GAP Q with rank r and volume $|Q| = n^{C-\frac{r}{2}}$. The dominating term in the calculation will be the number of multi-subsets of size n of Q, which is

$$|Q|^n = n^{(C-\frac{r}{2}+\epsilon)n} \leq n^{Cn} n^{-\frac{n}{2}+\epsilon n} = \rho^{-n} n^{-n(\frac{1}{2}-\epsilon)}. \quad (4)$$

Motivated by questions from random matrix theory, Tao and the second author obtained the following continuous analogue of this result.

Let n be a positive integer and β, p be positive numbers that may depend on n. Let $\mathcal{S}_{n,\beta,\rho}$ be the collection of all multiple sets $V = (v_1, \dots, v_n), v_i \in \mathbf{R}^2$ such that $\sum_{i=1}^n \|v_i\|^2 = 1$ and $\rho_{\beta,\eta}(V) \geq \rho$.

Theorem 1.11 (The β-net Theorem)**.** [28] *Let $0 < \epsilon \leq 1/3$ and $C > 0$ be constants. Then, for all sufficiently large n and $\beta \geq \exp(-n^\epsilon)$ and $\rho \geq n^{-C}$ there is a set $\mathcal{S} \subset (\mathbf{R}^2)^n$ of size at most*

$$\rho^{-n} n^{-n(\frac{1}{2}-\epsilon)} + \exp(o(n))$$

such that for any $V = \{v_1, \dots, v_n\} \in \mathcal{S}_{n,\beta,p}$ there is $V' = (v'_1, \dots, v'_n) \in \mathcal{S}$ such that $\|v_i - v'_i\|_2 \leq \beta$ for all i.

The theorem looks a bit cleaner if we use $\mathbf{C}$ instead of $\mathbf{R}^2$ (as in [28]). However, we prefer the current form as it is more suitable for generalization. The set $\mathcal{S}$ is usually referred to as a β-net of $\mathcal{S}_{n,\beta,p}$.

Theorem 1.11 is at the heart of the establishment of the Circular Law conjecture in random matrix theory (see [28, 24]). It also plays an important role in the study of condition number of randomly perturbed matrices (see [29]). Its proof in [28] is quite technical and occupies the bulk part of that paper.

On the other hand, given the above discussion, one would, obviously, expects to obtain Theorem 1.11 as a simple corollary of a continuous analogue of Theorem 1.10. However, the arguments in [28] have not yet provided such inverse theorem (although it did provide a sufficient amount of information about the set S that makes the estimate possible). The paper [16] of Rudelson and Vershynin also contained a characterization of the set S, but their characterization of somewhat different spirit than those discussed in this paper.

2. A new approach and new results

In this paper, we introduce a new approach to the inverse theorem. The core of this new approach is a (long range) variant of Freiman's famous inverse theorem.

[1] A more detailed version of Theorems 1.8 and 1.10 tells us that there are not too many ways to choose the generators of Q. In particular, if $N = n^{O(1)}$, the number of ways to fix these is negligible compared to the main term.

This new approach seems powerful. First of all, it enables us to remove the error term n^ϵ in Theorem 1.10, resulting in an optimal inverse theorem.

Theorem 2.1 (Optimal inverse Littlewood-Offord theorem, discrete case)**.** *Let C and $1 > \varepsilon$ be positive constants. There is a constant $c_1 = c_1(\varepsilon, C)$ such that the following holds. Assume that*

$$\rho(V) \geq n^{-C}.$$

Then there exists a proper symmetric GAP Q of rank $r = O_{C,\varepsilon}(1)$ which contains all but at most εn elements of V (counting multiplicity), where

$$|Q| = O_{C,\varepsilon}(\rho(V)^{-1} n^{-\frac{r}{2}}).$$

This immediately implies several forward theorems, such as Theorems 1.2 and 1.3. For example, we can prove Theorem 1.2 as follows.

Proof. (of Theorem 1.2) Assume, for contradiction, that there is a set V of n different number such that $\rho(V) \geq c_1 n^{-3/2}$ for some large constant c_1 to be chosen. Set $\varepsilon = .1, C = 3/2$. By Theorem 2.1, there is a GAP Q of rank r and size $O_{C,\epsilon}(\frac{1}{c_1} n^{C-\frac{r}{2}})$ that contains at least a $.9n$ elements from V. This implies $|Q| \geq .9n$. By setting c_1 sufficiently large and using the fact that $C = 3/2$ and $r \geq 1$, we can guarantee that $|Q| \leq .8n$, a contradiction. □

Theorem 1.3 can be proved similarly with the detail left as an exercise.

Similar to [25, 23], our method and results can be extended (rather automatically) to much more general settings.

General V. Instead of taking V to be a subset of $\mathbf{Z}$, we can take it to be a subset of any abelian torsion free group G (thanks to Freiman isomorphism). We can also replace $\mathbf{Z}$ by the finite field $\mathbf{F}_p$, where p is any sufficiently large prime. (In fact, the first step in our proof is to embed V into $\mathbf{F}_p$.)

General η. We can replace the Bernoulli random variables by independent random variables η_i satisfying the following condition. There is a constant $c > 0$ and an infinite sequence of primes p such that for any p in the sequence, any (multi)-subset V of size n of $\mathbf{F}_p$ and any $t \in \mathbf{F}_p$

$$\prod_{i=1}^n |\mathbf{E} e_p(\eta_i v_i t)| \leq \exp(-c \sum_{i=1}^n \|\frac{v_i t}{p}\|^2) \tag{5}$$

where $\|x\|$ denote the distance from x to the closest integer (we view the elements of $\mathbf{F}_p$ as integers between 0 and $p-1$) and $e_p(x) := \exp(2\pi\sqrt{-1}x/p)$.

Example 2.2. (Lazy random walks) Given a parameter $0 < \mu \leq 1$, let η_i^μ be iid copies of a random variable η^μ: $\eta^\mu = 1$ or -1 with probability $\mu/2$, and $\eta^\mu = 0$ with probability $1 - \mu$. The sum

$$S^\mu(V) := \sum_{i=1}^n \eta_i{}^\mu v_i,$$

can be viewed as a *lazy random walk* with steps in V. A simple calculation shows

$$\mathbf{E}e_p(\eta x) = (1-\mu) + \mu \cos \frac{2\pi x}{p}.$$

It is easy to show that there is a constant $c > 0$ depending on μ such that

$$|(1-\mu) + \mu \cos \frac{2\pi x}{p}| \leq \exp(-c\|\frac{x}{p}\|^2).$$

Example 2.3. (μ-bounded variables) It suffices to assume that there is some constant $0 < \mu \leq 1$ such that for all i

$$|\mathbf{E}e_p(\eta_i x)| \leq (1-\mu) + \mu \cos \frac{2\pi x}{p}. \tag{6}$$

Theorem 2.4. *The conclusion of Theorem 2.1 holds for the case when V is a multi-subset of an arbitrary torsion free abelian group G and $\eta_i, 1 \leq i \leq n$ are independent random variables satisfying* (5).

Remark 2.5. In the upcoming paper [14], by following the method used to prove Theorem 2.1, we are able to address the issues concerning Theorem 1.4. We prove that $\rho(V_0) = (\sqrt{\frac{24}{\pi}} + o(1))n^{-3/2}$. More importantly, we obtain a stable version of Theorem 1.4, which shows that if $\rho(V)$ is close to $(\sqrt{24/\pi}+o(1))n^{-3/2}$, then V is "close" to V_0 . As a byproduct, we obtain the first *non-algebraic* proof for the asymptotic version of Stanley theorem.

We now turn to the continuous setting. In this part, we consider real random variable z such that there exists a constant C_z so

$$\mathbf{P}(1 \leq z_1 - z_2 \leq C_z) \geq 1/2, \tag{7}$$

where z_1, z_2 are idd copies of z. We notice that Bernoulli random variables are clearly of this type. (Also, the interested reader may find (7) more general than the condition of κ-controlled second moment defined in [28] and the condition of bounded third moment in [16].) In the above C_z is not uniquely defined. In what follows, we will take the smallest value of C_z.

We say that a vector $v \in \mathbf{R}^d$ is δ-*close* to a set $Q \subset \mathbf{R}^d$ if there exists a vector $q \in Q$ such that $\|v - q\|_2 \leq \delta$. A set X is δ-close to a set Q if every element of X is δ-close to Q. The analogue of Example (1.7) is the following.

Example 2.6. Let Q be a proper symmetric GAP of rank r and volume N in $\mathbf{R}^d$. Let $v_1, \dots, v_n$ be (not necessarily distinct) vectors which are $O(\beta n^{-1/2})$-close to a GAP of rank r Q. If we set $|Q| = n^{C-\frac{r}{2}}$ for some constant $C \geq r/2$, then

$$\rho_{\beta,\eta}(V) = \Omega(\frac{1}{n^C}). \tag{8}$$

Thus, one would expect that if $\rho_{\beta,z}(V)$ is large, then (most of) V is $O(\beta n^{-1/2})$-close to a GAP with small volume. Confirming this intuition, we obtain the following continuous analogue of Theorem 2.1.

Theorem 2.7 (Optimal inverse Littlewood-Offord theorem, continuous case)**.** *Let $\delta, C > 0$ be arbitrary constants and $\beta > 0$ be a parameter that may depend on n. Suppose that $V = \{v_1, \dots, v_n\}$ is a (multi-) subset of $\mathbf{R}^d$ such that $\sum_{i=1}^n \|v_i\|_2^2 = 1$ and that V has large small ball probability*

$$\rho := \rho_{\beta,z}(V) \geq n^{-C},$$

where z is a real random variable satisfying (7)*. Then there exists a proper symmetric GAP Q of rank $r = O(1)$ so that all but at most δn elements of V (counting multiplicity) are $O(\beta \frac{\log n}{n^{1/2}})$-close to Q, where*

$$|Q| = O(\rho^{-1} n^{(-r+d)/2}).$$

The theorem is optimal in the sense that the exponent $(-r+d)/2$ in the bound on $|Q|$ cannot be improved in general (see Appendix B for more details). On the other hand, it is different from what one may predict based on Example 2.6, motivated by the discrete case.

Theorem 2.7 is a special case of the following more general theorem.

Theorem 2.8 (Continuous Inverse Littlewood-Offord theorem, general setting)**.** *Let $0 < \epsilon < 1; 0 < C$ be constants. Let $\beta > 0$ be a parameter that may depend on n. Suppose that $V = \{v_1, \dots, v_n\}$ is a (multi-) subset of $\mathbf{R}^d$ such that $\sum_{i=1}^n \|v_i\|_2^2 = 1$ and that V has large small ball probability*

$$\rho := \rho_{\beta,z}(V) \geq n^{-C},$$

where z is a real random variable satisfying (7)*. Then the following holds. For any number $n^\epsilon \leq n' \leq n$, there exists a proper symmetric GAP $Q = \{\sum_{i=1}^r x_i g_i : |x_i| \leq L_i\}$ such that*

- *(Full dimension) There exists $\sqrt{\frac{n'}{\log n}} \ll k \ll \sqrt{n'}$ such that the dilate $P := \beta^{-1} k \cdot Q$ contains the discrete hypercube $\{0,1\}^d$.*
- *(Approximation) At least $n - n'$ elements of V are $O(\frac{\beta}{k})$-close to Q.*
- *(Small rank and cardinality) Q has constant rank $d \leq r = O(1)$, and cardinality*

$$|Q| = O(\rho^{-1} n'^{(-r+d)/2}).$$

- *(Small generators) There is a non-zero integer $p = O(\sqrt{n'})$ such that all steps g_i of Q have the form $g_i = (g_{i1}, \dots, g_{id})$, where $g_{ij} = \beta \frac{p_{ij}}{p}$ with $p_{ij} \in \mathbf{Z}$ and $p_{ij} = O(\beta^{-1}\sqrt{n'})$.*

Theorem 2.8 implies the following corollary (see also Appendix B for a simple proof), from which one can derive Theorem 1.11 in a straightforward manner (similar to the discrete case discussed earlier).

Corollary 2.9. *Let $0 < \epsilon < 1; 0 < C$ be constants. Let $\beta > 0$ be a parameter that may depend on n. Suppose that $V = \{v_1, \dots, v_n\}$ is a (multi-) subset of $\mathbf{R}^d$ such that $\sum_{i=1}^n \|v_i\|_2^2 = 1$ and that V has large small ball probability*

$$\rho := \rho_{\beta,z}(V) \geq n^{-C},$$

where z is a real random variable satisfying (7). Then the following holds. For any number n' between n^ϵ and n, there exists a proper symmetric GAP $Q = \{\sum_{i=1}^r x_i g_i : |x_i| \leq L_i\}$ such that

- *At least $n - n'$ elements of V are β-close to Q.*
- *Q has small rank, $r = O(1)$, and small cardinality*

$$|Q| \leq \max\left(O(\frac{\rho^{-1}}{\sqrt{n'}}), 1\right).$$

- *There is a non-zero integer $p = O(\sqrt{n'})$ such that all steps g_i of Q have the form $g_i = (g_{i1}, \dots, g_{id})$, where $g_{ij} = \beta\frac{p_{ij}}{p}$ with $p_{ij} \in \mathbf{Z}$ and $p_{ij} = O(\beta^{-1}\sqrt{n'})$.*

In the above theorems, the hidden constants could depend on previously set constants ϵ, C, C_z, d. We could have written $O_{\epsilon,C,C_z,d}$ and $\ll_{\epsilon,C,C_z,d}$ everywhere, but these notations are somewhat cumbersome and this dependence is not our focus.

Proof. (of Theorem 1.11) Set $n' := n^{1-\frac{3\epsilon}{2}}$ (which is $\gg n^\epsilon$ as $\epsilon \leq 1/3$). Let $\mathcal{S}'$ be the collection of all subsets of size at least $n - n'$ of GAPs whose parameters satisfy the conclusion of Theorem 2.8.

Since each GAP is determined by its generators and dimensions, the number of such GAPs is bounded by $((\beta^{-1}\sqrt{n'})\sqrt{n'})^{O(1)}(\frac{\rho^{-1}}{\sqrt{n'}})^{O(1)} = \exp(o(n))$. (The term $(\frac{\rho^{-1}}{\sqrt{n'}})^{O(1)}$ bounds the number of choices of the dimensions M_i.) Thus $|\mathcal{S}'| = \left(O((\frac{\rho^{-1}}{\sqrt{n'}})^n) + 1\right)\exp(o(n))$.

We approximate each of the exceptional elements by a lattice point in $\beta \cdot (\mathbf{Z}/d)^d$. Thus if we let $\mathcal{S}''$ to be the set of these approximated tuples then $|\mathcal{S}''| \leq \sum_{i \leq n'}(O(\beta^{-1}))^i = \exp(o(n))$ (here we used the assumption $\beta \geq \exp(-n^\epsilon)$).

Set $\mathcal{S} := \mathcal{S}' \times \mathcal{S}''$. It is easy to see that $|\mathcal{S}| \leq O(n^{-1/2+\epsilon}\rho^{-1})^n + \exp(o(n))$. Furthermore, if $\rho(V) \geq n^{-O(1)}$ then V is β-close to an element of $\mathcal{S}$, concluding the proof. □

3. The long range inverse theorem

Let us first recall a famous theorem of Freiman [30, Chapter 5].

Theorem 3.1 (Freiman's inverse theorem)**.** *Let γ be a positive constant and X a subset of a torsion-free group such that $|2X| \leq \gamma|X|$. Then there is a proper symmetric GAP Q of rank at most $r = O_\gamma(1)$ and cardinality $O_\gamma(|X|)$ such that $X \subset Q$.*

In our analysis, we will need to deal with an assumption of the form $|kX| \leq k^\gamma |X|$, where γ is a constant, but k is not. (Typically, k will be a positive power of $|X|$.) We succeed to give a structure for X under this condition in the following theorem, which we will refer to as the long range inverse theorem.

Theorem 3.2 (long range inverse theorem)**.** *Let $\gamma > 0$ be constant. Assume that X is a subset of a torsion-free group such that $0 \in X$ and $|kX| \leq k^\gamma |X|$ for some positive integer $k \geq 2$. Then there is proper symmetric GAP Q of rank $r = O(\gamma)$ and cardinality $O_\gamma(k^{-r}|kX|)$ such that $X \subset Q$.*

Notice that for any given $\epsilon > 0$ and if k is large enough, it is implied from Theorem 3.2 that the rank of Q is at most $\gamma + \epsilon$. The implicit constant involved in the size of Q can be taken to be $2^{2^{2^{O(\gamma)}}}$, which is quite poor. Although we have not elaborated on this bound much, our method does not seem to say anything when the polynomial growth in size of kX is replaced by something faster.

Theorem 3.2 will serve as our main technical tool. This theorem can be proved by applying an earlier result of [26]. We give a short deduction in Appendix A.

4. Freiman isomorphism

We now introduce the concept of Freiman isomorphism, that allows us to transfer an additive problem to another group in a way which is more flexible than the usual notion of group isomorphism.

Definition 4.1 (Freiman isomorphism of order k)**.** Two sets V, V' of additive groups G, G' (not necessarily torsion-free) are Fremain-isomorphism of order k (in generalized form) if there is an injective map f from V to V' such that $f(v_1) + \cdots + f(v_k) = f(v'_1) + \cdots + f(v'_k)$ in G', if and only if $v_1 + \cdots + v_k = v'_1 + \cdots + v'_k$ in G.

The following theorem allows us to pass from an arbitrary torsion-free group to $\mathbf{Z}$ or cyclic groups of prime order (see [30, Chapter 5]).

Theorem 4.2. *Let V be a finite subset of a torsion-free additive group G. Then for any integer k, there is a Freiman isomorphism $\phi : V \to \phi(V)$ of order k to some finite subset $\phi(V)$ of the integers $\mathbf{Z}$. The same is true if we replace $\mathbf{Z}$ by $\mathbf{F}_p$, if p is sufficiently large depending on V.*

By following the same proof, we can show a stronger result below.

Theorem 4.3. *Let V be a finite subset of a torsion-free additive group G. Then for any integer k, there is a map $\phi : V \to \phi(V)$ to some finite subset $\phi(v)$ of the integers $\mathbf{Z}$ such that*

$$v_1 + \cdots + v_i = v'_1 + \cdots + v'_j \Leftrightarrow \phi(v_1) + \cdots + \phi(v_i) = \phi(v'_1) + \ldots \phi(v'_j)$$

For all $i, j \leq k$. The same is true if we replace $\mathbf{Z}$ by $\mathbf{F}_p$, if p is sufficiently large depending on V.

By Theorem 4.3, there exist a large prime p and a set $V_p \subset \mathbf{F}_p$ which is Freimain isomorphism of order n with V. We hence infer that

$$\rho(V) = \rho(V_p).$$

Thus instead of working with a subset V of a torsion-free group, it is sufficient to work with subset of $\mathbf{F}_p$, where p is large enough.

5. Proof of Theorem 2.1

Embedding. The first step is to embed the problem into the finite field $\mathbf{F}_p$ for some prime p. In the case when v_i are integers, we simply take p to be a large prime (for instance $p \geq 2^n(\sum_{i=1}^n |v_i| + 1)$ suffices). If V is a subset of a general torsion-free group G, one can use Theorem 4.3.

From now on, we can assume that v_i are elements of $\mathbf{F}_p$ for some large prime p. We view elements of $\mathbf{F}_p$ as integers between 0 and $p-1$. We use short hand ρ to denote $\rho(V)$.

Fourier Analysis. The main advantage of working in $\mathbf{F}_p$ is that one can make use of discrete Fourier analysis. Assume that

$$\rho = \rho(V) = \mathbf{P}(S = a),$$

for some $a \in \mathbf{F}_p$. Using the standard notation $e_p(x)$ for $\exp(2\pi\sqrt{-1}x/p)$, we have

$$\rho = \mathbf{P}(S = a) = \mathbf{E}\frac{1}{p}\sum_{\xi \in \mathbf{F}_p} e_p(\xi(S-a)) = \mathbf{E}\frac{1}{p}\sum_{\xi \in \mathbf{F}_p} e_p(\xi S)e_p(-\xi a). \tag{9}$$

By independence

$$\mathbf{E}e_p(\xi S) = \prod_{i=1}^n e_p(\xi \eta_i v_i) = \prod_{i=1}^n \cos\frac{2\pi\xi v_i}{p} \tag{10}$$

Since $|e_p(-\xi a)| = 1$, it follows that

$$\rho \leq \frac{1}{p}\sum_{\xi \in \mathbf{F}_p} |\cos\frac{2\pi v_i \xi}{p}| = \frac{1}{p}\sum_{\xi \in \mathbf{F}_p} |\frac{\cos \pi v_i \xi}{p}|. \tag{11}$$

By convexity, we have that $|\sin \pi z| \geq 2\|z\|$ for any $z \in \mathbf{R}$, where $\|z\| := \|z\|_{\mathbf{R}/\mathbf{Z}}$ is the distance of z to the nearest integer. Thus,

$$|\cos\frac{\pi x}{p}| \leq 1 - \frac{1}{2}\sin^2\frac{\pi x}{p} \leq 1 - 2\|\frac{x}{p}\|^2 \leq \exp(-2\|\frac{x}{p}\|^2), \tag{12}$$

where in the last inequality we used that fact that $1 - y \leq \exp(-y)$ for any $0 \leq y \leq 1$.

Consequently, we obtain a key inequality

$$\rho \leq \frac{1}{p} \sum_{\xi \in \mathbf{F}_p} \prod_i |\cos \frac{\pi v_i \xi}{p}| \leq \frac{1}{p} \sum_{\xi \in F_p} \exp(-2 \sum_{i=1}^n \|\frac{v_i \xi}{p}\|^2). \tag{13}$$

Large level sets. Now we consider the level sets $S_m := \{\xi | \sum_{i=1}^n \|v_i \xi/p\|^2 \leq m\}$. We have

$$n^{-C} \leq \rho \leq \frac{1}{p} \sum_{\xi \in \mathbf{F}_p} \exp(-2 \sum_{i=1}^n \|\frac{v_i \xi}{p}\|^2) \leq \frac{1}{p} + \frac{1}{p} \sum_{m \geq 1} \exp(-2(m-1))|S_m|.$$

Since $\sum_{m \geq 1} \exp(-m) < 1$, there must be is a large level set S_m such that

$$|S_m| \exp(-m+2) \geq \rho p. \tag{14}$$

In fact, since $\rho \geq n^{-C}$, we can assume that $m = O(\log n)$.

Double counting and the triangle inequality. By double counting we have

$$\sum_{i=1}^n \sum_{\xi \in S_m} \|\frac{v_i \xi}{p}\|^2 = \sum_{\xi \in S_m} \sum_{i=1}^n \|\frac{v_i \xi}{p}\|^2 \leq m|S_m|.$$

So, for most v_i

$$\sum_{\xi \in S_m} \|\frac{v_i \xi}{p}\|^2 \leq \frac{C_0 m}{n} |S_m| \tag{15}$$

for some large constant C_0.

Set $C_0 = \varepsilon^{-1}$. By averaging, the set of v_i satisfying (15) has size at least $(1-\varepsilon)n$. We call this set V'. The set $V \backslash V'$ has size at most εn and this is the exceptional set that appears in Theorem 2.1. In the rest of the proof, we are going to show that V' is a dense subset of a proper GAP.

Since $\|\cdot\|$ is a norm, by the triangle inequality, we have for any $a \in kV'$

$$\sum_{\xi \in S_m} \|\frac{a\xi}{p}\|^2 \leq k^2 \frac{C_0 m}{n} |S_m|. \tag{16}$$

More generally, for any $l \leq k$ and $a \in lV'$

$$\sum_{\xi \in S_m} \|\frac{a\xi}{p}\|^2 \leq k^2 \frac{C_0 m}{n} |S_m|. \tag{17}$$

Dual sets. Define $S_m^* := \{a | \sum_{\xi \in S_m} \|\frac{a\xi}{p}\|^2 \leq \frac{1}{200}|S_m|\}$ (the constant 200 is adhoc and any sufficiently large constant would do). S_m^* can be viewed as some sort of a *dual* set of S_m. In fact, one can show as far as cardinality is concerned, it does behave like a dual

$$|S_m^*| \leq \frac{8p}{|S_m|}. \tag{18}$$

To see this, define $T_a := \sum_{\xi \in S_m} \cos \frac{2\pi a \xi}{p}$. Using the fact that $\cos 2\pi z \geq 1 - 100\|z\|^2$ for any $z \in \mathbf{R}$, we have, for any $a \in S_m^*$

$$T_a \geq \sum_{\xi \in S_m} (1 - 100\|\frac{a\xi}{p}\|^2) \geq \frac{1}{2}|S_m|.$$

One the other hand, using the basic identity $\sum_{a \in \mathbf{F}_p} \cos \frac{2\pi a x}{p} = p\mathbf{I}_{x=0}$, we have

$$\sum_{a \in \mathbf{F}_p} T_a^2 \leq 2p|S_m|.$$

(18) follows from the last two estimates and averaging.

Set $k := c_1 \sqrt{\frac{n}{m}}$, for a properly chosen constant $c_1 = c_1(C_0)$. By (17) we have $\cup_{l=1}^k lV' \subset S_m^*$. Set $V'' = V' \cup \{0\}$; we have $kV'' \subset S_m^* \cup \{0\}$. This results in the critical bound

$$|kV''| = O(\frac{p}{|S_m|}) = O(\rho^{-1} \exp(-m + 2)). \tag{19}$$

The long range inverse theorem. The role of $\mathbf{F}_p$ is now no longer important, so we can view v_i as integers. The inequality (19) is exactly the assumption of our long range Inverse Theorem.

With this theorem in hand, we are ready to conclude the proof. A slight technical problem is that V'' is not a set but a multi-set, so we are going to apply Theorem 3.2 to X being the set of different elements of V''. Notice that $k = \Omega(\sqrt{\frac{n}{m}}) = \Omega(\sqrt{\frac{n}{\log n}})$, so $\rho^{-1} \leq n^C$ is bounded from above by k^{2C+1}.

It follows from Theorem 3.2 that X is a subset of a proper symmetric GAP Q of rank $r = O_{C,\epsilon}(1)$ and cardinality

$$O_{C,\epsilon}(k^{-r}|kX|) = O_{C,\epsilon}(k^{-r}|kV''|) = O_{C,\epsilon}\left(\rho^{-1}\exp(-m)(\sqrt{\frac{n}{m}})^{-r}\right)$$
$$= O_{C,\epsilon}(\rho^{-1}n^{-r}),$$

concluding the proof.

6. Proof of Theorem 2.8

We denote the z-norm of a real number to be

$$\|w\|_z := (\mathbf{E}\|w(z_1 - z_2)\|^2)^{1/2},$$

where z_1, z_2 are two iid copies of z. This proof follows pretty much the same steps as in the discrete case, with some additional simple arguments.

Fourier analysis. Our first step is to obtain the following analogue of (13), using the Fourier transform.

Lemma 6.1 (bounds for small ball probability)**.**

$$\rho_{r,z}(V) \leq \exp(\pi r^2)\int_{\mathbf{R}^d}\exp(-\sum_{i=1}^n \|\langle v_i, \xi\rangle\|_z^2/2 - \pi\|\xi\|_2^2)d\xi.$$

This lemma is basically from [28]; the proof is presented in Appendix C, for the reader's convenience.

Next, consider $V_\beta := \beta^{-1}\cdot V = \{\beta^{-1}v_1, \dots, \beta^{-1}v_n\}$. It is clear that

$$\rho_{\beta,z}(V) = \rho_{1,z}(V_\beta).$$

We now work with V_β. Thus $\rho_{1,z}(V_\beta) \geq n^{-O(1)}$ and $\sum_{v\in V_\beta}\|v\|^2 = \beta^{-2}$.

For short, we write ρ for $\rho_{1,z}(V_\beta)$. Set $M := 2A\log n$ where A is large enough. From Lemma 6.1 and that $\rho \geq n^{-O(1)}$ we easily obtain

$$\int_{\|\xi\|_2\leq M}\exp(-\frac{1}{2}\sum_{v\in V_\beta}\|\langle v,\xi\rangle\|_z^2 - \pi\|\xi\|_2^2)d\xi \geq \frac{\rho}{2}.$$

Large level sets. For each integer $0 \leq m \leq M$ we define the level set

$$S_m := \left\{\xi \in \mathbf{R}^d : \sum_{v\in V_\beta}\|\langle v,\xi\rangle\|_z^2 + \|\xi\|_2^2 \leq m\right\}.$$

Then it follows that $\sum_{m\leq M}\mu(S_m)\exp(-\frac{m}{2}+1)\geq\rho$, where $\mu(.)$ denotes the Lebesgue measure of a measurable set. Hence there exists $m\leq M$ such that $\mu(S_m)\geq\rho\exp(\frac{m}{4}-2)$.

Next, since $S_m\subset B(0,\sqrt{m})$, by pigeon-hole principle there exists a ball $B(x,\frac{1}{2})\subset B(0,\sqrt{m})$ such that

$$\mu(B(x,\frac{1}{2})\cap S_m)\geq c_d\mu(S_m)m^{-d/2}\geq c_d\rho\exp(\frac{m}{4}-2)m^{-d/2}.$$

Consider $\xi_1,\xi_2\in B(x,1/2)\cap S_m$. By Cauchy-Schwarz inequality, and notice that $\|.\|_z$ is a norm, we have

$$\sum_{v\in V_\beta}\|\langle v,(\xi_1-\xi_2)\rangle\|_z^2\leq 4m.$$

Since $\xi_1-\xi_2\in B(0,1)$ and $\mu(B(x,\frac{1}{2})\cap S_m-B(x,\frac{1}{2})\cap S_m)\geq\mu(B(x,\frac{1}{2})\cap S_m)$, if we put

$$T:=\{\xi\in B(0,1),\sum_{i=1}^n\|\langle\xi,v_i\rangle\|_z^2\leq 4m\},$$

then

$$\mu(T)\geq c_d\rho\exp(\frac{m}{4}-2)m^{-d/2}.$$

Discretization. Choose N to be a sufficiently large prime (depending on the set T). Define the discrete box

$$B_0:=\{(k_1/N,\dots,k_d/N):k_i\in\mathbf{Z},-N\leq k_i\leq N\}.$$

We consider all the shifted boxes $x+B_0$, where $x\in[0,1/N]^d$. By pigeon-hole principle, there exists x_0 such that the size of the discrete set $(x_0+B_0)\cap T$ is at least the expectation, $|(x_0+B_0)\cap T|\geq N^d\mu(T)$ (to see this, we first consider the case when T is a box itself).

Let us fix a $\xi_0\in(x_0+B_0)\cap T$. Then for any $\xi\in(x_0+B_0)\cap T$ we have

$$\sum_{v\in V_\beta}\|\langle v,\xi_0-\xi\rangle\|_z^2\leq 2\left(\sum_{v\in V_\beta}\|\langle v,\xi\rangle\|_z^2+\sum_{v\in V_\beta}\|\langle v,\xi_0\rangle\|_z^2\right)\leq 16m.$$

Notice that $\xi_0-\xi\in B_1:=B_0-B_0=\{(k_1/N,\dots,k_d/N):k_i\in\mathbf{Z},-2N\leq k_i\leq 2N\}$. Thus there exists a subset S of size at least $c_dN^d\rho\exp(\frac{m}{4}-2)m^{-d/2}$ of B_1 such that the following holds for any $s\in S$

$$\sum_{v\in V_\beta} \|\langle v,s\rangle\|_z^2 \le 16m.$$

Double counting. We let $y = z_1 - z_2$, where z_1, z_2 are iid copies of z. By definition of S, we have

$$\sum_{s\in S}\sum_{v\in V_\beta} \|\langle v,s\rangle\|_z^2 \le 16m|S|$$
$$\mathbf{E}_y \sum_{s\in S}\sum_{v\in V_\beta} \|y\langle v,s\rangle\|_{\mathbf{R}/\mathbf{Z}}^2 \le 16m|S|.$$

It is then implied that there exists $1 \le |y_0| \le C_z$ such that

$$\sum_{s\in S}\sum_{v\in V_\beta} \|y_0\langle v,s\rangle\|_{\mathbf{R}/\mathbf{Z}}^2 \le 16m|S|\mathbf{P}(1 \le |y| \le C_z)^{-1}.$$

On the other hand, by property (7) we have $\mathbf{P}(1 \le |y| \le C_z) \ge 1/2$. So

$$\sum_{s\in S}\sum_{v\in V_\beta} \|y_0\langle v,s\rangle\|_{\mathbf{R}/\mathbf{Z}}^2 \le 32m|S|.$$

Let n' be any number between n^ϵ and n. We say that $v \in V_\beta$ is *bad* if

$$\sum_{s\in S} \|y_0\langle v,s\rangle\|_{\mathbf{R}/\mathbf{Z}}^2 \ge \frac{32m|S|}{n'}.$$

Then the number of bad vectors is at most n'. Let V'_β be the set of remaining vectors. Thus V'_β contains at least $n - n'$ elements. In the rest of the proof, we are going to show that V'_β is close to a GAP, as claimed in the theorem.

Dual sets. Consider $v \in V'_\beta$, we have $\sum_{s\in S} \|y_0\langle s,v\rangle\|_{\mathbf{R}/\mathbf{Z}}^2 \le 32|K|/n'$.

Set $k := \sqrt{\frac{n'}{64\pi^2 m}}$ and let $V''_\beta := k(V'_\beta \cup \{0\})$. By Cauchy-Schwarz inequality (see (17)), for any $a \in V''_\beta$ we have

$$\sum_{s\in S} 2\pi^2 \|\langle s, y_0 a\rangle\|_{\mathbf{R}/\mathbf{Z}}^2 \le \frac{|S|}{2},$$

which implies

$$\sum_{s\in S}\cos(2\pi\langle s, y_0 a\rangle) \geq \frac{|S|}{2}.$$

Observe that for any $x \in B(0, \frac{1}{256d})$ and any $s \in S \subset B(0,2)$ we always have $\cos(2\pi\langle s,x\rangle) \geq 1/2$ and $\sin(2\pi\langle s,x\rangle) \leq 1/12$. Thus for any $x \in B(0, \frac{1}{256d})$,

$$\sum_{s\in S}\cos\left(2\pi\langle s, (y_0 a + x)\rangle\right) \geq \frac{|S|}{4} - \frac{|S|}{12} = \frac{|S|}{6}.$$

On the other hand,

$$\begin{aligned}\int_{x\in[0,N]^d}\left(\sum_{s\in S}\cos(2\pi\langle s,x\rangle)\right)^2 dx &\leq \sum_{s_1,s_2\in S}\int_{x\in[0,N]^d}\exp\left(2\pi\sqrt{-1}\langle s_1-s_2,x\rangle\right)dx\\ &\ll_d |S|N^d.\end{aligned}$$

Hence we deduce the following

$$\mu_{x\in[0,N]^d}\left((\sum_{s\in S}\cos(2\pi\langle s,x\rangle))^2 \geq (\frac{|S|}{6})^2\right) \ll_d \frac{|S|N^d}{(|S|/6)^2} \ll_d \frac{N^d}{|S|}.$$

Now we use the fact that S has large size, $|S| \gg_d N^d\rho\exp(\frac{m}{4}-2)m^{-d/2}$, and $y_0V''_\beta + B(0,\frac{1}{256d}) \subset [0,N]^d$,

$$\mu(y_0V''_\beta + B(0,\frac{1}{256d})) \ll_d \rho^{-1}\exp(-\frac{m}{4}+2)m^{d/2}.$$

Thus, we obtain the following analogue of (19)

$$\mu\left(k(V'_\beta\cup\{0\}) + B(0,\frac{1}{256dy_0})\right) \ll_d \rho^{-1}y_0^{-d}\exp(-\frac{m}{4}+2)m^{d/2}. \tag{20}$$

The long range inverse theorem. Our analysis again relies on the Long Range Inverse theorem. Let $D := 1024dy_0$. We approximate each vector v' of V'_β by a closest vector in $(\frac{\mathbf{Z}}{Dk})^d$,

$$\|v' - \frac{a}{Dk}\|_2 \leq \frac{\sqrt{d}}{Dk}, \text{ with } a\in\mathbf{Z}^d.$$

Let A_β be the collection of all such a. Since $\sum_{v' \in V'_\beta} \|v'\|_2^2 = O(\beta^{-2})$, we have

$$\sum_{a \in A_\beta} \|a\|_2^2 = O_{d,C_z}(k^2\beta^{-2}). \tag{21}$$

It follows from (20) that

$$\begin{aligned} |k(A_\beta + C(0,1))| &= O_{d,C_z}\left(\rho^{-1}(Dk)^d y_0^{-d} \exp(-\frac{m}{4}+2)m^{d/2}\right) \\ &= O_{d,C_z}\left(\rho^{-1}k^d \exp(-\frac{m}{4}+2)m^{d/2}\right), \end{aligned}$$

where $C(0,r)$ is the discrete cube $\{(z_1,\dots,z_d) \in \mathbf{Z}^d : |z_i| \le r\}$.

Now we apply Lemma 3.2 to the set $A_\beta + C(0,1)$ (notice that $0 \in A_\beta$). There exists a proper GAP $P = \{\sum_{i=1}^r x_i g_i : |x_i| \le N_i\} \subset \mathbf{Z}^d$ containing $A_\beta + C(0,1)$ which has small rank $r = O(1)$, and small size

$$\begin{aligned} |P| &= O_{d,C_z}\left((\rho^{-1}k^d \exp(-\frac{m}{4}+2)m^{d/2}k^{-r}\right) \\ &= O_{d,C_z}(\rho^{-1}n'^{(-r+d)/2}). \end{aligned}$$

Moreover, we have learned from the proof of Theorem 3.2 and Lemma **??** that kQ can be contained in a set $ck(A_\beta + C(0,1))$ for some $c = O(1)$. Using (21), we conclude that all the generators g_i of Q are bounded,

$$\|g_i\|_2 = O_{d,C_z}(k\beta^{-1}).$$

Next, since $C(0,1) \subset Q$, the rank r of P is at least d. It is a routine calculation to see that $Q := \frac{\beta}{Dk} \cdot P$ satisfies all required properties in the theorem.

Appendix A. Proof of the long range inverse theorem

A.1. **The main lemmas.** We first mention an earlier result of Tao and the second author ([26, Theorem 1.21])

Lemma A.2. *Let $\epsilon > 0, \gamma > 0$ be constants. Assume that X is a subset of integers such that $|kX| \le k^\gamma |X|$ for some number $k \ge 2$. Then kX is contained in a symmetric 2-proper GAP Q of rank $r = O_{\gamma,\epsilon}(1)$, and of cardinality $O_{\gamma,\epsilon}(|kX|)$.*

Next, if $kX \subset kQ$, where Q is a GAP, then it is natural to suspect that $X \subset Q$. This is, however, not always true. On the other hand, the conclusion holds if kQ is 2-proper and $0 \in X$.

Lemma A.3. *(Dividing sumsets relations) Assume that $0 \in X$ and that $P = \{\sum_{i=1}^r x_i a_i : |x_i| \leq N_i\}$ is a symmetric 2-proper GAP that contains kX. Then $X \subset \{\sum_{i=1}^r x_i a_i : |x_i| \leq 2N_i/k\}$.*

A good way to keep this lemma in mind is the following. Consider the relation $X \subset P$. It is trivial that this relation can always be *multiplied*, namely, for all integer $k \geq 1$, $kX \subset kP$. The above lemma asserts that under certain assumptions, the relation $kX \subset kP$ can be *divided*, giving $X \in P$.

Proof. (of Lemma A.3) Without loss of generality, we can assume that $k = 2^l$. It is enough to show that $2^{l-1}X \subset \{\sum_{i=1}^r x_i a_i : |x_i| \leq N_i/2\}$. Since $0 \in X$, $2^{l-1}X \subset 2^l X \subset P$, any element x of $2^{l-1}X$ can be written as $x = \sum_{i=1}^r x_i a_i$, with $|x_i| \leq N_i$. Now, since $2x \in P \subset 2P$ and $2P$ is proper (as P is 2-proper), we must have $0 \leq |2x_i| \leq N_i$. □

It is clear that Theorem 3.2 follows from Lemma A.2 and Lemma A.3.

Appendix B. Remarks on Theorem 2.8

Consider the set $U := [-2n, -n] \cup [n, 2n]$. Sample n points $v_1, \dots, v_n$, from U, independently with respect to the (continuous) uniform distribution and let A be the set of sampled points. Let ξ be the gaussian random variable $N(0, 1)$ and consider the sum

$$S := v_1\xi_1 + \cdots + v_n\xi_n,$$

where ξ_i are iid copies of ξ.

S has gaussian distribution with mean 0 and variance $\Theta(n^3)$, with probability one. Thus, for any interval I of length 1, $\mathbf{P}(S \in I) \leq Cn^{-3/2}$, for some constant C.

Set $n' = \delta n$, for some small positive constant δ. Theorem 2.8 states that (most of) A is $O(\frac{\log n}{\sqrt{n}})$-close to a GAP of rank r and volume $O(n^{2-\frac{r}{2}})$. We show that one cannot replace this bound by $O(n^{2-\frac{r}{2}-\epsilon})$. There are only three possible values for r: $r = 1, 2, 3$ and our claim follows from the following simple lemma, whose proof is left as an exercise.

Lemma B.1. *Let C, δ, ϵ be positive constants and $n \to \infty$. The followings hold with probability $1 - o(1)$ (with respect to the random choice of A).*

- *There is no subset A' of A of cardinality at least $(1-\delta)n$ and an AP Q of length at most $Cn^{3/2-\epsilon}$ such that A' is $\frac{C\log n}{\sqrt{n}}$-close to Q.*
- *There is no subset A' of A of cardinality at least $(1-\delta)n$ and a GAP Q of rank 2 and volume at most $Cn^{1-\epsilon}$ such that A' is $\frac{C\log n}{\sqrt{n}}$-close to Q.*
- *There is no subset A' of A of cardinality at least $(1-\delta)n$ and a GAP Q of rank 3 and volume at most $Cn^{1/2-\epsilon}$ such that A' is $\frac{C\log n}{\sqrt{n}}$-close to Q.*

The above construction can be generalized to higher dimensions as well, but we do not attempt to do so here. In the rest of this section, we prove Corollary 2.9.

Proof. We consider the following two cases.

Case 1: $r \geq d+1$. Consider the GAP P at the end of the proof of Theorem 2.8. Recall that $|P| = O_{d,C_z}(\rho^{-1} n'^{(d-r)/2}) = O_{d,C_z}(\rho^{-1}/\sqrt{n'})$. Let

$$Q := \frac{\beta}{Dk} \cdot P.$$

It is clear that Q satisfies all the conditions of Corollary 2.9. (Notice that in this case we obtain a stronger approximation: almost all elements of V are $O(\frac{\beta \log n'}{\sqrt{n'}})$-close to Q.)

Case 2: $r = d$. Because the unit vectors $e_j = (0, \dots, 1, \dots, 0)$ belong to $P = \{\sum_{i=1}^d x_i g_i : |x_i| \leq N_i\} \subset \mathbf{Z}^d$, the set of generators $g_i, i = 1, \dots, d$ forms a base with unit determinant of $\mathbf{R}^d$. In P, consider the set of lattice points with all coordinates divisible by k. We observe that (for instance by [30, Theorem 3.36]) this set can be contained in a GAP P' of rank d and cardinality $\max\left(O(\frac{1}{k^r}|P|, 1\right) = \max\left(O(\rho^{-1}/n'^{r/2}), 1\right)$ (here we use the bound $|P| = O(\rho^{-1} \exp(-\frac{m}{4}) m^{d/2})$). Next, define

$$Q := \frac{\beta}{Dk} \cdot P'.$$

It is easy to verify that Q satisfies all the conditions of Corollary 2.9. (Notice that in this case we obtain a stronger bound on the size of Q.) □

Appendix C. Proof of Lemma 6.1

We have

$$\begin{aligned}
\mathbf{P}(\sum_{i=1}^n z_i v_i \in B(x,r)) &= \mathbf{P}(\| \sum_{i=1}^n z_i v_i - x\|_2^2 \leq r^2) \\
&= \mathbf{P}\left(\exp(-\pi \| \sum_{i=1}^n z_i v_i - x\|_2^2) \geq \exp(-\pi r^2) \right) \\
&\leq \exp(\pi r^2) \mathbf{E} \exp(-\pi \| \sum_{i=1}^n z_i v_i - x\|_2^2).
\end{aligned}$$

Notice that

$$\exp(-\pi\|x\|_2^2) = \int_{\mathbf{R}^d} e(\langle x,\xi\rangle)\exp(-\pi\|\xi\|_2^2)d\xi.$$

We thus have

$$\mathbf{P}(\sum_{i=1}^n z_i v_i \in B(x,r)) \leq \exp(\pi r^2)\int_{\mathbf{R}^d} \mathbf{E}e(\langle\sum_{i=1}^n z_i v_i,\xi\rangle)e(-\langle x,\xi\rangle)\exp(-\pi\|\xi\|_2^2)d\xi.$$

Using

$$|\mathbf{E}e(\langle\sum_{i=1}^n z_i v_i,\xi\rangle)| = \prod_{i=1}^n |\mathbf{E}e(z_i\langle v_i,\xi\rangle)|,$$

and

$$|\mathbf{E}e(z_i\langle v_i,\xi\rangle)| \leq |\mathbf{E}e(z_i\langle v_i,\xi\rangle)|^2/2 + 1/2 \leq \exp(-\|\langle v_i,\xi\rangle\|_z^2/2),$$

we obtain

$$\rho_{r,z}(V) \leq \exp(\pi r^2)\int_{\mathbf{R}^d}\exp(-\sum_{i=1}^n \|\langle v_i,\xi\rangle\|_z^2/2 - \pi\|\xi\|_2^2)d\xi.$$

References

[1] Y. Bilu, *Structure of sets with small sumset*, Structure theory of set addition, Asterisque 258 (1999), xi, 77-108.

[2] M. C. Chang, *Generalized arithmetical progressions and sumsets*, Acta Math. Hungar. 65 (1994), no. 4, 379-388.

[3] P. Erdős, *On a lemma of Littlewood and Offord*, Bull. Amer. Math. Soc. 51 (1945), 898-902.

[4] P. Erdős and L. Moser, *Elementary Problems and Solutions*: Solutions: E736. Amer. Math. Monthly, 54 (1947), no. 4, 229-230.

[5] P. Frankl and Z. Füredi, *Solution of the Littlewood-Offord problem in high dimensions*, Ann. of Math. (2) 128 (1988), no. 2, 259-270.

[6] J. Griggs, *Database Security and the Distribution of Subset Sums in* $\mathbf{R}^m$, Graph Theory and Combinatorial Biology, Balatonlelle 1996 , Bolyai Math. Studs. 7 (1999), 223–252.

[7] J. Griggs, J. Lagarias, A. Odlyzko and J. Shearer, *On the tightest packing of sums of vectors*, European J. Combin. 4 (1983), no. 3, 231–236.

[8] G. Halász, *Estimates for the concentration function of combinatorial number theory and probability*, Period. Math. Hungar. 8 (1977), no. 3-4, 197-211.

[9] J. Kahn, J. Komlós and E. Szemerédi, *On the probability that a random* ± 1 *matrix is singular*, J. Amer. Math. Soc. 8 (1995), 223-240.

[10] G. Katona, *On a conjecture of Erdős and a stronger form of Sperner's theorem.* Studia Sci. Math. Hungar 1 (1966), 59–63.

[11] D. Kleitman, *On a lemma of Littlewood and Offord on the distributions of linear combinations of vectors*, Advances in Math. 5 (1970), 155-157.

[12] V. Lev, *Optimal representations by sumsets and subset sums*, J. Number Theory 62 (1997), 127–143.

[13] J. E. Littlewood and A. C. Offord, *On the number of real roots of a random algebraic equation*. III. Rec. Math. Mat. Sbornik N.S. 12 , (1943). 277–286.
[14] H. Nguyen and V. Vu, *A new approach to an old result of Stanley*, in preparation.
[15] R. A. Proctor, *Solution of two difficult combinatorial problems with linear algebra*. Amer. Math. Monthly 89 (1982), no. 10, 721-734.
[16] M. Rudelson and R. Vershynin, *The Littlewood-Offord problem and the condition number of random matrices*, Advances in Mathematics 218 (2008), no 2, 600-633.
[17] A. Sárközy, *Finite addition theorems I*, J. Num. Thy. 32 (1989), 114–130.
[18] A. Sárközy and E. Szemerédi, *Über ein Problem von Erdős und Moser*, Acta Arithmetica, 11 (1965) 205-208.
[19] R. Stanley, *Weyl groups, the hard Lefschetz theorem, and the Sperner property*, SIAM J. Algebraic Discrete Methods 1 (1980), no. 2, 168–184.
[20] E. Szemerédi and V. Vu, *Finite and Infinite Arithmetic Progressions in Sumsets*, Annals of Mathematics, 163 (2006), no 1, 1-35.
[21] E. Szemerédi and V. Vu, *Long arithmetic progressions in sumsets: thresholds and bounds*, J. Amer. Math. Soc. 19 (2006), 119–169.
[22] T. Tao, *Freiman's theorem in solvable groups*, `http://arxiv.org/abs/0906.3535`
[23] T. Tao and V. Vu, *A sharp inverse Littlewood-Offord theorem*, to appear in Random Structures and Algorithms.
[24] T. Tao and V. Vu, *From the Littlewood-Offord problem to the circular law: universality of the spectral distribution of random matrices,* Bull. Amer. Math. Soc. (N.S.) 46 (2009), no. 3, 377–396.
[25] T. Tao and V. Vu, *Inverse Littlewood-Offord theorems and the condition number of random matrices*, Annals of Mathematics (2) 169 (2009), no 2, 595-632.
[26] T. Tao and V. Vu, *John-type theorems for generalized arithmetic progressions and iterated sumsets,* Adv. Math. 219 (2008), no. 2, 428–449.
[27] T. Tao and V. Vu, *On the singularity probability of random Bernoulli matrices*, Journal of the A. M. S 20 (2007), 603-673.
[28] T. Tao and V. Vu, *Random matrices: The Circular Law*, Communication in Contemporary Mathematics 10 (2008), 261-307.
[29] T. Tao and V. Vu, *Smooth analysis of the condition number and the least singular value*, (to appear in Mathematics of Computation).
[30] T. Tao and V. Vu, *Additive Combinatorics*, Cambridge Univ. Press, 2006.

Department of Mathematics, Rutgers, Piscataway, NJ 08854

E-mail address: `hoi@math.rutgers.edu`

Department of Mathematics, Rutgers, Piscataway, NJ 08854

E-mail address: `vanvu@math.rutgers.edu`