Freiman homomorphisms on sparse random sets

D. Conlon* W. T. Gowers†

Abstract

A result of Fiz Pontiveros shows that if A is a random subset of $\mathbb{Z}_N$ where each element is chosen independently with probability $N^{-1/2+o(1)}$, then with high probability every Freiman homomorphism defined on A can be extended to a Freiman homomorphism on the whole of $\mathbb{Z}_N$. In this paper we improve the bound to $CN^{-2/3}(\log N)^{1/3}$, which is best possible up to the constant factor.

1 Introduction

A *Freiman homomorphism* from a subset $A \subset \mathbb{Z}_N$ to $\mathbb{Z}_N$ is a function $\phi : A \to \mathbb{Z}_N$ with the property that $\phi(a) + \phi(b) = \phi(c) + \phi(d)$ whenever a, b, c and d are elements of A with $a + b = c + d$. If ϕ is an affine function, meaning that there exist r and s such that $\phi(a) = ra + s$ for every $a \in A$, then clearly ϕ is a Freiman homomorphism. It is also easy to prove that every Freiman homomorphism from $\mathbb{Z}_N$ to $\mathbb{Z}_N$ is affine. From this it follows that a Freiman homomorphism defined on A is affine if and only if it can be extended to a Freiman homomorphism on the whole of $\mathbb{Z}_N$.

Over the last ten years or so, a number of results have been proved that show that important combinatorial and additive properties of structures such as complete graphs and finite Abelian groups are preserved when one passes to random subsets of those structures of surprisingly low density. We will not attempt an exhaustive summary of this area here, referring the reader instead to the papers [2, 4, 5, 7, 10, 11] and the surveys [3, 9]. In the light of these developments, it is natural to ask whether with high probability every Freiman homomorphism defined on a random subset of $\mathbb{Z}_N$ must be affine. It turns out to be an easy exercise to show that if the elements of A are chosen independently with probability $CN^{-1/3}$, then this is indeed the case. Recently, Fiz Pontiveros [6] proved the significantly harder result that this remains true if the elements of A are chosen with probability $N^{-1/2+o(1)}$.

In the other direction, it is easy to see that the result is false if the elements of A are chosen with probability $N^{-2/3}/2$. Indeed, for each element a, the expected number of triples $(b, c, d) \in A^3$ with $a + b = c + d$ and no two of a, b, c and d equal is at most $N^2(N^{-2/3}/2)^3 = 1/8$, from which it follows that there is a high probability that there will be an "isolated" element $a \in A$ that belongs to no non-trivial quadruple $a + b = c + d$. But that implies that however we choose $\phi(a)$ we will not violate the condition for ϕ to be a Freiman homomorphism.

With a little more effort, one can show that even if we choose elements independently with probability $c(\log N)^{1/3}N^{-2/3}$, for a sufficiently small positive constant c, there is still a significant probability

*Mathematical Institute, Woodstock Road, Oxford OX2 6GG, UK. E-mail: david.conlon@maths.ox.ac.uk. Research supported by a Royal Society University Research Fellowship and by ERC Starting Grant 676632.

†Department of Pure Mathematics and Mathematical Statistics, Wilberforce Road, Cambridge CB3 0WB, UK. Email: w.t.gowers@dpmms.cam.ac.uk. Research supported by a Royal Society 2010 Anniversary Research Professorship.

that some of the elements of A will be isolated in this sense. Therefore, the best result that one can hope for is a probability on the order of $(\log N)^{1/3}N^{-2/3}$. We shall obtain a bound of this form.

Theorem 1.1 *There exists a positive constant C such that if A is a random subset of $\mathbb{Z}_N$ where each element is chosen independently with probability $C(\log N)^{1/3}N^{-2/3}$, then, with high probability, every Freiman homomorphism from A to $\mathbb{Z}_N$ is affine.*

It is also possible to show by similar methods that if the elements of A are chosen with probability $CN^{-2/3}$, then, with high probability, for every Freiman homomorphism ϕ defined on A there is an affine function that agrees with ϕ on most of A. We have not shown this in detail, because the additional factor of $(\log N)^{1/3}$ that is needed for our main result to be true also simplifies other aspects of the proof.

Theorem 1.1 is somewhat unexpected because the bound obtained by Fiz Pontiveros is a natural boundary for his method and could lead one to think that his result is best possible. His proof is based on the following idea. If ϕ is a Freiman homomorphism defined on A, then one can define a function ψ on $A-A$ by taking $\psi(a-b)$ to be $\phi(a)-\phi(b)$. This is well-defined because if $a-b=c-d$, then $\phi(a)-\phi(b)=\phi(c)-\phi(d)$. Fiz Pontiveros proves that ϕ is an affine function by noting that $A-A$ is, with high probability, the whole of $\mathbb{Z}_N$ and then proving that ψ is an additive function: that is, $\psi(x+y)=\psi(x)+\psi(y)$ for every x and y. This implies that ψ is linear and hence that ϕ is affine.

However, the difference set $A-A$ trivially ceases to be the whole of $\mathbb{Z}_N$ once p goes below $N^{-1/2}$, which is why $N^{-1/2}$ is as far as Fiz Pontiveros's argument will go. Indeed, one can say more. For any non-zero x, the expected number of pairs $(a,b)\in A^2$ such that $a-b=x$ is p^2N, so if p is substantially less than $N^{-1/2}$, as it will be for us, then this expectation is substantially less than 1. Moreover, even when x *is* an element of $A-A$, this will almost always happen in just one way, so it seems as though it ought to be very hard for the values of ϕ on one part of A to influence the values it takes on other parts. However, all is not quite lost, because if $p=C(\log N)^{1/3}N^{-2/3}$, then for each $a\in\mathbb{Z}_N$, and in particular for each $a\in A$, the expected number of triples $(b,c,d)\in A^3$ such that $a+b=c+d$ and no two of a, b, c and d are equal is roughly $C^3\log N$. Therefore, a typical element of A will feel the influence from the rest of the set. Nevertheless, it is quite surprising that one can obtain a global result when the average number of such triples is so small.

The proof of Theorem 1.1 has three parts. To begin, we prove a transference principle similar to that used in [4] to prove a number of analogues of combinatorial theorems in sparse random sets. One of the main tools we used there was the finite-dimensional Hahn–Banach separation theorem. In this paper, we shall also use the Hahn–Banach theorem, but this time it is the complex version of the theorem that will be useful to us, and the way we use it will be different. We then use this complex-valued transference principle to prove that any Freiman homomorphism from A to $\mathbb{Z}_N$ agrees on most of A with an affine function. In [4], our results were usually straightforward corollaries of the transference principle. Here, a number of additional ideas are needed to make the argument work. Finally, we conclude with a short argument showing that any function which is affine on most of A is actually affine on all of A. Since our proof does not depend in a serious way on the structure of $\mathbb{Z}_N$, we shall prove the following more general result.

Theorem 1.2 *There exists a positive constant C such that if G is an Abelian group of order n and U is a random subset of G where each element is chosen independently with probability $C(\log n)^{1/3}n^{-2/3}$,*

then, with high probability, every Freiman homomorphism from U to an Abelian group H can be extended to a Freiman homomorphism defined on all of G.

Note that a Freiman homomorphism defined on all of G is simply an affine homomorphism: that is, a map of the form $g \mapsto \phi(g)+h$ for some group homomorphism ϕ and some $h \in H$. Thus, the result is saying that all Freiman homomorphisms defined on U are restrictions of affine homomorphisms.

2 Consequences of the complex finite-dimensional Hahn–Banach theorem

The version of the Hahn–Banach theorem that we shall rely on is the following. Recall that, given a norm $\|.\|$ on a vector space X, the dual norm $\|.\|^*$ of $\|.\|$ is a norm on the collection of linear functionals ϕ acting on X, given by

$$\|\phi\|^* = \sup\{|\phi(x)| : \|x\| \leq 1\}.$$

Theorem 2.1 *Let $\|.\|$ be a norm on $\mathbb{C}^n$ for some positive integer n and let ϕ be a linear functional defined on a subspace X of $\mathbb{C}^n$. Then ϕ can be extended to a linear functional ψ on $\mathbb{C}^n$ with $\|\psi\|^* = \|\phi\|^*$.*

In particular, this theorem has the following corollary. Recall that a subset of $\mathbb{C}^n$ is *absolutely convex* if it is convex and closed under multiplication by scalars of modulus 1.

Corollary 2.2 *Let K and L be two closed bounded absolutely convex subsets of $\mathbb{C}^n$ and suppose that 0 belongs to the interior of $K+L$. Let $v \in \mathbb{C}^n$ be a vector that does not belong to $K+L$. Then there exists a linear functional ψ such that $\psi(v) > 1$ and $|\psi(w)| \leq 1$ for every $w \in K \cup L$.*

Proof. Define a norm $\|.\|$ on $\mathbb{C}^n$ by $\|v\| = \inf\{|\lambda| + |\mu| : v \in \lambda K + \mu L\}$. The absolute convexity of K and L guarantees that this is a norm. Then $\|v\| > 1$, since otherwise for every $\epsilon > 0$ we would be able to find $x \in K$ and $y \in L$ and complex scalars λ and μ with $|\lambda| + |\mu| \leq 1 + \epsilon$ such that $v = \lambda x + \mu y$. By the compactness of K and L and the fact that both are closed under multiplication by scalars of modulus at most 1, it would then follow that $v \in K + L$.

Since $\|v\| > 1$, the linear functional ϕ defined on the 1-dimensional subspace generated by v given by $\phi(\lambda v) = \lambda\|v\|$ has dual norm 1, and $\phi(v) = \|v\| > 1$. By Theorem 2.1, we can extend ϕ to a linear functional ψ defined on all of $\mathbb{C}^n$ such that $\|\psi\|^* \leq 1$. From this last property, we see that if $w \in K$ then $|\psi(w)| \leq \|w\| \leq 1$, and similarly if $w \in L$. This proves the lemma. □

The main result of this section is the following further corollary. We will use the fact that every linear functional on a Hilbert space is of the form $\langle \cdot, \phi\rangle$ for some ϕ, where the inner product of two functions $f, g : X \to \mathbb{C}$ is given by $\langle f, g\rangle = \mathbb{E}_x f(x)\overline{g(x)}$. By saying that a function is a *measure* on a finite set X, we mean that it is a non-negative function from X to $\mathbb{R}$.

Corollary 2.3 *Let μ and ν be measures on a finite set X and let $\|.\|$ be a norm on $\mathbb{C}^X$. Suppose that $|\langle \mu - \nu, |\phi|\rangle| \leq \epsilon$ for every function ϕ such that $\|\phi\|^* \leq \eta^{-1}$. Let f be a function such that $|f| \leq \mu$. Then there exists a function g such that $0 \leq |g| \leq (1-\epsilon)^{-1}\nu$ and $\|f - g\| \leq \eta$.*

Proof. Suppose that we cannot find such a g. Then $f \notin K+L$, where $K = \{g : 0 \leq |g| \leq (1-\epsilon)^{-1}\nu\}$ and $L = \{h : \|h\| \leq \eta\}$. Since both K and L are closed, bounded, and absolutely convex, it follows from Corollary 2.2 that there exists ϕ such that $\langle f, \phi\rangle > 1$, $|\langle g, \phi\rangle| \leq (1-\epsilon)$ whenever $0 \leq |g| \leq \nu$, and $|\langle h, \phi\rangle| \leq 1$ whenever $\|h\| \leq \eta$. The third condition tells us that $\|\phi\|^* \leq \eta^{-1}$. The second tells us that $\langle \nu, |\phi|\rangle \leq 1-\epsilon$, since the function g that maximizes $|\langle g, \phi\rangle|$ subject to the constraint that $|g| \leq \nu$ is the function $g(x) = \nu(x)e^{i\arg(\phi(x))}$, and for that g we have $\langle g, \phi\rangle = \langle \nu, |\phi|\rangle$. From these facts and our hypothesis we deduce that

$$1 < \langle f, \phi\rangle \leq \langle |f|, |\phi|\rangle \leq \langle \mu, |\phi|\rangle \leq \epsilon + \langle \nu, |\phi|\rangle \leq 1,$$

a contradiction. □

3 A norm to which the Hahn–Banach argument will be applied

Let G be an Abelian group of order n and let Γ be the set of all non-degenerate additive quadruples in G. That is, Γ is the set of all quadruples (x, y, z, w) such that $x + y = z + w$ and neither x nor y is equal to z or w. For every function f we define a quantity $M(f)$ to be $\mathbb{E}_{(x,y,z,w)\in\Gamma} f(x)f(y)\overline{f(z)f(w)}$.

Before we continue, we make a quick remark. It is important to consider non-degenerate additive quadruples only, since in a random set of density $n^{-2/3}$, the non-degenerate additive quadruples are swamped by the degenerate ones. Indeed, the number of non-degenerate additive quadruples is approximately $n^{3-8/3} = n^{1/3}$, whereas the number of degenerate additive quadruples is around $(n^{1/3})^2 = n^{2/3}$. (The density at which there are roughly equal numbers of degenerate and non-degenerate additive quadruples is $n^{1/2}$, which is another reason for the natural-seeming barrier there.) At first this appears to be a serious problem, because the quantity $M(f)$ that we have defined is not the fourth power of a norm. However, it turns out not to matter, because the norm we use is constructed in a different way from the U^2 norm.

Let U be a random subset of G with characteristic measure μ. (The *characteristic measure* of a set is its characteristic function divided by its density. Thus, μ is zero outside U, constant inside U, and has average value 1.) We shall show that, with high probability, for every function f from G to $\mathbb{C}$ such that $|f| \leq \mu$ there exists a function g with the following three properties:

1. $\|g\|_\infty \leq 1$;
2. $M(g) \approx M(f)$;
3. $\langle g, \tau\rangle \approx \langle f, \tau\rangle$ for every character $\tau : G \to \mathbb{C}$.

The approach we use to obtain this transference result is closely related to the approach we used in [4], though in that paper we used the real Hahn–Banach theorem. We begin by defining a norm $\|.\|$ with the property that if $\|f - g\|$ is small, then conclusions 2 and 3 above hold. For this, we begin with a simple observation. Write $M(f, g, h, k)$ for the quantity $\mathbb{E}_{(x,y,z,w)\in\Gamma} f(x)g(y)\overline{h(z)k(w)}$. (Thus, $M(f)$ can be thought of as shorthand for $M(f, f, f, f)$.) Then M is additive in all four variables, so

$$M(f) - M(g) = M(f-g, f, f, f) + M(g, f-g, f, f) + M(g, g, f-g, f) + M(g, g, g, f-g).$$

Each of the four expressions on the right-hand side can be regarded as the inner product of $f - g$ with another function. For example, $M(g, g, f-g, f) = \langle h, f-g\rangle$, where $h(z) = \mathbb{E}_{(x,y,z,w)\in\Gamma} g(x)g(y)\overline{f(w)}$.

This motivates the following definitions. First, for any $x \in \mathbb{Z}_N$, let us define Γ_x to be the set of triples (y,z,w) such that $(x,y,z,w) \in \Gamma$. These triples satisfy the equation $z+w-y=x$, but they also satisfy the non-degeneracy condition.

Now define a *basic anti-uniform function* to be any function from G to $\mathbb{C}$ of the form

$$u(x) = \mathbb{E}_{(y,z,w)\in\Gamma_x}\overline{h_1(y)}h_2(z)h_3(w),$$

where each h_i either satisfies $\|h_i\|_\infty \leq 1$ or $|h_i(v)| \leq \mu(v)$ for every v. Then each of the four terms on the right-hand side of the equation above is the inner product of $f-g$ with some basic anti-uniform function (with $f-g$ either on the left or on the right). Therefore, if all such inner products are small, then $M(f) \approx M(g)$. If $h_1 = h_2 = h_3 = \tau$ for some character τ, then $u = \tau$, since for every $(y,z,w) \in \Gamma_x$ we have $\overline{\tau(y)}\tau(z)\tau(w) = \tau(x)$. Therefore, all characters are basic anti-uniform functions and we also obtain the third condition.

We now define a norm $\|.\|$ by setting $\|f\|$ to be the maximum of $|\langle f,u\rangle|$ over all basic anti-uniform functions u. From the discussion above, the following lemma follows easily.

Lemma 3.1 *Let G be a finite Abelian group, let f and g be functions from G to $\mathbb{C}$, and let $\|.\|$ be the norm just defined. Then $|M(f)-M(g)| \leq 4\|f-g\|$.*

Proof. As commented above, $M(f)-M(g)$ can be written as a sum of four terms, each of which is an inner product of $f-g$ with a basic anti-uniform function. □

We now want to apply Corollary 2.3. To do that, we need an expression for the dual norm of the norm $\|.\|$ we have just defined. The following lemma is a standard fact, which can be proved easily with the help of the Hahn–Banach theorem.

Lemma 3.2 *Let X be a finite set and let $\|.\|$ be a norm on $\mathbb{C}^X$ defined by a formula of the form $\|f\| = \max\{|\langle f,\psi\rangle| : \psi \in \Psi\}$. Then the dual norm $\|.\|^*$ is given by the formula*

$$\|\phi\|^* = \inf\{\sum_{i=1}^n |a_i| : \phi = \sum_{i=1}^n a_i\psi_i,\ \psi_i \in \Psi\}.$$

The fact that $\mathbb{C}^X$ is finite-dimensional implies also that this infimum is attained. So the condition in Corollary 2.3 that $\|\phi\|^* \leq \eta^{-1}$ implies that ϕ can be written as a linear combination of basic anti-uniform functions with the absolute values of the coefficients summing to at most η^{-1}.

We want to apply Corollary 2.3 with μ being the characteristic measure of a sparse random set U and with ν being the constant function 1. Therefore, we need to establish that $|\langle \mu-1, |\phi|\rangle|$ is small whenever ϕ is of the form just described. Before we start on this, we must find a way to deal with the fact that we are looking at $|\phi|$, which cannot be described as easily as ϕ. This we do with the help of the following lemma, which is a special case of the Stone–Weierstrass theorem.

Lemma 3.3 *For every pair of real numbers $C,\epsilon > 0$ there exists a polynomial P in z and $\overline{z}$ that uniformly approximates the function $|z|$ to within ϵ on the disc $\{z : |z| \leq C\}$.*

From this we obtain a further reduction of what we hope to prove.

Corollary 3.4 *For every $\epsilon > 0$ and every real number C there exist $\delta > 0$ and a positive integer k with the following property. Let μ be a measure on G with $\mathbb{E}\mu = 1$ and let $\|.\|$ be the norm defined earlier on $\mathbb{C}^G$. Suppose that $|\langle \mu - 1, \xi\rangle| \leq \delta$ for every ξ that can be written as a product of at most k basic anti-uniform functions. Suppose also that $\|u\|_\infty \leq 2$ for every basic anti-uniform function. Then $|\langle \mu - 1, |\phi|\rangle| \leq \epsilon$ for every function ϕ such that $\|\phi\|^* \leq C$.*

Proof. Let P be a polynomial in z and $\overline{z}$ that approximates $|z|$ to within $\epsilon/3$ on the closed disc of radius $2C$. Then $\|P \circ \phi - |\phi|\|_\infty \leq \epsilon/3$ for every function $\phi : G \to \mathbb{C}$ with $\|\phi\|_\infty \leq 2C$. By assumption, $\|u\|_\infty \leq 2$ for every basic anti-uniform function u, so, by Lemma 3.2, $\|\phi\|_\infty \leq 2C$ whenever $\|\phi\|^* \leq C$.

By Lemma 3.2 again, and also the remark following it, if $\|\phi\|^* \leq C$ then we can express ϕ as a linear combination $\sum_{i=1}^n a_i\psi_i$, where the ψ_i are basic anti-uniform functions. Since the pointwise complex conjugate of a basic anti-uniform function is also a basic anti-uniform function, it follows that $P \circ \phi$ is a linear combination of products of at most k basic anti-uniform functions, where k is the degree of P and the sum of the absolute values of the coefficients in the linear combination is bounded above by a constant M that depends on P and C only. (A bound that works is $Q(C)$, where Q is the polynomial obtained from P by replacing each coefficient by its absolute value.)

If we set $\delta = \epsilon/3M$, then we obtain the upper bound $|\langle \mu - 1, P \circ \phi\rangle| \leq \epsilon/3$ for every ϕ with $\|\phi\|^* \leq C$. Since $\|P \circ \phi - |\phi|\|_\infty \leq \epsilon/3$ and $\|\mu\|_1 = \|1\|_1 = 1$, it follows that $|\langle \mu - 1, |\phi|\rangle| \leq \epsilon$, as claimed. □

The main task ahead of us, therefore, is to prove that under suitable circumstances we have the hypothesis that $|\langle \mu - 1, \xi\rangle| \leq \delta$ for every product ξ of at most k basic anti-uniform functions. For this we shall need some probabilistic lemmas, proved in the next section, followed by an argument similar to one in [4] that established a result of this type. However, the argument in this paper is substantially simpler, because the factor of $(\log n)^{1/3}$ in our probability allows us to prove that certain functions are uniformly bounded rather than bounded almost everywhere. (We stress that we do not introduce the logarithmic factor merely to simplify the proof: for this problem it gives the correct bound up to a constant factor.)

4 Some probabilistic lemmas

We begin with the standard Chernoff bound.

Lemma 4.1 *Let X be a set of size n, let $0 < \delta \leq 1$, let $p \in [0, 1]$ and let U be a random subset of X where each element is chosen independently with probability p. Then*

$$\mathbb{P}\big[\big||U| - pn\big| > \delta pn\big] \leq 2\exp(-\delta^2 pn/3).$$

Amongst other things, this lemma allows us to ignore the difference between the characteristic measure μ of U, defined by $\mu(x) = |X|/|U|$ if $x \in U$ and 0 otherwise, and its associated measure, defined by $\mu'(x) = p^{-1}$ if $x \in U$ and 0 otherwise. For example, in the next section, we will actually prove that with high probability $|\langle \mu' - 1, \xi\rangle| \leq \delta$ for every product ξ of at most k basic anti-uniform functions built relative to μ', but the result for the characteristic measure is an immediate corollary.

Lemma 4.2 *Let U_1 and U_2 be two random subsets of an Abelian group G of size n, each with elements chosen independently with probability p. Then the probability that there exists an element of G that can be written in s ways as $u_1 + u_2$ with $u_1 \in U_1$ and $u_2 \in U_2$ is at most $n^{s+1}p^{2s}$.*

Proof. Fix $x \in G$ and for each u let $E(u)$ be the event that $u \in U_1$ and $x - u \in U_2$. Then the events $E(u)$ are independent and each holds with probability p^2, so the expected number of s-tuples $(u_1, \dots, u_s)$ with distinct elements such that $E(u_1), \dots, E(u_s)$ all hold is at most $n^s p^{2s}$. Therefore, the probability that x can be written in s ways as the sum of an element of U_1 and an element of U_2 is at most $n^s p^{2s}$, so the probability that *some* x can be written in s ways is at most $n^{s+1}p^{2s}$, as claimed. □

For the next lemma, we recall that the convolution of two functions $f, g : G \to \mathbb{C}$ is given by $f * g(x) = \mathbb{E}_{y+z=x} f(y)g(z)$.

Lemma 4.3 *For every $0 < \epsilon \leq 1$, there exists a constant C with the following property. Let G be an Abelian group of order n and let U_1, U_2 and U_3 be independent random subsets of G with each element chosen (for each U_i) independently with probability $p = Cn^{-2/3}(\log n)^{1/3}$. For each i, let $\mu_i = p^{-1}\chi_{U_i}$, where χ_{U_i} is the characteristic function of U_i, and let $\mu_i^-(x) = \mu_i(-x)$. Then, with probability $1 - o(n^{-3})$, every value of the function $\mu_1 * \mu_2 * \mu_3^-$ lies between $1 - \epsilon$ and $1 + \epsilon$.*

Proof. Let $x \in G$. Then $\mu_1 * \mu_2 * \mu_3^-(x)$ is equal to p^{-3} times the number of ways of writing x as $u_1 + u_2 - u_3$ with $u_i \in U_i$ for each i. By Lemma 4.2, there is a probability of $1 - o(n^{-3})$ that the maximum value of $\chi_{U_1} * \chi_{U_2}$ is at most 12. For convenience, write η for $\epsilon/4$. Then, by Lemma 4.1, there is a probability of $1 - o(n^{-3})$ that both U_1 and U_2 have sizes between $(1-\eta)pn$ and $(1+\eta)pn$. Let us fix U_1 and U_2 with these properties and consider our random choice of U_3.

The number of ways of writing x as $u_1 + u_2 - u_3$ is equal to $\sum_{u \in U_3} \chi_{U_1} * \chi_{U_2}(x + u)$. This we can represent as a sum of independent random variables X_u, where $X_u = \chi_{U_1} * \chi_{U_2}(x + u)$ with probability p and 0 with probability $1 - p$. The expectation of $\sum_u X_u$ is $n^{-1}|U_1||U_2|\mathbb{E}|U_3|$, which lies between $(1-\eta)^2 p^3 n^2$ and $(1+\eta)^2 p^3 n^2$. From this and the fact that each X_u is bounded above by 12, $\sum_u \mathrm{var}(X_u) \leq 12(1+\eta)^2 p^3 n^2 \leq 20 p^3 n^2$. Therefore, by Bernstein's inequality,

$$\mathbb{P}\big[\big|\sum_u X_u - \mathbb{E}\sum_u X_u\big| > \eta p^3 n^2\big] \leq 2\exp(-\eta^2 p^6 n^4/2(20p^3n^2 + 4\eta p^3 n^2)) \leq 2\exp(-\eta^2 p^3 n^2/48).$$

But $p^3 n^2 = C^3 \log n$, so if we set $C = 8/\eta^{2/3}$ then this upper bound is at most $\exp(-8\log n) = o(n^{-4})$.

We are more or less done, but we need to renormalize. Note that

$$\mathbb{E}_{u_1+u_2-u_3=x}\mu_1(u_1)\mu_2(u_2)\mu_3(u_3) = n^{-2}p^{-3}\sum_u X_u,$$

so we have shown that with probability $1 - o(n^{-3})$, every value of $\mu_1 * \mu_2 * \mu_3^-$ lies between $(1-\eta)^2 - \eta$ and $(1+\eta)^2 + \eta$. Since $\eta \leq 1$, these values lie between $1 - 4\eta$ and $1 + 4\eta$, which proves the result. □

5 Basic anti-uniform functions are approximately contained in the convex hull of a small set

We now return to our task of proving that with high probability $\langle \mu - 1, \xi \rangle$ is small whenever ξ is a product of not too many basic anti-uniform functions. A difficulty with doing this is that the definition of a basic anti-uniform function depends on the measure μ, but this turns out to be less of a problem than it looks: it will be enough to prove the result when the random measure μ and all the measures

used to define the different basic anti-uniform functions are independent. We shall do this first and then give a standard argument that shows why it is enough.

As for the result when all the measures are independent, the technique here, as in [4], is to prove that there is a set $\mathcal{F}$ of bounded functions that is not too large such that every product ξ of basic anti-uniform functions can be approximated by a convex combination ω of functions in $\mathcal{F}$. In [4] we had to give a somewhat complicated definition of "can be approximated by", but here it is simply a uniform approximation. If $\mathcal{F}$ is small enough and if for every bounded function g the inner product $\langle \mu - 1, g\rangle$ is small with high probability, then a union bound will tell us that with high probability $\langle \mu - 1, \phi\rangle$ is small for every $\phi \in \mathcal{F}$ and hence for every ϕ in the convex hull of $\mathcal{F}$.

The proof is slightly complicated by the fact that our basic anti-uniform functions are convolutions of two kinds of functions: functions that are bounded above by 1 and functions that are bounded above by the characteristic measure of a sparse random set. Let us consider first the functions of the latter kind. For these, our argument is based on a simple observation. Suppose that f is a function defined on a finite Abelian group and $\theta \in [0,1]$. Then we define the *θ-random restriction* $R_\theta f$ of f to be a random function g where for each x we have $g(x) = f(x)/\theta$ with probability θ and $g(x) = 0$ with probability $1-\theta$, where all these events are independent. It is important that $R_\theta f$ is not a fixed function but a random one. We shall adopt as a notational convention that if we have functions $R_\theta f_1, \dots, R_\theta f_k$ then all the random values $R_\theta f_i(x)$ are independent. That is, the random restrictions of the different f_i are made independently. We adopt this convention even if some of the f_i are equal. For example, $R_\theta f * R_\theta f$ denotes the convolution of a random restriction of f with another, independently chosen, random restriction of f.

We shall want to take averages over random functions. To avoid confusion with averages of the form $\mathbb{E}_x f(x)$ we shall write $\mathbb{E}_\Sigma$ for the average over the functions themselves.

Our simple observation is that many quantities defined in terms of random restrictions average to the corresponding quantities for the original functions. The underlying reason for this is that for each x we have $\mathbb{E}_\Sigma R_\theta f(x) = f(x)$. That is, $\mathbb{E}_\Sigma R_\theta f = f$. From this it follows that if f, g and h are functions defined on G, then $\mathbb{E}_\Sigma R_\theta f * R_\theta g * R_\theta h = f * g * h$. (Here we are using independence, so that expectations of products are products of expectations.) We even have that

$$\mathbb{E}_\Sigma \prod_{i=1}^k R_\theta f_i * R_\theta g_i * R_\theta h_i = \prod_{i=1}^k f_i * g_i * h_i$$

for any functions $f_1, \dots, f_k, g_1, \dots, g_k, h_1, \dots, h_k$.

This implies that a product of basic anti-uniform functions is a convex combination of products of functions built like basic anti-uniform functions but out of random restrictions instead. Since the random restrictions have smaller support, there are fewer of them – or rather, they can be approximated by a smaller net – which gives us the small set we are looking for.

Unfortunately, when we have $0 \leq f \leq 1$ rather than $0 \leq f \leq \mu$, the number of random restrictions is too large for our argument to work. However, in this case (which we would expect in advance to be easier) there is something else we can do. We first choose three independent random sets $A, B, C \subset G$ of size qn. Given a function f with $0 \leq |f| \leq 1$, we then define $R_A f$ as follows. We choose a random $u \in G$ and we then set $R_A f(x) = q^{-1} f(x)$ if $x \in A + u$ and 0 otherwise. We make similar definitions for $R_B f$ and $R_C f$. We have that $\mathbb{E}_\Sigma R_A f = f$, where here A is fixed and $\mathbb{E}_\Sigma$ is the average over the different $R_A f$ corresponding to different choices of u, and similarly for B and C. With the same

convention that different random restrictions are chosen independently, we also have that

$$\mathbb{E}_\Sigma \prod_{i=1}^k R_A f_i * R_B g_i * R_C h_i = \prod_{i=1}^k f_i * g_i * h_i$$

for any functions $f_1, \dots, f_k, g_1, \dots, g_k, h_1, \dots, h_k$. More generally, let us take Rf to be $R_\theta f$ if $0 \leq f \leq \mu$ and $R_A f$, $R_B f$ or $R_C f$ if $0 \leq f \leq 1$, depending on whether f appears first, second or third in the convolution. Then we have the identity

$$\mathbb{E}_\Sigma \prod_{i=1}^k Rf_i * Rg_i * Rh_i = \prod_{i=1}^k f_i * g_i * h_i$$

regardless of which kinds of functions the f_i, g_i and h_i are.

The reason this observation does not instantly prove what we want is that in order to prove that $\langle \mu - 1, \phi \rangle$ is small for every $\phi \in \mathcal{F}$ we shall need the functions in $\mathcal{F}$ to be bounded. This is almost always true when we take convolutions of random restrictions, but not quite always. So we need to prove that the functions that are not bounded form a sufficiently small proportion of the functions in $\mathcal{F}$ that we can remove them from the convex combination above and still have an approximate equality, where the approximation is in the uniform norm. This we shall do with the help of the following definition and lemma.

Definition 5.1 *Let G be an Abelian group of order n, let $q \in [0,1]$ and let W_1, W_2, W_3 be subsets of G. Then the triple (W_1, W_2, W_3) is* (ϵ, q)-good *if it has the following properties:*

1. *Each W_i has size $(1 + o(1))qn$.*
2. *Let V_1, V_2, V_3 be independent random subsets of G, where each element of each set is chosen with probability q. For each i and each $u_i \in G$, let $\omega_{i,u_i} = q^{-1}\chi_{W_i + u_i}$ and $\nu_i = q^{-1}\chi_{V_i}$. Then, with probability $1 - o(1)$, all convolutions $\beta_1 * \beta_2 * \beta_3^-$ where each β_i is either ω_{i,u_i} or ν_i have the property that every value they take lies between $1 - \epsilon$ and $1 + \epsilon$.*

We shall say that the triple (V_1, V_2, V_3) complements *the triple (W_1, W_2, W_3) if V_1, V_2, V_3 satisfy the conclusion of condition 2.*

Lemma 5.2 *For every $0 < \epsilon \leq 1$ there exists a constant D with the following property. Let G be an Abelian group of order n and let W_1, W_2 and W_3 be independent random subsets of G with elements chosen with probability q, where $q = Dn^{-2/3}(\log n)^{1/3}$. Then the triple (W_1, W_2, W_3) is (ϵ, q)-good with probability at least $1 - o(1)$.*

Proof. Let D be the constant given by Lemma 4.3 (where it is called C), let (V_1, V_2, V_3) be chosen according to the same distribution as (W_1, W_2, W_3) and let β_1, β_2 and β_3 be one of the possible choices for the β_i as in the definition above. Then, by Lemma 4.3, the probability that the values of $\beta_1 * \beta_2 * \beta_3^-$ all lie between $1 - \epsilon$ and $1 + \epsilon$ is $1 - o(n^{-3})$. Therefore, the probability is $1 - o(1)$ that this is true for all $(n+1)^3$ possible choices of $\beta_1, \beta_2, \beta_3$. □

Now let us define a set of bounded functions Ψ_1 and prove that every product of at most k basic anti-uniform functions can be approximated by a convex combination of functions in Ψ_1. We will then find a fairly small subset $\Psi \subset \Psi_1$ with the same property.

Let q be as in Lemma 5.2 and let (W_1, W_2, W_3) be a good triple. Let $p = Cn^{-2/3}(\log n)^{1/3} = Cq/D$ and let U_1, U_2, U_3 be independent sets where each element is chosen independently with probability p. Let $\mu_i = p^{-1}\chi_{U_i}$ for each i.

Given any function $f : G \to \mathbb{C}$, write $\mathrm{supp}(f)$ for the support of f and define $\sigma(f)$ to be the function that takes the value q^{-1} on $\mathrm{supp}(f)$ and 0 elsewhere. This is a kind of "normalized support" of f. Given a function h, let h^* be the function given by $h^*(x) = \overline{h(-x)}$. In particular, $\mathrm{supp}(h^*) = -\mathrm{supp}(h)$. We shall now let Ψ_1 consist of all functions of the form $\prod_{i=1}^k f_i * g_i * h_i^*$ with the following properties:

- For each i, $\mathrm{supp}(f_i)$ is contained in either U_1 or a translate of W_1, $\mathrm{supp}(g_i)$ is contained in either U_2 or a translate of W_2, and $\mathrm{supp}(h_i)$ is contained in either U_3 or a translate of W_3.
- For each i, $\|f_i\|_\infty$, $\|g_i\|_\infty$ and $\|h_i\|_\infty$ are all at most q^{-1}.
- For each i, all of $\mathrm{supp}(f_i)$, $\mathrm{supp}(g_i)$ and $\mathrm{supp}(h_i)$ have size at most $2qn$.
- $\left\|\prod_{i=1}^k \sigma(f_i) * \sigma(g_i) * \sigma(h_i^*)\right\|_\infty \leq 3/2$.

We also need to change slightly the notion of a basic anti-uniform function. We now define it to be a convolution $u_1 * u_2 * u_3^*$ such that for each i we either have $0 \leq |u_i(x)| \leq \mu_i(x)$ for every x or we have $0 \leq |u_i(x)| \leq 1$ for every x. (The earlier definition had $U_1 = U_2 = U_3$.)

Corollary 5.3 *Suppose that (W_1, W_2, W_3) is (ϵ, q)-good, with $q = Dn^{-2/3}(\log n)^{1/3}$ and $\epsilon = 1/4k$, and let U_1, U_2, U_3 and Ψ_1 be as defined above. Then, with probability $1 - o(1)$, every product of at most k basic anti-uniform functions can be approximated up to η in the uniform norm by a convex combination of functions in Ψ_1.*

Proof. Since (W_1, W_2, W_3) is a good triple, the probability that a random triple (V_1, V_2, V_3), where the elements of each V_i are chosen independently with probability q, complements (W_1, W_2, W_3) is $1 - o(1)$. But we can choose the triple (V_1, V_2, V_3) by first choosing a triple (U_1, U_2, U_3), with elements chosen independently with probability p, and then letting each V_i be a random subset of U_i with elements chosen independently with probability $q/p = D/C$. Therefore, with probability $1 - o(1)$ the triple (U_1, U_2, U_3) is such that with probability $1 - o(1)$ a random triple of subsets (V_1, V_2, V_3) chosen in this way complements the triple (W_1, W_2, W_3). In particular, this also implies that, with probability $1 - o(1)$, all convolutions $\beta_1 * \beta_2 * \beta_3^-$, where each β_i is either ω_{i,u_i} or $\mu_i = p^{-1}\chi_{U_i}$, satisfy $\|\beta_1 * \beta_2 * \beta_3^-\|_\infty \leq 2$.

Now let us fix (U_1, U_2, U_3) such that this is the case. Let $\prod_{i=1}^k f_i * g_i * h_i^*$ be a product of basic anti-uniform functions, where each f_i either satisfies $0 \leq f_i \leq 1$ or $0 \leq f_i \leq \mu_1$, and similarly for g_i and h_i with μ_2 and μ_3. If $0 \leq f_i \leq 1$, then let the random restriction Rf_i be defined as follows. We choose u uniformly at random from G and then set $Rf_i(x)$ to be $q^{-1}f_i(x)$ if $x \in W_1 + u$ and 0 otherwise. If $0 \leq f_i \leq \mu_1$, then let $Rf_i(x) = (p/q)f_i(x)$ with probability q/p and 0 otherwise, with the choices being independent. Define the random restrictions $Rg_i(x)$ and $Rh_i(x)$ in the obvious corresponding ways.

As remarked earlier, we then have that

$$\prod_{i=1}^k f_i * g_i * h_i^* = \mathbb{E}_\Sigma \prod_{i=1}^k Rf_i * Rg_i * Rh_i^*.$$

This does not yet finish the proof, since the functions $\prod_{i=1}^k Rf_i * Rg_i * Rh_i^*$ do not necessarily belong to Ψ_1. However, because (W_1, W_2, W_3) is a good triple, we see that with probability $1 - o(1)$ we have $\|\sigma(Rf_i) * \sigma(Rg_i) * \sigma(Rh_i^*)\|_\infty \leq 1 + \epsilon$ for every i, which, since $(1+\epsilon)^k \leq 3/2$, implies that $F = \prod_{i=1}^k Rf_i * Rg_i * Rh_i^*$ belongs to Ψ_1.

Let Σ_1 be the subset of the probability space Σ for which the random function F belongs to Ψ_1 and let $\Sigma_2 = \Sigma \setminus \Sigma_1$. We know that $\|Rf_i * Rg_i * Rh_i^*\|_\infty \leq (p/q)^3 \|\beta_1 * \beta_2 * \beta_3^-\|_\infty$, where each β_i is equal to μ_i or ω_{i,u_i}, so $\|Rf_i * Rg_i * Rh_i^*\|_\infty \leq 2(p/q)^3 = 2(C/D)^3$. Therefore, we always have the bound $\|F\|_\infty \leq (2C/D)^{3k}$.

By the law of total probability, we have

$$\prod_{i=1}^k f_i * g_i * h_i^* = \mathbb{P}[F \in \Psi_1]\, \mathbb{E}_{\Sigma_1} F + \mathbb{P}[F \notin \Psi_1]\, \mathbb{E}_{\Sigma_2} F.$$

The first term is a convex combination of functions in Ψ_1 and the second has ℓ_∞ norm $o(1)(2C/D)^{3k} = o(1)$. This proves the result. □

It remains to show that Ψ_1 has a subset Ψ that is not too large, such that every convex combination of functions in Ψ_1 can be uniformly approximated by a convex combination of functions in Ψ. This we do in a crude way. Given $\delta > 0$, let Δ be a δ-net of the unit disc in $\mathbb{C}$ that includes 0 and has size at most $16/\delta^2$ and let Ψ consist of all functions $\prod_{i=1}^k f_i * g_i * h_i^* \in \Psi_1$ such that every value of every f_i, g_i and h_i is of the form $q^{-1}z$ for some $z \in \Delta$.

Now let $\prod_{i=1}^k f_i * g_i * h_i^* \in \Psi_1$. For each i we can choose functions f_i', g_i' and h_i' taking values in $q^{-1}\Delta$ with $|f_i - f_i'| \leq \delta\sigma(f_i)$, $|g_i - g_i'| \leq \delta\sigma(g_i)$ and $|h_i - h_i'| \leq \delta\sigma(h_i)$. By telescoping, it follows from the triangle inequality that

$$\|f_i * g_i * h_i^* - f_i' * g_i' * h_i'^*\|_\infty \leq 3\delta \|\sigma(f_i) * \sigma(g_i) * \sigma(h_i^*)\|_\infty$$

and, more generally, that

$$\Big\|\prod_i f_i * g_i * h_i^* - \prod_i f_i' * g_i' * h_i'^*\Big\|_\infty \leq 3k\delta \Big\|\prod_i \sigma(f_i) * \sigma(g_i) * \sigma(h_i^*)\Big\|_\infty \leq 9k\delta/2.$$

We are now ready to prove the main result of this section.

Lemma 5.4 *For every $\epsilon > 0$ and every positive integer k there exist constants C and D with the following property. Let $p = Cn^{-2/3}(\log n)^{1/3}$ and let $q = Dp/C = Dn^{-2/3}(\log n)^{1/3}$. Let U be a random set where each element is chosen independently with probability p. Let $\mu = p^{-1}\chi(U)$. Let U_1, U_2, U_3 be further random sets chosen independently in the same way and for each i let $\mu_i = p^{-1}\chi(U_i)$. Then, with probability $1 - o(1)$, $|\langle \mu - 1, \xi\rangle| \leq \epsilon$ for every product ξ of at most k basic anti-uniform functions (as defined just before Corollary 5.3).*

Proof. By Corollary 5.3, with probability $1 - o(1)$ we can express each ξ as a convex combination of functions in Ψ_1 plus an error term that is uniformly bounded by $\epsilon/4$. And then, by the remarks above, if we set $\delta = \epsilon/18k$, we can express every convex combination of functions in Ψ_1 by a convex combination of functions in Ψ plus an error term that is again uniformly bounded by $\epsilon/4$. Thus, $\xi = \xi_1 + \xi_2$, where ξ_1 is a convex combination of functions in Ψ and $\|\xi_2\|_\infty \leq \epsilon/2$.

Since

$$|\langle \mu - 1, \xi \rangle| \leq |\langle \mu - 1, \xi_1 \rangle| + |\langle \mu - 1, \xi_2 \rangle| \leq |\langle \mu - 1, \xi_1 \rangle| + p^{-1}|U|\epsilon/2n,$$

and since $|U| \leq 3pn/2$ with probability $1 - o(1)$, it is enough to prove that with probability $1 - o(1)$, $|\langle \mu - 1, \psi \rangle| \leq \epsilon/4$ for every $\psi \in \Psi$. For this we can use a union bound. That is, we need to obtain upper bounds for the number of functions in Ψ and for the probability that $|\langle \mu - 1, \psi \rangle| > \epsilon/4$ for any individual $\psi \in \Psi$.

For any given set A of size at most $2qn$, the number of functions supported in A and taking values in $q^{-1}\Delta$ is $|\Delta|^{2qn}$. Since we are taking δ to be $\epsilon/18k$, this is at most $(10^4 k^2/\epsilon^2)^{2qn}$. The number of possible supports for each function involved in a product of convolutions in Ψ is at most $n + \binom{2pn}{2qn}$, which is at most $(4C/D)^{2qn}$. Therefore, the number of functions in Ψ is at most $(10^5 k^2 C/\epsilon^2 D)^{6kqn}$. Since each $\psi \in \Psi$ has ℓ_∞ norm at most 2, the probability that $|\langle \mu - 1, \psi \rangle| > \epsilon/4$ is, by Bernstein's inequality, at most $2\exp(-\epsilon^2 n^2/32(4p^{-1}n + \frac{1}{6}p^{-1}\epsilon n)) \leq 2\exp(-\frac{1}{160}\epsilon^2 pn)$, since $\langle \mu - 1, \psi \rangle$ is an average of n random variables, each of mean zero, second moment at most $4p^{-1}$, and maximum at most $2p^{-1}$.

We are therefore done provided that $(10^5 k^2 C/\epsilon^2 D)^{6kqn} \exp(-\frac{1}{160}\epsilon^2 pn) = o(1)$. Taking logs, we require

$$\frac{1}{160}\epsilon^2 pn - 6kqn \log(10^5 k^2 C/\epsilon^2 D)$$

to tend to infinity, for which it is enough if $C/D > 10^3 k\epsilon^{-2} \log(10^5 k^2 C/\epsilon^2 D)$. If we let D be the constant given by Lemma 5.2, then we are done. □

We are not quite done, since we have assumed that our basic anti-uniform functions were built from functions supported on random sets that were independent of U. To recover the statement for basic anti-uniform functions built from functions supported on U itself, we think of U as being the union of t independent random sets $U_1, \dots, U_t$, each chosen with probability $\frac{1}{t}Cn^{-2/3}(\log n)^{1/3}$, with U_i having associated measure μ_i. With high probability, these sets are all disjoint, so $\mu = \mathbb{E}_i \mu_i$ and any function f with $0 \leq |f| \leq \mu$ can be written as an expectation $\mathbb{E}_i f_i$ where $0 \leq |f_i| \leq \mu_i$ for each i. In particular, any expression of the form $\langle \mu - 1, \xi \rangle$ can be rewritten as the expectation over expressions of the form $\langle \mu_i - 1, \xi_{j_1,\dots,j_{3k}} \rangle$, where the indices $j_1, \dots, j_{3k}$ indicate which of the sets $U_1, \dots, U_t$ the basic anti-uniform function $\xi_{j_1,\dots,j_{3k}}$ is built over.

Since $i, j_1, \dots, j_{3k}$ all vary over $1, 2, \dots, t$, the probability that any two of them are equal is at most $\binom{3k+1}{2}/t$. If they are all different, Lemma 5.4 applied with C/t and $\epsilon/2$ implies that $|\langle \mu_i - 1, \xi_{j_1,\dots,j_{3k}} \rangle| \leq \epsilon/2$, while in general, using that $\|\mu_a * \mu_b * \mu_c^-\|_\infty \leq 2$ with high probability for all a, b, c, we have the bound $|\langle \mu_i - 1, \xi_{j_1,\dots,j_{3k}} \rangle| \leq 2^{k+1}$. Therefore,

$$\begin{aligned} |\langle \mu - 1, \xi \rangle| &= |\mathbb{E}_{i,j_1,\dots,j_{3k}} \langle \mu_i - 1, \xi_{j_1,\dots,j_{3k}} \rangle| \\ &\leq \frac{\epsilon}{2} + 2^{k+1} \frac{\binom{3k+1}{2}}{t} \leq \epsilon, \end{aligned}$$

for an appropriate t. This completes the proof of our transference principle. In the next section, we will show how to apply it.

6 Freiman homomorphisms are affine almost everywhere

In this section, we shall use the main result of the previous section to prove that if G is an Abelian group of order n and U is a random subset of G chosen with probability $C(\log n)^{1/3} n^{-2/3}$, then, with

high probability, for every finitely generated Abelian group H and every Freiman homomorphism $\phi : U \to H$ there is an affine homomorphism $\psi : G \to H$ such that $\psi(u) = \phi(u)$ for at least 99.99% of the elements $u \in U$. Then in the next section we shall get from 99.99% to 100%.

The condition here that H is finitely generated is not important, since we can always restrict to the subgroup generated by the image of U, but it will be convenient for us to assume it so that we can take expectations over the characters of H.

Suppose now that $\phi : U \to H$ is a Freiman homomorphism. To apply our transference principle, we need a function from U to $\mathbb{C}$, and the obvious way of producing such a function is to compose ϕ with a character, except that we shall normalize this function by multiplying it by the characteristic measure μ of U. Accordingly, if χ is a character on H, let us define f_χ to be the function $\mu(\chi \circ \phi)$. Note that $\chi \circ \phi$ is a Freiman homomorphism from U to the unit circle.

Let us now fix χ and write f for f_χ. Let $\eta > 0$ be a constant to be chosen later. By the transference principle, we can find a function $g : G \to \mathbb{C}$ such that $\|g\|_\infty \leq 1$, $|M(g) - M(f)| \leq \eta$, and $|\langle g - f, \tau\rangle| \leq \eta$ for every character $\tau : G \to \mathbb{C}$.

We shall now show that $M(f) \approx 1$, which implies that $M(g) \approx 1$, which implies that g is approximately equal to a character τ, which implies that $\langle f, \tau\rangle \approx 1$. This will show that $\|\hat{f}_\chi\|_{12} \approx 1$ for every character $\chi : H \to \mathbb{C}$, which (as later calculations will reveal) effectively allows us to prove that ϕ coincides with a homomorphism of order 6 on almost all of U. But since $3U = G$ with high probability, a homomorphism of order 6 on almost all of U gives us a homomorphism of order 2 on almost all of G.

In the rest of this section we shall give the details. To avoid a lot of repetition, let us establish once and for all that G is a finite Abelian group of order n, $\eta > 0$ is a small positive constant, U is a random subset of G chosen with probability $C(\log n)^{1/3}n^{-2/3}$, and μ is the characteristic measure of U. In addition, we shall adopt the following convention. When we make a statement that involves U, it is to be understood that that statement is valid with probability $1 - o(1)$. Moreover, if that statement involves a universal quantifier (either explicitly, or via the word "let"), it is to be understood that the "with probability $1 - o(1)$" comes *before* the quantifier. For instance, if we write, "Let A be such that $P(A)$. Then $Q(A, U)$," that is shorthand for, "With probability $1 - o(1)$, $Q(A, U)$ for every A such that $P(A)$."

We begin with a simple probabilistic lemma. Though it has a much easier proof, we note here that it follows as a corollary of the fact that $\langle \mu - 1, \xi\rangle$ is small for all basic anti-uniform functions ξ.

Lemma 6.1 $M(\mu) \geq 1 - \eta$.

Recall that for each character $\tau : G \to \mathbb{C}$ and any function $f : G \to \mathbb{C}$, the Fourier transform is given by $\hat{f}(\tau) = \langle f, \tau\rangle = \mathbb{E}_x f(x)\overline{\tau(x)}$. We also have the Fourier inversion formula, $f(x) = \sum_\tau \hat{f}(\tau)\tau(x)$. In keeping with these normalisations, we write $\|f\|_p = (\mathbb{E}_x|f(x)|^p)^{1/p}$, while $\|\hat{f}\|_p = (\sum_\tau |\hat{f}(\tau)|^p)^{1/p}$. This allows us to state Parseval's formula $\|f\|_2 = \|\hat{f}\|_2$ and a number of related identities in a clean fashion.

Lemma 6.2 *Let H be a finitely generated Abelian group, let $\phi : U \to H$ be a Freiman homomorphism, let $\chi : H \to \mathbb{C}$ be a character, and let $f = \mu(\chi \circ \phi)$. Then $\|\hat{f}\|_{12}^{12} \geq 1 - 36\eta$.*

Proof. As mentioned above, there exists a function $g : G \to \mathbb{C}$ such that $\|g\|_\infty \leq 1$, $M(g) \geq M(f) - \eta$,

and $\langle f,\tau\rangle \geq \langle g,\tau\rangle - \eta$ for every character $\tau : G \to \mathbb{C}$. But

$$M(f) = \mathbb{E}_{(x,y,z,w)\in\Gamma}\mu(x)\mu(y)\mu(z)\mu(w)\chi(\phi(x)+\phi(y)-\phi(z)-\phi(w)) = M(\mu),$$

since ϕ is a Freiman homomorphism. Therefore, by Lemma 6.1, $M(f) \geq 1-\eta$. It follows that $M(g) \geq 1-2\eta$.

Now the sum of $g(x)g(y)\overline{g(z)g(w)}$ over degenerate additive quadruples $x+y=z+w$ (that is, ones where $x=z$ and $y=w$ or $x=w$ and $y=z$) is a non-negative real number. Therefore, $\|g\|_{U^2}^4 \geq 1-2\eta$. Therefore, $\|\hat{g}\|_4^4 \geq 1-2\eta$. Since $\|\hat{g}\|_2 = \|g\|_2 \leq 1$, it follows that $\|\hat{g}\|_\infty^2 \geq 1-2\eta$ and therefore that $\|\hat{g}\|_\infty \geq 1-2\eta$.

Since $\langle f,\tau\rangle \geq \langle g,\tau\rangle - \eta$ for every character $\tau : G \to \mathbb{C}$, it follows that $\|\hat{f}\|_\infty \geq 1-3\eta$ and therefore that $\|\hat{f}\|_{12}^{12} \geq 1-36\eta$, as claimed. □

The proof of the above lemma is the only place where we need to use the transference result of the previous section. (However, the lemma is crucial to our main argument.)

Let us now regard H and ϕ as fixed (though it is important that the statement we are proving about U is one that with high probability holds for all H and ϕ). We shall now convert this statement about ℓ_{12} norms of Fourier transforms into a statement about Freiman homomorphisms of order 6. To begin with, we obtain *near* homomorphisms. The rough meaning of the next statement is that for almost every additive 12-tuple $(x_1,\dots,x_{12})$ in U, we also have $\phi(x_1)+\cdots+\phi(x_6) = \phi(x_7)+\cdots+\phi(x_{12})$.

Lemma 6.3 *Let Γ be the set of all $(x_1,\dots,x_{12}) \in G^{12}$ such that $x_1+\cdots+x_6 = x_7+\cdots+x_{12}$ and let Φ be the set of all $(x_1,\dots,x_{12}) \in \Gamma$ such that $\phi(x_1)+\cdots+\phi(x_6) = \phi(x_7)+\cdots+\phi(x_{12})$. Then*

$$\mathbb{E}_{(x_1,\dots,x_{12})\in\Gamma}\ \mu(x_1)\dots\mu(x_{12})\mathbf{1}_{(x_1,\dots,x_{12})\in\Phi} \geq 1-36\eta.$$

Proof. Let us write f_χ for the function $\mu(\chi\circ\phi)$. (Previously we just wrote f, but now we need to consider all such functions.) Then, by Lemma 6.2, we know that $\|\hat{f}_\chi\|_{12}^{12} \geq 1-36\eta$ for every character χ, and therefore that $\mathbb{E}_\chi\|\hat{f}_\chi\|_{12}^{12} \geq 1-36\eta$. Note that

$$\|\hat{f}_\chi\|_{12}^{12} = \sum_\tau |\hat{f}(\tau)|^{12} = \langle \hat{f}^6, \hat{f}^6\rangle = \langle f*\cdots*f, f*\cdots*f\rangle = \mathbb{E}_{(x_1,\dots,x_{12})\in\Gamma} f(x_1)\dots f(x_6)\overline{f(x_7)\dots f(x_{12})}$$

and, therefore, $\mathbb{E}_\chi\|\hat{f}_\chi\|_{12}^{12}$ is equal to

$$\mathbb{E}_\chi\mathbb{E}_{(x_1,\dots,x_{12})\in\Gamma}\ \mu(x_1)\dots\mu(x_{12})\chi(\phi(x_1)+\cdots+\phi(x_6)-\phi(x_7)-\cdots-\phi(x_{12})).$$

The expectation over χ is 1 if $(x_1,\dots,x_{12}) \in \Phi$ and 0 otherwise, so we have precisely the expectation on the left-hand side of the inequality we are trying to prove. □

The rough idea of what we want to do next is to define a function $\psi : 3U \to H$ by taking $\psi(u)$ to be the most popular value of $\phi(x_1)+\phi(x_2)+\phi(x_3)$ when $x_1+x_2+x_3 = u$. Lemma 6.3 implies that most of the time one value is predominant and that the function ψ thus defined has the property that $u_1+u_2 = u_3+u_4$ almost always implies that $\psi(u_1)+\psi(u_2) = \psi(u_3)+\psi(u_4)$. This in turn implies that ψ agrees almost everywhere with an affine homomorphism.

However, arguments of the above kind can get quite messy, so for the sake of tidiness we shall instead postpone for as long as we can the moment where we have to use an averaging argument to

commit ourselves to a particular function. The price we pay for this is that we must consider functions whose values are probability distributions on H rather than single elements of H. However, this is a very natural thing to do: instead of defining $\psi(u)$ to be the most popular value of $\phi(x_1)+\phi(x_2)+\phi(x_3)$ such that $x_1+x_2+x_3=u$, we define it to be something like the probability distribution where the probability that $\psi(u)=h$ is the probability that $\phi(x_1)+\phi(x_2)+\phi(x_3)=h$ given that $x_1+x_2+x_3=u$. This is not quite accurate, because we need to take account of the fact that $\mu*\mu*\mu$ is not quite constant, so here is the precise definition.

Definition 6.4 *Let $\phi:U\to H$ and let $\pi(H)$ be the set of all non-negative real-valued finitely supported functions on H. Then, for each $x\in G$, let $\psi(x)\in\pi(H)$ be the function defined by the formula*

$$\psi(x)(h)=\mathbb{E}\{\mu(x_1)\mu(x_2)\mu(x_3):x_1+x_2+x_3=x,\phi(x_1)+\phi(x_2)+\phi(x_3)=h\}.$$

Typically, $\psi(x)(h)$ will be almost 1 for one $h\in H$ and the sum over h will be asymptotically equal to 1. In other words, it will be concentrated at one h, so we can think of ψ as like a function from G to H but slightly fuzzy.

We also need to be able to make sense of expressions such as $\psi(x)+\psi(y)$. That is easy enough: we simply convolve $\psi(x)$ with $\psi(y)$. To avoid confusion, we shall write $\psi(x)*\psi(y)$, which we define formally as follows. Note that the normalization is different from that used earlier: it is more appropriate for convolutions of probability distributions.

Definition 6.5 *Let $p,q\in\pi(H)$. The* convolution *$p*q$ of p and q is defined by the formula $(p*q)(h)=\sum_{h_1+h_2=h}p(h_1)q(h_2)$.*

If $\psi(x)$ is concentrated at h_1 and $\psi(y)$ is concentrated at h_2, then $\psi(x)*\psi(y)$ is concentrated (but not quite as strongly) at h_1+h_2. In this situation, convolution can be thought of as a fuzzy version of addition.

We would also like a fuzzy version of subtraction, so we define $p_-(h)$ to be $p(-h)$. Then the fuzzy analogue of $\psi(x)-\psi(y)$ is $\psi(x)*\psi_-(y)$.

Finally, we need a fuzzy version of equality: there are no circumstances under which we can reasonably expect two expressions such as $\psi(x_1)*\psi(x_2)$ and $\psi(x_3)*\psi(x_4)$ to be exactly equal, but we can certainly expect them to be roughly equal. To measure this, we define an inner product in an obvious way. Again, this is defined with a different normalization from earlier.

Definition 6.6 *Let p and q be two non-negative finitely supported functions on H. Then their* inner product *$\langle p,q\rangle$ is defined to be $\sum_{h\in H}p(h)q(h)$.*

If p and q both sum to 1 (or approximately 1), then we regard p and q as close if $\langle p,q\rangle$ is close to 1. This is a stronger statement than merely that p and q are similar functions: it implies also that p and q are both fairly concentrated at a single value. However, that is exactly what we want to show, so this notion of closeness is useful.

The next lemma is just a reformulation of Lemma 6.3.

Lemma 6.7 $\mathbb{E}_{z_1+z_2=z_3+z_4}\langle\psi(z_1)*\psi(z_2),\psi(z_3)*\psi(z_4)\rangle\geq 1-36\eta$.

Proof. The inner product expands to

$$\sum_{h_1+h_2=h_3+h_4} \psi(z_1)(h_1)\psi(z_2)(h_2)\psi(z_3)(h_3)\psi(z_4)(h_4).$$

But $\psi(z_i)(h_i)$ can be written as

$$\mathbb{E}\{\mu(x_{3i-2})\mu(x_{3i-1})\mu(x_{3i}) : x_{3i-2}+x_{3i-1}+x_{3i}=z_i, \phi(x_{3i-2})+\phi(x_{3i-1})+\phi(x_{3i})=h_i\}.$$

Therefore, taking the expectation of the inner product over all additive quadruples $z_1+z_2=z_3+z_4$, we obtain

$$\mathbb{E}\{\mu(x_1)\dots\mu(x_{12}) : x_1+\dots+x_6=x_7+\dots+x_{12}, \phi(x_1)+\dots+\phi(x_6)=\phi(x_7)+\dots+\phi(x_{12})\}$$

which is another way of writing $\mathbb{E}_{(x_1,\dots,x_{12})\in\Gamma}\ \mu(x_1)\dots\mu(x_{12})\mathbf{1}_{(x_1,\dots,x_{12})\in\Phi}$. Thus, the result follows from Lemma 6.3. □

Corollary 6.8 $\mathbb{E}_{z_1-z_2=z_3-z_4}\langle\psi(z_1)*\psi_-(z_2),\psi(z_3)*\psi_-(z_4)\rangle\geq 1-36\eta$.

Proof. The inner product expands to

$$\sum_{h_1+h_2=h_3+h_4} \psi(z_1)(h_1)\psi(z_2)(-h_2)\psi(z_3)(h_3)\psi(z_4)(-h_4)$$

which equals

$$\sum_{h_1-h_2=h_3-h_4} \psi(z_1)(h_1)\psi(z_2)(h_2)\psi(z_3)(h_3)\psi(z_4)(h_4)$$

which equals

$$\sum_{h_1+h_4=h_3+h_2} \psi(z_1)(h_1)\psi(z_2)(h_2)\psi(z_3)(h_3)\psi(z_4)(h_4)$$

which equals $\langle\psi(z_1)*\psi(z_4),\psi(z_2)*\psi(z_3)\rangle$. We are taking the expectation of this quantity over quadruples (z_1,z_2,z_3,z_4) such that $z_1+z_4=z_2+z_3$, so the result follows from Lemma 6.7. □

We are about to define $\theta : G\to\pi(H)$ as the convolution of ψ with ψ_-^-, where $\psi_-^-(x)(h)$ is defined to be $\psi(-x)(-h)$. However, we must first say what "convolution" means here.

Definition 6.9 *Let $\phi,\psi : G\to\pi(H)$. The* convolution $\phi*\psi : G\to\pi(H)$ *is defined by the formula*

$$(\phi*\psi)(x)=\mathbb{E}_{x_1+x_2=x}\phi(x_1)*\psi(x_2).$$

Let $\zeta : G\to H$ be a Freiman homomorphism and let $\phi : G\to\pi(H)$ be defined by $\phi(x)(h)=1$ if $h=\zeta(x)$ and 0 otherwise. In that case, $\phi*\phi_-^-(x)(h)=1$ if for every $x_1-x_2=x$ we have $\zeta(x_1)-\zeta(x_2)=h$ and 0 otherwise. In other words, $\phi*\phi_-^-$ is essentially the well-defined function from $G-G$ to H that is induced by the fact that ζ is a Freiman homomorphism. Lemma 6.7 can be thought of as saying that ψ is an *approximate* homomorphism. We therefore expect $\theta=\psi*\psi_-^-$ to be "approximately well-defined". We shall show that for every x the function $\theta(x)$ is concentrated at some value $\gamma(x)\in H$, and that γ is a group homomorphism.

First, we need a lemma that can be thought of as a kind of triangle inequality, with $1-\langle p,q\rangle$ being the "distance" between p and q. This distance is similar to the Ruzsa distance between two sets: in particular, a function need not be close to itself.

Lemma 6.10 *Let p, q and r be elements of $\pi(H)$, each of which sums to at most $1+\beta$, where $\beta \leq 1$. Then*

$$1 - \langle p, r\rangle \leq 1 - \langle p, q\rangle + 1 - \langle q, r\rangle + 3\beta.$$

Proof. For every $h \in H$ we have the inequality $(1+\beta-p(h))(1+\beta-r(h)) \geq 0$, since p and r take values in $[0, 1+\beta]$. It follows that $p(h)r(h) \geq (1+\beta)(p(h)+r(h)) - (1+\beta)^2$. Therefore, since q also takes values in $[0, 1+\beta]$,

$$\begin{aligned}
1 - \langle p, r\rangle &= 1 - \sum_h p(h)r(h) \\
&\leq 1 - (1+\beta)^{-1} \sum_h q(h)p(h)r(h) \\
&\leq 1 - \sum_h q(h)(p(h)+r(h)-(1+\beta)) \\
&= 1 - \langle p, q\rangle - \langle q, r\rangle + (1+\beta)\sum_h q(h) \\
&\leq 1 - \langle p, q\rangle + 1 - \langle q, r\rangle + 3\beta,
\end{aligned}$$

as claimed. □

Let us write $d(p, q)$ for $1 - \langle p, q\rangle$. Then Lemma 6.10 tells us that $d(p, r) \leq d(p, q) + d(q, r) + \beta$. Note that $d(p, q)$ can be negative, but it cannot be smaller than $-2\beta - \beta^2$. The fact that it can be negative turns out not to matter.

Now we need an estimate that can be used to give us a β to use in the previous lemma.

Lemma 6.11 $\|\mu * \mu * \mu - 1\|_\infty \leq \eta$.

Proof. It will suffice to show that for any fixed x, the number of distinct ways S of writing x as a sum of three elements in U satisfies $|S - \mathbb{E}S| \leq \eta p^3 n^2$ with probability $1 - o(1/n)$. A straightforward application of Janson's inequality (see [1]) gives the required estimate for the probability that $S < \mathbb{E}S - \eta p^3 n^2$. We will therefore focus on estimating the probability that $S > \mathbb{E}S + \eta p^3 n^2$.

We will use the method described in Section 2.3.4 of [8]. Let S' be the random variable counting the maximum number of disjoint three element sets each of which sum to x. We claim that $S \leq S' + 28$ with probability $1 - o(1/n)$. To prove the claim, let $\mathcal{S}_i$ be the collection of i element subsets of U giving rise to a triple summing to x. Form a graph J whose vertices are the elements of $\mathcal{S}_3$, where two elements are joined if and only if they intersect. Then, with probability $1 - o(1/n)$, it is straightforward to verify that the maximum degree of J is at most 4 and the largest induced matching has size at most 3. Note that S' is the order of the largest independent set in J. Since at most 24 vertices are joined to a maximal induced matching and the remaining set is independent, we have

$$S' \geq |J| - 24 = |\mathcal{S}_3| - 24.$$

Since it is also easy to verify that $|\mathcal{S}_2| \leq 3$ with probability $1 - o(1/n)$ and $\mathcal{S}_1 \leq 1$ (unless G has characteristic 3 and $x = 0$), the claim follows.

The required conclusion now follows from the inequality

$$\mathbb{P}[S' \geq \mathbb{E}S + t] \leq \exp\left(\frac{-t^2}{2(\mathbb{E}S + t/3)}\right),$$

which is Lemma 2 of [8]. □

Remark. It is possible to prove the above lemma in a slightly more elementary way, by deducing it from Lemma 4.3. To do this, one must split μ into several independent parts and use the fact that most of the terms that arise involve different parts. (We used a similar idea at the end of the previous section.)

There is a simple way of measuring the well-definedness of θ: we look at how small the distances $d(\theta(x), \theta(x))$ are.

Lemma 6.12 *For every $x \in G$, $d(\theta(x), \theta(x)) \leq 75\eta$.*

Proof. Fix $x \in G$. We need to show that

$$\mathbb{E}_{z_1 - z_2 = z_3 - z_4 = x} \langle \psi(z_1) * \psi_-(z_2), \psi(z_3) * \psi_-(z_4) \rangle \geq 1 - 75\eta.$$

From Corollary 6.8, we know both that

$$\mathbb{E}_{w_1 - w_2 = w_3 - w_4}(1 - \langle \psi(w_1) * \psi_-(w_2), \psi(w_3) * \psi_-(w_4) \rangle) \leq 36\eta$$

and that

$$\mathbb{E}_{w_1 - w_2 = w_3 - w_4}(1 - \langle \psi(w_1 + x) * \psi_-(w_2 + x), \psi(w_3) * \psi_-(w_4) \rangle) \leq 36\eta.$$

Now $\sum_{h \in H} \psi(x)(h) = \mu * \mu * \mu(x)$, which is at most $1 + \eta$, by Lemma 6.11. It follows from Lemma 6.10 that

$$\mathbb{E}_{w_1, w_2}(1 - \langle \psi(w_1) * \psi_-(w_2), \psi(w_1 + x) * \psi_-(w_2 + x) \rangle) \leq 75\eta.$$

But

$$\langle \psi(w_1) * \psi_-(w_2), \psi(w_1 + x) * \psi_-(w_2 + x) \rangle = \langle \psi(w_1 + x) * \psi_-(w_1), \psi(w_2 + x) * \psi_-(w_2) \rangle.$$

Therefore,

$$\mathbb{E}_{w_1, w_2} \langle \psi(w_1 + x) * \psi_-(w_1), \psi(w_2 + x) * \psi_-(w_2) \rangle \geq 1 - 75\eta,$$

which is (a slightly rewritten version of) what we needed to show. □

Corollary 6.13 *For every x, there exists h such that $\theta(x)(h) \geq 1 - 77\eta$.*

Proof. We know that $\sum_h \theta(x)(h)^2 \geq 1 - 75\eta$. Also, since $\sum_h \psi(y)(h) \leq 1 + \eta$ for every y, $\sum_h \theta(x)(h) \leq \mathbb{E}_{x_1 - x_2 = x} \sum_{h_1, h_2} \psi(x_1)(h_1)\psi(x_2)(-h_2)$ is at most $(1+\eta)^2$. It follows that $\max_h \theta(x)(h) \geq (1 - 75\eta)(1 + \eta)^{-2} \geq 1 - 77\eta$, as claimed. □

Corollary 6.14 $\theta(0)(0) \geq 1 - 77\eta$.

Proof. For every h,

$$\theta(0)(h) = \mathbb{E}_x \sum_{h_1 - h_2 = h} \psi(x)(h_1)\psi(x)(h_2).$$

Let h' be such that $\psi(x)(h')$ is the largest value over all h of $\psi(x)(h)$. Then, if $h \neq 0$,

$$\sum_{h_1 - h_2 = h} \psi(x)(h_1)\psi(x)(h_2) \leq 2\psi(x)(h') \sum_{h_1 \neq h'} \psi(x)(h_1) \leq \frac{(1+\eta)^2}{2},$$

where we used that $\psi(x)(h') \sum_{h_1 \neq h'} \psi(x)(h_1)$ is bounded by an expression of the form $x((1+\eta)-x) \leq (1+\eta)^2/4$. Therefore, if $h \neq 0$, $\theta(0)(h) \leq (1+\eta)^2/2$. Since this is less than $1 - 77\eta$ for η sufficiently small, the only way that Corollary 6.13 can be true is if $\theta(0)(0) \geq 1 - 77\eta$. □

Lemma 6.15 *Let $\beta, \gamma \geq 0$ and let $p, q \in \pi(H)$ with $\sum_h p(h)$ and $\sum_h q(h)$ at most $1 + \beta$. Suppose that $d(p, q) \leq \gamma$. Then there exists h such that $p(h)q(h) \geq (1 - \beta - \gamma)^2$.*

Proof. We know that

$$\sum_h p(h)q(h) \leq (\max_h p(h)^{1/2} q(h)^{1/2}) \sum_h p(h)^{1/2} q(h)^{1/2}.$$

If the result is false, then $\max_h p(h)^{1/2}q(h)^{1/2}$ is less than $1 - \beta - \gamma$, and, by the Cauchy–Schwarz inequality and the assumptions on p and q, the inner sum is at most $1 + \beta$. It follows that $\langle p, q\rangle < (1 - \beta - \gamma)(1 + \beta) \leq 1 - \gamma$, and therefore that $d(p, q) > \gamma$, a contradiction. □

The next lemma tells us that inner products are "approximately Lipschitz" functions of their arguments.

Lemma 6.16 *Let $0 \leq \beta \leq 1/5$ and let $p, q, r \in \pi(H)$ be such that $\sum_h p(h)$, $\sum_h q(h)$ and $\sum_h r(h)$ are all at most $1 + \beta$. Then $|\langle p, r\rangle - \langle q, r\rangle| \leq 5d(p, q) + 10\beta$.*

Proof. For any element $p \in \pi(H)$ and any $s \in [1, \infty)$, write $\|p\|_s$ for $\left(\sum_{h \in H} p(h)^s\right)^{1/s}$ and $\|p\|_\infty$ for $\max_h p(h)$. Then

$$\begin{aligned} |\langle p, r\rangle - \langle q, r\rangle| &= |\langle p - q, r\rangle| \\ &\leq \|p - q\|_1 \|r\|_\infty \\ &\leq (1 + \beta)\|p - q\|_1. \end{aligned}$$

Let us write γ for $d(p, q)$. By Lemma 6.15, there exists h such that $p(h)q(h) \geq (1 - \beta - \gamma)^2$, which implies that $p(h) + q(h) \geq 2(1 - \beta - \gamma)$. Therefore, $\sum_{h' \neq h}(p(h') + q(h')) \leq 4\beta + 2\gamma$. Also, $|p(h) - q(h)|$ is at most $1 + \beta - (1 - \beta - \gamma)^2/(1 + \beta) \leq 4\beta + 2\gamma$. Therefore, $\|p - q\|_1 \leq 8\beta + 4\gamma$. Since $\beta \leq 1/5$, this gives us the desired estimate. □

We would also like to know that convolutions are approximately Lipschitz.

Corollary 6.17 *Let $0 \leq \beta \leq 1/2$ and let $p, q, r, s \in \pi(H)$ be such that $\|p\|_1, \|q\|_1, \|r\|_1, \|s\|_1 \leq 1 + \beta$. Then*

$$d(p * r, q * s) \leq (1 + \beta)^2(d(p, q) + d(r, s)) + 8\beta^2.$$

Proof. We shall use the fact that $f * g_-(0) = \langle f, g\rangle$ for any two functions $f, g \in \pi(H)$. That implies that

$$\begin{aligned}\langle p * r, q * s\rangle &= \langle p * q_-, s * r_-\rangle \\ &\geq \langle p, q\rangle\langle r, s\rangle.\end{aligned}$$

Since $((1+\beta)^2 - \langle p, q\rangle)((1+\beta)^2 - \langle r, s\rangle) \geq 0$,

$$\langle p, q\rangle\langle r, s\rangle \geq (1+\beta)^2(\langle p, q\rangle + \langle r, s\rangle) - (1+\beta)^4.$$

The result follows after a quick calculation. □

The next lemma tells us that θ is close to a group homomorphism. Here and in what follows, we write δ_a for the function taking value 1 at a and 0 everywhere else.

Lemma 6.18 *For every $x_1, x_2 \in G$, $d(\theta(x_1 + x_2), \theta(x_1) * \theta(x_2)) \leq 1100\eta$.*

Proof. By definition,

$$\theta(x_1 + x_2) = \mathbb{E}_x \psi(x + x_1 + x_2) * \psi_-(x).$$

We would like to begin by "adding and subtracting $\psi(x + x_1)$". More precisely, we would like to approximate $\theta(x_1 + x_2)$ by

$$\mathbb{E}_x \psi(x + x_1 + x_2) * \psi_-(x + x_1) * \psi(x + x_1) * \psi_-(x).$$

Lemma 6.12 tells us that $d(\theta(x_1 + x_2), \theta(x_1 + x_2)) \leq 75\eta$. Since the distance is a bilinear function (in the sense that it commutes with expectations), this implies that

$$\mathbb{E}_{x,y} d(\psi(x + x_1 + x_2) * \psi_-(x), \psi(y + x_1 + x_2) * \psi_-(y)) \leq 75\eta.$$

We are trying to estimate the distance

$$d(\mathbb{E}_x \psi(x + x_1 + x_2) * \psi_-(x), \mathbb{E}_y \psi(y + x_1 + x_2) * \psi_-(y + x_1) * \psi(y + x_1) * \psi_-(y)). \tag{1}$$

By the bilinearity of d, it equals

$$\mathbb{E}_{x,y} d(\psi(x + x_1 + x_2) * \psi_-(x), \psi(y + x_1 + x_2) * \psi_-(y + x_1) * \psi(y + x_1) * \psi_-(y)),$$

which equals

$$\mathbb{E}_{x,y} d(\psi(x + x_1 + x_2) * \psi_-(x) * \psi_-(y + x_1 + x_2) * \psi(y), \psi_-(y + x_1) * \psi(y + x_1)).$$

By Lemma 6.16, this is at most the sum of

$$\mathbb{E}_{x,y} d(\psi(x + x_1 + x_2) * \psi_-(x) * \psi_-(y + x_1 + x_2) * \psi(y), \delta_0),$$

and

$$5\mathbb{E}_y d(\psi_-(y + x_1) * \psi(y + x_1), \delta_0) + 10\beta,$$

where $\beta = (1+\eta)^4 - 1$. The first of these terms is equal to

$$\mathbb{E}_{x,y} d(\psi(x + x_1 + x_2) * \psi_-(x), \psi(y + x_1 + x_2) * \psi_-(y)),$$

which, as we have already remarked, is at most 75η. We also have that

$$\mathbb{E}_y d(\psi_-(y + x_1) * \psi(y + x_1), \delta_0) = d(\mathbb{E}_y \psi_-(y + x_1) * \psi(y + x_1), \delta_0) = d(\theta(0), \delta_0),$$

which, by Corollary 6.14, is at most 77η. Therefore, the distance (1) is at most $75\eta + 385\eta + 10\beta < 700\eta$.

We shall now further approximate

$$\mathbb{E}_x \psi(x + x_1 + x_2) * \psi_-(x + x_1) * \psi(x + x_1) * \psi_-(x)$$

by

$$\theta(x_2) * \theta(x_1) = \mathbb{E}_{u,v} \psi(u + x_1 + x_2) * \psi_-(u + x_1) * \psi(v + x_1) * \psi_-(v).$$

By Corollary 6.17 with $\beta = 2\eta + \eta^2$, and using the bilinearity of d, the distance between them is at most

$$\begin{aligned}(1+\beta)^2 \mathbb{E}_{x,u} d(\psi(x + x_1 + x_2) * \psi_-(x + x_1), \psi(u + x_1 + x_2) * \psi_-(u + x_1)) \\ + (1+\beta)^2 \mathbb{E}_{x,v} d(\psi(x + x_1) * \psi_-(x), \psi(v + x_1) * \psi_-(v)) + 8\beta^2.\end{aligned}$$

But this equals $(1+\beta)^2 (d(\theta(x_2), \theta(x_2)) + d(\theta(x_1), \theta(x_1))) + 8\beta^2$, which, by Lemma 6.12, is at most $(1+\beta)^2 .150\eta + 8\beta^2$. Since $(1+\beta)^2 = (1+\eta)^4 \leq 2$, this is at most 400η. □

Now let us put together what we have proved so far.

Corollary 6.19 *There exists a group homomorphism $\gamma : G \to H$ such that $d(\theta(x), \delta_{\gamma(x)}) \leq 80\eta$ for every $x \in G$.*

Proof. Corollary 6.13 is the statement that there exists a *function* $\gamma : G \to H$ with the property stated. It remains to show that γ is a group homomorphism.

Let $\beta = (1+\eta)^2 - 1$. Then, by Corollary 6.17,

$$d(\theta(x_1) * \theta(x_2), \delta_{\gamma(x_1)} * \delta_{\gamma(x_2)}) \leq (1+\beta)^2 \big(d(\theta(x_1), \delta_{\gamma(x_1)}) + d(\theta(x_2), \delta_{\gamma(x_2)}) \big) + 8\beta^2,$$

which is at most $(1+\beta)^2 .160\eta + 8\beta^2 \leq 400\eta$.

By Corollary 6.13, Lemma 6.18, Lemma 6.16 (twice) and this calculation,

$$\begin{aligned} d(\delta_{\gamma(x_1+x_2)}, \delta_{\gamma(x_1)} * \delta_{\gamma(x_2)}) &\leq d(\theta(x_1 + x_2), \delta_{\gamma(x_1)} * \delta_{\gamma(x_2)}) + 5d(\theta(x_1 + x_2), \delta_{\gamma(x_1+x_2)}) + 10\beta \\ &\leq d(\theta(x_1 + x_2), \delta_{\gamma(x_1)} * \delta_{\gamma(x_2)}) + 400\eta + 10\beta \\ &\leq d(\theta(x_1 + x_2), \theta(x_1) * \theta(x_2)) + 5d(\theta(x_1) * \theta(x_2), \delta_{\gamma(x_1)} * \delta_{\gamma(x_2)}) + 400\eta + 20\beta \\ &\leq 1100\eta + 2000\eta + 400\eta + 20\beta \leq 4000\eta. \end{aligned}$$

But $\delta_{\gamma(x_1)} * \delta_{\gamma(x_2)} = \delta_{\gamma(x_1)+\gamma(x_2)}$, so if $\eta < 1/4000$, the only way this estimate can be true is if $\gamma(x_1 + x_2) = \gamma(x_1) + \gamma(x_2)$. □

It remains to relate the homomorphism γ to the original function $\phi : U \to H$. This we do with a standard averaging argument.

Theorem 6.20 *There exists an affine homomorphism $\alpha : G \to H$ such that $\phi(x) = \alpha(x)$ for all but $80\eta|U|$ elements $x \in U$.*

Proof. First, let us write an expression for $\theta(x)$. It is given by

$$\theta(x)(h) = \mathbb{E}_{x_1+x_2+x_3-x_4-x_5-x_6=x}\mu(x_1)\dots\mu(x_6)\mathbf{1}_{\phi(x_1)+\phi(x_2)+\phi(x_3)-\phi(x_4)-\phi(x_5)-\phi(x_6)=h}.$$

Therefore, we have just shown that

$$\mathbb{E}_{x_1+x_2+x_3-x_4-x_5-x_6=x}\mu(x_1)\dots\mu(x_6)\mathbf{1}_{\phi(x_1)+\phi(x_2)+\phi(x_3)-\phi(x_4)-\phi(x_5)-\phi(x_6)=\gamma(x)}$$

is at least $1 - 80\eta$.

Taking the expectation over x, we deduce that

$$\mathbb{E}_{x_1,\dots,x_6}\mu(x_1)\dots\mu(x_6)\mathbf{1}_{\phi(x_1)+\phi(x_2)+\phi(x_3)-\phi(x_4)-\phi(x_5)-\phi(x_6)=\gamma(x_1+x_2+x_3-x_4-x_5-x_6)}$$

is also at least $1 - 80\eta$. Therefore (since μ is a probability measure on G), there exist $x_2, \dots, x_6$ such that, writing $z = x_2 + x_3 - x_4 - x_5 - x_6$ and $h = \phi(x_2) + \phi(x_3) - \phi(x_4) - \phi(x_5) - \phi(x_6)$, we have

$$\mathbb{E}_{x_1}\mu(x_1)\mathbf{1}_{\phi(x_1)=\gamma(x_1+z)-h} \geq 1 - 80\eta.$$

That is, $\phi(x) = \gamma(x) + \gamma(z) - h$ for all but $80\eta|U|$ elements of U. Thus, we may take $\alpha(x) = \gamma(x) + \gamma(z) - h$. □

7 Freiman homomorphisms that are affine almost everywhere are affine everywhere

The final step in the argument is to prove that if U is a random set where every element is chosen independently with probability $Cn^{-2/3}(\log n)^{1/3}$, then with high probability every Freiman homomorphism defined on U that is affine on at least 99.99% of U is in fact affine on all of U.

Definition 7.1 *Let U be a subset of an Abelian group G. An* affine homomorphism *$\phi : U \to H$ from U to an Abelian group H is a function of the form $\phi(u) = a + \psi(u)$, where $\psi : U \to H$ is the restriction to U of a group homomorphism. The set U has the* $(1-\eta)$-extension property *if every Freiman homomorphism from U to an Abelian group H that coincides with an affine homomorphism ϕ on a subset $V \subset U$ of size at least $(1-\eta)|U|$ is in fact equal to ϕ.*

Definition 7.2 *Let U be a subset of a finite Abelian group G, let $W \subset U$ and let $V = U \setminus W$. Then W is* additively isolated in U *if $(V + V - V) \cap W = \emptyset$. Let $\eta > 0$. The set U is* $(1-\eta)$-additively connected *if no subset of U of size at most $\eta|U|$ is additively isolated in U.*

The definitions we have given are not standard, but they are the ones that we shall need for our argument. Let us prove a simple lemma that illustrates their usefulness.

Lemma 7.3 *Let G be a finite Abelian group and let $\eta > 0$. Then every $(1-\eta)$-additively connected subset U of G has the $(1-\eta)$-extension property.*

Proof. Let ϕ be a Freiman homomorphism defined on U, let V_0 be a subset of U of size at least $(1-\eta)|U|$ and suppose that the restriction of ϕ to V_0 is also the restriction to V_0 of an affine homomorphism $\psi : G \to H$. Let V be the set of all points $u \in U$ such that $\phi(u) = \psi(u)$.

If V is not the whole of U, then we have an easy contradiction, since $|V| \geq (1-\eta)|U|$, from which it follows by hypothesis that its complement W is not additively isolated in U. Therefore, we can find $v_1, v_2, v_3 \in V$ such that $v_1 + v_2 - v_3 = w \in W$. Since ϕ is a Freiman homomorphism, it follows that $\phi(w) = \phi(v_1) + \phi(v_2) - \phi(v_3)$, which equals $\psi(v_1) + \psi(v_2) - \psi(v_3)$, and since ψ is an affine homomorphism, this is equal to $\psi(w)$. □

The converse of this statement is not quite true, because it is also possible for $W + W - W$ to intersect V. We leave it as an exercise to find a counterexample.

Our strategy, then, is to prove that if G is a finite Abelian group of order n, then a binomial random subset $U \subset G$ where each element is chosen independently with probability $Cn^{-2/3}(\log n)^{1/3}$ is $(1-\eta)$-additively connected with high probability for some absolute constant $\eta > 0$. In what follows, we will work in a slightly different probabilistic model, proving that random subsets U of G with fixed size $C(n \log n)^{1/3}$ are $(1-\eta)$-additively connected with high probability for some appropriate $\eta > 0$. However, this easily implies that the same holds in the binomial model.

We begin with a preparatory lemma.

Lemma 7.4 *Let G be a finite Abelian group of order n, let K be a subset of G of size k and let A be a random subset of G of size m. Then, if $km \leq n$, $\mathbb{P}[|A+K| < km/4] \leq e^{-m/32}$, and if $n \leq km \leq 2n$, $\mathbb{P}[|A+K| < n/2] \leq e^{-m/800}$.*

Proof. Let x be any element of G. The probability that $x \in A + K$ is at least $1 - (1-k/n)^m \geq 1 - e^{-km/n}$. Therefore, the expectation of $|A+K|$ is at least $n(1-e^{-km/n})$. If we alter a single element of A, then the change to $|A+K|$ is at most k. It follows by Azuma's inequality that

$$\mathbb{P}[|A+K| < \mathbb{E}|A+K| - tk] \leq e^{-t^2/2m}.$$

Now $e^{-x} \leq 1 - x + x^2/2$ for every $x \geq 0$, so if $km \leq n$, then $n(1-e^{-km/n}) \geq n(km/n - k^2m^2/2n^2) \geq nkm/2n = km/2$. Therefore, $\mathbb{P}[|A+K| < km/4] \leq e^{-(m/4)^2/2m}$, which establishes the first bound. If $km \geq n$, then $n(1-e^{-km/n}) \geq n(1-1/e) > 3n/5$, so $\mathbb{P}[|A+K| < n/2] \leq e^{-(n/10k)^2/2m} \leq e^{-(m/20)^2/2m}$, which establishes the second. □

Now let X be the set $\{1, 2, \dots, t\}$ for a t to be chosen later (which will be of the form $C(n \log n)^{1/3}$). Let $k \leq \eta t$ for some $\eta > 0$ also to be chosen later (which will be an absolute constant). For each $B' \subset X$ of size k, let $A_1'(B'), A_2'(B')$ and $A_3'(B')$ be three sets of equal size that partition $X \setminus B'$, with the exception of at most two elements, with these sets chosen arbitrarily. Now let ϕ be a random function from X to G. It is easy to show that for sufficiently large n the probability that the restrictions of ϕ to the sets $B', A_1'(B'), A_2'(B')$ and $A_3'(B')$ are all injections is at least $1/2$. If we condition on this event, then the images B, A_1, A_2, A_3 of those four sets are independent random subsets of G of the appropriate cardinalities. If we condition further on ϕ being an injection (so now we ask for the images to be disjoint), which again is true with probability at least $1/2$, then their union U is a random set of size t and B is a random subset of U of size k.

For each $B' \subset X$, let $P(B')$ be the probability that $A_1 + A_2 - A_3$ is disjoint from B, given that the restrictions of ϕ to the sets $B', A_1'(B'), A_2'(B')$ and $A_3'(B')$ are all injections. If $\sum_{|B'| \leq \eta t} P(B') = p$,

then the probability that there exists $B' \subset X$ of size at most ηt such that $A_1 + A_2 - A_3$ is disjoint from B is at most p. If we now condition on ϕ being an injection, this probability goes up to at most $2p$. Therefore, with probability at least $1 - 2p$, no subset of U of size at most $\eta|U|$ is additively isolated in U. In other words, with probability at least $1 - 2p$, U is $(1 - \eta)$-additively connected.

We shall therefore concentrate our attention on estimating $P(B')$.

Lemma 7.5 *Let G be a finite Abelian group of order n, let B be a fixed subset of G of size k, and let A_1, A_2 and A_3 be random subsets of G of size $s = C(n \log n)^{1/3}$. Let $t = k + 3s$ and suppose that $k \leq t/10^5$. Then there exists an absolute constant $C > 0$ such that the probability that $A_1 + A_2 - A_3$ and B are disjoint is at most $(2n)^{-2}\binom{t}{k}^{-1}$.*

Proof. Observe first that $A_1 + A_2 - A_3$ and B are disjoint if and only if $A_1 + A_2 - B$ and A_3 are disjoint. Next, note that by Lemma 7.4 and the fact that $ks \leq n$ (we are assuming throughout that n is sufficiently large), we have that $|A_1 - B| \geq ks/4$ with probability at least $1 - e^{-s/32}$.

Let $K = A_1 - B$. The rest of the proof splits into two cases. If $|K|s \leq n$, then Lemma 7.4 implies that $|A_2 + K| \geq |K|s/4 \geq ks^2/16$ with probability at least $1 - e^{-s/32}$. If this event happens, then the probability that $A_1 + A_2 - B$ is disjoint from A_3 is at most $(1 - ks^2/16n)^s \leq \exp(-ks^3/16n) = n^{-C^3k/16}$. Since $k \leq t/10^5$, standard estimates for binomial coefficients give us, with room to spare, that $\binom{t}{k} \leq (et/k)^k \leq e^{t/4000} \leq e^{s/1000}$. Also, when $C = 4$, $n^{-C^3k/16}$ is much less than $\binom{n}{k}^{-1}$. Together, these estimates easily suffice to show that

$$2e^{-s/32} + n^{-C^3k/16} \leq (2n)^{-2}\binom{t}{k}^{-1}.$$

If $|K|s > n$, then let L be a subset of K such that $n \leq |L|s \leq 2n$. The second part of Lemma 7.4 implies that $\mathbb{P}[|A_2 + L| < n/2] \leq e^{-s/800}$. If $|A_2 + L| \geq n/2$, then the probability that $A_2 + L$ is disjoint from A_3 is at most 2^{-s}, which is much smaller than $\binom{t}{k}^{-1}$ when $k = t/10^5$, by the estimates in the previous paragraph. Moreover, by the same estimates, $e^{-s/800}$ is much less than $\binom{t}{k}^{-1}$, so we are again done in this case provided C is sufficiently large. □

Combining Lemma 7.5 with the preceding remarks, we obtain the main result of this section.

Lemma 7.6 *Let G be a finite Abelian group of order n and let U be a random subset of G of size $C(n \log n)^{1/3}$. Then there exists an absolute constant $C > 0$ such that the probability that U is $(1 - 10^{-5})$-additively connected is at least $1 - 1/n$.*

Since this was what we needed to complete the proof of our main theorem, we are now done.

8 Concluding remarks

Ultimately, one might hope to prove a more precise result still. Define the *Freiman dimension* of a subset A of an Abelian group to be one less than the dimension of the vector space of all Freiman homomorphisms from A to $\mathbb{R}$. For example, $\mathbb{Z}_N$ has Freiman dimension 0, since every Freiman homomorphism from $\mathbb{Z}_N$ to $\mathbb{R}$ is constant, and an arithmetic progression $P \subset \mathbb{Z}$ has Freiman dimension 1, since a Freiman homomorphism from P to $\mathbb{R}$ is determined by the values it takes at the first two

points. If A is a subset of $\mathbb{Z}_N$ such that every Freiman homomorphism from A to an Abelian group H extends to a Freiman homomorphism defined on all of $\mathbb{Z}_N$, then A has Freiman dimension 0, since if $H = \mathbb{R}$, then the extension, and therefore the original homomorphism, must be constant. Therefore, Theorem 1.2 implies that if a random subset A of $\mathbb{Z}_N$ is chosen with probability $C(\log N)^{1/3}N^{-2/3}$, then it has Freiman dimension 0. It would be interesting to understand how the Freiman dimension of the random set A decreases as the probability moves from $CN^{-2/3}$ to $C(\log N)^{1/3}N^{-2/3}$. For example, it might be the case that a hitting time result holds. That is, if we build up our random set one element at a time, it may be that the Freiman dimension drops to 0 at precisely the same moment when all elements are contained in an additive quadruple.

References

[1] N. Alon and J. H. Spencer, **The Probabilistic Method**, 3rd edition, John Wiley & Sons, Inc., Hoboken, NJ, 2008.

[2] J. Balogh, R. Morris and W. Samotij, Independent sets in hypergraphs, *J. Amer. Math. Soc.* **28** (2015), 669–709.

[3] D. Conlon, Combinatorial theorems relative to a random set, in Proceedings of the International Congress of Mathematicians 2014, Vol. 4, 303–328, Kyung Moon Sa, Seoul, 2014.

[4] D. Conlon and W. T. Gowers, Combinatorial theorems in sparse random sets, *Ann. of Math.* **184** (2016), 367–454.

[5] D. Conlon, W. T. Gowers, W. Samotij and M. Schacht, On the KŁR conjecture in random graphs, *Israel J. Math.* **203** (2014), 535–580.

[6] G. Fiz Pontiveros, Freiman homomorphisms of random subsets of $\mathbb{Z}_N$, *Combin. Probab. Comput.* **22** (2013), 592–611.

[7] E. Friedgut, V. Rödl and M. Schacht, Ramsey properties of random discrete structures, *Random Structures Algorithms* **37** (2010), 407–436.

[8] S. Janson and A. Ruciński, The infamous upper tail, *Random Structures Algorithms* **20** (2002), 317–342.

[9] V. Rödl and M. Schacht, Extremal results in random graphs, in Erdős Centennial, 535–583, Bolyai Soc. Math. Stud., Vol. 25, János Bolyai Math. Soc., Budapest, 2013.

[10] D. Saxton and A. Thomason, Hypergraph containers, *Invent. Math.* **201** (2015), 925–992.

[11] M. Schacht, Extremal results for random discrete structures, *Ann. of Math.* **184** (2016), 333–365.