

# The Inverse Theorem for the $U^3$ Gowers Uniformity Norm on Arbitrary Finite Abelian Groups: Fourier-analytic and Ergodic Approaches

Asgar Jamneshan\*      Terence Tao†

*Received 17 March 2022; Published 31 July 2023*

**Abstract:** We state and prove a quantitative inverse theorem for the Gowers uniformity norm  $U^3(G)$  on an arbitrary finite abelian group  $G$ ; the cases when  $G$  was of odd order or a vector space over  $\mathbf{F}_2$  had previously been established by Green and the second author and by Samorodnitsky respectively by Fourier-analytic methods, which we also employ here. We also prove a qualitative version of this inverse theorem using a structure theorem of Host–Kra type for ergodic  $\mathbb{Z}^\omega$ -actions of order 2 on probability spaces established recently by Shalom and the authors.

## 1 Introduction

In this paper we investigate the inverse theory for the third Gowers norm  $U^3(G)$  on arbitrary finite abelian groups  $G = (G, +)$ , with a particular focus on results that apply uniformly for *all* such groups  $G$  without additional hypotheses.

---

\*Supported by DFG-research fellowship JA 2512/3-1.

†Supported by a Simons Investigator grant, the James and Carol Collins Chair, the Mathematical Analysis & Application Research Fund Endowment, and by NSF grant DMS-1764034.

## 1.1 Gowers uniformity norms

We first recall the definition of the Gowers uniformity norms, first introduced in [10] to obtain a new proof of Szemerédi’s theorem [32] on arithmetic progressions. Throughout this paper we use the usual averaging notation

$$\mathbb{E}_{x \in A} f(x) := \frac{1}{|A|} \sum_{x \in A} f(x)$$

for a function  $f: A \rightarrow \mathbb{C}$  defined on a finite non-empty set  $A$  of some cardinality  $|A|$ .

**Definition 1.1** (Uniformity norms). Let  $G = (G, +)$  be a finite additive<sup>1</sup> group.

- (i) If  $d \geq 0$  and  $(f_\omega)_{\omega \in \{0,1\}^d}$  is a tuple of functions  $f_\omega: G \rightarrow \mathbb{C}$ , we define the *Gowers inner product*  $\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d(G)}$  by the formula

$$\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d(G)} := \mathbb{E}_{x, h_1, \dots, h_d \in G} \prod_{\omega \in \{0,1\}^d} \mathcal{C}^{|\omega|} f_\omega(x + \omega \cdot (h_1, \dots, h_d)) \quad (1)$$

where  $\omega = (\omega_1, \dots, \omega_d)$ ,  $|\omega| := \omega_1 + \dots + \omega_d$ ,  $\mathcal{C}: z \mapsto \bar{z}$  is the complex conjugation map, and we use the dot product notation

$$(n_1, \dots, n_d) \cdot (v_1, \dots, v_d) := n_1 v_1 + \dots + n_d v_d$$

whenever  $n_1, \dots, n_d$  are integers and  $v_1, \dots, v_d$  lie in an additive group.

- (ii) If  $d \geq 1$  and  $f: G \rightarrow \mathbb{C}$  is a function, the *Gowers uniformity norm*  $\|f\|_{U^d(G)}$  is defined by the formula

$$\|f\|_{U^d(G)} := \langle (f)_{\omega \in \{0,1\}^d} \rangle_{U^d(G)}^{1/2^d}.$$

**Example 1.2.** One has

$$\begin{aligned} \|f\|_{U^1(G)} &= |\mathbb{E}_{x, h \in G} f(x) \bar{f}(x+h)|^{1/2} \\ &= |\mathbb{E}_{x \in G} f(x)| \end{aligned} \quad (2)$$

and

$$\|f\|_{U^2(G)} = |\mathbb{E}_{x, h_1, h_2 \in G} f(x) \bar{f}(x+h_1) \bar{f}(x+h_2) f(x+h_1+h_2)|^{1/4}.$$

We also have the recursive formula

$$\|f\|_{U^{d+1}(G)} = \left( \mathbb{E}_{h \in G} \|\Delta_h^\times f\|_{U^d(G)}^{2^d} \right)^{1/2^{d+1}} \quad (3)$$

for any  $d \geq 1$ , where  $\Delta_h^\times f(x) := f(x) \overline{f(x-h)}$  denotes the “multiplicative derivative” of  $f$  in the direction  $h$ . In a similar spirit, we observe the recursive formula

$$\langle (f_\omega)_{\omega \in \{0,1\}^{d+1}} \rangle_{U^{d+1}(G)} = \mathbb{E}_{h^0, h^1 \in G} \langle (f_{\omega,0}(\cdot - h^0) \overline{f_{\omega,1}(\cdot - h^1)})_{\omega \in \{0,1\}^d} \rangle_{U^d(G)} \quad (4)$$

which follows after expanding out both sides and applying a change of variables.

<sup>1</sup>In this paper all additive groups are assumed to be abelian; multiplicative groups  $G = (G, \cdot)$  are not required to be abelian.

One can verify that  $\|\cdot\|_{U^d(G)}$  is a semi-norm for  $d = 1$  and a norm for  $d > 1$ ; see for instance [35, §11.1] for the basic properties of these norms. Our focus in this paper will primarily be on the  $U^3(G)$  norm.

If a function  $f: G \rightarrow \mathbb{C}$  is 1-bounded in the sense that  $|f(x)| \leq 1$  for all  $x \in G$ , then clearly we have

$$\|f\|_{U^{s+1}(G)} \leq 1$$

for any  $s \geq 1$ . For many applications to additive combinatorics, it is of interest to establish an *inverse theorem* that describes when a converse inequality

$$\|f\|_{U^{s+1}(G)} \geq \eta$$

holds for some  $0 < \eta \leq 1$ . Here one is primarily interested in the regime where  $\eta$  is fixed, but the order  $|G|$  of the group can be extremely large. The inverse theorem for  $U^{s+1}(G)$  is a foundational result in “degree  $s$  Fourier analysis”; see for instance [13], [35, §11] for further discussion.

In the case  $s = 1$  of linear Fourier analysis, the inverse theorem is easy to state and prove. Recall that the *Pontraguin dual group*  $\hat{G}$  of a finite additive group is the collection of all additive homomorphisms  $\xi: G \rightarrow \mathbb{R}/\mathbb{Z}$  from  $G$  to  $\mathbb{R}/\mathbb{Z}$ ; we write  $\xi \cdot x$  for  $\xi(x)$ , and refer to elements  $\xi$  of  $\hat{G}$  as *frequencies*.

**Theorem 1.3** (Inverse theorem for  $U^2(G)$ ). *Let  $G$  be a finite additive group, let  $0 < \eta \leq 1$ , and let  $f: G \rightarrow \mathbb{C}$  be a 1-bounded function such that  $\|f\|_{U^2(G)} \geq \eta$ . Then there exists  $\xi \in \hat{G}$  such that*

$$|\mathbb{E}_{x \in G} f(x) e(-\xi \cdot x)| \geq \eta^2$$

where  $e: \mathbb{R}/\mathbb{Z} \rightarrow \mathbb{C}$  is the standard character  $e(\theta) := e^{2\pi i \theta}$ .

*Proof.* See e.g., [14, Proposition 2.2]. □

## 1.2 $U^3(G)$ inverse theorem via quadratic Fourier analysis

The inverse theorems in the higher order case  $s > 1$  are harder to describe. In the  $s = 2$  case, one way to formulate an inverse theorem (following [14, Theorem 2.7]) is to replace the linear phases  $x \mapsto e(\xi \cdot x)$  in Theorem 1.3 by “locally quadratic phase functions”. To describe this formulation more precisely, we need to recall some standard definitions.

**Definition 1.4** (Bohr sets). (See e.g., [35, §4.4]) Let  $G$  be a finite additive group.

- (i) If  $S \subset \hat{G}$  is a set of frequencies, we define a “seminorm”  $\|x\|_S: G \rightarrow [0, 1/2]$  on  $G$  by

$$\|x\|_S := \sup_{\xi \in S} \|\xi \cdot x\|_{\mathbb{R}/\mathbb{Z}}$$

where  $\|\theta\|_{\mathbb{R}/\mathbb{Z}}$  denotes the distance to the nearest integer. The *Bohr set*  $\text{Bohr}(S, \rho) \subset G$  is then defined as

$$\text{Bohr}(S, \rho) := \{x \in G : \|x\|_S < \rho\}.$$

(ii) A Bohr set  $\text{Bohr}(S, \rho)$  is *regular* if one has

$$(1 - 100|S||\kappa|)|\text{Bohr}(S, \rho)| \leq |\text{Bohr}(S, (1 + \kappa)\rho)| \leq (1 + 100|S||\kappa|)|\text{Bohr}(S, \rho)|$$

$$\text{whenever } \kappa \in \left[-\frac{1}{100|S|}, \frac{1}{100|S|}\right].$$

**Definition 1.5** (Locally quadratic functions). Let  $U$  be a subset of an additive group  $G$ , and let  $H$  be another additive group.

- (i) If  $\phi: U \rightarrow H$  is a map from  $U$ , and  $h \in G$ , the derivative  $\partial_h \phi: U \cap (U + h) \rightarrow H$  is defined by  $\phi(x) := \phi(x) - \phi(x - h)$ . (For some values of  $h$ ,  $\partial_h \phi$  may be the empty function.)
- (ii) A map  $\phi: U \rightarrow H$  is said to be *locally linear* (resp. *locally quadratic*) on  $U$  if the second derivatives  $\partial_h \partial_k \phi$  (resp. the third derivatives  $\partial_h \partial_k \partial_l \phi$ ) vanish identically for all  $h, k, l \in G$ . (Here we adopt the convention that the empty function vanishes identically.)

The first main result of our paper is the following inverse theorem for the  $U^3(G)$  norm. Our conventions for asymptotic notation are laid out in Section 1.5 below.

**Theorem 1.6** (Inverse theorem for  $U^3(G)$ , locally quadratic version). *Let  $G$  be a finite additive group, let  $0 < \eta \leq 1/2$ , and let  $f: G \rightarrow \mathbb{C}$  be a 1-bounded function with  $\|f\|_{U^3(G)} \geq \eta$ . Then there exists a regular Bohr set  $B(S, \rho)$  with  $|S| \ll \eta^{-O(1)}$  and  $\exp(-\eta^{-O(1)}) \ll \rho \leq 1/100$ , a locally quadratic function  $\phi: \text{Bohr}(S, \rho) \rightarrow \mathbb{R}/\mathbb{Z}$ , and a function  $\xi: G \rightarrow \hat{G}$  such that*

$$\mathbb{E}_{x \in G} |\mathbb{E}_{h \in \text{Bohr}(S, \rho)} f(x + h) e(-\phi(h) - \xi(x) \cdot h)| \gg \eta^{O(1)}. \quad (5)$$

In the case where  $|G|$  is odd, this theorem was essentially established in [14, Theorem 2.7]; when instead  $G$  is of the form  $G = \mathbb{F}_2^n$  for some  $n$ , this result was essentially established in [28, Theorem 2.3]. In both cases the result was established by the methods of “quadratic Fourier analysis” [13], which in particular utilised tools from additive combinatorics such as the Balog–Szemerédi–Gowers theorem. In these arguments, the restrictions on  $G$  arise at a key step when one wishes to express a certain symmetric (locally) bilinear form in terms of a (locally) quadratic phase function, and this is achieved either by taking advantage of the ability to divide by two (in the case when  $|G|$  is odd) or by working with an explicit basis for  $G$  (in the case when  $G = \mathbb{F}_2^n$ ). In Section 2 we establish Theorem 1.6 in full generality. The proof uses the same quadratic Fourier analysis methods employed in [14], [28]; to achieve the key step for arbitrary  $G$ , we introduce a lifting lemma that converts a locally bilinear form on  $G$  to a globally bilinear form on a larger group  $G_S$ , and then works using an explicit basis in the latter group to describe the latter bilinear form in terms of a (globally) quadratic form, which then descends to give the locally quadratic function  $\phi$  in the above theorem.

### 1.3 Nilmanifold formulation

It is now generally accepted that the most natural formulation of  $U^{s+1}(G)$  inverse theorems for higher values of  $s$  involves nilmanifolds and related objects. In order to state such a formulation, we recall some key definitions.

**Definition 1.7** (Host-Kra spaces). (See e.g., [19, Appendix B].)

- (i) A *prefiltration* on a group  $G = (G, \cdot)$  is a collection of subgroups  $G_\bullet = (G_i)_{i=0}^\infty$  of  $G$  such that  $G_0 \geq G_1 \geq \dots$  and  $[G_i, G_j] \subset G_{i+j}$  for all  $i, j \geq 0$ . We say that the prefiltration has *degree at most  $s$*  if  $G_i$  is trivial for  $i > s$ . We refer to the pair  $(G, G_\bullet)$  as a *prefiltered group*. If  $G_0 = G_1 = G$ , we say that the prefiltration is a *filtration*, and that  $(G, G_\bullet)$  is a *filtered group*. Similarly if the group  $G$  is written additively instead of multiplicatively. Note that filtered groups of degree at most  $s$  are automatically nilpotent of nilpotency step at most  $s$ . If all the groups  $G_i$  in a filtered group are locally compact, we call  $G$  a *filtered locally compact group*. A subgroup  $\Gamma$  of a filtered locally compact group  $G$  is a *filtered lattice* if  $\Gamma$  is discrete and  $\Gamma_i := \Gamma \cap G_i$  is a cocompact subgroup of  $G_i$  for all  $i$ ; thus  $(\Gamma, (\Gamma_i)_{i=0}^\infty)$  is a filtered discrete group.

- (ii) If  $G = (G, +)$  is an additive group, we define the *abelian filtration* on  $G$  by setting  $G_0 = G_1 := G$  and  $G_i = \{0\}$  for  $i > 1$ .

- (iii) If  $G = (G, G_\bullet)$  is a prefiltered group, and  $k \geq 0$ , we define the *Host-Kra group*  $\mathrm{HK}^k(G) = \mathrm{HK}^k(G, G_\bullet)$  to be the subgroup of  $G^{\{0,1\}^k}$  generated by the elements

$$[g_{\omega_0}]_{\omega_0} := \left( g_{\omega_0}^{1_{\omega \geq \omega_0}} \right)_{\omega \in \{0,1\}^k} \quad (6)$$

for  $\omega_0$  in  $\{0,1\}^k$  and  $g_{\omega_0} \in G_{|\omega_0|}$ , where  $g_{\omega_0}^{1_{\omega \geq \omega_0}}$  is defined to equal  $g_{\omega_0}$  when  $\omega \geq \omega_0$  (in the product order on  $\{0,1\}^k$ ) and the identity 1 otherwise.

- (iv) If  $G = (G, G_\bullet)$  is a prefiltered group,  $k \geq 0$ , and  $\Gamma$  is a subgroup of  $G$ , we define the *Host-Kra space*  $\mathrm{HK}^k(G/\Gamma)$  to be the set

$$\mathrm{HK}^k(G/\Gamma) := \pi^{\{0,1\}^k}(\mathrm{HK}^k(G)) \subset (G/\Gamma)^{\{0,1\}^k},$$

where  $\pi: G \rightarrow G/\Gamma$  is the quotient map and  $\pi^{\{0,1\}^k}: G^{\{0,1\}^k} \rightarrow (G/\Gamma)^{\{0,1\}^k}$  is the map defined by pointwise evaluation of  $\pi$ , thus

$$\pi^{\{0,1\}^k}((h_\omega)_{\omega \in \{0,1\}^k}) := (\pi(h_\omega))_{\omega \in \{0,1\}^k}.$$

Note that we have a canonical identification

$$\mathrm{HK}^k(G/\Gamma) \equiv \mathrm{HK}^k(G)/\mathrm{HK}^k(\Gamma)$$

and also

$$\mathrm{HK}^k(\Gamma) = \mathrm{HK}^k(G) \cap \Gamma^{\{0,1\}^k}.$$

- (v) If  $G = (G, G_\bullet)$  and  $H = (H, H_\bullet)$  are prefiltered groups, and  $\Gamma, \Lambda$  are subgroups of  $G, H$  respectively, a *nilspace morphism*<sup>2</sup>  $\phi: G/\Gamma \rightarrow H/\Lambda$  is a function such that for every  $k \geq 0$ , the map  $\phi^{\{0,1\}^k}: (G/\Gamma)^{\{0,1\}^k} \rightarrow (H/\Lambda)^{\{0,1\}^k}$  defined by the pointwise application of  $\phi$ , thus

$$\phi^{\{0,1\}^k}((x_\omega)_{\omega \in \{0,1\}^k}) := (\phi(x_\omega))_{\omega \in \{0,1\}^k}$$

<sup>2</sup>For a systematic treatment of nilspaces, of which the filtered groups  $G$  and quotients  $H/\Lambda$  are special cases, see [4]. However, we will not require the general theory of nilspaces here, as we will work only with filtered groups and their quotients.

maps  $\mathrm{HK}^k(G/\Gamma)$  to  $\mathrm{HK}^k(H/\Lambda)$ . A nilspace morphism from a prefiltered group  $G$  to a quotient  $H/\Lambda$  will be called a *polynomial map*. (For instance, the projection  $\pi: H \rightarrow H/\Lambda$  from (iv) is a polynomial map.)

For future reference we note that if  $G = (G, +)$  has the abelian filtration and  $H = (H, \cdot)$  is a filtered group, then a map  $\phi: G \rightarrow H$  is a polynomial map if and only if one has the relations

$$\partial_{h_1} \dots \partial_{h_k} \phi(x) \in H_k \quad (7)$$

for all  $k \geq 0$  and  $x, h_1, \dots, h_k \in G$ , where the difference operators  $\partial_h$  are defined as

$$\partial_h \phi(x) := \phi(x) \phi(x - h)^{-1};$$

see for instance [19, Theorem B.3].

**Definition 1.8** (Nilmanifolds). [16] Let  $s \geq 1$ . A *filtered nilmanifold* of degree at most  $s$  is a quotient  $H/\Gamma$ , where  $H = (H, H_\bullet)$  is a filtered Lie group of degree at most  $s$ , and  $\Gamma = (\Gamma, \Gamma_\bullet)$  is a filtered lattice in  $H$ . A *prefiltered nilmanifold* is defined similarly, but with all filtrations replaced by prefiltrations.

It is not difficult to see that a filtered nilmanifold  $H/\Gamma$  has the structure of a smooth compact manifold (and the Host–Kra spaces  $\mathrm{HK}^k(H/\Gamma)$  are also smooth compact manifolds); see for instance [20, §2]. On the other hand, we stress that we do *not* require a nilmanifold to be connected. Nilsequences  $x \mapsto F(g(x))$  arising from prefiltered nilmanifolds, with  $g: G \rightarrow H/\Gamma$  a polynomial map and  $F: H/\Gamma \rightarrow \mathbb{C}$  a Lipschitz map can be converted to nilsequences on a filtered nilmanifold by observing that  $g(x) \in H_1 g(0)$  for all  $x \in G$ , so if one lets  $\{g(0)\} \in H$  be a bounded element of  $H$  that maps  $g(0)$  to the origin, then  $g(x) = \{g(0)\} \tilde{g}(x)$  for some polynomial map  $\tilde{g}: G \rightarrow H_1/(\Gamma \cap H_1)$  into the filtered nilmanifold  $H_1/(\Gamma \cap H_1)$ . One can then write  $F(g(x)) = \tilde{F}(\tilde{g}(x))$  where  $\tilde{F}: H_1/(\Gamma \cap H_1) \rightarrow \mathbb{C}$  is the Lipschitz map  $\tilde{F}(z) := F(\{g(0)\}z)$ .

For the  $U^3$  inverse theory, we will work with the following specific construction of a degree 2 (pre-)filtered nilmanifold, which one can view as a variant of the Heisenberg nilmanifold:

**Definition 1.9** (Specific nilmanifold construction). Let  $N \geq 0$  be a natural number. For any  $d \geq 0$ , we let  $\mathrm{Poly}_{\leq d}(\mathbb{R}^N \rightarrow \mathbb{R})$  denote the vector space of real valued polynomials on  $\mathbb{R}^N$  of degree at most  $d$ . These spaces have a translation action of  $\mathbb{R}^N$ , and in particular we can form the semidirect product

$$H(\mathbb{R}^N) := \mathbb{R}^N \ltimes \mathrm{Poly}_{\leq 2}(\mathbb{R}^N \rightarrow \mathbb{R})$$

which is the space of pairs  $(x, \phi)$  with  $x \in \mathbb{R}^N$  and  $\phi: \mathbb{R}^N \rightarrow \mathbb{R}$  a quadratic polynomial, with group law

$$(x, \phi)(y, \psi) := (x + y, T^y \phi + \psi)$$

where  $T^x$  denotes the translation operation  $T^x \psi(y) := \psi(y + x)$ . This is a nilpotent Lie group (of nilpotency class three), and can be given the degree 2 prefiltration

$$H(\mathbb{R}^N)_0 := \mathbb{R}^N \ltimes \mathrm{Poly}_{\leq 2}(\mathbb{R}^N \rightarrow \mathbb{R})$$

$$H(\mathbb{R}^N)_1 := \mathbb{R}^N \ltimes \mathrm{Poly}_{\leq 1}(\mathbb{R}^N \rightarrow \mathbb{R})$$

$$H(\mathbb{R}^N)_2 := 0 \ltimes \mathrm{Poly}_{\leq 0}(\mathbb{R}^N \rightarrow \mathbb{R})$$

$$H(\mathbb{R}^N)_i := 0 \text{ for all } i > 2.$$

Inside  $H(\mathbb{R}^N)$  we have a filtered lattice

$$H(\mathbb{Z}^N) := \mathbb{Z}^N \ltimes \text{Poly}_{\leq 2}(\mathbb{Z}^N \rightarrow \mathbb{Z})$$

defined and filtered in the analogous fashion (replacing all occurrences of  $\mathbb{R}$  with  $\mathbb{Z}$ ), using interpolation to identify  $\text{Poly}_{\leq 2}(\mathbb{Z}^N \rightarrow \mathbb{Z})$  with a subgroup of  $\text{Poly}_{\leq 2}(\mathbb{R}^N \rightarrow \mathbb{R})$ . Thus  $H(\mathbb{R}^N)/H(\mathbb{Z}^N)$  is a prefiltered nilmanifold (of dimension  $\frac{N^2+5N+2}{2}$ ). We can give it a (somewhat artificial) metric by identifying the Lie algebra of  $H(\mathbb{R}^N)$  (as a vector space) with  $\mathbb{R}^N \times \text{Poly}_{\leq 2}(\mathbb{R}^N \rightarrow \mathbb{R})$  and using the standard basis of  $\mathbb{R}^N$  and the monomial basis of  $\text{Poly}_{\leq 2}(\mathbb{R}^N \rightarrow \mathbb{R})$  as an orthonormal basis for this Lie algebra, which then defines a right-invariant Riemannian metric on  $H(\mathbb{R}^N)$ , and hence a Riemannian metric on  $H(\mathbb{R}^N)/H(\mathbb{Z}^N)$ .

In Section 3 we will use Theorem 1.6 (or more precisely a slight strengthening of this result that involves a globally quadratic function on a lift of  $G$ ) to establish the following inverse theorem, whose formulation is inspired by the recent work of Candela and Szegedy [6] (which we will discuss shortly):

**Theorem 1.10** (Inverse theorem for  $U^3(G)$ , nilmanifold version). *Let  $G$  be a finite additive group, let  $0 < \eta \leq 1/2$ , and let  $f: G \rightarrow \mathbb{C}$  be a 1-bounded function with  $\|f\|_{U^3(G)} \geq \eta$ . Then there exists a natural number  $N \ll \eta^{-O(1)}$ , a polynomial map  $g: G \rightarrow H(\mathbb{R}^N)/H(\mathbb{Z}^N)$ , and a Lipschitz function  $F: H(\mathbb{R}^N)/H(\mathbb{Z}^N) \rightarrow \mathbb{C}$  of norm  $O(\exp(\eta^{-O(1)}))$  such that*

$$|\mathbb{E}_{x \in G} f(x) \overline{F(g(x))}| \gg \exp(-\eta^{-O(1)}). \quad (8)$$

By the preceding discussion one can represent the prefiltered nilsequence  $F(g(x))$  as a filtered nilsequence. This theorem is somewhat weaker than Theorem 1.6 in that the correlation in (8) is only exponential in  $\eta$  instead of polynomial. With a bit more effort one could create a more complicated “averaged correlation” variant of (8) (resembling (5)), in which the lower bound in the correlation is polynomial in  $\eta$ , but we will not do so here.

## 1.4 $U^3(G)$ inverse theorem via ergodic theory

Based on Theorem 1.6, we now formulate the following conjecture<sup>3</sup> for general  $G$  and  $s$ :

**Conjecture 1.11** (Inverse theorem for  $U^{s+1}(G)$ ). *Let  $G$  be a finite additive group, let  $\eta > 0$ , let  $s \geq 1$ , and let  $f: G \rightarrow \mathbb{C}$  be a 1-bounded function with  $\|f\|_{U^{s+1}(G)} \geq \eta$ . Then there exists a degree  $s$  filtered nilmanifold  $H/\Gamma$ , drawn from some finite collection  $\mathcal{N}_{s,\eta}$  of such nilmanifolds that depends only on  $s, \eta$  but not on  $G$  (and each such nilmanifold in  $\mathcal{N}_{s,\eta}$  is endowed arbitrarily with a smooth Riemannian metric), a Lipschitz function  $F: H/\Gamma \rightarrow \mathbb{C}$  of Lipschitz norm  $O_{\eta,s}(1)$ , and a polynomial map  $g: G \rightarrow H/\Gamma$  such that*

$$|\mathbb{E}_{x \in G} f(x) \overline{F(g(x))}| \gg_{\eta,s} 1. \quad (9)$$

There is substantial evidence already towards this conjecture:

- (i) Theorem 1.10 clearly establishes the  $s = 2$  case of this conjecture, and Theorem 1.3 similarly establishes the  $s = 1$  case.

---

<sup>3</sup>One could similarly formulate a conjectural qualitative analogue of Theorem 1.6 to higher values of  $s$ , but we will not do so here, in part because one would have to decide precisely how to define the notion of a regular higher order Bohr set.

- (ii) Conjecture 1.11 is known<sup>4</sup> when  $G$  is a cyclic group [18], [19], [31], [30], [27], [6]. We remark that this special case has applications to number theory, for instance in obtaining asymptotics for linear equations in primes [15].
- (iii) Conjecture 1.11 is known when  $G$  is of the form  $\mathbb{F}_p^n$  for some finite field  $\mathbb{F}_p$  of fixed prime order [36], [2], [37], [31], [11], [12], [3], [5]. In this case one can take the nilmanifold  $H/\Gamma$  to be the unit circle  $\mathbb{R}/\mathbb{Z}$ , or even a cyclic group  $\frac{1}{p^j}\mathbb{Z}/\mathbb{Z}$  for some  $j$  depending only on  $p, s$ , equipped with a suitable degree  $s$  filtration.
- (iv) The inverse theorem in [6, Theorem 1.6] implies (as a special case) a weaker version of Conjecture 1.11 for all  $G, s$ , in which the nilmanifold  $H/\Gamma$  is replaced by the more general notion of a “CFR nilspace”. The main difficulty in passing from this theorem to the full strength version of Conjecture 1.11 is that the CFR nilspaces produced by that theorem need not be connected (or “toral”) without further hypotheses<sup>5</sup> on the group  $G$ .
- (v) In the converse direction, one can adapt standard arguments (e.g., [14, Proposition 12.6]) to show that if  $f: G \rightarrow \mathbb{C}$  is 1-bounded and  $|\mathbb{E}_{x \in G} f(x) \overline{F(g(x))}| \geq \delta$  for some  $\delta > 0$ , degree  $s$  filtered nilmanifold  $H/\Gamma$ , some Lipschitz function  $F: H/\Gamma \rightarrow \mathbb{C}$  of norm at most  $1/\delta$ , and some polynomial map  $g: G \rightarrow H/\Gamma$ , then  $\|f\|_{U^{s+1}(G)} \gg_{\delta, H/\Gamma, s} 1$ .

**Remark 1.12.** From item (v) we see that Conjecture 1.11 is an equivalence at the qualitative level. However, it may be still possible to “strengthen” this conjecture by placing additional properties on  $H/\Gamma$  and  $g$  in certain cases. For instance, the results in [16], [6] strongly suggest that the map  $g$  can be chosen to obey a suitable “balance” or “equidistribution” property. If the group  $G$  is the product of boundedly many cyclic factors, one should be able to place additional connectedness or “toral” hypotheses on the nilmanifold  $H/\Gamma$ , in analogy with the analysis in [22] or [6]. In the case  $G = \mathbb{F}_p^n$ , it was recently established in [5] that the nilmanifold  $H/\Gamma$  can be chosen to obey a “ $p$ -homogeneity” property. We will not pursue the question of what the optimal properties one can place on  $H/\Gamma$  or  $g$  further in this paper.

We will not establish this conjecture in full in this paper; however in the  $s = 2$  case we can establish it via an ergodic theory result (Theorem 1.14 below) that was recently established in [23] by Shalom and the authors. In order to state this result, we will need some definitions.

**Definition 1.13** ( $\mathbb{Z}^\omega$ -systems).

- (i) We define the additive group  $\mathbb{Z}^\omega$  to be the free abelian group by a countable sequence of generators  $e_1, e_2, \dots$ . This is an amenable group; it will be convenient to use the specific Følner sequence  $\Phi_1, \Phi_2, \dots \subset \mathbb{Z}^\omega$  defined by

$$\Phi_n := \{a_1 e_1 + \dots + a_{2^n} e_{2^n} : a_1, \dots, a_{2^n} \in \{0, \dots, n\}\}. \quad (10)$$

<sup>4</sup>In the earlier references an ostensibly weaker version of the conjecture is established in which one works with “non-periodic nilsequences” rather than “periodic nilsequences”, but the two formulations can be shown to be equivalent: see [26] for further discussion.

<sup>5</sup>For instance, as remarked previously, when  $G = \mathbb{F}_p^n$ , the CFR nilspace can be chosen to be isomorphic to the disconnected nilmanifold  $\frac{1}{p^j}\mathbb{Z}/\mathbb{Z}$ ; see for instance [37].

- (ii) A  $\mathbb{Z}^\omega$ -system  $(X, \mu, T)$  is a measure space  $(X, \mu) = (X, \mathcal{X}, \mu)$ , together with a group homomorphism  $T: h \mapsto T^h$  from  $\mathbb{Z}^\omega$  to the automorphism group  $\text{Aut}(X, \mu)$  of  $(X, \mu)$  (the group of invertible measure-preserving transformations of  $(X, \mu)$ ). Note that any sequence  $T^{e_1}, T^{e_2}, \dots$  of commuting automorphisms of  $(X, \mu)$  will generate such a system. For any  $f \in L^\infty(X, \mu)$  and  $h \in \mathbb{Z}^\omega$ , we write  $T^h f := f \circ T^{-h}$ . A  $\mathbb{Z}^\omega$ -system is *ergodic* if the invariant factor  $L^\infty(X, \mu)^T := \bigcap_{h \in \mathbb{Z}^\omega} \{f \in L^\infty(X, \mu) : T^h f = f\}$  consists only of the constants. If the  $\sigma$ -algebra  $\mathcal{X}$  is countably generated modulo null sets, we say that the  $\mathbb{Z}^\omega$ -system is *separable*.
- (iii) Let  $(X, \mu)$  be a  $\mathbb{Z}^\omega$ -system,  $d \geq 0$ , and  $(f_\omega)_{\omega \in \{0,1\}^d}$  be a tuple of functions  $f_\omega \in L^\infty(X, \mu)$ . For any natural numbers  $n_1, \dots, n_d$ , we define the local Host–Kra inner product

$$\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U_{n_1, \dots, n_d}^d(X)} := \mathbb{E}_{h_i^0, h_i^1 \in \Phi_{n_i} \forall i=1, \dots, d} \int_X \prod_{\omega \in \{0,1\}^d} \mathbb{C}^{|\omega|} T^{h_1^{\omega_1} + \dots + h_d^{\omega_d}} f_\omega d\mu \quad (11)$$

and the global Host–Kra inner product

$$\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d(X)} := \lim_{(n_1, \dots, n_d) \rightarrow \infty} \langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U_{n_1, \dots, n_d}^d(X)} \quad (12)$$

(for a proof of convergence of this limit and its independence of the choice of Følner sequence  $\Phi_n$ , see [38, Theorem 2.1, Remark 1.14].) For any  $f \in L^\infty(X, \mu)$  and  $d \geq 1$ , we define the *Host–Kra seminorm*

$$\|f\|_{U^d(X)} := \langle (f)_{\omega \in \{0,1\}^d} \rangle_{U^d(X)}^{1/2^d}.$$

- (iv) A  $\mathbb{Z}^\omega$ -morphism  $\pi: X \rightarrow Y$  from one  $\mathbb{Z}^\omega$ -system  $(X, \mu_X, T_X)$  to another  $(Y, \mu_Y, T_Y)$  is a measure-preserving transformation  $\pi: X \rightarrow Y$  such that  $T_Y^h \circ \pi = \pi \circ T_X^h$  for all  $h \in \mathbb{Z}^\omega$ .

One can show that  $\|\cdot\|_{U^d(X)}$  is indeed a seminorm on  $L^\infty(X)$ ; see for instance [38, §2], where it is also shown that these seminorms agree with the ones defined in [22] using cubic measures. From (11), (12) we observe the identity

$$\langle (f_\omega)_{\omega \in \{0,1\}^{d+1}} \rangle_{U^{d+1}(X)} = \lim_{n \rightarrow \infty} \mathbb{E}_{h^0, h^1 \in \Phi_n} \langle (T^{h^0} f_{\omega,0} \overline{T^{h^1} f_{\omega,1}})_{\omega \in \{0,1\}^d} \rangle_{U^d(X)} \quad (13)$$

for any  $s \geq 0$  and  $f_\omega \in L^\infty(X)$ ,  $\omega \in \{0,1\}^{d+1}$ ; compare with (4).

We can now give

**Theorem 1.14** (Conze–Lesigne inverse theorem for  $\mathbb{Z}^\omega$ -actions). *Let  $(X, \mu, T)$  be an ergodic separable  $\mathbb{Z}^\omega$ -system, and let  $f \in L^\infty(X, \mu)$  be such that  $\|f\|_{U^3(X)} > 0$ . Then there exists a filtered locally compact group  $H$  of degree at most 2, a filtered lattice  $\Gamma$  of  $H$ , a translation action  $T_{H/\Gamma}: \mathbb{Z}^\omega \rightarrow \text{Aut}(H/\Gamma)$  (equipping  $H/\Gamma$  with Haar measure) defined by*

$$T_{H/\Gamma}^h(x) := \phi(h)x \quad (14)$$

*for some group homomorphism  $\phi: \mathbb{Z}^\omega \rightarrow H$ , a  $\mathbb{Z}^\omega$ -morphism  $\Pi: X \rightarrow H/\Gamma$ , and a continuous function  $F: H/\Gamma \rightarrow \mathbb{C}$  such that*

$$\int_X f(x) \overline{F(\Pi(x))} d\mu \neq 0. \quad (15)$$

*Proof.* For the definition of any unexplained term in this argument we refer to [23]. The hypothesis  $\|f\|_{U^3(X)} > 0$  is equivalent to  $\mathbb{E}(f|Z^2(X)) \neq 0$ , where  $Z^2(X)$  is the second Host–Kra–Ziegler factor; see e.g., [2, Appendix A]. By [23, Lemma 1.2],  $Z^2(X)$  is a Conze–Lesigne system. By [23, Theorem 1.7],  $Z^2(X)$  is then the inverse limit of translational systems  $H/\Gamma$ , with  $H = (H, H_2)$  a filtered locally compact Polish group of degree at most 2,  $\Gamma$  a filtered lattice in  $H$ , and with the action given by (14). Thus we have  $\mathbb{E}(f|H/\Gamma) \neq 0$  for at least one such system. By Lusin’s theorem this implies the existence of a continuous function  $F: H/\Gamma \rightarrow \mathbb{C}$  such that

$$\int_{H/\Gamma} \mathbb{E}(f|H/\Gamma)(y) \overline{F(y)} d\mu_{H/\Gamma}(y) \neq 0 \quad (16)$$

where  $d\mu_{H/\Gamma}$  is the Haar probability measure on  $H/\Gamma$ . As  $H/\Gamma$  is a factor of  $X$ , there is an (abstract) factor map  $\Pi: X \rightarrow H/\Gamma$ , which can be upgraded to a concrete probability-preserving map by [23, Proposition A.2(i)]. Pulling back (16) by  $\Pi$  we obtain (15), and the claim follows.  $\square$

**Remark 1.15.** Note in the above theorem that  $H$  is merely locally compact rather than a Lie group. This is necessary in order for the theorem to hold; see the example presented after [29, Conjecture 2.14] (in the discussion of [29, Theorem 4.3]). In particular, we cannot necessarily take  $H/\Gamma$  to be a nilmanifold, although we shall later see that it is still an inverse limit of nilmanifolds (in the category of compact nilspaces, not the category of  $\mathbb{Z}^d$ -systems). On the other hand, the requirement that  $X$  be separable can be easily removed in practice; see [23] for further discussion.

In Sections 4–7 we establish the following connection between Theorem 1.14 and Conjecture 1.11:

**Theorem 1.16** (Conze–Lesigne inverse theorem implies Gowers inverse theorem). *Theorem 1.14 implies the  $s = 2$  case of Conjecture 1.11.*

Our arguments are an adaptation of those in [36], in which a random sampling method is used to obtain a correspondence principle that allows one to connect the combinatorial setting of Gowers uniformity to the ergodic setting of Host–Kra uniformity, and a rigidity result for polynomial maps is used to obtain the desired combinatorial conclusion. Similar strategies (omitting the random sampling step as one no longer seeks to introduce dynamics) have also been employed in the nilspace framework [31], [30], [6]. Our arguments are in fact valid for any  $s$ , in that the obvious generalization of Theorem 1.14 to a higher value of  $s$  (replacing the  $U^3$  seminorm with the  $U^{s+1}$  norm, and making  $H$  of degree  $s$  rather than degree 2) would imply the corresponding case of Conjecture 1.11. However, we suspect that these higher degree analogues of Theorem 1.14 may be false as stated, and may need some further modification in order to be able to prove Conjecture 1.11 for general  $s$ . We hope to report further on this question in future work.

## 1.5 Notation

We use  $X \ll Y$ ,  $Y \gg X$ ,  $X = O(Y)$  to denote the estimate  $|X| \leq CY$  for some constant  $C$ , and write  $X \asymp Y$  for  $X \ll Y \ll X$ . If we need the implied constant  $C$  to depend on additional parameters, we indicate this by subscripts, thus for instance  $X \gg_{\eta,s} 1$  denotes an estimate of the form  $X \geq c_{\eta,s}$  for some  $c_{\eta,s} > 0$  depending only on  $\eta, s$ .

In some sections we will switch to nonstandard asymptotic notation, in particular using  $o(1)$  to refer to an infinitesimal quantity; see Appendix A for details.

Given a measure space  $X = (X, \mathcal{X}, \mu)$ , we define the Lebesgue spaces  $L^p(X) = L^p(X, \mu) = L^p(X, \mathcal{X}, \mu)$  for  $1 \leq p \leq \infty$  in the usual fashion, in particular identifying any two functions that agree  $\mu$ -almost everywhere.

## 2 Quadratic Fourier analysis arguments

### 2.1 The structure of locally bilinear forms

Our proof of Theorem 1.6 uses the same quadratic Fourier analysis arguments that were also used in [14], [28]. The main new ingredient in addition to these arguments is a finer analysis of the structure of locally bilinear forms. We first give a definition of these forms.

**Definition 2.1.** Let  $U$  be a subset of an additive group  $G$ , and let  $H$  be another additive group. A *locally bilinear form*  $B: U \times U \rightarrow H$  is a function such that

$$B(x_1 + x_2, y) = B(x_1, y) + B(x_2, y)$$

whenever  $x_1, x_2, x_1 + x_2, y \in U$ , and similarly

$$B(x, y_1 + y_2) = B(x, y_1) + B(x, y_2)$$

whenever  $x, y_1, y_2, y_1 + y_2 \in U$ . If  $U = G$ , we refer to  $B$  as a *globally bilinear form*. If  $B(x, y) = B(y, x)$  for all  $x, y \in U$ , we call  $B$  *symmetric*.

Consider a Bohr set  $\text{Bohr}(S, \rho)$  of an additive group  $G$ . Clearly, any global bilinear form  $\tilde{B}: G \times G \rightarrow H$  restricts to a local bilinear form  $B: \text{Bohr}(S, \rho) \times \text{Bohr}(S, \rho) \rightarrow H$ . Unfortunately, the converse is not true; it is not difficult to find local bilinear forms  $B: \text{Bohr}(S, \rho) \times \text{Bohr}(S, \rho) \rightarrow H$  that do not extend to a global bilinear form (e.g., see Example 2.3). However, it turns out that global extensions (or more precisely, local lifts) do exist as long as one also lifts the group  $G$  to a larger group. The construction is as follows. If  $G$  is a finite additive group and  $S \subset \hat{G}$  is a set of frequencies, we let  $(\xi)_{\xi \in S}: G \rightarrow (\mathbb{R}/\mathbb{Z})^S$  denote the homomorphism

$$(\xi)_{\xi \in S}(x) := (\xi \cdot x)_{\xi \in S}.$$

Thus for instance  $\text{Bohr}(S, \rho)$  is the inverse image of the cube  $(-\rho, \rho)^S$  under this homomorphism  $(\xi)_{\xi \in S}$ , where we view  $(-\rho, \rho)$  as a subset of  $\mathbb{R}/\mathbb{Z}$ . We then define a new group  $G_S \leq G \times \mathbb{R}^S$  by the formula

$$G_S := \{(x, \theta) \in G \times \mathbb{R}^S : (\xi)_{\xi \in S}(x) = \theta \pmod{\mathbb{Z}^S}\},$$

that is to say  $G_S$  consists of tuples  $(x, (\theta_\xi)_{\xi \in S})$  where  $x \in G$  and for each  $\xi \in S$ ,  $\theta_\xi$  is a real number with  $\theta_\xi = \xi \cdot x \pmod{1}$ . Clearly,  $G_S$  is a lattice, and thus  $G_S$  is a finitely generated abelian group, and that one has a short exact sequence

$$0 \rightarrow \mathbb{Z}^S \rightarrow G_S \rightarrow G \rightarrow 0 \tag{17}$$

where the inclusion homomorphism from  $\mathbb{Z}^S$  to  $G_S$  is given by  $\theta \mapsto (0, \theta)$ , and the projection homomorphism from  $G_S$  to  $G$  is given by  $(x, \theta) \mapsto x$ . Thus  $G_S$  is an extension of  $G$  by the lattice  $\mathbb{Z}^S$ .

If we endow  $\mathbb{R}^S$  with the  $\ell^\infty$  norm  $\|\theta\|_{\mathbb{R}^S} := \sup_{\xi \in S} |\theta_\xi|$  (with the convention that the supremum is identically zero when  $S$  is empty) and define the cubes

$$B_{\mathbb{R}^S}(0, \rho) := \{\theta \in \mathbb{R}^S : \|\theta\|_{\mathbb{R}^S} < \rho\} = (-\rho, \rho)^S$$

then by chasing definitions, we see that we have an identification

$$G_S \cap (G \times B_{\mathbb{R}^S}(0, \rho)) \equiv \text{Bohr}(S, \rho) \quad (18)$$

for  $0 < \rho < 1/2$ , where the identification map is given by restricting the projection homomorphism  $\pi: G_S \rightarrow G$  to the set  $G_S \cap (G \times B_{\mathbb{R}^S}(0, \rho))$ .

The following useful lemma lets us (locally) lift a locally bilinear form to a globally bilinear one:

**Lemma 2.2** (Globalising a locally bilinear form). *Let  $G$  be a finite additive group, let  $S \subset \hat{G}$  be a set of frequencies, and let  $0 < \rho < 1/2$ . Suppose that we have a locally bilinear form  $B: \text{Bohr}(S, \rho) \times \text{Bohr}(S, \rho) \rightarrow H$  into some additive group  $H$ . Then there exists a globally bilinear form  $\tilde{B}: G_S \times G_S \rightarrow H$  such that*

$$\tilde{B}((x, \theta), (y, \sigma)) = B(x, y) \quad (19)$$

whenever  $(x, \theta), (y, \sigma) \in G_S$  are such that  $\|\theta\|_{\mathbb{R}^S}, \|\sigma\|_{\mathbb{R}^S} \leq \exp(-C|S|^C)\rho$  for some absolute constant  $C > 0$ . Furthermore, if  $B$  is symmetric, then one can choose  $\tilde{B}$  to also be symmetric.

**Example 2.3.** Let  $G = \mathbb{Z}/N\mathbb{Z}$  be a cyclic group, let  $\alpha$  be a real number, and let  $0 < \rho < 1/4$ . Let  $\phi: \mathbb{Z}/N\mathbb{Z} \rightarrow \mathbb{Z}$  be the lift  $\phi(x = [n]) = n \pmod{N}$ , i.e., a group homomorphism such that  $\pi \circ \phi = \text{id}_{\mathbb{Z}/N\mathbb{Z}}$ , where  $\pi: \mathbb{Z} \rightarrow \mathbb{Z}/N\mathbb{Z}$  is the canonical projection. Let  $\frac{1}{N}: \mathbb{Z} \rightarrow \mathbb{R}/\mathbb{Z}$  be the operation of division by  $N$ . Let  $S = \{\psi = \frac{1}{N} \circ \phi\}$  and  $\text{Bohr}(S, \rho) = \{x \in \mathbb{Z}/N\mathbb{Z} : \|\psi(x)\|_{\mathbb{R}/\mathbb{Z}} < \rho\}$ . The form  $B: \text{Bohr}(S, \rho) \times \text{Bohr}(S, \rho) \rightarrow \mathbb{R}$  defined by

$$B([n], [m]) := \alpha \phi(n) \phi(m)$$

is locally bilinear, but will not in general be extendable to a global bilinear form on  $\mathbb{Z}/N\mathbb{Z}$ . The problem will be that we may have carry-over effects on  $\mathbb{Z}/N\mathbb{Z}$  which would destroy linearity of  $B$ . However, on the infinite cyclic subgroup

$$G_S := \left\{ (x, \theta) \in \mathbb{Z}/N\mathbb{Z} \times \mathbb{R} : \frac{x}{N} = \theta \pmod{1} \right\} = \left\{ \left( n \pmod{N}, \frac{n}{N} \right) : n \in \mathbb{Z} \right\}$$

of  $\mathbb{Z}/N\mathbb{Z} \times \mathbb{R}$ , we can define the global bilinear form  $\tilde{B}: G_S \times G_S \rightarrow \mathbb{R}$  by

$$\tilde{B}\left(\left(n \pmod{N}, \frac{n}{N}\right), \left(m \pmod{N}, \frac{m}{N}\right)\right) := \alpha nm$$

and then it is clear that (19) holds for all  $(x, \theta), (y, \sigma) \in G_S$  with  $|\theta|, |\sigma| < \rho$ . By creating a second coordinate, we could honestly record the effects of carry-over.

*Proof.* We modify the proof of [35, Lemma 4.22]. We may assume that  $|S| \geq 1$ , since the claim is trivial for  $|S| = 0$ . Using the equivalence (18), we have a locally bilinear form

$$B' : (G_S \cap (G \times B_{\mathbb{R}^S}(0, \rho))) \times (G_S \cap (G \times B_{\mathbb{R}^S}(0, \rho))) \rightarrow H$$

such that

$$B'((x, \theta), (y, \sigma)) = B(x, y)$$

whenever  $(x, \theta), (y, \sigma) \in G_S$  are such that  $\|\theta\|_{\mathbb{R}^S}, \|\sigma\|_{\mathbb{R}^S} \leq \rho$ , with  $B'$  symmetric whenever  $B$  is. Our task is thus to locate a globally bilinear form  $\tilde{B} : G_S \times G_S \rightarrow H$  such that

$$\tilde{B}((x, \theta), (y, \sigma)) = B'((x, \theta), (y, \sigma)) \quad (20)$$

whenever  $\|\theta\|_{\mathbb{R}^S}, \|\sigma\|_{\mathbb{R}^S} \leq \exp(-C|S|^C)\rho$ , with  $\tilde{B}$  symmetric whenever  $B'$  is.

Let  $\Gamma := \{\theta : (x, \theta) \in G_S\}$ , then  $\Gamma$  is a subgroup of  $\mathbb{R}^S$  that contains  $\mathbb{Z}^S$  as a finite index subgroup, and is thus a lattice. Applying the discrete John's theorem [34, Lemma 1.6] (see also [1] for recent refinements), we can find linearly independent vectors  $w_\xi \in \Gamma$  and real numbers  $N_\xi > 0$  for  $\xi \in S$  such that

$$B_{\mathbb{R}^S} \left( O(|S|)^{-3|S|/2} t \right) \cap \Gamma \subset (-tN, tN) \cdot w \subset B_{\mathbb{R}^S}(t) \cap \Gamma$$

for any  $t > 0$ , where

$$(-tN, tN) \cdot w := \{n_1 w_1 + \dots + n_{|S|} w_{|S|} : n_1, \dots, n_{|S|} \in \mathbb{Z}; |n_i| < tN_i \text{ for all } i = 1, \dots, |S|\}. \quad (21)$$

In other words, the  $w_1, \dots, w_{|S|}$  generate  $\Gamma$ , and one has the inequalities

$$O(|S|)^{-3|S|/2} \sup_{1 \leq i \leq |S|} \frac{|n_i|}{N_i} \leq \|n_1 w_1 + \dots + n_{|S|} w_{|S|}\|_{\mathbb{R}^S} \leq \sup_{1 \leq i \leq |S|} \frac{|n_i|}{N_i}$$

for all integers  $n_1, \dots, n_{|S|}$ .

We may relabel so that  $N_1 \geq \dots \geq N_{|S|}$ . We can find  $v_1, \dots, v_{|S|} \in G_S$  such that each  $v_i$  is of the form  $v_i = (g_i, w_i)$  for some  $g_i \in G$ . Every element  $(x, \theta)$  of  $G_S$  can thus be uniquely represented in the form

$$(x, \theta) = (y, 0) + n_1 v_1 + \dots + n_{|S|} v_{|S|}$$

where  $n_1, \dots, n_{|S|}$  are integers and  $y \in K := \{y \in G : (y, 0) \in G_S\}$ . If we let  $j$  be the largest index for which  $N_j > 1$  (or  $j = 0$  if no such  $N_j$  exists), we conclude that the form

$$((y, n_1, \dots, n_j), (y', n'_1, \dots, n'_j)) \mapsto B'((y, 0) + n_1 v_1 + \dots + n_j v_j, (y', 0) + n'_1 v_1 + \dots + n'_j v_j)$$

is a locally bilinear form on

$$\{(y, n_1, \dots, n_j) \in K \times \mathbb{Z}^j : |n_i| < N_i \forall 1 \leq i \leq j\}.$$

In particular, this form can be expanded in coordinates as

$$B'((y, 0), (y', 0)) + \sum_{i=1}^j n_i B'(v_i, (y', 0)) + n'_i B'((y, 0), v'_i) + \sum_{1 \leq i, i' \leq j} n_i n'_{i'} B'(v_i, v'_{i'}).$$

This latter expression is well-defined in  $H$  for all  $y, y' \in K$  and  $v_1, \dots, v_{|S|}, v'_1, \dots, v'_{|S|} \in \mathbb{Z}$  (those  $v_i, v'_i$  with  $i > j$  play no role in this form). We can thus extend this form to a globally bilinear form on  $K \times \mathbb{Z}^{|S|}$ , which we can identify with  $G_S$ ; by inspection we see that this form is symmetric if  $B'$  is. From (21) we see that (20) holds whenever  $\|\theta\|_{\mathbb{R}^S}, \|\sigma\|_{\mathbb{R}^S} \leq (C|S|)^{-3|S|/2} \rho$  for some absolute constant  $C$ , and the claim follows.  $\square$

As our main application of this globalisation lemma we can express symmetric bilinear forms in terms of quadratic polynomials.

**Lemma 2.4** (Integrating a symmetric bilinear form). *Let  $G$  be a finitely generated additive group.*

- (i) *If  $B: G \times G \rightarrow \mathbb{R}/\mathbb{Z}$  is a symmetric globally bilinear form, then there exists a globally quadratic function  $\phi: G \rightarrow \mathbb{R}/\mathbb{Z}$  such that  $\phi(x+y) = \phi(x) + \phi(y) + B(x, y)$  for all  $x, y \in G$ .*
- (ii) *If  $G$  is finite,  $S \subset \hat{G}$ ,  $\rho > 0$ , and  $B: \text{Bohr}(S, \rho) \times \text{Bohr}(S, \rho) \rightarrow \mathbb{R}/\mathbb{Z}$  is a symmetric locally bilinear form, then there exists  $\exp(-|S|^{O(1)})\rho \ll \rho' \leq \rho$  and a locally quadratic function  $\phi: \text{Bohr}(S, \rho') \rightarrow \mathbb{R}/\mathbb{Z}$  such that  $\phi(x+y) = \phi(x) + \phi(y) + B(x, y)$  for all  $x, y \in \text{Bohr}(S, \rho'/2)$ . Furthermore, there is a globally quadratic lift  $\tilde{\phi}: G_S \rightarrow \mathbb{R}/\mathbb{Z}$  such that  $\tilde{\phi}(x, \theta) = \phi(x)$  whenever  $(x, \theta) \in G_S$  and  $\|\theta\|_{\mathbb{R}^S} < \rho'$ .*

It is possible that the finite generation or finiteness hypotheses on  $G$  here can be relaxed, but we will not need to do so here. If  $|G|$  were odd, then one proceed by setting  $\phi(x) := B(x, \frac{1}{2}x)$  since one now has the ability to divide by two in  $G$ ; this observation was implicit in the arguments in [14]. Similarly, if  $G = \mathbb{F}_2^N$ , then (for part (i) at least) one could proceed by inspecting the matrix coefficients of  $B$  and extracting the upper triangular half to define  $\phi$ ; this observation was implicit in the arguments in [28]. The fact that the above lemma holds for arbitrary finite additive groups  $G$  is the main reason why we are able to establish Theorem 1.6 without additional hypotheses on  $G$ .

*Proof.* We begin with (i). We first observe that if the claim is true for two finitely generated abelian groups  $G_1, G_2$  then it is true for the direct product  $G_1 \times G_2$ . Indeed, if  $B: (G_1 \times G_2) \times (G_1 \times G_2) \rightarrow \mathbb{R}/\mathbb{Z}$  is a symmetric bilinear form and the claim is already established for  $G_1, G_2$ , then one can find quadratic functions  $\phi_1: G_1 \rightarrow \mathbb{R}/\mathbb{Z}, \phi_2: G_2 \rightarrow \mathbb{R}/\mathbb{Z}$  such that

$$\phi_1(x_1 + y_1) = \phi_1(x_1) + \phi_1(y_1) + B((x_1, 0), (y_1, 0))$$

and

$$\phi_2(x_2 + y_2) = \phi_2(x_2) + \phi_2(y_2) + B((0, x_2), (0, y_2))$$

for all  $x_1, y_1 \in G_1$  and  $x_2, y_2 \in G_2$ . One then checks from direct calculation (using the symmetric bilinear nature of  $B$ ) that the function  $\phi: G_1 \times G_2 \rightarrow \mathbb{R}/\mathbb{Z}$  defined by

$$\phi(x_1, x_2) := \phi_1(x_1) + \phi_2(x_2) + B((x_1, 0), (0, x_2))$$

is a quadratic form on  $G_1 \times G_2$  that obeys the required property

$$\phi(x_1 + y_1, x_2 + y_2) = \phi(x_1, y_1) + \phi(x_2, y_2) + B((x_1, x_2), (y_1, y_2)).$$

Since every finitely generated abelian group is the direct product of finitely many copies of the integers  $\mathbb{Z}$  or cyclic groups  $\mathbb{Z}/N\mathbb{Z}$ , it thus suffices to verify the claim when  $G$  is either the integers or a cyclic group. If  $G = \mathbb{Z}$ , then the bilinear form  $B$  takes the form

$$B(x, y) = \alpha xy \pmod{1}$$

for some real number  $\alpha \in \mathbb{R}$  (cf. Example 2.3), and one can take  $\phi(x) := \frac{\alpha x^2}{2} \pmod{1}$ . If  $G = \mathbb{Z}/N\mathbb{Z}$ , then the bilinear form  $B$  similarly takes the form

$$B(x, y) = \frac{axy}{N} \pmod{1}$$

for some integer  $a$ , and then one can take

$$\phi(x \pmod{N}) := \frac{a\binom{x}{2}}{N} - \frac{ax\binom{N}{2}}{N^2} \pmod{1}$$

for all integers  $N$ , observing from the identity  $\binom{x+N}{2} = \binom{x}{2} + \binom{N}{2} + Nx$  that the right-hand side is  $N$ -periodic (regardless of whether  $N$  is even or odd). This establishes (i).

Now we establish (ii). By Lemma 2.2, we can find a symmetric globally bilinear form  $\tilde{B}: G_S \times G_S \rightarrow \mathbb{R}/\mathbb{Z}$  such that

$$\tilde{B}((x, \theta), (y, \sigma)) = B(x, y) \tag{22}$$

whenever  $(x, \theta), (y, \sigma) \in G_S$  are such that  $\|\theta\|_{\mathbb{R}^S}, \|\sigma\|_{\mathbb{R}^S} \leq \exp(-C|S|^C)\rho$  for some absolute constant  $C > 0$ . By part (i), we can then find a globally quadratic function  $\tilde{\phi}: G_S \rightarrow \mathbb{R}/\mathbb{Z}$  such that

$$\tilde{\phi}(x+y, \theta+\sigma) = \tilde{\phi}(x, \theta) + \tilde{\phi}(y, \sigma) + \tilde{B}((x, \theta), (y, \sigma)) \tag{23}$$

for any  $(x, \theta), (y, \sigma) \in G_S$ . Using (18), we can now define  $\phi: \text{Bohr}(S, \rho) \rightarrow \mathbb{R}/\mathbb{Z}$  so that

$$\tilde{\phi}(x, \theta) = \phi(x) \tag{24}$$

whenever  $(x, \theta) \in G_S$  are such that  $\|\theta\|_{\mathbb{R}^S} \leq \rho$ . In particular,  $\phi$  is locally quadratic on  $\text{Bohr}(S, \rho)$ . Combining (22), (23), (24), we obtain the claim (setting  $\rho' := \frac{1}{100} \exp(-C|S|^C)\rho$ , say).  $\square$

## 2.2 Proof of inverse theorem

We can now prove Theorem 1.6. For the rest of the section, we let  $G, f, \eta$  be as in the hypotheses of that theorem. In this section, we adopt the convention from [14] of using  $b()$  to denote various 1-bounded functions of the indicated variables, with the functions allowed to vary from occurrence to occurrence.

We now execute some standard arguments from [9], [14], [28] in which we obtain an increasing amount of control on various objects relating to  $f$ . We begin with attaching a frequency  $Mh$  depending in a locally linear fashion on  $h$  to all  $h$  in a Bohr set:

**Lemma 2.5.** *There exists a regular Bohr set  $B(S, \rho)$  with  $|S| \ll \eta^{-O(1)}$  and  $1 \ll \rho \leq 1/8$  and a locally linear function  $M: \text{Bohr}(S, 2\rho) \rightarrow \hat{G}$  such that*

$$\left| \mathbb{E}_{\substack{x \in G \\ h \in \text{Bohr}(S, \rho)}} b(h)b(x)f(x+h)e(-Mh \cdot x) \right| \gg \eta^{O(1)}.$$

*Proof.* This is a slight modification of [14, Proposition 9.3], with some factors of 2 removed to account for the fact that  $|G|$  is no longer required<sup>6</sup> to be odd. More specifically, one repeats the arguments in [14, Proposition 9.1], until one reaches the conclusion that  $2\Gamma'' - 2\Gamma''$  is a graph. In [14], the (implicit) assumption that  $|G| = |\hat{G}|$  was odd was used to write the portion of this graph over  $B_0$  as<sup>7</sup>  $\{(h, 2M(h)) : h \in B_0\}$  for some function  $M : B_0 \rightarrow \hat{G}$ . But for general values of  $|G|$ , we can instead write this portion as  $\{(h, M(h)) : h \in B_0\}$ . If one then continues the arguments of [14, Proposition 9.1] and [14, Proposition 9.3], replacing all occurrences of  $2M$  by  $M$ , one obtains the claim. We remark that these arguments rely heavily on tools from additive combinatorics, such as the Balog–Szemerédi–Gowers lemma, the Plünnecke inequalities, and the Bogolybov lemma.  $\square$

Next, we replace the locally bilinear form  $(h, x) \mapsto Mh \cdot x$  with a symmetric locally bilinear form:

**Lemma 2.6.** *There exists a regular Bohr set  $B(S, \rho)$  with  $|S| \ll \eta^{-O(1)}$  and  $\eta^{O(1)} \ll \rho \leq 1/8$  and a symmetric locally bilinear form  $B : \text{Bohr}(S, 2\rho) \times \text{Bohr}(S, 2\rho) \rightarrow \mathbb{R}/\mathbb{Z}$  such that*

$$\left| \mathbb{E}_{\substack{x \in G \\ h, k \in \text{Bohr}(S, \rho)}} b(x, h) b(x, k) f(x + h + k) e(-B(h, k)) \right| \gg \eta^{O(1)}.$$

*Proof.* Let  $M$  be the locally linear function from Lemma 2.5. We follow the “symmetry argument” used to establish [14, Lemma 9.4], but without the additional factor of 2 (and in particular deleting the arguments in the final paragraph devoted to removing this power of 2) to conclude (after adjusting the Bohr set  $\text{Bohr}(S, \rho)$  from Lemma 2.5 appropriately) that

$$\|Mx \cdot z - Mz \cdot x\|_{\mathbb{R}/\mathbb{Z}} \ll \eta^{-O(1)} \|x\|_S$$

whenever  $x, z \in \text{Bohr}(S, \rho)$ . In particular, if  $\rho' \sim C^{-1} \eta^C \rho$  for a sufficiently large absolute constant  $C$  with  $\text{Bohr}(S, \rho')$  regular (the existence of which is guaranteed by [14, Lemma 8.2]), one has

$$\|Mx \cdot z - Mz \cdot x\|_{\mathbb{R}/\mathbb{Z}} < \eta^{C/2} < \frac{1}{10}$$

(say) for  $x, z \in \text{Bohr}(S, \rho')$ . We can thus uniquely define a “midpoint”  $B(x, z) \in \mathbb{R}/\mathbb{Z}$  and a difference  $\Delta(x, z) \in [-\eta^C, \eta^C] \subset [-1/10, 1/10]$  for  $x, z \in \text{Bohr}(S, \rho')$  such that

$$Mx \cdot z = B(x, z) + \Delta(x, z); \quad Mz \cdot x = B(x, z) - \Delta(x, z) \quad (25)$$

for  $x, z \in \text{Bohr}(S, \rho')$ . One then easily verifies that  $B$  is a symmetric locally bilinear form on  $\text{Bohr}(S, \rho')$ . From Lemma 2.5 and a change of variables (using [14, Lemma 4.2(ii)]) we have (for  $C$  large enough)

$$\left| \mathbb{E}_{\substack{x \in G \\ h \in \text{Bohr}(S, \rho) \\ y, z \in \text{Bohr}(S, \rho')}} b(h + y) b(x + z) f(x + z + h + y) e(-M(h + y) \cdot (x + z)) \right| \gg \eta^{O(1)}$$

<sup>6</sup>This hypothesis was omitted by mistake in the hypotheses of that proposition, as well as in [14, Proposition 9.1].

<sup>7</sup>There is a typo at this step of the proof in [14], in that  $2H'' - 2H''$  should be  $2\Gamma'' - 2\Gamma''$ .

which after collecting terms can be rearranged as

$$\left| \mathbb{E}_{\substack{x \in G \\ h \in \text{Bohr}(S, \rho) \\ y, z \in \text{Bohr}(S, \rho')}} b(x, h, y) b(x, h, z) f(x + h + y + z) e(-My \cdot z) \right| \gg \eta^{O(1)};$$

using (25) to approximate  $My \cdot z$  by  $B(y, z)$  and using the triangle inequality, we obtain (for  $C$  large enough)

$$\left| \mathbb{E}_{\substack{x \in G \\ h \in \text{Bohr}(S, \rho) \\ y, z \in \text{Bohr}(S, \rho')}} b(x, h, y) b(x, h, z) f(x + h + y + z) e(-B(y, z)) \right| \gg \eta^{O(1)}.$$

Making the substitution  $x' = x + h$  and then pigeonholing in  $h$  we conclude that

$$\left| \mathbb{E}_{\substack{x' \in G \\ y, z \in \text{Bohr}(S, \rho')}} b(x', y) b(x', z) f(x' + y + z) e(-B(y, z)) \right| \gg \eta^{O(1)}$$

and the claim follows after a relabeling.  $\square$

Now we can prove Theorem 1.6. Let  $S, \rho$  be as in Lemma 2.6. By Lemma 2.4 (and [14, Lemma 8.2]), we can find  $\exp(-\eta^{-O(1)})\rho \ll \rho_1 < \rho/100$  with  $\text{Bohr}(S, \rho_1)$  regular, and a locally quadratic function  $\phi: \text{Bohr}(S, 2\rho_1) \rightarrow \mathbb{R}/\mathbb{Z}$  such that

$$\phi(x + y) = \phi(x) + \phi(y) + B(x, y) \quad (26)$$

for all  $x, y \in \text{Bohr}(S, \rho_1)$ , as well as a globally quadratic lift  $\tilde{\phi}: G_S \rightarrow \mathbb{R}/\mathbb{Z}$  such that  $\tilde{\phi}(x, \theta) = \phi(x)$  whenever  $(x, \theta) \in G_S$  and  $\|\theta\|_{\mathbb{R}^S} < 10\rho_1$ .

Let  $C$  be a sufficiently large constant, then by another application of [14, Lemma 8.2] we may find  $\rho_2 \asymp C^{-1}\eta^C\rho_1$  such that  $\text{Bohr}(S, \rho_2)$  is regular. From Lemma 2.6 and a change of variables (using [14, Lemma 4.2(ii)]) we have

$$\left| \mathbb{E}_{\substack{x \in G \\ h, k \in \text{Bohr}(S, \rho) \\ y \in \text{Bohr}(S, \rho_1) \\ z \in \text{Bohr}(S, \rho_2)}} b(x, h + y + z) b(x, k - y) f(x + h + k + z) e(-B(h + y + z, k - y)) \right| \gg \eta^{O(1)}.$$

From the locally bilinear nature of  $B$ , followed by (26), we may write

$$\begin{aligned} B(h + y + z, k - y) &= B(h, k) - B(h, y) + B(y + z, k) - B(y, y) - B(y, z) \\ &= B(h, k) - B(h, y) + B(y + z, k) - B(y, y) - \phi(y + z) + \phi(y) + \phi(z) \end{aligned}$$

so on collecting terms we obtain

$$\left| \mathbb{E}_{\substack{x \in G \\ h, k \in \text{Bohr}(S, \rho) \\ y \in \text{Bohr}(S, \rho_1) \\ z \in \text{Bohr}(S, \rho_2)}} b(x, h, k, y + z) b(x, h, k, y) f(x + h + k + z) e(-\phi(z)) \right| \gg \eta^{O(1)}.$$

Making the change of variables  $x' := x + h + k$  and pigeonholing in  $h, k$ , we conclude that

$$\left| \mathbb{E}_{\substack{x' \in G \\ y \in \text{Bohr}(S, \rho_1) \\ z \in \text{Bohr}(S, \rho_2)}} b(x', y + z) b(x', y) f(x' + z) e(-\phi(z)) \right| \gg \eta^{O(1)}.$$

Thus, we have

$$\left| \mathbb{E}_{\substack{y \in \text{Bohr}(S, \rho_1) \\ z \in \text{Bohr}(S, \rho_2)}} b(x', y + z) b(x', y) f(x' + z) e(-\phi(z)) \right| \gg \eta^{O(1)}$$

for  $\gg \eta^{O(1)}|G|$  values of  $x' \in G$ . Applying [14, Lemma 4.4], we conclude for each such  $x'$  that there exists  $\xi(x') \in \hat{G}$  such that

$$\left| \mathbb{E}_{z \in \text{Bohr}(S, \rho_1)} f(x' + z) e(-\phi(z) - \xi(x') \cdot z) \right| \gg \eta^{O(1)}$$

and Theorem 1.6 follows.

For future reference we observe that in fact we have demonstrated the following slight refinement of Theorem 1.6:

**Theorem 2.7.** *In the conclusion of Theorem 1.6, one can also ensure that there exists a globally quadratic lift  $\tilde{\phi}: G_S \rightarrow \mathbb{R}/\mathbb{Z}$  of  $\phi$  such that  $\tilde{\phi}(x, \theta) = \phi(x)$  whenever  $(x, \theta) \in G_S$  and  $\|\theta\|_{\mathbb{R}^S} < 10\rho$ .*

### 3 From locally quadratic functions to nilmanifolds

We now prove Theorem 1.10. The algebraic constructions here share some similarity with those in [26], in particular taking advantage of the semidirect product construction to convert the outer automorphisms arising from a translation action into inner automorphisms.

We turn to the details. By Theorem 1.6 and the pigeonhole principle, there exists a regular Bohr set  $\text{Bohr}(S, \rho)$  with  $|S| \ll \eta^{-O(1)}$  and  $\exp(-\eta^{-O(1)}) \ll \rho < 1/2$  and  $x_0 \in G$  such that

$$\left| \mathbb{E}_{h \in \text{Bohr}(S, \rho)} f(x_0 + h) e(-\phi(h)) \right| \gg \eta^{O(1)}$$

for some locally quadratic  $\phi$ , which can be lifted to a globally quadratic function  $\tilde{\phi}: G_S \rightarrow \mathbb{R}/\mathbb{Z}$  in the sense that  $\tilde{\phi}(x, \theta) = \phi(x)$  whenever  $(x, \theta) \in G_S$  and  $\|\theta\|_{\mathbb{R}^S} < 10\rho$ . By (18) we then have

$$\left| \sum_{(h, \theta) \in G_S} f(x_0 + h) 1_{\|\theta\|_{\mathbb{R}^S} < \rho} e(-\tilde{\phi}(h, \theta)) \right| \gg \exp(-\eta^{-O(1)})|G|.$$

Smoothing out the cutoff  $1_{\|\theta\|_{\mathbb{R}^S} < \rho}$  at a scale  $\exp(-C\eta^{-C})$  for some sufficiently large absolute constant  $C$ , and using the regularity of the Bohr set, we conclude that

$$\left| \sum_{(h, \theta) \in G_S} f(x_0 + h) \varphi(\theta) e(-\tilde{\phi}(h, \theta)) \right| \gg \exp(-\eta^{-O(1)})|G| \quad (27)$$

for some Lipschitz function  $\varphi : \mathbb{R}^S \rightarrow \mathbb{R}$  supported on (say) the unit ball with Lipschitz norm at most  $\exp(\eta^{-O(1)})$ .

For any  $(h, \theta) \in G_S$ , the map  $n \mapsto \tilde{\phi}(h, \theta + n)$  is a quadratic map from  $\mathbb{Z}^S$  to  $\mathbb{R}/\mathbb{Z}$ . By performing a Taylor expansion of this map and then lifting the coefficients from  $\mathbb{R}/\mathbb{Z}$  to  $\mathbb{R}$ , we may then find a quadratic map  $\Phi(h, \theta) : \mathbb{R}^S \rightarrow \mathbb{R}$  with the property that

$$\tilde{\phi}(h, \theta + n) = \Phi(h, \theta)(n) \pmod{1} \quad (28)$$

for all  $n \in \mathbb{Z}^S$ . This map is well defined up to an element of  $\text{Poly}_{\leq 2}(\mathbb{Z}^S \rightarrow \mathbb{Z})$ , and so by abuse of notation one can view  $\Phi(h, \theta)$  as lying in the quotient group  $\text{Poly}_{\leq 2}(\mathbb{R}^S \rightarrow \mathbb{R})/\text{Poly}_{\leq 2}(\mathbb{Z}^S \rightarrow \mathbb{Z})$ . From the globally quadratic nature of  $\tilde{\phi}$ , we see (using (7)) that the map  $\Phi : G_S \rightarrow \text{Poly}_{\leq 2}(\mathbb{R}^S \rightarrow \mathbb{R})/\text{Poly}_{\leq 2}(\mathbb{Z}^S \rightarrow \mathbb{Z})$  is a polynomial map, using the the abelian filtration on  $G_S$  and the degree 2 prefiltration

$$\begin{aligned} \text{Poly}_{\leq 2}(\mathbb{R}^S \rightarrow \mathbb{R})/\text{Poly}_{\leq 2}(\mathbb{Z}^S \rightarrow \mathbb{Z}) &\geq \text{Poly}_{\leq 1}(\mathbb{R}^S \rightarrow \mathbb{R})/\text{Poly}_{\leq 1}(\mathbb{Z}^S \rightarrow \mathbb{Z}) \\ &\geq \text{Poly}_{\leq 0}(\mathbb{R}^S \rightarrow \mathbb{R})/\text{Poly}_{\leq 0}(\mathbb{Z}^S \rightarrow \mathbb{Z}) \\ &\geq 0 \geq 0 \dots \end{aligned}$$

on  $\text{Poly}_{\leq 2}(\mathbb{R}^S \rightarrow \mathbb{R})/\text{Poly}_{\leq 2}(\mathbb{Z}^S \rightarrow \mathbb{Z})$ . We then define the map  $g : G_S \rightarrow H(\mathbb{R}^S)/H(\mathbb{Z}^S)$  by defining

$$g(h, \theta) := (\theta, \Phi(h, \theta)) \pmod{H(\mathbb{Z}^S)};$$

this is well-defined since  $\Phi(h, \theta)$  can be viewed as an element of  $\text{Poly}_{\leq 2}(\mathbb{R}^S \rightarrow \mathbb{R})$  modulo an element of  $\text{Poly}_{\leq 2}(\mathbb{Z}^S \rightarrow \mathbb{Z})$ , and  $0 \times \text{Poly}_{\leq 2}(\mathbb{Z}^S \rightarrow \mathbb{Z})$  is a subgroup of  $H(\mathbb{Z}^S)$ . For any  $n \in \mathbb{Z}^S$ , we have

$$\begin{aligned} g(h, \theta + n) &= (\theta + n, \Phi(h, \theta + n)) \pmod{H(\mathbb{Z}^S)} \\ &= (\theta, T^{-n}\Phi(h, \theta + n))(n, 0) \pmod{H(\mathbb{Z}^S)} \\ &= (\theta, T^{-n}\Phi(h, \theta + n)) \pmod{H(\mathbb{Z}^S)} \\ &= (\theta, \Phi(h, \theta)) \pmod{H(\mathbb{Z}^S)} \\ &= g(h, \theta) \end{aligned}$$

where we use the fact (from (28)) that  $T^{-n}\Phi(h, \theta + n)$  and  $\Phi(h, \theta)$  differ by an element of  $\text{Poly}_{\leq 2}(\mathbb{Z}^S \rightarrow \mathbb{Z})$ . Thus by the short exact sequence (17),  $g$  actually descends to a map  $g : G \rightarrow H(\mathbb{R}^S)/H(\mathbb{Z}^S)$ , where we define  $H(\mathbb{R}^S), H(\mathbb{Z}^S)$  as in Definition 1.9 but using  $S$  as the index set instead of  $\{1, \dots, N\}$ . If we define the map  $F : H(\mathbb{R}^S) \rightarrow \mathbb{C}$  by

$$F(\theta, \psi) := \sum_{n \in \mathbb{Z}^S} \varphi(\theta + n)e(\psi(n))$$

for  $\theta \in \mathbb{R}^S$  and  $\psi \in \text{Poly}_{\leq 2}(\mathbb{R}^S \rightarrow \mathbb{R})$ , one can check that

$$F((\theta, \psi)(m, \lambda)) = F(\theta, \psi)$$

for all  $m \in \mathbb{Z}^S$  and  $\lambda \in \text{Poly}_{\leq 2}(\mathbb{Z}^S \rightarrow \mathbb{Z})$ , so that  $F$  descends to a function on  $H(\mathbb{R}^S)/H(\mathbb{Z}^S)$ . From chasing the definitions we see that for every  $(h, \theta) \in G_S$  we have

$$\begin{aligned} F(g(h)) &= F(\theta_0, \Phi(h, \theta_0)) \\ &= \sum_{n \in \mathbb{Z}^S} \varphi(\theta_0 + n) e(\Phi(h, \theta_0)(n)) \\ &= \sum_{n \in \mathbb{Z}^S} \varphi(\theta_0 + n) e(\tilde{\phi}(h, \theta_0 + n)) \\ &= \sum_{\theta: (h, \theta) \in G_S} \varphi(\theta) e(\tilde{\phi}(h, \theta)) \end{aligned}$$

whenever  $(h, \theta_0) \in G_S$ . From (27) we conclude that

$$|\mathbb{E}_{h \in G} f(x_0 + h) \overline{F}(g(h))| \gg \exp(-\eta^{-O(1)}).$$

It is not difficult to verify that  $F$  is Lipschitz of norm at most  $\exp(\eta^{-O(1)})$ . Since  $H(\mathbb{R}^S)/H(\mathbb{Z}^S)$  is isomorphic to  $H(\mathbb{R}^N)/H(\mathbb{Z}^N)$  with  $N = |S|$ , the only remaining task is to show that  $g$  is a polynomial map from  $G$  to  $H(\mathbb{R}^S)/H(\mathbb{Z}^S)$ . As  $G_S$  is a finitely generated abelian group, we may write  $G_S$  (non-uniquely) as  $G_S = \Pi(\mathbb{Z}^D)$  for some lattice  $\mathbb{Z}^D$  and surjective homomorphism  $\Pi$ . Then the map  $\Phi \circ \Pi$  is a polynomial map from  $\mathbb{Z}^D$  (with the abelian filtration) to  $\text{Poly}_{\leq 2}(\mathbb{R}^S \rightarrow \mathbb{R})/\text{Poly}_{\leq 2}(\mathbb{Z}^S \rightarrow \mathbb{Z})$  and hence (by lifting Taylor coefficients) can be lifted to a polynomial map  $\Psi: \mathbb{Z}^D \rightarrow \text{Poly}_{\leq 2}(\mathbb{R}^S \rightarrow \mathbb{R})$ , and then  $(0, \Psi): \mathbb{Z}^D \rightarrow H(\mathbb{R}^S)$  is also a polynomial map. Meanwhile, the map

$$n \mapsto (\pi_\theta(\Pi(n)), 0)$$

is also easily verified to be a polynomial map, where  $\pi_\theta: G_S \rightarrow \mathbb{R}^S$  is the coordinate projection  $\pi_\theta(h, \theta) := \theta$ . Multiplying these two maps together (using the Lazard–Leibman theorem, see e.g., [19, Corollary B.4]), we conclude that the map  $\tilde{g}: \mathbb{Z}^D \rightarrow H(\mathbb{R}^S)$  defined by

$$\tilde{g}(n) := (\pi_\theta(\Pi(n)), \Psi(n))$$

is also a polynomial map. On the other hand, by comparing definitions we see that

$$\tilde{g}(n) = g(\Pi(n)) \mod H(\mathbb{Z}^S)$$

for all  $n \in \mathbb{Z}^D$ . Since  $\Pi: \mathbb{Z}^D \rightarrow G_S$  is surjective, we conclude that  $g: G_S \rightarrow H(\mathbb{R}^S)/H(\mathbb{Z}^S)$  is a polynomial map, and hence on descending to  $G$  we conclude that  $g: G \rightarrow H(\mathbb{R}^S)/H(\mathbb{Z}^S)$  as well, as desired. This concludes the proof of Theorem 1.10.

## 4 Nonstandard formulation

We now begin the proof of Theorem 1.16. As is common in several other approaches to the inverse conjecture (e.g., [19], [6]), the first step is to translate Conjecture 1.11 into the language of nonstandard analysis, in order to access tools such as Loeb measure. The nonstandard analysis formalism we require is summarized in Appendix A for convenience.

Let  $G = \prod_{n \rightarrow p} G_n$  be a nonstandard finite additive group, that is to say an ultraproduct of standard finite additive groups  $G_n$ . As discussed in [38, §5], the standard Gowers uniformity structures on  $G_n$  then induce analogous (external) structures on  $G$ ; in particular, there is an inner product  $\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d(G)} \in \mathbb{C}$  one can define for any  $d \geq 1$  and any (external) functions  $f_\omega \in L^\infty(G)$ , with the property that whenever the internal functions  $f_\omega := \lim_{n \rightarrow p} f_{\omega,n}$  are ultralimits of uniformly bounded standard functions  $f_{\omega,n}: G_n \rightarrow \mathbb{C}$ , one has

$$\langle (\text{st} f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d(G)} = \text{st} \langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d(G)} \quad (29)$$

where the *internal Gowers inner product*  $\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d(G)}$  of the  $f_\omega$  is defined to be the nonstandard complex number

$$\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d(G)} := \lim_{n \rightarrow p} \langle (f_{\omega,n})_{\omega \in \{0,1\}^d} \rangle_{U^d(G_n)}.$$

In particular, we can define an (external) seminorm  $\| \cdot \|_{U^d(G)}$  on  $L^\infty(G)$  by the formula

$$\|f\|_{U^d(G)} := \langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d(G)}^{1/2^d};$$

one can verify from Hölder's inequality that these norms and inner products are well-defined, with

$$|\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d(G)}| \leq \prod_{\omega \in \{0,1\}^d} \|f_\omega\|_{L^{2^d}(G)}$$

and

$$\|f\|_{U^d(G)} \leq \|f\|_{L^{2^d}(G)}.$$

We can now formulate

**Conjecture 4.1** (Inverse theorem for  $U^{s+1}(G)$ , nonstandard formulation). *Let  $G$  be a nonstandard finite additive group (with the Loeb probability space structure), let  $s \geq 1$  be standard, and let  $f \in L^\infty(G)$  be such that  $\|f\|_{U^{s+1}(G)} > 0$ . Then there exists a standard degree  $s$  filtered nilmanifold  $H/\Gamma$ , a standard continuous function  $F: H/\Gamma \rightarrow \mathbb{C}$ , and an internal polynomial map  $g: G \rightarrow {}^*(H/\Gamma)$  such that*

$$\int_G f \overline{F(\text{st} g)} d\mu_G \neq 0. \quad (30)$$

We claim that Conjecture 1.11 follows from Conjecture 4.1 for any given choice of  $s$ . It suffices to establish this theorem in the standard universe. Suppose for contradiction that Conjecture 1.11 failed in that universe for some standard  $\eta, s$ . As is well known (using for instance the theory of Mal'cev bases [25]), there are only countably many isomorphism classes of (standard) degree  $s$  filtered nilmanifolds. If we let  $H_n/\Lambda_n$  be an enumeration of (standard) representatives of these classes, then from the axiom of choice we see that we may obtain a sequence  $G_n$  of standard finite additive groups and standard 1-bounded functions  $f_n: G_n \rightarrow \mathbb{C}$  with the property that  $\|f_n\|_{U^{s+1}(G_n)} \geq \eta$ , and furthermore that there does *not* exist a degree  $s$  filtered nilmanifold  $H/\Gamma = H_{n'}/\Lambda_{n'}$  for some  $n' \leq n$ , some standard Lipschitz function  $F_n: H/\Gamma \rightarrow \mathbb{C}$  with norm at most  $n$ , and some standard polynomial map  $g_n: G_n \rightarrow H/\Gamma$  such that

$$|\mathbb{E}_{x \in G_n} f_n(x) \overline{F_n(g_n(x))}| \geq \frac{1}{n}.$$

Now let  $G := \prod_{n \rightarrow p} G_n$  be the ultraproduct of the  $G_n$ , and  $f := \text{stlim}_{n \rightarrow p} f_n$  be the standard part of the ultralimit of the  $f_n$ . Then we can endow  $G$  with the Loeb probability space structure, and  $f \in L^\infty(G)$ . From (29) we have

$$\|f\|_{U^{s+1}(G)} \geq \eta > 0.$$

We may then apply Conjecture 4.1 to conclude that there exists a standard degree  $s$  filtered nilmanifold  $H/\Gamma = H_{n_0}/\Lambda_{n_0}$ , a standard continuous function  $F: H/\Gamma \rightarrow \mathbb{C}$ , and an internal polynomial map  $g: G \rightarrow {}^*(H/\Gamma)$  such that (30) holds. By the Stone–Weierstrass theorem,  $F$  is the uniform limit of Lipschitz functions, so we may assume without loss of generality that  $F$  is Lipschitz. Observe that  $F(\text{st}g) = \text{st}F(g)$ . Writing  $g = \lim_{n \rightarrow p} g_n$  as the ultralimit of standard polynomial maps  $g_n: G_n \rightarrow H/\Gamma$ , we conclude that

$$\overline{fF(\text{st}g)} = \text{st} \lim_{n \rightarrow p} \overline{f_n F(g_n)}$$

and hence by (43)

$$\text{st} \lim_{n \rightarrow p} \mathbb{E}_{x \in G_n} f_n(x) \overline{F(g_n(x))} \neq 0.$$

In particular, there is a  $p$ -large set of  $n$  such that

$$|\mathbb{E}_{x \in G_n} f_n(x) \overline{F(g_n(x))}| \geq \frac{1}{n}.$$

But this contradicts the construction of the  $f_n$  for  $n$  sufficiently large, giving the claim.

**Remark 4.2.** It is not difficult to also conversely establish that Conjecture 4.1 follows from Conjecture 1.11 for a given choice of  $s$ ; we leave the details to the interested reader.

To establish Theorem 1.16, it thus suffices to show that Theorem 1.14 implies the  $s = 2$  case of Conjecture 4.1. This is the purpose of the next three sections of the paper.

## 5 Correspondence principle

In this section we create a separable  $\mathbb{Z}^\omega$ -system associated to the data of Conjecture 4.1, whose dynamics are related to the additive combinatorial structure of the internal function  $f$  appearing in that conjecture, in the spirit of the Furstenberg correspondence principle [7]. To create this correspondence we follow [36] and rely on random sampling methods, though we will use the language of nonstandard analysis in order to simplify some of the “epsilon management” present in the arguments of [36].

More precisely, we will establish the following claim.

**Proposition 5.1** (Correspondence principle). *Let  $G$  be a nonstandard finite additive group, and let  $\mathcal{F}$  be an at most countable subset of  $L^\infty(G)$ . Then there exists an ergodic separable  $\mathbb{Z}^\omega$ -system  $(X, \mathcal{X}, \mu_X, T)$ , where  $(X, \mathcal{X}, \mu_X)$  is a factor of  $(G, \mathcal{L}_G, \mu_G)$  (thus  $X = G$  as a set,  $\mathcal{X}$  is a subalgebra of  $\mathcal{L}_G$ , and  $\mu_X$  is the restriction of  $\mu_G$  to  $\mathcal{X}$ , and in particular  $L^\infty(X) \subset L^\infty(G)$ ) with the following properties:*

- (i) (Equivalence of Host–Kra and Gowers inner products) *The Gowers inner products on  $L^\infty(G)$  restrict to the Host–Kra inner products on  $L^\infty(X)$ . In other words, for any  $d \geq 0$  and any tuple  $(f_\omega)_{\omega \in \{0,1\}^d}$  of functions  $f_\omega \in L^\infty(X) \subset L^\infty(G)$  one has the identity*

$$\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d(X)} = \langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d(G)}. \quad (31)$$

(ii) ( $\mathcal{F}$  is modeled) We have  $\mathcal{F} \subset L^\infty(X)$ . In particular, from (31) one has  $\|f\|_{U^d(X)} = \|f\|_{U^d(G)}$  for all  $d \geq 1$  and  $f \in \mathcal{F}$ .

**Remark 5.2.** The arguments in [36, §3,4], when translated to the nonstandard setting, provide an analogue of this proposition in which  $G$  is now a internally finite dimensional vector space over a standard finite field  $\mathbb{F}_p$  of prime order, and the group  $\mathbb{Z}^\omega$  is replaced by  $\mathbb{F}_p^\omega$ . The proposition may also be compared with [6, Proposition 3.12], in which  $G$  is replaced by an ultraproduct of CFR coset nilspaces and  $X$  is similarly replaced by a nilspace (with no dynamics). A closely related correspondence principle was also established by Towsner [39].

We now prove this proposition. By splitting the elements of  $\mathcal{F}$  into real and imaginary parts we may assume without loss of generality that all the elements of  $\mathcal{F}$  are real-valued.

The idea is as follows. The shift maps  $T^h: G \rightarrow G$  will be defined on the entire group  $G$  by the formula

$$T^h(x) := x + \sum_{j=1}^{\infty} h_j \mathbf{g}_j \quad (32)$$

whenever  $h = \sum_{j=1}^{\infty} h_j e_j$  and  $x \in G$ , for some suitable shifts  $\mathbf{g}_j \in G$ ; note that these sums contain only finitely many non-zero terms. In order to be able to verify the identity (31), as well as the ergodicity of the system  $(X, \mathcal{X}, \mu_X, T)$ , it will be necessary that these shifts  $\mathbf{g}_j$  are “generic” in a suitable sense relative to the function  $f$  and its shifts. To locate such generic shifts we shall use a probabilistic method, generating the shifts  $\mathbf{g}_j$  randomly (i.e., as measurable functions of an (external) probability space  $\Omega$ ) and verifying that these shifts obey the required properties almost surely. This is analogous to how one can generate a uniquely ergodic system on the circle  $\mathbb{R}/\mathbb{Z}$  by choosing a shift  $\alpha$  uniformly at random from  $\mathbb{R}/\mathbb{Z}$  and then considering the shift map  $x \mapsto x + \alpha$ , noting that the unique ergodicity property will hold for the resulting system almost surely (since  $\alpha$  will almost surely be irrational).

We begin by setting up the probability space  $\Omega$ . Write  $G = \prod_{n \rightarrow p} G_n$  as the ultraproduct of finite additive groups  $G_n$ . By the Carathéodory extension theorem, for each  $n$  the (standard) product space  $G_n^{\mathbb{N}}$  of sequences  $(g_{j,n})_{j \in \mathbb{N}}$  can be given the structure of a (standard) probability space (the product of the probability space structures on  $G_n$  arising from the discrete  $\sigma$ -algebra and uniform probability measure). One can view these standard probability spaces as modeling a standard sequence of independent uniformly distributed random elements  $g_{j,n}$  of  $G_n$ . We then let  $\Omega := \prod_{n \rightarrow p} (G_n^{\mathbb{N}})$  be the ultraproduct of these spaces; by the Loeb measure construction (see Appendix A), we can view  $\Omega$  as an (external) probability space. We caution that we have the inequality

$$\Omega = \prod_{n \rightarrow p} (G_n^{\mathbb{N}}) \neq \left( \prod_{n \rightarrow p} G_n \right)^{\mathbb{N}} = G^{\mathbb{N}}$$

in general; however there is certainly a measure-preserving map  $\pi: \Omega \rightarrow G^{\mathbb{N}}$  (giving  $G^{\mathbb{N}}$  the external probability space structure coming from taking the product of the probability spaces  $G$ ) defined by

$$\pi \left( \lim_{n \rightarrow p} (g_{j,n})_{j \in \mathbb{N}} \right) := \left( \lim_{n \rightarrow p} g_{j,n} \right)_{j \in \mathbb{N}}$$

that one can easily verify to be a well-defined measure-preserving map by chasing all the definitions. If one writes  $\pi = (\mathbf{g}_j)_{j \in \mathbb{N}}$ , one can then view the  $\mathbf{g}_j: \Omega \rightarrow G$  as independent random variables taking values in  $G$ . By construction, we see that if  $J$  is a standard natural number and  $F_n: G_n^J \rightarrow \mathbb{C}$  are a uniformly bounded sequence of standard functions, with limit  $F: G^J \rightarrow \mathbb{C}$  defined by  $F := \text{st} \lim_{n \rightarrow p} F_n$ , then

$$\mathbb{E}F(\mathbf{g}_1, \dots, \mathbf{g}_J) = \text{st} \lim_{n \rightarrow p} \mathbb{E}_{g_{1,n}, \dots, g_{J,n} \in G_n} F_n(g_{1,n}, \dots, g_{J,n}) \quad (33)$$

where on the left-hand side the symbol  $\mathbb{E}$  denotes expectation with respect to the probability space  $\Omega$ .

**Remark 5.3.** We caution that the limit function  $F: G^J \rightarrow \mathbb{C}$  in (33) is *not* necessarily measurable with respect to the product  $\sigma$ -algebra  $\mathcal{L}_G^J$  of the Loeb  $\sigma$ -algebra  $\mathcal{L}_G$  of  $G$ ; instead, it is merely measurable with respect to the larger  $\sigma$ -algebra  $\mathcal{L}_{G^J}$  formed by applying the Loeb  $\sigma$ -algebra construction directly to  $G^J = \prod_{n \rightarrow p} (G_n^J)$ . To put it another way, the Loeb construction does *not* commute with Cartesian products; see also Appendix A for further discussion. However, this lack of measurability will not be a problem in practice since the identity (33) holds in all cases of interest. Roughly speaking, this identity tells us that the random variables  $\mathbf{g}_j$  are in fact “better” than merely independent copies of a single random variable  $\mathbf{g}$  on  $G$ , because they can measure events in a larger  $\sigma$ -algebra  $\mathcal{L}_{G^J}$  of  $G^J$  than just the product  $\sigma$ -algebra of  $J$  copies of  $\mathcal{L}_G$ .

We now use these random variables  $\mathbf{g}_j$  to define a (random)  $\mathbb{Z}^\omega$ -system  $(X, \mu, T) = (X, \mathcal{X}, \mu_X, T)$  as follows. We define a shift  $T^h: G \rightarrow G$  for each  $h \in \mathbb{Z}^\omega$  by the formula (32)

$$T^h(x) := x + \sum_{j=1}^{\infty} h_j \mathbf{g}_j$$

whenever  $h = \sum_{j=1}^{\infty} h_j e_j$  and  $x \in G$ ; note that these sums contain only finitely many non-zero terms. That is to say, each generator  $T^{e_j}$  is a shift in the random direction  $\mathbf{g}_j$ . As usual we then define  $T^h f := f \circ T^{-h}$  for any  $f \in L^\infty(G)$ . It is then easy to verify that  $(G, \mathcal{L}_G, \mu_G, T)$  is a  $\mathbb{Z}^\omega$ -system, but it will not be separable in general. We therefore restrict attention to the separable factor  $(X, \mathcal{X}, \mu_X, T)$  of  $(G, \mathcal{L}_G, \mu_G, T)$  generated by  $\mathcal{F}$ ; thus  $\mathcal{X}$  is the (complete)  $\sigma$ -algebra generated by the  $T^h f$  for  $f \in \mathcal{F}$  and  $h \in \mathbb{Z}^\omega$ , and  $\mu_X$  is the restriction of  $\mu_G$  to  $\mathcal{X}$ .

From construction it is clear that  $(X, \mathcal{X}, \mu_X, T)$  is a separable  $\mathbb{Z}^\omega$ -system (since there are only countably many functions of the form  $T^h f$  up to  $\mu_X$ -almost everywhere equivalence), and property (ii) of Theorem 5.1 is obeyed. It remains to (almost surely) establish ergodicity, as well as property (i).

We begin with (i). Fix  $d \geq 0$ . By homogeneity we may restrict the  $f_\omega$  to be 1-bounded. From Hölder’s inequality one easily checks that

$$\langle (f_\omega)_{\omega \in \{0,1\}^d} \rangle_{U^d(X)} \leq \prod_{\omega \in \{0,1\}^d} \|f_\omega\|_{L^{2^d}(X)}$$

for all  $f_\omega \in L^\infty(X)$  (cf., [38, §5.1]). Thus, the left-hand side of (31) is jointly continuous in the  $f_\omega$  in the  $L^{2^d}(X)$  topology. A similar argument shows that the right-hand side of (31) is also continuous in this topology. Thus we may restrict the  $f_\omega$  to any  $L^{2^d}(X)$ -dense subset of the unit ball of  $L^\infty(X)$ . By the Stone–Weierstrass theorem, an example of such a set is given by polynomial combinations of finitely

many of the shifts  $T^h f$  with rational coefficients with  $f \in \mathcal{F}$ . This set is countable (modulo  $\mu_X$ -almost everywhere equivalence), so to show that (i) holds almost surely, it suffices to verify (i) for fixed choices of  $f_\omega$ , each of which is a polynomial combination of finitely many of the  $T^h f$  with rational coefficients. In particular, each  $f_\omega$  takes the form

$$f_\omega(x) = \text{st}F_\omega(x, \mathbf{g}_1, \dots, \mathbf{g}_J)$$

for some standard  $J$  and some (deterministic) internal 1-bounded function  $F_\omega: G \times G^J \rightarrow {}^*\mathbb{C}$ . It thus suffices to establish the identity

$$\langle (\text{st}F_\omega(\cdot, \mathbf{g}_1, \dots, \mathbf{g}_J))_{\omega \in \{0,1\}^d} \rangle_{U^d(X)} = \langle (\text{st}F_\omega(\cdot, \mathbf{g}_1, \dots, \mathbf{g}_J))_{\omega \in \{0,1\}^d} \rangle_{U^d(G)} \quad (34)$$

almost surely for any standard  $J$  and any deterministic internal 1-bounded functions  $F_\omega: G \times G^J \rightarrow {}^*\mathbb{C}$  for  $\omega \in \{0,1\}^d$ .

We establish this by induction on  $d$ . For  $d = 0$  both sides are equal to

$$\text{st}\mathbb{E}_{x \in G} F_\omega(x, \mathbf{g}_1, \dots, \mathbf{g}_J)$$

so the claim is immediate in this case. Now suppose that  $d \geq 1$  and that the claim has already been proven for  $d - 1$ . From (13), the left-hand side of (34) can be written as

$$\lim_{n \rightarrow \infty} \mathbb{E}_{h^0, h^1 \in \Phi_n} \langle (\text{st}F_{\omega, h^0, \mathbf{g}, h^1, \mathbf{g}, \mathbf{g}_1, \dots, \mathbf{g}_J})_{\omega \in \{0,1\}^{d-1}} \rangle_{U^{d-1}(X)}$$

where

$$F_{\omega, a^0, a^1, \mathbf{g}_1, \dots, \mathbf{g}_J}(x) := F_{\omega, 0}(x - a^0, \mathbf{g}_1, \dots, \mathbf{g}_J) \overline{F_{\omega, 1}}(x - a^1, \mathbf{g}_1, \dots, \mathbf{g}_J), \quad (35)$$

for  $\omega \in \{0,1\}^{d-1}$  with  $h^j = \sum_{i=1}^\infty h_i^j e_i$ , and

$$h^j \cdot \mathbf{g} := \sum_{i=1}^\infty h_i^j \mathbf{g}_i$$

for  $j = 0, 1$ . Note that each  $F_{\omega, h^0, \mathbf{g}, h^1, \mathbf{g}, \mathbf{g}_1, \dots, \mathbf{g}_J}(x)$  is a 1-bounded internal function of  $x$  and finitely many of the  $\mathbf{g}_i$ . Applying the induction hypothesis, we may write the above expression as

$$\lim_{n \rightarrow \infty} \mathbb{E}_{h^0, h^1 \in \Phi_n} \langle (\text{st}F_{\omega, h^0, \mathbf{g}, h^1, \mathbf{g}, \mathbf{g}_1, \dots, \mathbf{g}_J})_{\omega \in \{0,1\}^{d-1}} \rangle_{U^{d-1}(G)}.$$

We can rewrite this as

$$\lim_{n \rightarrow \infty} \text{st}\mathbb{E}_{h^0, h^1 \in \Phi_n} H(h^0 \cdot \mathbf{g}, h^1 \cdot \mathbf{g}, \mathbf{g}_1, \dots, \mathbf{g}_J)$$

where  $H$  is the internal Gowers inner product

$$H(a^0, a^1, \mathbf{g}_1, \dots, \mathbf{g}_J) := \langle (F_{\omega, a^0, a^1, \mathbf{g}_1, \dots, \mathbf{g}_J})_{\omega \in \{0,1\}^{d-1}} \rangle_{U^{d-1}(G)}. \quad (36)$$

Note that  $H: G^2 \times G^J \rightarrow {}^*\mathbb{C}$  is an internal 1-bounded function. We will shortly show that one almost surely has the sampling identity

$$\lim_{n \rightarrow \infty} \text{st}\mathbb{E}_{h^0, h^1 \in \Phi_n} H(h^0 \cdot \mathbf{g}, h^1 \cdot \mathbf{g}, \mathbf{g}_1, \dots, \mathbf{g}_J) = \text{st}^*\mathbb{E}_{a^0, a^1 \in G} H(a^0, a^1, \mathbf{g}_1, \dots, \mathbf{g}_J) \quad (37)$$

for any given 1-bounded internal function  $H : G^2 \times G^J \rightarrow {}^*\mathbb{C}$ , where we recall that  ${}^*\mathbb{E}_{a^0, a^1 \in G}$  denotes the internal average over  $G \times G$ . Assuming this identity holds for the moment, we can now write the left-hand side of (34) as

$$\text{st} {}^*\mathbb{E}_{a^0, a^1 \in G} H(a^0, a^1, \mathbf{g}_1, \dots, \mathbf{g}_J).$$

But from (36), (35), (4) one has the internal identity

$${}^*\mathbb{E}_{a^0, a^1 \in G} H(a^0, a^1, g_1, \dots, g_J) = \langle F_\omega(\cdot, g_1, \dots, g_J) \rangle_{{}^*U^d(G)}$$

for any  $g_1, \dots, g_J \in G$ , as can be seen by first establishing the corresponding identity on each  $G_n$  and then taking ultralimits. The claim (34) now follows from (29).

It remains to establish (37) almost surely for a given internal function  $H$ . If  $H(a^0, a^1, g_1, \dots, g_J)$  does not actually depend on the first two inputs  $a^0, a^1$  then the claim is trivial. Subtracting off the internal marginal mean  ${}^*\mathbb{E}_{a^0, a^1 \in G} H(a^0, a^1, g_1, \dots, g_J)$  and using linearity, it thus suffices to verify (37) under the additional assumption of vanishing marginal mean

$${}^*\mathbb{E}_{a^0, a^1 \in G} H(a^0, a^1, g_1, \dots, g_J) = 0 \quad (38)$$

for all  $g_1, \dots, g_J$ , in which case our task is to show that

$$\lim_{n \rightarrow \infty} \text{st} \mathbb{E}_{h^0, h^1 \in \Phi_n} H(h^0 \cdot \mathbf{g}, h^1 \cdot \mathbf{g}, \mathbf{g}_1, \dots, \mathbf{g}_J) = 0$$

almost surely. By the Borel–Cantelli lemma, it suffices to show that

$$\sum_{n=1}^{\infty} \mathbb{E} \text{st} |\mathbb{E}_{h^0, h^1 \in \Phi_n} H(h^0 \cdot \mathbf{g}, h^1 \cdot \mathbf{g}, \mathbf{g}_1, \dots, \mathbf{g}_J)|^2 < \infty.$$

The left-hand side can be expanded as

$$\sum_{n=1}^{\infty} \mathbb{E}_{h^0, h^1, h^2, h^3 \in \Phi_n} \mathbb{E} \text{st} H(h^0 \cdot \mathbf{g}, h^1 \cdot \mathbf{g}, \mathbf{g}_1, \dots, \mathbf{g}_J) \overline{H}(h^2 \cdot \mathbf{g}, h^3 \cdot \mathbf{g}, \mathbf{g}_1, \dots, \mathbf{g}_J) \quad (39)$$

Now we take advantage of the specific form (10) of the Følner sequence  $\Phi_n$ , which gives  $h^i = \sum_{j=1}^{2^n} h_j^i e_j$  for some  $h_j^i \in \{0, \dots, n\}$  for each  $i = 0, 1, 2, 3$ . Suppose that there exists an odd number  $J < j < 2^n$  such that  $h_{j+i'}^i = 1_{i=i'}$  for all  $i = 0, 1, 2, 3$  and  $i' = 0, 1$ . Then for each  $i' = 0, 1$ , the random variable  $\mathbf{g}_{j+i'}$  only makes an appearance in the expression

$$\mathbb{E} \text{st} H(h^0 \cdot \mathbf{g}, h^1 \cdot \mathbf{g}, \mathbf{g}_1, \dots, \mathbf{g}_J) \overline{H}(h^2 \cdot \mathbf{g}, h^3 \cdot \mathbf{g}, \mathbf{g}_1, \dots, \mathbf{g}_J)$$

only through the  $h^i \cdot \mathbf{g}$  term, which can be written as the sum of  $\mathbf{g}_{j+i'}$  plus another expression not depending on either of the  $\mathbf{g}_j, \mathbf{g}_{j+1}$ . From (38) and a change of variable, we then have

$${}^*\mathbb{E}_{g \in G^{2^n}} H(h^0 \cdot g, h^1 \cdot g, g_1, \dots, g_J) \overline{H}(h^2 \cdot g, h^3 \cdot g, g_1, \dots, g_J) = 0$$

(writing  $g = (g_1, \dots, g_{2^n})$  and defining the inner products  $h^i \cdot g$  in the obvious fashion), and hence by (33)

$$\mathbb{E} \text{st} H(h^0 \cdot \mathbf{g}, h^1 \cdot \mathbf{g}, \mathbf{g}_1, \dots, \mathbf{g}_J) \overline{H}(h^2 \cdot \mathbf{g}, h^3 \cdot \mathbf{g}, \mathbf{g}_1, \dots, \mathbf{g}_J) = 0.$$

Thus in the expression (39) we may restrict attention to those  $h^0, h^1, h^2, h^3$  for which there do *not* exist any odd  $J < j < 2^n$  such that  $h^i_{j+i'} = 1_{i=i'}$  for all  $i = 0, 1, 2, 3$  and  $i' = 0, 1$ . For each given  $j$ , the proportion of  $h^0, h^1, h^2, h^3$  for which this property fails is  $1 - \frac{1}{(n+1)^8}$ , and the events are independent for the  $\lfloor \max(\frac{2^n-j}{2}, 0) \rfloor$  different choices of  $j$ , and hence we can upper bound (39) by

$$\sum_{n=1}^{\infty} \left( 1 - \frac{1}{(n+1)^8} \right)^{\lfloor \max(\frac{2^n-j}{2}, 0) \rfloor}.$$

This sum is easily seen to be finite (bounding  $1 - \frac{1}{(n+1)^8}$  by  $\exp(-\frac{1}{(n+1)^8})$ , and the claim follows. This concludes the proof of Proposition 5.1(i).

It remains to establish that  $X$  is ergodic. If this is not the case, then there exists a non-trivial  $\mathbb{Z}^\omega$ -invariant function  $F \in L^\infty(X)$  of mean zero, which we can take to be 1-bounded. By the properties of Loeb measure, one can write  $F = \text{st}\tilde{F}$  for some internal 1-bounded function  $\tilde{F}$ . From the  $d = 1$  case of Proposition 5.1(i) and (29), (2), (44) we have

$$\begin{aligned} \langle F, F \rangle_{U^1(X)} &= \langle F, F \rangle_{U^1(G)} \\ &= \text{st} \langle \tilde{F}, \tilde{F} \rangle_{*U^1(G)} \\ &= \text{st} |\mathbb{E}_{x \in G} \tilde{F}(x)|^2 \\ &= \left| \int_X F \, d\mu \right|^2 \\ &= 0. \end{aligned}$$

On the other hand, from (11), (12) and the hypothesis that  $F$  is invariant, we easily compute that

$$\langle F, F \rangle_{U^1(X)} = \int_X |F|^2 \, d\mu.$$

Since  $F$  is non-trivial, we obtain the required contradiction. This completes the proof of Proposition 5.1.

## 6 The structure of nilpotent quotient spaces

Theorem 1.14 produces a quotient space  $H/\Gamma$  where  $H$  is a filtered locally compact group of degree  $s = 2$ , and  $\Gamma$  is a filtered lattice of  $H$ . This quotient space is then compact, and the Host–Kra spaces  $\text{HK}^k(H/\Gamma) \subset (H/\Gamma)^{\{0,1\}^k}$  are similarly compact (this follows readily from the proof of [15, Lemma E.10]). In the language of nilspaces,  $H/\Gamma$  in fact has the structure of a compact nilspace; see [20, Appendix A] for further discussion.

In this section we explore the nilspace structure of  $H/\Gamma$  further (“forgetting” for now about the dynamical structure of the shift  $T$ ), in particular establishing that  $H/\Gamma$  is an inverse limit (in the category of compact nilspaces) of nilmanifolds. Our arguments will not be restricted to the  $s = 2$  case. If the  $H_i$  were connected then one could argue using tools such as [6, Theorem 6.2] to establish this claim, but in our setting we do not have any connectedness properties on  $H$  and we will need to rely instead on the general structural theory of nilpotent locally compact groups that are not necessarily connected.

We first need a simple algebraic lemma concerning the ability to “factor” Host–Kra groups. Given two subsets  $A, B$  of a multiplicative group  $G$ , we define the product set  $AB := \{ab : a \in A, b \in B\}$ . (In particular, the product  $HK$  of two subgroups  $H, K$  of  $G$  need not be a subgroup if neither of  $H, K$  are normal.)

**Lemma 6.1** (Factoring Host–Kra groups). *Let  $G = (G, \cdot)$  be a prefiltered group, and let  $H, K$  be prefiltered subgroups of  $G$  such that  $G_i \subset H_i K_i$  for all  $i \geq 0$ . Then  $\mathrm{HK}^k(G) \subset \mathrm{HK}^k(H) \mathrm{HK}^k(K)$  for all  $k$ .*

*Proof.* For any  $0 \leq i \leq k+1$ , let  $S_i$  denote the set of all elements  $(g_\omega)_{\omega \in \{0,1\}^k}$  of  $\mathrm{HK}^k(G)$  such that  $g_\omega = 1$  whenever  $|\omega| < i$ . We claim the inclusion

$$S_i \subset \mathrm{HK}^k(H) S_{i+1} \mathrm{HK}^k(K) \quad (40)$$

for all  $0 \leq i \leq k$ . Iterating this and noting that  $S_0 = \mathrm{HK}^k(G)$  and  $S_{k+1} = \{1\}$ , we obtain the claim.

It remains to establish (40). Let  $(g_\omega)_{\omega \in \{0,1\}^k}$  be an element of  $S_i$ . We order the elements of  $\{0,1\}^k$  as  $\omega_1, \dots, \omega_{2^k}$  in such a way that  $|\omega_1|, \dots, |\omega_{2^k}|$  is non-decreasing. Applying [20, Proposition A.5] (or [15, (E.1)]), we may factor

$$(g_\omega)_{\omega \in \{0,1\}^k} = \prod_{j=1}^{2^k} [x_{\omega_j}]_{\omega_j}$$

as an ordered product for some  $x_{\omega_j} \in G_{|\omega_j|}$ , where the notation  $[x_{\omega_j}]_{\omega_j}$  was defined in (6). Since this product lies in  $S_i$ , we have  $x_{\omega_j} = 1$  whenever  $|\omega_j| < i$ . When  $|\omega_j| = i$ , we may use the hypothesis  $G_i \subset H_i K_i$  to factor  $x_{\omega_j} = y_{\omega_j} z_{\omega_j}$  where  $y_{\omega_j} \in H_i$  and  $z_{\omega_j} \in K_i$ . We can then factor

$$(g_\omega)_{\omega \in \{0,1\}^k} = \left( \prod_{1 \leq j \leq 2^k: |\omega_j|=i} [y_{\omega_j}]_{\omega_j} \right) (g'_\omega)_{\omega \in \{0,1\}^k} \left( \prod_{1 \leq j \leq 2^k: |\omega_j|=i} [z_{\omega_j}]_{\omega_j} \right)$$

for some  $(g'_\omega)_{\omega \in \{0,1\}^k}$  that one readily verifies to lie in  $S_{i+1}$ . The claim follows.  $\square$

Using this lemma we can replace the filtered group  $H$  by a compactly generated filtered group  $H'$  without affecting the Host–Kra structure.

**Corollary 6.2** (Replacing with a compactly generated group). *Let  $H$  be a filtered locally compact group of degree  $s$ , and let  $\Gamma$  be a filtered lattice in  $H$ . Then there exists a compactly generated open subgroup  $H'$  of  $H$  (which has the structure of a locally compact filtered group with  $H'_i := H' \cap H_i$ ) such that  $\Gamma \cap H'$  is a filtered lattice of  $H'$  with  $H/\Gamma$  and  $H'/(H' \cap \Gamma)$  isomorphic as compact topological spaces, with the identification also being a nilspace isomorphism.*

*Proof.* Since  $H_i/\Gamma_i$  is compact and  $\Gamma_i$  is discrete, there exists a non-empty precompact open subset  $K_i$  of  $H_i$  such that  $H_i = K_i \Gamma_i$ . If we let  $H'$  be the group generated by  $\bigcup_{i=0}^s K_i$ , then  $H'$  is a compactly generated open subgroup of  $H$  and  $H_i = H'_i \Gamma_i$  for all  $i$ . In particular the obvious identification between  $H_i/\Gamma_i$  and  $H'_i/(H'_i \cap \Gamma_i)$  is a homeomorphism, which implies that  $H'_i \cap \Gamma_i$  is cocompact in  $H'_i$ . Thus  $H' \cap \Gamma$  is a lattice in  $H'$ . From Lemma 6.1 we have

$$\mathrm{HK}^k(H) = \mathrm{HK}^k(H') \mathrm{HK}^k(\Gamma).$$

Now we apply the canonical projection  $\pi: H^{\{0,1\}^k} \rightarrow (H/\Gamma)^{\{0,1\}^k} \equiv (H'/H' \cap \Gamma)^{\{0,1\}^k}$  to both sides of this identity. The image of  $\mathrm{HK}^k(H)$  under  $\pi$  is  $\mathrm{HK}^k(H/\Gamma)$  by definition, while the image of  $\mathrm{HK}^k(H')\mathrm{HK}^k(\Gamma)$  is

$$\pi(\mathrm{HK}^k(H')\mathrm{HK}^k(\Gamma)) = \pi(\mathrm{HK}^k(H')) \equiv \mathrm{HK}^k(H'/(H' \cap \Gamma)),$$

noting (from the group isomorphism theorems) that the restriction of  $\pi$  to  $(H')^{\{0,1\}^k}$  agrees (after the obvious identifications) with the projection from  $(H')^{\{0,1\}^k}$  to  $(H'/H' \cap \Gamma)^{\{0,1\}^k}$ . Thus we have  $\mathrm{HK}^k(H/\Gamma) \equiv \mathrm{HK}^k(H'/(H' \cap \Gamma))$  for all  $k$ , giving the required nilspace isomorphism.  $\square$

The advantage of passing to a compactly generated setting is provided by the following variant of the Gleason–Yamabe theorem.

**Proposition 6.3** (Gleason–Yamabe for compactly generated nilpotent groups). *Let  $H$  be a compactly generated locally compact nilpotent group. Then  $H$  is the inverse limit of Lie groups. In other words, every neighborhood  $U$  of the identity in  $H$  contains a compact normal subgroup  $N$  such that  $H/N$  is isomorphic to a Lie group.*

*Proof.* This follows from [21, Theorem 9]. We remark that this theorem is simpler to prove than the general Gleason–Yamabe theorem [8], [40], which asserts that any locally compact group contains an open subgroup that is an inverse limit of Lie groups. The point is that (as shown in [21]) the passage to an open subgroup is unnecessary in the compactly generated nilpotent case.  $\square$

**Remark 6.4.** The hypothesis of compact generation in Proposition 6.3 is necessary. For instance, if  $p$  is a prime, the Heisenberg group  $\begin{pmatrix} 1 & \mathbb{Q}_p & \mathbb{Q}_p \\ 0 & 1 & \mathbb{Q}_p \\ 0 & 0 & 1 \end{pmatrix}$  over the  $p$ -adic field  $\mathbb{Q}_p$  is locally compact nilpotent, but not the inverse limit of Lie groups, basically because all compact normal subgroups of this group lie in the center  $\begin{pmatrix} 1 & 0 & \mathbb{Q}_p \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ . For similar reasons the nilpotency hypothesis is necessary, as can be seen by considering the compactly generated solvable group  $\mathbb{Z} \ltimes \mathbb{Q}_p$ , where the action of  $\mathbb{Z}$  on  $\mathbb{Q}_p$  is generated by multiplication by  $p$ .

We can combine Corollary 6.2 and Proposition 6.3 to approximate continuous functions on a general degree  $s$  quotient  $H/\Gamma$  by (pullbacks of) continuous functions of degree  $s$  nilmanifolds  $\tilde{H}/\tilde{\Gamma}$ .

**Corollary 6.5** (Approximation by nilmanifolds). *Let  $H$  be a filtered locally compact group of degree  $s$ , and let  $\Gamma$  be a filtered lattice in  $H$ . Let  $F: H/\Gamma \rightarrow \mathbb{C}$  be a continuous map, and let  $\varepsilon > 0$ . Then there exists a filtered nilmanifold  $\tilde{H}/\tilde{\Gamma}$  of degree  $s$ , a continuous function  $\tilde{F}: \tilde{H}/\tilde{\Gamma} \rightarrow \mathbb{C}$ , and a continuous nilspace morphism  $\pi: H/\Gamma \rightarrow \tilde{H}/\tilde{\Gamma}$  such that*

$$|F(x) - \tilde{F}(\pi(x))| \leq \varepsilon$$

for all  $x \in H/\Gamma$ .

Another way of expressing this corollary is that every quotient  $H/\Gamma$  of a filtered locally compact group of degree  $s$  by a filtered lattice is an inverse limit (in the category of compact nilspaces) of degree  $s$  filtered nilmanifolds.

*Proof.* We may assume without loss of generality that  $F$  is real-valued. By Corollary 6.2 we may assume without loss of generality that  $H$  is compactly generated, and is thus an inverse limit of Lie groups by Proposition 6.3, which are nilpotent since  $H$  is nilpotent. Since  $F$  is continuous on the compact space  $H/\Gamma$ , it is uniformly continuous, thus there exists an open neighborhood  $U$  of the identity in  $H$  such that  $|F(hx) - F(x)| \leq \varepsilon$  for all  $h \in U$ . By Proposition 6.3, we can thus find a compact normal subgroup  $N$  of  $H$  in  $U$  such that  $H/N$  is a nilpotent Lie group. Since  $N \subset U$ , we have  $|F(hx) - F(x)| \leq \varepsilon$  for all  $h \in N$ . The map<sup>8</sup>  $x \mapsto \min_{h \in N} F(hx)$  is then a continuous  $N$ -invariant function that lies within  $\varepsilon$  of  $F$  in the uniform topology. This function descends to a continuous function on the quotient  $(H/N)/(\Gamma N/N)$ , which is a filtered nilmanifold by Proposition 6.3. The claim follows.  $\square$

## 7 Concluding the implication of the Gowers norm inverse theorem from the Host–Kra inverse theorem

We can now conclude the proof of Theorem 1.16. By the discussion in Section 4, it suffices to show that Conjecture 4.1 holds for  $s = 2$ . Let  $G, f$  be as in that conjecture. We invoke Proposition 5.1 (with  $\mathcal{F} = \{f\}$ ) to construct an ergodic  $\mathbb{Z}^\omega$ -system  $(X, \mu_X, T)$  obeying the conclusions of that proposition. In particular

$$\|f\|_{U^{s+1}(X)} = \|f\|_{U^{s+1}(G)} > 0.$$

Applying Theorem 1.14, we can find a filtered locally compact group  $H$  of degree at most  $s = 2$ , a filtered lattice  $\Gamma$  of  $H$ , a translation action  $T_{H/\Gamma}: \mathbb{Z}^\omega \rightarrow \text{Aut}(H/\Gamma)$  defined by

$$T_{H/\Gamma}^h(x) := \phi(h)x$$

for some group homomorphism  $\phi: \mathbb{Z}^\omega \rightarrow H$ , a  $\mathbb{Z}^\omega$ -morphism  $\Pi: X \rightarrow H/\Gamma$ , and a continuous function  $F: H/\Gamma \rightarrow \mathbb{C}$  such that

$$\int_X f(x) \overline{F(\Pi(x))} d\mu_X \neq 0.$$

Applying Corollary 6.5 for a sufficiently small  $\varepsilon > 0$  and using the triangle inequality, we can thus find a filtered nilmanifold  $\tilde{H}/\tilde{\Gamma}$  of degree  $s$ , a continuous function  $\tilde{F}: \tilde{H}/\tilde{\Gamma} \rightarrow \mathbb{C}$ , and a continuous nilspace morphism  $\pi: H/\Gamma \rightarrow \tilde{H}/\tilde{\Gamma}$  such that

$$\int_X f(x) \overline{\tilde{F}(\pi \circ \Pi(x))} d\mu_X \neq 0. \quad (41)$$

We now make a key definition:

---

<sup>8</sup>One could also average here using the Haar measure on  $N$  if desired.

**Definition 7.1.** A Loeb-measurable map  $\tilde{g}: G \rightarrow \tilde{H}/\tilde{\Gamma}$  is *almost polynomial* if, for every standard  $k \geq 0$ , one has

$$(\tilde{g}(x_\omega))_{\omega \in \{0,1\}^k} \in \text{HK}^k(\tilde{H}/\tilde{\Gamma})$$

for  $\mu_{\text{HK}^k(G)}$ -almost all  $(x_\omega)_{\omega \in \{0,1\}^k} \in \text{HK}^k(G)$ , where  $\mu_{\text{HK}^k(G)}$  is Loeb measure on  $\text{HK}^k(G)$ .

**Lemma 7.2.** *The map  $\pi \circ \Pi: G \rightarrow \tilde{H}/\tilde{\Gamma}$  is almost polynomial.*

*Proof.* Fix  $k \geq 0$ . By the second countable nature of  $(\tilde{H}/\tilde{\Gamma})^{\{0,1\}^k}$ , it suffices to show that for any open sets  $U_\omega \subset \tilde{H}/\tilde{\Gamma}$ ,  $\omega \in \{0,1\}^k$  with  $\prod_{\omega \in \{0,1\}^k} U_\omega$  disjoint from  $\text{HK}^k(\tilde{H}/\tilde{\Gamma})$ , that the set

$$\{(x_\omega)_{\omega \in \{0,1\}^k} \in \text{HK}^k(G) : \pi(\Pi(x_\omega)) \in U_\omega \forall \omega \in \{0,1\}^k\}$$

is  $\mu_{\text{HK}^k(G)}$ -null, or equivalently that

$$\int_{\text{HK}^k(G)} \prod_{\omega \in \{0,1\}^k} 1_{\pi^{-1}(U_\omega)}(\Pi(x_\omega)) d\mu_{\text{HK}^k(G)}((x_\omega)_{\omega \in \{0,1\}^k}) = 0. \quad (42)$$

From (29), we can write the left-hand side of (42) as

$$\langle (1_{\pi^{-1}(U_\omega)} \circ \Pi)_{\omega \in \{0,1\}^k} \rangle_{U^k(G)}$$

which by (31) is equal to

$$\langle (1_{\pi^{-1}(U_\omega)} \circ \Pi)_{\omega \in \{0,1\}^k} \rangle_{U^k(X)}.$$

Since  $\Pi: X \rightarrow H/\Gamma$  is a  $\mathbb{Z}^\omega$ -morphism, we may write this as

$$\langle (1_{\pi^{-1}(U_\omega)})_{\omega \in \{0,1\}^k} \rangle_{U^k(H/\Gamma)}.$$

By (11), (12), it thus suffices to show that

$$\int_{H/\Gamma} \prod_{\omega \in \{0,1\}^k} T_{H/\Gamma}^{h_1^{\omega_1} + \dots + h_k^{\omega_k}} 1_{\pi^{-1}(U_\omega)} = 0$$

for any  $h_1^0, h_1^1, \dots, h_k^0, h_k^1 \in \mathbb{Z}^\omega$ . By (14), it thus suffices to show that

$$(\phi(h_1^{\omega_1} + \dots + h_k^{\omega_k})x)_{\omega \in \{0,1\}^k} \notin \prod_{\omega \in \{0,1\}^k} \pi^{-1}(U_\omega).$$

But the left-hand side lies in  $\text{HK}^k(H/\Gamma)$ , which is in  $(\pi^{\{0,1\}^k})^{-1}(\tilde{H}/\tilde{\Gamma})$ , and the claim then follows since  $\prod_{\omega \in \{0,1\}^k} U_\omega$  is disjoint from  $\text{HK}^k(\tilde{H}/\tilde{\Gamma})$ .  $\square$

We now invoke the following stability result:

**Lemma 7.3** (Stability of polynomials). *If  $\tilde{g}: G \rightarrow \tilde{H}/\tilde{\Gamma}$  is almost polynomial, then there exists an internal polynomial map  $g: G \rightarrow {}^*(\tilde{H}/\tilde{\Gamma})$  such that  $\tilde{g}(x) = \text{stg}(x)$  for  $\mu_G$ -almost all  $x$ .*

This is a special case of [6, Theorem 4.2], translated to the language of nonstandard analysis; the argument is valid for arbitrary degrees  $s$ . For the convenience of the reader we supply a self-contained proof of this result in Appendix B. In the special case where  $G$  is a vector space over a field  $\mathbb{F}_p$  of prime order  $p$ , this result (phrased in the language of standard analysis) was established in [36, Lemma 4.5]. We remark that the converse implication of Lemma 7.3 (that functions that agree almost everywhere with the standard part of an internal polynomial are almost polynomial) is trivial.

Combining these two lemmas, we obtain an internal polynomial map  $g: G \rightarrow {}^*(\tilde{H}/\tilde{\Gamma})$  such that

$$\pi \circ \Pi(x) = \text{stg}(x)$$

for  $\mu_G$ -almost all  $x$ , and (30) now follows from (41) (recalling that  $\mu_X$  is the restriction of  $\mu_G$  to  $\mathcal{X}$ ). This concludes the proof of Theorem 1.16.

## A A review of nonstandard analysis

In this appendix we briefly review some basic definitions and tools from nonstandard analysis that we will use in this paper. This formalism is frequently used in this subject (see e.g., [19], [6], [38]); our formalism here is particularly close to that in [38, §5].

We assume the existence of a *standard universe*  $\mathcal{U}$  that contains all the objects implicitly discussed in results such as Theorem 1.11, which will henceforth be referred to as *standard objects*. For instance, the group  $G$  in Theorem 1.11 can be assumed to be a standard finite additive group, the nilmanifold  $H/\Gamma$  can be assumed to be a standard filtered nilmanifold, and so forth. We will shortly also introduce a nonstandard universe  ${}^*\mathcal{U}$  that contains the nonstandard objects, and refer to *external objects* as objects that do not necessarily lie in either the standard or nonstandard universes. (For instance, a standard or nonstandard object can either be viewed internally with respect to the universe it lies in, or externally without explicit mention of any universe.)

Throughout this paper we fix an (external) non-principal ultrafilter  $p \in \beta\mathbb{N} \setminus \mathbb{N}$ , which induces notions of ultraproduct, ultrapower, and ultralimit as defined for instance in [38, §5]. Given some standard notion of a space (e.g., a finite abelian group, or a nilmanifold), a *nonstandard space* will be an ultraproduct  $X = \prod_{n \rightarrow p} X_n$  of standard spaces  $X_n$ , whose elements are ultralimits  $x = \lim_{n \rightarrow p} x_n$  of standard objects  $x_n \in X_n$  for a  $p$ -large set of  $n$ . Thus for instance a nonstandard finite additive group is an ultraproduct  $G = \prod_{n \rightarrow p} G_n$  of standard finite additive groups  $G_n$ . There is an obvious identification

$$\left( \prod_{n \rightarrow p} X_n \right) \times \left( \prod_{n \rightarrow p} Y_n \right) \equiv \prod_{n \rightarrow p} (X_n \times Y_n)$$

for any nonstandard spaces  $\prod_{n \rightarrow p} X_n, \prod_{n \rightarrow p} Y_n$ ; by abuse of notation, we shall make frequent use of this identification in this paper without further comment, so that the Cartesian product of finitely many nonstandard spaces is again a nonstandard space. (We caution however that such an identification breaks down for infinite Cartesian products.)

By Łos's theorem, every structural property of standard spaces that is expressible as a sentence in first-order logic is inherited by their nonstandard counterparts. For instance, a nonstandard finite additive group will remain an additive group in the external sense, though it will usually not be externally finite.

The ultrapower  $\prod_{n \rightarrow p} X$  of a standard space  $X$  will be denoted  ${}^*X$ , and contains  $X$  as a subspace via the identification  $x \equiv \lim_{n \rightarrow p} x$ . For instance,  ${}^*\mathbb{R}$  will denote the nonstandard real numbers (also known as the hyperreals),  ${}^*\mathbb{C}$  will denote the nonstandard complex numbers, and all nonstandard objects lie in the nonstandard universe  ${}^*\mathcal{U}$ . The nonstandard reals  ${}^*\mathbb{R}$  form an ordered field that contains the standard reals  $\mathbb{R}$  as a subfield.

Let  $X = \prod_{n \rightarrow p} X_n$  and  $Y = \prod_{n \rightarrow p} Y_n$  be nonstandard sets. An *internal function*  $f: X \rightarrow Y$  between these sets is a function that is an ultralimit  $f = \lim_{n \rightarrow p} f_n$  of standard functions  $f_n: X_n \rightarrow Y_n$ . One can similarly define internal homomorphisms between nonstandard groups, or internal polynomial maps between nonstandard prefiltered groups, in the obvious fashion. Given a standard function  $f: X \rightarrow Y$  between standard sets, we define its nonstandard counterpart  ${}^*f: {}^*X \rightarrow {}^*Y$  between the associated ultrapowers by  ${}^*f := \lim_{n \rightarrow p} f$ . In some cases we may abuse notation by abbreviating  ${}^*f$  as  $f$ .

A nonstandard complex number  $z$  is said to be *bounded* if  $|z| \leq C$  for some standard real  $C > 0$ , and *infinitesimal* if  $|z| \leq \varepsilon$  for all standard reals  $\varepsilon > 0$ . We denote infinitesimal complex numbers by  $o(1)$ . By a well-known adaptation of the proof of the Bolzano–Weierstrass theorem, we see that to every bounded nonstandard complex number  $z$  there is a unique standard complex number  $\text{st}(z)$  (known as the *standard part* of  $z$ ) such that  $z = \text{st}(z) + o(1)$ . The map  $z \mapsto \text{st}(z)$  is a  $*$ -homomorphism from the external algebra of bounded nonstandard complex numbers to  $\mathbb{C}$ . In a similar spirit, given any compact metric space  $X$ , there is a standard part map  $\text{st}: {}^*X \rightarrow X$  which is the unique map such that  ${}^*d(x, \text{st}(x)) = o(1)$  for all  $x \in {}^*X$ , where  ${}^*d: {}^*X \times {}^*X \rightarrow {}^*\mathbb{R}$  is the nonstandard counterpart of the metric  $d: X \times X \rightarrow \mathbb{R}$ . In fact the standard part map is well defined for any compact *metrizable* space, since one easily checks that any two metrics that generate the same topology also generate the same standard part map.

An important construction for us will be the *Loeb measure construction* [24]. Given a sequence  $(X_n, \mathcal{X}_n, \mu_n)$  of standard probability spaces, this construction yields an (external) complete probability space  $(X, \mathcal{X}, \mu_X)$  where  $X = \prod_{n \rightarrow p} X_n$  is the ultraproduct of the  $X_n$ , with the following two properties:

- (i) Whenever the internal function  $f = \lim_{n \rightarrow p} f_n$  is the ultralimit of uniformly bounded standard measurable functions  $f_n: X_n \rightarrow \mathbb{C}$ , the function  $\text{st}f: X \rightarrow \mathbb{C}$  is bounded and measurable in  $(X, \mathcal{X}, \mu)$ , and

$$\int_X \text{st}f \, d\mu_X = \text{st} \int_X f \, d{}^*\mu_X \quad (43)$$

where the *internal integral*  $\int_X f \, d{}^*\mu_X$  of  $f$  is defined by the formula

$$\int_X f \, d{}^*\mu_X := \lim_{n \rightarrow p} \int_{X_n} f_n \, d\mu_n.$$

- (ii) Conversely, whenever<sup>9</sup>  $f \in L^\infty(X) = L^\infty(X, \mu_X)$ , there exists a sequence  $f_n: X_n \rightarrow \mathbb{C}$  of uniformly bounded standard measurable functions such that  $f = \text{st} \lim_{n \rightarrow p} f_n$   $\mu_X$ -almost everywhere.

See for instance [38, §5] for the construction of Loeb measure (based on the Carathéodory extension theorem). One can then define the Lebesgue spaces  $L^p(X, \mu_X)$  for  $1 \leq p \leq \infty$  in the usual fashion. We caution that these measure spaces will not be separable in general.

As a key special case of the Loeb measure construction, if  $G$  is a nonstandard finite additive group, thus  $G = \prod_{n \rightarrow p} G_n$  is the ultraproduct of standard finite additive groups  $G_n$ , then one can give  $G$  the

<sup>9</sup>Here we adopt the usual convention of identifying two elements on  $L^\infty(X)$  whenever they agree  $\mu_X$ -almost everywhere.

structure of an (external) probability space  $(G, \mathcal{L}_G, \mu_G)$ , by endowing each of the standard finite groups  $G_n$  with the discrete  $\sigma$ -algebra and uniform measure, and applying the Loeb construction; we refer to  $(G, \mathcal{L}_G, \mu_G)$  as the *Loeb probability space* associated with  $G$ , with  $\mathcal{L}_G$  the *Loeb  $\sigma$ -algebra* and  $\mu_G$  the *Loeb measure* on  $G$ . Thus for instance if the internal function  $f = \lim_{n \rightarrow p} f_n$  is the ultralimit of uniformly bounded standard functions  $f_n: G_n \rightarrow \mathbb{C}$ , then

$$\int_G \text{st} f \, d\mu_G = \text{st}^* \mathbb{E}_{x \in G} f(x) \quad (44)$$

where the *internal average*  $^* \mathbb{E}_{x \in G} f(x)$  of  $f$  in  $G$  is defined as

$$^* \mathbb{E}_{x \in G} f(x) := \lim_{n \rightarrow p} \mathbb{E}_{x \in G_n} f_n(x).$$

Because counting measure on each  $G_n$  is translation-invariant, Loeb measure on  $G$  is also translation-invariant. Thus Loeb measure  $\mu_G$  is very analogous to Haar measure, though we caution that Loeb measure is not actually a special case of Haar measure since  $G$  does not have the structure of a locally compact group (and  $\mathcal{L}_G$  is not a Borel  $\sigma$ -algebra).

If  $G, H$  are nonstandard non-empty finite sets, then we can construct Loeb measure spaces  $(G, \mathcal{L}_G, \mu_G)$ ,  $(H, \mathcal{L}_H, \mu_H)$  on the nonstandard sets  $G, H$ , as well as a Loeb measure space  $(G \times H, \mathcal{L}_{G \times H}, \mu_{G \times H})$  on the product space  $G \times H$ . We caution that in general, the Loeb measure space is *not* the product  $(G \times H, \mathcal{L}_G \times \mathcal{L}_H, \mu_G \times \mu_H)$  of the individual Loeb measure spaces  $(G, \mathcal{L}_G, \mu_G)$ ,  $(H, \mathcal{L}_H, \mu_H)$ , even after completing that product measure; see [33, Remark 2.10.4]. Instead, the former space is an *extension* of the latter:  $\mathcal{L}_G \times \mathcal{L}_H$  is a subalgebra of  $\mathcal{L}_{G \times H}$ , and  $\mu_G \times \mu_H$  is the restriction of  $\mu_{G \times H}$  to  $\mathcal{L}_G \times \mathcal{L}_H$ . Furthermore, we have the pleasant fact that the familiar Fubini theorem for  $\mu_G \times \mu_H$  extends to  $\mu_{G \times H}$ :

**Lemma A.1** (Keisler–Fubini theorem). *Let  $G, H$  be nonstandard non-empty finite sets.*

(i) *If  $f: G \times H \rightarrow \mathbb{C}$  is a  $\mathcal{L}_{G \times H}$ -measurable bounded function, then*

$$\begin{aligned} \int_{G \times H} f(g, h) \, d\mu_{G \times H}(g, h) &= \int_G \left( \int_H f(g, h) \, d\mu_H(h) \right) d\mu_G(g) \\ &= \int_H \left( \int_G f(g, h) \, d\mu_G(g) \right) d\mu_H(h). \end{aligned}$$

(ii) *If  $P(g, h)$  is an  $\mathcal{L}_{G \times H}$ -measurable property of elements  $(g, h)$  of  $G \times H$  (i.e.,  $\{(g, h) \in G \times H : P(g, h) \text{ holds}\}$  is measurable in  $\mathcal{L}_{G \times H}$ ), then  $P(g, h)$  holds for  $\mu_{G \times H}$ -almost all  $(g, h) \in G \times H$  if and only if, for  $\mu_G$ -almost all  $g \in G$ ,  $P(g, h)$  holds for  $\mu_H$ -almost all  $h \in H$ .*

*Similarly with the roles of  $G$  and  $H$  reversed.*

In this paper we will only need part (ii) of this lemma; part (i) is only stated here in order to prove part (ii).

*Proof.* For part (i), see for instance [33, Theorem 2.10.3]. As with the usual Fubini theorem, one can weaken “bounded” here to “absolutely integrable” (and there is a Tonelli-type variant in which  $f$  instead takes values in  $[0, +\infty]$ ), but we will not need these extensions of this theorem here. Part (ii) follows by applying part (i) to the indicator function of  $P$ .  $\square$

## B Proof of stability of polynomials

In this appendix we establish Lemma 7.3. We need some preparatory lemmas. First, we show that an “almost homomorphism” can always be uniquely completed to an actual homomorphism.

**Lemma B.1** (Repairing an almost homomorphism). *Let  $G = (G, \cdot), K = (K, \cdot)$  be nonstandard groups with  $G$  nonstandard finite, and let  $\phi: G \rightarrow K$  be an internal map which is an “almost homomorphism” in the sense that  $\phi(gh) = \phi(g)\phi(h)$  for  $\mu_{G^2}$ -almost all  $(g, h) \in G^2$ . Then there exists a unique internal homomorphism  $\tilde{\phi}: G \rightarrow K$  such that  $\phi(g) = \tilde{\phi}(g)$  for  $\mu_G$ -almost all  $g \in G$ . Finally, if  $\phi$  takes values in an external subgroup  $K'$  of  $K$ , then so does  $\tilde{\phi}$ .*

*Proof.* Let  $g \in G$ . We claim that for  $\mu_{G^2}$ -almost all  $a, b \in G$ , one has

$$\phi(a)^{-1}\phi(ag) = \phi(b)^{-1}\phi(bg).$$

Indeed, writing  $b = ca$ , we see from Lemma A.1(ii) and translation invariance that for  $\mu_{G^2}$ -almost all  $a, b \in G$ , one has

$$\phi(b) = \phi(c)\phi(a)$$

and

$$\phi(bg) = \phi(c)\phi(ag)$$

giving the claim. Thus, if we define  $\tilde{\phi}(g)$  to be the mode of  $\phi(a)^{-1}\phi(ag)$  for  $a \in G$  (that is to say, the unique value that this quantity attains the most often),  $\tilde{\phi}$  is an internal function with the property that for all  $g \in G$ , one has

$$\phi(a)^{-1}\phi(ag) = \tilde{\phi}(g) \tag{45}$$

for  $\mu_G$ -almost every  $a \in G$ . In particular, for any  $g, h \in G$ , one has

$$\begin{aligned} \phi(a)^{-1}\phi(agh) &= \tilde{\phi}(gh) \\ \phi(a)^{-1}\phi(ag) &= \tilde{\phi}(g) \\ \phi(ag)^{-1}\phi(agh) &= \tilde{\phi}(h) \end{aligned}$$

for at least one  $a \in G$ , hence  $\tilde{\phi}(gh) = \tilde{\phi}(g)\tilde{\phi}(h)$ . Hence  $\tilde{\phi}$  is a homomorphism. Finally, from Lemma A.1(ii) and (45) we have  $\tilde{\phi}(g) = \phi(g)$  for  $\mu_G$ -almost all  $g$ . This gives existence. For uniqueness, observe that if we had two internal homomorphisms  $\tilde{\phi}, \tilde{\phi}'$  obeying the conclusions of this lemma, then for any  $g \in G$ , we see from invariance of Loeb measure that

$$\tilde{\phi}(g) = \tilde{\phi}(gh)\tilde{\phi}(h)^{-1} = \phi(gh)\phi(h)^{-1}$$

and

$$\tilde{\phi}'(g) = \tilde{\phi}'(gh)\tilde{\phi}'(h)^{-1} = \phi(gh)\phi(h)^{-1}$$

for almost all  $h \in G$ , hence  $\tilde{\phi} = \tilde{\phi}'$ . This argument also shows that  $\tilde{\phi}$  takes values in any external group  $K'$  that  $\phi$  does.  $\square$

Next, we establish some rigidity properties of near-polynomial maps.

**Lemma B.2** (Infinitesimal near-polynomials are constant). *Let  $G$  be a nonstandard finite additive group, let  $H$  be a standard prefiltered Lie group of degree  $s$  for some  $s \geq 0$ , and  $\phi: G \rightarrow {}^*H$  be an internal function such that for all standard  $k \geq 0$ , one has*

$$\partial_{h_1} \dots \partial_{h_k} \phi(x) \in {}^*H_k$$

*for  $\mu_{G^{k+1}}$ -almost all  $(x, h_1, \dots, h_k) \in G^{k+1}$ , and such that  $\text{st}\phi(x) = 1$  for  $\mu_G$ -almost all  $x \in G$ . Then there exists  $c \in {}^*H$  with  $\text{st}c = 1$  such that  $\phi(x) = c$  for  $\mu_G$ -almost all  $x \in G$ .*

*Proof.* We may assume that the claim is already established for any smaller value of  $s$ . By Lemma A.1(ii), we see that if  $s \geq 1$ , then for  $\mu_G$ -almost all  $h \in G$ , the function  $\partial_h \phi: G \rightarrow {}^*H$  obeys all the hypotheses of the lemma, but with  $s$  replaced by  $s - 1$  and the prefiltration  $(H_i)_{i=0}^\infty$  replaced by the shifted filtration  $(H_{i+1})_{i=0}^\infty$ . By induction hypothesis, this implies that for such  $h$  there exists  $\varepsilon_h \in {}^*H$  with  $\text{st}\varepsilon_h = 1$  such that  $\partial_h \phi(x) = \varepsilon_h$  for  $\mu_G$ -almost all  $x \in G$ . The same claim also holds in the  $s = 0$  case (with  $\varepsilon_h = 1$  in this case).

Clearly for each  $h$  the constant  $\varepsilon_h$ , if it exists, is unique, and depends internally on  $h$  (it is the mode of  $\partial_h \phi$ ). From the cocycle equation

$$\partial_{h+k} \phi(x) = \partial_h \phi(x) \partial_k \phi(x - h)$$

and Lemma A.1(ii), we see that

$$\varepsilon_{h+k} = \varepsilon_h \varepsilon_k$$

for  $\mu_{G^2}$ -almost all  $(h, k) \in G^2$ . By Lemma B.1, this means that there is an internal homomorphism  $\tilde{\varepsilon}: G \rightarrow {}^*H$  such that  $\varepsilon_h = \tilde{\varepsilon}_h$  for  $\mu_G$ -almost all  $(h, g) \in G^2$ . Thus  $\text{st}\tilde{\varepsilon}_h = 1$  for  $\mu_G$ -almost all  $h \in G$ , and hence for all  $h \in G$  by the homomorphism property; thus  $\tilde{\varepsilon}_h$  always stays infinitesimally close to the identity. Since the Lie group  $H$  contains no small subgroups (e.g., see [33, Exercise 1.4.19]), the nonstandard group  ${}^*H$  contains no infinitesimal internal subgroups, thus the internal homomorphism  $h \mapsto \tilde{\varepsilon}_h$  must be trivial. We conclude that  $\partial_h \phi(x) = 1$  for  $\mu_G$ -almost all  $x \in G$ , hence by Lemma A.1(ii) and a change of variables we have  $\phi(x) = \phi(y)$  for  $\mu_{G^2}$ -almost all  $x, y \in G$ . By another application of Lemma A.1(ii), we conclude that there exists  $y \in G$  such that  $\phi(x) = \phi(y)$  for  $\mu_G$ -almost all  $y \in G$ , and such that  $\text{st}\phi(y) = 1$ , and the claim follows.  $\square$

**Corollary B.3** (Nearby polynomial maps differ by a constant). *Let  $G$  be a nonstandard finite additive group, and  $H/\Gamma$  a standard filtered nilmanifold. If  $\phi_1, \phi_2: G \rightarrow {}^*(H/\Gamma)$  are internal polynomial maps such that  $\text{st}\phi_1(x) = \text{st}\phi_2(x)$  for  $\mu_G$ -almost all  $x \in G$ , then there exists  $c \in {}^*H$  with  $\text{st}c = 1$  such that  $\phi_1 = c\phi_2$ . In particular, if  $\phi_1, \phi_2$  agree at one point, then they are identical.*

*Proof.* By hypothesis, we can find an internal map  $\phi: G \rightarrow {}^*H$  such that  $\phi_1(x) = \phi(x)\phi_2(x)$  and  $\text{st}\phi(x) = 1$  for  $\mu_G$ -almost all  $x \in G$ . We claim that for any standard  $k \geq 0$ , one has

$$\partial_{h_1} \dots \partial_{h_k} \phi(x) \in {}^*H_k$$

for  $\mu_{G^{k+1}}$ -almost all  $(x, h_1, \dots, h_k) \in G^{k+1}$ . This in turn will follow if we can show that the tuple

$$\Phi_{x,h} := (\phi(x + h \cdot \omega))_{\omega \in \{0,1\}^k} \in ({}^*H)^{\{0,1\}^k}$$

lies in  $\text{HK}^k(*H)$  for  $\mu_{G^{k+1}}$ -almost all  $(x, h_1, \dots, h_k) \in G^{k+1}$ . But by Lemma A.1(ii), we see that for  $\mu_{G^{k+1}}$ -almost all  $(x, h_1, \dots, h_k) \in G^{k+1}$ , this tuple  $\Phi_{x,h}$  is infinitesimally close to the identity and obeys the identity

$$\Phi_{1,x,h} = \Phi_{x,h} \Phi_{2,x,h} \quad (46)$$

in  $(*H)^{\{0,1\}^k} / (*\Gamma)^{\{0,1\}^k}$ , where for  $i = 1, 2$  we define the points  $\Phi_{i,x,h} \in (*H)^{\{0,1\}^k} / (*\Gamma)^{\{0,1\}^k}$  by the formula

$$\Phi_{i,x,h} := (\phi_i(x + h \cdot \omega))_{\omega \in \{0,1\}^k}.$$

Since  $\phi_1, \phi_2$  are internal polynomials, In particular,  $\Phi_{1,x,h}$  and  $\Phi_{2,x,h}$  are infinitesimally close in  $\text{HK}^k(*H)/\text{HK}(*\Gamma)$  (which we identify with a subset of  $(*H)^{\{0,1\}^k} / (*\Gamma)^{\{0,1\}^k}$  in the obvious fashion); since  $\Gamma$  is cocompact in  $H$ , we can then write

$$\Phi_{2,x,h} = \Psi_{x,h}(*\Gamma)^{\{0,1\}^k}; \quad \Phi_{1,x,h} = \varepsilon_{x,h} \Psi_{x,h}(*\Gamma)^{\{0,1\}^k}$$

for some  $\varepsilon_{x,h}, \Psi_{x,h} \in \text{HK}^k(*H)$  with  $\varepsilon_{x,h}$  infinitesimally close to the identity and  $\Psi_{x,h}$  bounded. Inserting this back into (46) and rearranging, we conclude that

$$\Psi_{x,h}^{-1} \varepsilon_{x,h}^{-1} \Phi_{x,h} \Psi_{x,h} \in (*\Gamma)^{\{0,1\}^k}.$$

The left-hand side is infinitesimally close to the identity. Since  $\Gamma$  is discrete, the left-hand side must therefore equal the identity. Thus  $\Phi_{x,h} = \varepsilon_{x,h}$ , and thus  $\Phi_{x,h}$  lies in  $\text{HK}^k(*H)$  as claimed.

Applying Lemma B.2, we conclude that there exists  $c \in *H$  with  $\text{st}c = 1$  such that  $\phi_1(x) = c\phi_2(x)$  for  $\mu_G$ -almost all  $x \in G$ . By Lemma A.1(ii), this implies that for all  $x \in G$ , we have for  $\mu_{G^{s+1}}$ -almost all  $h_1, \dots, h_{s+1} \in G$  that

$$\phi_1(x + \omega \cdot h) = c\phi_2(x + \omega \cdot h)$$

for all  $\omega \in \{0, 1\}^{s+1} \setminus \{0\}^{s+1}$ . Since  $\phi_1, c\phi_2$  are both polynomial maps, one has

$$\partial_{h_1} \dots \partial_{h_{s+1}} \phi_1(x) = \partial_{h_1} \dots \partial_{h_{s+1}} c\phi_2(x) = 1$$

and hence  $\phi_1(x) = \phi_2(x)$  for all  $x \in G$ , giving the claim.  $\square$

**Lemma B.4** (Cocycle triviality on vector spaces). *Let  $V$  be a standard finite-dimensional vector space, let  $G$  be a nonstandard finite additive group, let  $E$  be an internal subset of  $G$  of full Loeb measure. Let*

$$c: \{(h, k) \in E \times E : h + k \in E\} \rightarrow *V$$

*be an internal function obeying the cocycle equation*

$$c(h, k) + c(h + k, l) = c(h, k + l) + c(k, l) \quad (47)$$

*whenever  $h, k, l \in E$  are such that  $h + k, k + l, h + k + l \in E$ . Then there exists an internal function  $f: E \rightarrow *V$  such that*

$$c(h, k) = f(h + k) - f(h) - f(k). \quad (48)$$

Furthermore, if we place a norm  $\|\cdot\|_V$  on  $V$  (and hence an internal norm  $\|\cdot\|_{*V}$  on  ${}^*V$ ) we can choose  $f$  so that  $\|f\|_\infty \leq 2\|c\|_\infty$ , where

$$\|f\|_\infty := \sup_{h \in E} \|f(h)\|_{*V}$$

and

$$\|c\|_\infty := \sup_{h,k \in E: h+k \in E} \|c(h,k)\|_{*V}$$

and we use the internal supremum.

*Proof.* Fix  $V$ , and let  $\varepsilon > 0$  be a sufficiently small standard constant. It suffices to prove the following standard version of the above statement: if  $G$  is a standard finite additive group,  $E$  is a subset of  $G$  of density at least  $1 - \varepsilon$ , and  $c: \{(h,k) \in E \times E : h+k \in E\} \rightarrow V$  obeys (47) whenever  $h,k,l \in E$  are such that  $h+k, k+l, h+k+l \in E$ , then there exists  $f: E \rightarrow V$  with  $\|f\|_\infty \leq 2\|c\|_\infty$  obeying (48).

Recall that in this paper we use the standard asymptotic notation of using  $O(X)$  to denote a quantity bounded in magnitude by  $CX$  for some absolute constant  $C$ . Define the averaged function  $f: E \rightarrow V$  by the formula

$$f(h) := \frac{1}{|E|} \sum_{k \in E} c(h,k).$$

For any  $h,k \in E$  with  $h+k \in E$ , we average (47) over all  $l \in E$  with  $k+l, h+k+l \in E$ , which is a subset of  $E$  that omits only  $O(\varepsilon|E|)$  of the elements of  $E$ , to conclude that

$$c(h,k) + f(h+k) = f(h) + f(k) + O(\varepsilon\|c\|_\infty)$$

for all such  $h,k$ . In particular, if  $\varepsilon$  is small enough, the function

$$c'(h,k) := c(h,k) + f(h+k) - f(h) - f(k)$$

is such that  $\|c'\|_\infty \leq \frac{1}{2}\|c\|_\infty$ ; also  $\|f\|_\infty \leq \|c\|_\infty$ . Because  $c$  obeys (47), the function  $c'$  does also. If one iterates this construction and sums the telescoping series, one obtains the claim.  $\square$

## B.1 Main argument

We now establish the following proposition, which implies Lemma 7.3 as a special case:

**Proposition B.5** (Stability of polynomials). *Let  $G$  be a nonstandard finite additive group, let  $H/\Gamma$  be a prefiltered nilmanifold of some degree  $s \geq 0$ , and let  $\tilde{g}: G \rightarrow H/\Gamma$  be an almost polynomial map. Then there exists an internal polynomial map  $g: G \rightarrow {}^*(H/\Gamma)$  such that*

$$\tilde{g}(x) = \text{stg}(x) \tag{49}$$

for  $\mu_G$ -almost all  $x$ .

The rest of this section is devoted to a proof of this proposition. We induct on  $s$ . When  $s = 0$  then  $\tilde{g}$  is almost always constant in the sense that  $\partial_h g(x) = 1$  for  $\mu_{G^2}$ -almost all  $h,x$  and the claim is clear by repeating the final part of the proof of Lemma B.2, so suppose  $s \geq 1$  and the claim has already been proven for  $s - 1$ .

It is convenient to make some reductions regarding the nature of the prefiltered nilmanifold  $H/\Gamma$ :

**Proposition B.6.** *To establish Proposition B.5 for a given choice of  $s$ , we may assume without loss of generality that  $H = H_0 = H_1$  (i.e., that the prefiltration on  $H$  is a filtration) and that  $\Gamma_s = \Gamma \cap H_s$  is trivial (and hence  $H_s$  is now a central compact subgroup of  $H$ ).*

*Proof.* By construction,  $\tilde{g}$  almost always takes values in  $H_0/\Gamma$ , so after adjusting  $\tilde{g}$  on a  $\mu_G$ -null set and redefining  $H$  we may assume without loss of generality that  $H = H_0$ . Quotienting out the group  $H_1$  (which is now normal), we conclude that  $\tilde{g}$  is almost always constant modulo  $H_1$ , so by modification on a further null set we may assume that  $\tilde{g}$  takes values on a single coset of  $H_1$ . Applying a translation we can move this coset to the origin, and so after redefining  $H = H_0$  we may assume that  $H = H_0 = H_1$ . Among other things, this implies that  $H_s$  is now central. One can now quotient both  $H$  and  $\Gamma$  by the central group  $\Gamma \cap H_s$  without affecting the polynomial structure of  $H/\Gamma$ , so one may assume without loss of generality that  $\Gamma_s = \Gamma \cap H_s$  is trivial, as claimed.  $\square$

Henceforth we assume that the conclusions of the above proposition hold. We now use the induction hypothesis to reduce the task of proving Proposition B.5 to that of solving a certain lifting problem. Let  $\pi: H/\Gamma \rightarrow H/H_s\Gamma$  be the projection map, which is a nilspace morphism. We can identify  $H/H_s\Gamma$  with a prefiltered nilmanifold of degree  $s-1$ ; one can think of  $H/\Gamma$  as a principal  $H_s$ -bundle over  $H/H_s\Gamma$ . The function  $\pi \circ \tilde{g}$  is an almost polynomial map from  $G$  into this nilmanifold, thus by induction hypothesis there is an internal polynomial map  $\bar{g}: G \rightarrow {}^*(H/H_s\Gamma)$  such that

$$\pi \circ \tilde{g}(x) = \text{st}\bar{g}(x) \quad (50)$$

for  $\mu_G$ -almost all  $x \in G$ . Thus, if for every  $x \in G$  we define the nonstandard  $H_s$ -fibre  $F_x \subset {}^*(H/\Gamma)$  as

$$F_x := \{y \in {}^*(H/\Gamma) : {}^*\pi(y) = \bar{g}(x)\}$$

then  $F_x$  depends internally on  $x$ , and from (50) we see that  $\tilde{g}(x) \in \text{st}(F_x)$  for  $\mu_G$ -almost all  $x \in G$ . Observe that each  $F_x$  is a free orbit (or *torsor*) of  ${}^*H_s$ : thus if  $y, z \in F_x$  then there is a unique  $h_s \in {}^*H_s$  such that  $h_sy = z$ . Let  $\mathcal{S}$  denote the (internal) set of all internal functions  $g: G \rightarrow {}^*H/\Gamma$  that are “lifts” of  $\bar{g}$  in the sense that  $g(x) \in F_x$  for all  $x$ . This is a torsor of the internal functions from  $G$  to  ${}^*H_s$ ; indeed, if  $g \in \mathcal{S}$  and  $h_s: G \rightarrow {}^*H_s$  is internal then  $h_sg \in \mathcal{S}$ , and conversely if  $g, g' \in \mathcal{S}$  then there is a unique internal  $h_s: G \rightarrow {}^*H_s$  such that  $g' = h_sg$ . It will suffice to find an element  $g \in \mathcal{S}$  that is an internal polynomial and such that  $\text{st}g(x) = \tilde{g}(x)$  for  $\mu_G$ -almost all  $x \in G$ .

To find this lift we proceed in several stages. We first establish a weaker claim in which the requirement of being an (internal) polynomial is removed:

**Lemma B.7** (Existence of a non-polynomial lift near  $\tilde{g}$ ). *There exists  $g^{**} \in \mathcal{S}$  (not necessarily a polynomial) such that  $\text{st}g^{**}(x) = \tilde{g}(x)$  for  $\mu_G$ -almost all  $x \in G$ .*

*Proof.* As  $\tilde{g}$  is Loeb measurable, we can find an internal function  $g^*: G \rightarrow {}^*H/\Gamma$  such that  $\tilde{g}(x) = \text{st}g^*(x)$  for  $\mu_G$ -almost all  $x$ , and hence

$$\text{st}\bar{g}(x) = \text{st}{}^*\pi(g^*(x))$$

for  $\mu_G$ -almost all  $x \in G$ . Since  $\bar{g}$  and  ${}^*\pi \circ g^*$  are both internal functions, we can thus write

$$\bar{g}(x) = \bar{c}(x) {}^*\pi(g^*(x))$$

for all  $x \in G$ , where  $\bar{c}: G \rightarrow {}^*(H/H_s)$  is an internal function with  $\text{st}\bar{c}(x) = 1$  for  $\mu_G$ -almost all  $x \in G$ . We can lift this up to an internal function  $c: G \rightarrow {}^*H$  with  $\text{st}c(x) = 1$  for  $\mu_G$ -almost all  $x \in G$ ; if we then define  $g^{**} := cg^*$ , we see that  $g^{**}$  obeys the stated properties.  $\square$

Next, we identify the set  $\mathcal{S}$  with a certain type of group homomorphism. For any  $x, h \in G$ , let  $\text{Hom}(F_x \rightarrow F_{x-h})$  denote the space of internal functions  $\phi: F_x \rightarrow F_{x-h}$  which are  ${}^*H_s$ -equivariant in the sense that  $\phi(h_s y) = h_s \phi(y)$  for any  $h_s \in {}^*H_s$  and  $y \in F_x$ . These spaces give the fibration  $(F_x)_{x \in G}$  the structure of a groupoid, since we obviously have a composition law

$$\circ: \text{Hom}(F_{x-h} \rightarrow F_{x-h-k}) \times \text{Hom}(F_x \rightarrow F_{x-h}) \rightarrow \text{Hom}(F_x \rightarrow F_{x-h-k})$$

for any  $x, h, k \in {}^*G$ , obeying the following groupoid axioms:

- (Identity) The identity map  $\text{id}_x \in \text{Hom}(F_x \rightarrow F_x)$  is such that  $\phi \circ \text{id}_x = \text{id}_{x-h} \circ \phi = \phi$  for all  $x, h \in G$  and  $\phi \in \text{Hom}(F_x \rightarrow F_{x-h})$ .
- (Associativity) For any  $x, h, k, l \in G$  and  $\phi_1 \in \text{Hom}(F_{x-h-k} \rightarrow F_{x-h-k-l})$ ,  $\phi_2 \in \text{Hom}(F_{x-h} \rightarrow F_{x-h-k})$ ,  $\phi_3 \in \text{Hom}(F_x \rightarrow F_{x-h})$ , one has  $\phi_1 \circ (\phi_2 \circ \phi_3) = (\phi_1 \circ \phi_2) \circ \phi_3$ .
- (Inverse) For any  $x, h \in G$  and  $\phi \in \text{Hom}(F_x \rightarrow F_{x-h})$ , there exists  $\phi^{-1} \in \text{Hom}(F_{x-h} \rightarrow F_x)$  such that  $\phi \circ \phi^{-1} = \text{id}_{x-h}$  and  $\phi^{-1} \circ \phi = \text{id}_x$ .

. For each  $x \in G$ , one can canonically identify  $\text{Hom}(F_x \rightarrow F_x)$  with the group  ${}^*H_s$ , since the  ${}^*H_s$ -equivariant automorphisms of  $F_x$  are given precisely by left multiplication by elements of  ${}^*H_s$ . Note that if  $y \in F_x$  and  $z \in F_{x-h}$  there is a unique element of  $\text{Hom}(F_x \rightarrow F_{x-h})$  that maps  $y$  to  $z$ ; we will denote this map by  ${}^*\pi^\square(y, z)$  (the reason for the notation  $\pi^\square$  will be clearer later). Thus we have

$${}^*\pi^\square(y, z) = {}^*\pi^\square(h_s y, h_s z) \quad (51)$$

for any  $h_s \in {}^*H_s$ ,  $y \in F_x$ , and  $z \in F_{x-h}$ , and

$${}^*\pi^\square(z, w) \circ {}^*\pi^\square(y, z) = {}^*\pi^\square(y, w) \quad (52)$$

whenever  $y \in F_x$ ,  $z \in F_{x-h}$ , and  $w \in F_{x-h-k}$ .

Now let  $\mathcal{G}$  denote the collection of all pairs  $(h, \phi)$  where  $h \in G$  and  $\phi: G \mapsto \bigsqcup_{x \in G} \text{Hom}(F_x \rightarrow F_{x-h})$  is an internal function such that  $\phi(x) \in \text{Hom}(F_x \rightarrow F_{x-h})$  for all  $x \in G$ . We can define a group law

$$(k, \psi)(h, \phi) := (h + k, \psi(\cdot - h) \circ \phi). \quad (53)$$

One easily verifies that this gives  $\mathcal{G}$  the structure of an (internal) group. Observe that for any  $g \in \mathcal{S}$ , one can define an internal group homomorphism  $\partial g: G \rightarrow \mathcal{G}$  by the formula

$$\partial g(h) := (h, \partial_h g)$$

where the function  $\partial_h g$  is defined by  $\partial_h g(x) := {}^*\pi^\square(g(x), g(x-h))$  for any  $x, h \in G$ . From (52) one easily verifies that  $\partial g$  is indeed an internal homomorphism; from (51) we also see that  $\partial(h_s g) = \partial(g)$  whenever  $h_s \in {}^*H_s$  is a constant. In the converse direction, if one has an internal group homomorphism  $h \mapsto (h, \phi_h)$

from  $G$  to  $\mathcal{G}$ , then we claim that this homomorphism must be of the form  $\partial g$  for some  $g \in \mathcal{S}$ , which is unique up to left-multiplication by constants  $h_s \in {}^*H_s$ . Indeed, if we let  $y_0$  be an arbitrary element of  $F_0$ , we can define  $g(h)$  to be the unique element of  $F_h$  such that

$$\phi_{-h}(0) = \pi^\square(y_0, g(h)) \quad (54)$$

(i.e.,  $g(h)$  is the image of  $y_0$  under the equivariant map  $\phi_{-h}(0) \in \text{Hom}(F_0 \rightarrow F_h)$ ). For any  $h, k \in G$ , one sees from (53) and the group homomorphism hypothesis that

$$\phi_{h+k}(0) = \phi_k(-h) \circ \phi_h(0)$$

and thus by (52), (54)

$$\phi_k(-h) = {}^*\pi^\square(g(-h), g(-h-k)).$$

From this and a change of variables we see that  $\partial g(h) = (h, \phi_h)$ , giving existence. If we have  $\partial g = \partial g'$  for some other  $g' \in \mathcal{S}$ , then writing  $g' = h_s g$  for some internal  $h_s: G \rightarrow {}^*H_s$ , we see from (the converse of) (51) that  $h_s(x-h) = h_s(x)$  for all  $x, h \in G$ , and hence  $h_s$  is constant, giving the uniqueness claim.

Let  $g^{**}$  be as in Lemma B.7. Define the external subgroup  $\mathcal{G}^{**}$  of  $\mathcal{G}$  to be the set of all pairs  $(h, \phi) \in \mathcal{G}$  such that  $\text{st}\phi(x) = \text{st}\partial_h g^{**}(x)$  for  $\mu_G$ -almost all  $x \in G$ . Thus for instance  $(h, \partial_h g^{**}) \in \mathcal{G}^{**}$  for every  $h \in G$ . It is easy to verify that  $\mathcal{G}^{**}$  is an external subgroup of  $\mathcal{G}$ .

It will now suffice to show

**Proposition B.8** (Existence of an exact group homomorphism). *There exists an internal group homomorphism from  $G$  to  $\mathcal{G}^{**}$  of the form  $h \mapsto (h, \phi_h)$ , which is also “exact” in the sense that it is of the form  $\partial g$  for some internal polynomial  $g \in \mathcal{S}$ .*

Indeed,  $\phi_h, g$  obeyed the conclusions of Proposition B.8, then by Lemma A.1(ii) there exists  $x_0 \in G$  such that  $\text{st}\partial_h g(x_0) = \text{st}\phi_h(x_0) = \text{st}\partial_h g^{**}(x_0)$  for  $\mu_G$ -almost all  $h \in G$ . Writing  $g = h_s g^{**}$  for some internal  $h_s: G \rightarrow {}^*H_s$ , this implies that  $\text{st}h_s(x_0 - h) = \text{st}h_s(x_0)$  for  $\mu_G$ -almost all  $h \in G$ , thus the internal polynomial map  $g' := h_s(x_0)^{-1}g$  is an element of  $\mathcal{S}$  with  $\text{st}g'(x) = \text{st}g^{**}(x) = \tilde{g}(x)$  for  $\mu_G$ -almost all  $x \in G$ .

It remains to establish Proposition B.8. The next step is to interpret the internal maps  $\phi: G \rightarrow \biguplus_{x \in G} \text{Hom}(F_x \rightarrow F_{x-h})$  appearing in the definition of  $\mathcal{G}$  as maps between filtered spaces, using the algebraic formalism from [17, §7], in order to reformulate the exactness condition in a more tractable form. We form the prefiltered group  $H^\square$  to be the group  $H \times H$  equipped with the prefiltration

$$(H^\square)_i := \{(h_i, h_{i+1}h_i) : h_i \in H_i, h_{i+1} \in H_{i+1}\}.$$

The significance of this group is that we have the identity

$$\text{HK}^{k+1}(H) = \text{HK}^k(H^\square) \quad (55)$$

for all  $k \geq 0$ , by using the identification

$$(h_\omega)_{\omega \in \{0,1\}^{k+1}} \equiv ((h_{\omega,0}, h_{\omega,1}))_{\omega \in \{0,1\}^k}.$$

Indeed, this can be checked by noting that every generator of  $\text{HK}^{k+1}(H)$  can be expressed (using this identification) as a finite combination of generators of  $\text{HK}^k(H^\square)$ , and conversely. We define  $\Gamma^\square$  similarly.

It is easy to verify that  $H^\square/\Gamma^\square$  is a prefiltered nilmanifold of degree  $s$  (as a nilmanifold it is just the Cartesian square of  $H/\Gamma$ ). Note that  $H_s^\square = \{(h_s, h_s) : h_s \in H_s\}$  is a central subgroup of  $H^\square$  (and  $\Gamma_s^\square$  is trivial). We let  $\pi^\square : H^\square/\Gamma^\square \rightarrow H^\square/(\Gamma^\square)_s$  be the projection map; this is a nilspace morphism to a prefiltered nilmanifold  $H^\square/(\Gamma^\square)_s$  of degree  $s-1$ .

Observe that if  $y \in F_x$  and  $z \in F_{x-h}$  for some  $x, h \in G$ , then  ${}^*\pi^\square(y, z)$ , by definition, is the equivalence class  $\{(h_s y, h_s z) : h_s \in {}^*H_s\}$ . By abuse of notation we may therefore identify  ${}^*\pi^\square(y, z)$  with the unique element of  $\text{Hom}(F_x \rightarrow F_{x-h})$  mapping  $y$  to  $z$ , thus making this use of the notation  ${}^*\pi^\square(y, z)$  consistent with the one used earlier. In particular, we can view  $\bigsqcup_{x \in G} \text{Hom}(F_x \rightarrow F_{x-h})$  as an internal subset of the internal filtered nilmanifold  ${}^*\pi^\square(H^\square/\Gamma^\square)$ , and it now makes sense to ask when the map  $\phi : G \mapsto \bigsqcup_{x \in G} \text{Hom}(F_x \rightarrow F_{x-h})$  is an internal polynomial or not.

We define  $\mathcal{G}_0$  to be the internal subgroup of  $\mathcal{G}$  consisting of those pairs  $(h, \phi) \in \mathcal{G}$  with  $\phi$  an internal polynomial. It is easy to see that this is indeed an internal subgroup.

We now make the following key observations, which generalise the familiar fact that a (differentiable) function of one variable is a polynomial of degree  $s$  if and only if its derivative is a polynomial of degree  $s-1$ :

- (i) If  $g \in \mathcal{S}$  is a polynomial, then  $\partial g$  takes values in  $\mathcal{G}_0$  (i.e.,  $\partial_h g$  is a polynomial for every  $h \in G$ ).
- (ii) Conversely, if  $g \in \mathcal{S}$  is such that  $\partial g$  takes values in  $\mathcal{G}_0$ , then  $g$  is a polynomial.

The direction (i) is immediate from (55) (and its counterpart for  $\Gamma$ ). For the converse direction (ii), suppose that  $\partial_h g$  is a polynomial for every  $h \in G$ . Expanding out the definitions, we conclude in particular that

$$\left({}^*\pi^\square((g(x + \omega \cdot h), g(x + \omega \cdot h + h_{s+1})))\right)_{\omega \in \{0,1\}^s} \in \text{HK}^s({}^*\pi^\square(H^\square/\Gamma^\square))$$

for all  $x, h_1, \dots, h_{s+1} \in G$ , where  $h := (h_1, \dots, h_s)$ . Since  $\pi^\square$  quotients out  $H^\square/\Gamma^\square$  by the central order  $s$  group  $(H^\square)_s$ , we conclude from the definition of  $\text{HK}^s$  that

$$((g(x + \omega \cdot h), g(x + \omega \cdot h + h_{s+1})))_{\omega \in \{0,1\}^s} \in \text{HK}^s(H^\square/\Gamma^\square).$$

By (55) this implies that

$$(g(x + \omega' \cdot h'))_{\omega' \in \{0,1\}^{s+1}} \in \text{HK}^{s+1}(H/\Gamma)$$

for all  $x, h_1, \dots, h_{s+1} \in G$ , where  $h' := (h_1, \dots, h_{s+1})$ . Because  $H/\Gamma$  is a nilspace of degree at most  $s$  (see [20, Proposition 3.10]), this implies that

$$(g(x + \omega_1 h_1 + \dots + \omega_k h_k))_{\omega_1, \dots, \omega_k \in \{0,1\}} \in \text{HK}^k(H/\Gamma)$$

for any standard  $k \geq 0$  and  $x, h_1, \dots, h_k \in G$ . Thus  $g$  is a polynomial as required.

In view of the above equivalence, the task of proving Proposition B.8 (and hence Proposition B.5) now reduces to establishing

**Proposition B.9** (Existence of a group homomorphism). *There exists an internal group homomorphism from  $G$  to  $\mathcal{G}_0 \cap \mathcal{G}^{**}$  of the form  $h \mapsto (h, \phi_h)$ .*

By hypothesis,  $\tilde{g}$  is an almost polynomial, hence  $\text{st}g^{**}$  is also (since these two maps agree  $\mu_G$ -almost everywhere). From (55) and Lemma A.1(ii) we conclude that  $\text{st}\partial_h g^{**}: G \rightarrow \pi^\square(H^\square/\Gamma^\square)$  is an almost polynomial map for  $\mu_G$ -almost all  $h \in G$ . Applying the induction hypothesis (and the axiom of choice), we conclude

**Proposition B.10** (First approximation to  $\phi_h$ ). *For  $\mu_G$ -almost all  $h \in G$ , there exists an internal polynomial  $\phi'_h: G \rightarrow {}^*\pi^\square(H^\square/\Gamma^\square)$  such that  $\text{st}\partial_h g^{**}(x) = \text{st}\phi'_h(x)$  for  $\mu_G$ -almost all  $x \in G$ . In particular,  $(h, \phi'_h) \in \mathcal{G}_0$ .*

This begins to look like what we want; however, there are currently three undesirable features of this map  $\phi'_h$  that prevent us from simply taking  $\phi_h$  to equal  $\phi'_h$ :

- (i) It is not necessarily the case that  $(h, \phi'_h)$  lies in  $\mathcal{G}_0$ , because  $\phi'_h(x)$  need not lie in  $\text{Hom}(F_x \rightarrow F_{x-h})$ .
- (ii) Even if (i) was resolved, it is not necessarily the case that  $\phi'_h$  depends in an internal fashion on  $h$  (due to the use of the axiom of choice).
- (iii) Even if (i) and (ii) were resolved, it is not necessarily the case that the map  $h \mapsto (h, \phi'_h)$  is a group homomorphism.

We now address each of these issues in turn. We begin by resolving (i):

**Proposition B.11** (Second approximation to  $\phi_h$ ). *For almost all  $h \in G$ , there exists an internal polynomial  $\phi''_h \in \text{Hom}(F_x \rightarrow F_{x-h})$  such that  $\text{st}\partial_h g^{**}(x) = \text{st}\phi''_h(x)$  for  $\mu_G$ -almost all  $x \in G$ . In particular,  $(h, \phi''_h) \in \mathcal{G}_0 \cap \mathcal{G}^{**}$ .*

*Proof.* Let  $\phi'_h$  be as in Proposition B.10. By Lemma A.1(ii), we know that for  $\mu_{G^2}$ -almost all  $(x, h) \in G$ , that  $\text{st}\partial_h g^{**}(x) = \text{st}\phi'_h(x)$ . Since  $\partial_h g^{**}(x) \in \text{Hom}(F_x \rightarrow F_{x-h})$ , we thus see that

$$\text{st}^* \tilde{\pi}(\phi'_h(x)) = \text{st}^* \tilde{\pi}(\partial_h g^{**}(x)) = \text{st}(\bar{g}(x), \bar{g}(x-h))$$

where  $\tilde{\pi}: \pi^\square(H^\square/\Gamma^\square) \rightarrow \pi(H/\Gamma) \times \pi(H/\Gamma)$  is the projection map

$$\tilde{\pi}: \pi^\square(y, z) \mapsto (\pi(y), \pi(z)).$$

We conclude that

$${}^* \tilde{\pi}(\phi'_h(x)) = (\bar{c}_1(x, h)\bar{g}(x), \bar{c}_2(x, h)\bar{g}(x-h))$$

for all  $x, h \in G$  and some internal functions  $\bar{c}_1, \bar{c}_2: G \times G \rightarrow {}^*\pi(H)$  with the property that  $\text{st}\bar{c}_1(x, h) = \text{st}\bar{c}_2(x, h) = 1$  for  $\mu_{G^2}$ -almost all  $(x, h) \in G$ . Using some arbitrary continuous local lift of  $\pi(H)$  up to  $H$  around the identity, we can write  $\bar{c}_i = \pi(c_i)$  for  $i = 1, 2$  and some internal functions  $c_1, c_2: G \times G \rightarrow {}^*H$  with  $\text{st}c_1(x, h) = \text{st}c_2(x, h) = 1$  for  $\mu_{G^2}$ -almost all  $(x, h) \in G$ . If we then set

$$\phi''_h(x) := {}^*\pi^\square(c_1(x, h), c_2(x, h))^{-1} \phi'_h(x)$$

then  $(h, \phi''_h)$  is now an element of  $\mathcal{G}_0$  for every  $h$ , and  $\text{st}\partial_h g^{**}(x) = \text{st}\phi'_h(x) = \text{st}\phi''_h(x)$  for  $\mu_{G^2}$ -almost all  $(x, h) \in G^2$ , so that  $(h, \phi''_h) \in \mathcal{G}^{**}$  for  $\mu_G$ -almost all  $h \in G$ . The claim follows.  $\square$

Now we address the difficulty (ii).

**Proposition B.12** (Third approximation to  $\phi_h$ ). *There exists an internal subset  $E$  of  $G$  of full Loeb measure and an internal map  $h \mapsto (h, \phi_h''')$  from  $E$  to  $\mathcal{G}_0 \cap \mathcal{G}^{**}$ .*

*Proof.* Let  $\phi_h''$  be as in Proposition B.11. By Lemma A.1(ii), there exists  $x_0 \in G$  such that  $\text{st} \partial_h g^{**}(x_0) = \text{st} \phi_h''(x_0)$  for  $\mu_G$ -almost all  $h \in G$ . Since  $\partial_h g^{**}(x_0)$  and  $\phi_h''(x_0)$  both lie in  $\text{Hom}(F_{x_0} \rightarrow F_{x_0-h})$ , there exists an internal function  $c: G \rightarrow {}^*H_s$  such that  $\partial_h g^{**}(x_0) = c(h) \phi_h''(x_0)$  for all  $h \in G$ ; by the construction of  $x_0$ , we have  $\text{st} c(h) = 1$  for  $\mu_G$ -almost all  $h \in G$ .

Fix a metric  $d$  on  $\pi^\square(H^\square/\Gamma^\square)$ , and for any  $\varepsilon > 0$ , let  $E_\varepsilon \subset G$  be the set of all  $h \in G$  such that there exists  $(h, \phi_h''') \in \mathcal{G}_0$  with  $\phi_h'''(x_0) = \partial_h g^{**}(x_0)$ , and such that

$${}^*d(\phi_h'''(x), \partial_h g^{**}(x)) \leq \varepsilon \quad (56)$$

for a set of  $x$  in  $G$  of Loeb measure at least  $1 - \varepsilon$ . Clearly  $E_\varepsilon$  is an internal subset of  $G$ ; from the above discussion (choosing  $\phi_h''' = c(h) \phi_h''$ ) we see that  $E_\varepsilon$  has internal measure at least  $1 - \varepsilon$  for every standard  $\varepsilon > 0$ , hence by the underspill principle (see e.g., [38, Corollary 5.2(iv)]) we can find an infinitesimal  $\varepsilon > 0$  such that  $E = E_\varepsilon$  has internal measure at least  $1 - \varepsilon$ , and hence full Loeb measure. Fixing such an  $\varepsilon$ , we now see from Corollary B.3 that if  $h \in E$ , then the function  $\phi_h'''$  defined above is unique, so  $\phi_h'''$  depends internally on  $h$  on this domain. By construction  $h \mapsto (h, \phi_h''')$  is now an internal map from  $E$  to  $\mathcal{G}_0 \cap \mathcal{G}^{**}$ , and the claim follows.  $\square$

Now we address the issue (iii).

**Proposition B.13** (Fourth approximation to  $\phi_h$ ). *There exists an internal subset  $E$  of  $G$  of full Loeb measure and internal homomorphism  $h \mapsto (h, \phi_h''')$  from  $E$  to  $\mathcal{G}_0 \cap \mathcal{G}^{**}$ .*

*Proof.* Let  $\phi_h'''$  and  $E$  be as in Proposition B.12. If  $h, k, h+k \in E$ , we have

$$(h+k, \phi_{h+k}''') = (0, c(h, k))(h, \phi_h''')(k, \phi_k''') \quad (57)$$

for some  $(0, c(h, k)) \in \mathcal{G}_0 \cap \mathcal{G}^{**}$ , and from (56) we see that  $\text{st} c(h, k)(x) = 1$  for  $\mu_G$ -almost all  $x \in G$ . Applying Corollary B.3, we conclude that  $c(h, k) \in {}^*H_s$  is a constant independent of  $x$  with  $\text{st} c(h, k) = 1$ . In particular the group element  $(0, c(h, k))$  appearing in (57) is central in  $\mathcal{G}_0$ .

Now suppose that  $x, h, k, l$  are such that  $h, k, l, h+k, k+l, h+k+l \in E$ . By several applications of (57) we conclude that  $(h+k+l, \phi_{h+k+l}''')$  is equal to both

$$(0, c(h, k)c(h+k, l))(l, \phi_l''')(k, \phi_k''')(h, \phi_h''')$$

and to

$$(0, c(h, k+l)c(k, l))(l, \phi_l''')(k, \phi_k''')(h, \phi_h''')$$

leading to the cocycle equation

$$c(h, k)c(h+k, l) = c(h, k+l)c(k, l)$$

whenever  $h, k, l, h+k, k+l, h+k+l \in E$ . Taking logarithms (using the fact that  $\text{stc}(h, k) = 1$  and that  $H_s$  is a abelian Lie group) and using Lemma B.4, we conclude that there exists an internal function  $f: E \rightarrow {}^*H_s$  such that

$$c(h, k) = f(h+k)f(h)^{-1}f(k)^{-1}$$

whenever  $h, k, h+k \in E$ , and also  $\text{st}f(h) = 1$  for all  $h \in E$ . If one then defines

$$\phi_h'''(x) := f(h)^{-1}\phi_h'''(x)$$

for  $h \in E$ , we now see that  $h \mapsto (h, \phi_h''')$  is an internal map from  $E_\varepsilon$  to  $\mathcal{G}_0 \cap \mathcal{G}^{**}$  with the property that

$$(h+k, \phi_{h+k}''') = (k, \phi_k''')(h, \phi_h''')$$

whenever  $h, k, h+k \in E$ . □

By Proposition B.13 and Lemma B.1 we thus obtain the required internal homomorphism  $h \mapsto (h, \phi_h)$  from  $G$  to  $\mathcal{G}_0 \cap \mathcal{G}^{**}$ , giving Proposition B.8. The proof of Proposition B.5 is now complete.

## Acknowledgments

We thank Uwe Stroinski and James Leng for corrections, and the anonymous referee for useful suggestions.

## References

- [1] S. L. Berg and M. Henk. Discrete analogues of John’s theorem. *Mosc. J. Comb. Number Theory*, 8(4):367–378, 2019. [13](#)
- [2] V. Bergelson, T. Tao, and T. Ziegler. An inverse theorem for the uniformity seminorms associated with the action of  $\mathbb{F}_p^\infty$ . *Geom. Funct. Anal.*, 19(6):1539–1596, 2010. [8](#), [10](#)
- [3] A. Berger, A. Sah, M. Sawhney, and J. Tidor. Non-classical polynomials and the inverse theorem, 2021. [8](#)
- [4] P. Candela. Notes on nilspaces: algebraic aspects. *Discrete Anal.*, pages Paper No. 15, 59, 2017. [5](#)
- [5] P. Candela, D. González-Sánchez, and B. Szegedy. On higher-order fourier analysis in characteristic  $p$ , 2021. [8](#)
- [6] P. Candela and B. Szegedy. Regularity and inverse theorems for uniformity norms on compact abelian groups and nilmanifolds, 2019. [7](#), [8](#), [10](#), [20](#), [23](#), [27](#), [32](#)
- [7] H. Furstenberg. Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Anal. Math.*, 31:204–256, 1977. [22](#)
- [8] A. M. Gleason. The structure of locally compact groups. *Duke Math. J.*, 18:85–104, 1951. [29](#)

- [9] W. T. Gowers. A new proof of Szemerédi’s theorem for arithmetic progressions of length four. *Geom. Funct. Anal.*, 8(3):529–551, 1998. [15](#)
- [10] W. T. Gowers. A new proof of Szemerédi’s theorem. *Geom. Funct. Anal.*, 11(3):465–588, 2001. [2](#)
- [11] W. T. Gowers and L. Milićević. A quantitative inverse theorem for the  $u^4$  norm over finite fields, 2017. [8](#)
- [12] W. T. Gowers and L. Milićević. An inverse theorem for freiman multi-homomorphisms, 2020. [8](#)
- [13] B. Green. Montréal notes on quadratic Fourier analysis. In *Additive combinatorics*, volume 43 of *CRM Proc. Lecture Notes*, pages 69–102. Amer. Math. Soc., Providence, RI, 2007. [3](#), [4](#)
- [14] B. Green and T. Tao. An inverse theorem for the Gowers  $U^3(G)$  norm. *Proc. Edinb. Math. Soc. (2)*, 51(1):73–153, 2008. [3](#), [4](#), [8](#), [11](#), [14](#), [15](#), [16](#), [17](#), [18](#)
- [15] B. Green and T. Tao. Linear equations in primes. *Ann. of Math. (2)*, 171(3):1753–1850, 2010. [8](#), [27](#), [28](#)
- [16] B. Green and T. Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. *Ann. of Math. (2)*, 175(2):465–540, 2012. [6](#), [8](#)
- [17] B. Green and T. Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. *Ann. of Math. (2)*, 175(2):465–540, 2012. [41](#)
- [18] B. Green, T. Tao, and T. Ziegler. An inverse theorem for the Gowers  $U^4$ -norm. *Glasg. Math. J.*, 53(1):1–50, 2011. [8](#)
- [19] B. Green, T. Tao, and T. Ziegler. An inverse theorem for the Gowers  $U^{s+1}[N]$ -norm. *Ann. of Math. (2)*, 176(2):1231–1372, 2012. [5](#), [6](#), [8](#), [20](#), [32](#)
- [20] Y. Gutman, F. Manners, and P. P. Varjú. The structure theory of nilspaces I. *J. Anal. Math.*, 140(1):299–369, 2020. [6](#), [27](#), [28](#), [42](#)
- [21] K. H. Hofmann, J. R. Liukkonen, and M. W. Mislove. Compact extensions of compactly generated nilpotent groups are pro-Lie. *Proc. Amer. Math. Soc.*, 84(3):443–448, 1982. [29](#)
- [22] B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. Math.*, 161(1):397–488, 2005. [8](#), [9](#)
- [23] A. Jamneshan, O. Shalom, and T. Tao. A Host-Kra structure theorem for arbitrary Conze-Lesigne systems. *forthcoming*, 2021. [8](#), [10](#)
- [24] P. A. Loeb. Conversion from nonstandard to standard measure spaces and applications in probability theory. *Trans. Amer. Math. Soc.*, 211:113–122, 1975. [33](#)
- [25] A. I. Mal’cev. On a class of homogeneous spaces. *Izvestiya Akad. Nauk. SSSR. Ser. Mat.*, 13:9–32, 1949. [21](#)

- [26] F. Manners. Periodic nilsequences and inverse theorems on cyclic groups, 2014. [8](#), [18](#)
- [27] F. Manners. Quantitative bounds in the inverse theorem for the gowers  $u^{s+1}$ -norms over cyclic groups, 2018. [8](#)
- [28] A. Samorodnitsky. Low-degree tests at large distances. In *STOC'07—Proceedings of the 39th Annual ACM Symposium on Theory of Computing*, pages 506–515. ACM, New York, 2007. [4](#), [11](#), [14](#), [15](#)
- [29] O. Shalom. Host-kra factors for  $\bigoplus_{p \in P} \mathbb{Z}/p\mathbb{Z}$  actions and finite dimensional nilpotent systems. *Preprint, arXiv:2105.00446*, 2021. [10](#)
- [30] B. Szegedy. Gowers norms, regularization and limits of functions on abelian groups, 2010. [8](#), [10](#)
- [31] B. Szegedy. On higher order fourier analysis, 2012. [8](#), [10](#)
- [32] E. Szemerédi. On sets of integers containing no  $k$  elements in arithmetic progression. *Acta. Arith.*, 27:199–245, 1975. [2](#)
- [33] T. Tao. *Hilbert’s fifth problem and related topics*, volume 153 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, RI, 2014. [34](#), [36](#)
- [34] T. Tao and V. Vu. John-type theorems for generalized arithmetic progressions and iterated sumsets. *Adv. Math.*, 219(2):428–449, 2008. [13](#)
- [35] T. Tao and V. H. Vu. *Additive combinatorics*, volume 105 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 2010. Paperback edition [of MR2289012]. [3](#), [13](#)
- [36] T. Tao and T. Ziegler. The inverse conjecture for the Gowers norm over finite fields via the correspondence principle. *Anal. PDE*, 3(1):1–20, 2010. [8](#), [10](#), [22](#), [23](#), [32](#)
- [37] T. Tao and T. Ziegler. The inverse conjecture for the Gowers norm over finite fields in low characteristic. *Ann. Comb.*, 16(1):121–188, 2012. [8](#)
- [38] T. Tao and T. Ziegler. Concatenation theorems for anti-Gowers-uniform functions and Host-Kra characteristic factors. *Discrete Anal.*, pages Paper No. 13, 60, 2016. [9](#), [21](#), [24](#), [32](#), [33](#), [44](#)
- [39] H. Towsner. A correspondence principle for the Gowers norms. *J. Log. Anal.* Paper 4, 11 pp., 4, 2012. [23](#)
- [40] H. Yamabe. A generalization of a theorem of Gleason. *Ann. of Math. (2)*, 58:351–365, 1953. [29](#)

AUTHORS

Asgar Jamneshan  
Koç University  
College of Sciences  
Rumeli Feneri Yolu  
34450 Sarıyer  
İstanbul, Türkiye  
ajamneshan@ku.edu.tr  
<https://asgarjam.wixsite.com/home>

Terence Tao  
Department of Mathematics, UCLA  
Los Angeles, CA 90095  
USA  
tao@math.ucla.edu  
<https://www.math.ucla.edu/~tao>