SPHERICAL HIGHER ORDER FOURIER ANALYSIS OVER FINITE FIELDS III: A SPHERICAL GOWERS INVERSE THEOREM

WENBO SUN

Abstract. This paper is the third part of the series *Spherical higher order Fourier analysis over finite fields*, aiming to develop the higher order Fourier analysis method along spheres over finite fields, and to solve the geometric Ramsey conjecture in the finite field setting.

In this paper, we prove an inverse theorem over finite field for spherical Gowers norms, i.e. a local Gowers norm supported on a sphere. We show that if the $(s+1)$-th spherical Gowers norm of a 1-bounded function $f\colon \mathbb{F}_p^d \to \mathbb{C}$ is at least ϵ and if d is sufficiently large depending only on s, then f correlates on the sphere with a p-periodic s-step nilsequence, where the bounds for the complexity and correlation depend only on d and ϵ. This result will be used in later parts of the series to prove the geometric Ramsey conjecture in the finite field setting.

Contents

2020 *Mathematics Subject Classification.* 11T99, 37A99.
The author was partially supported by the NSF Grant DMS-2247331.

1. Introduction

1.1. **The spherical Gowers inverse theorem.**

This paper is the third part of the series *Spherical higher order Fourier analysis over finite fields* [19, 20, 21]. The purpose of this paper is to use the tools developed in [19, 20] to prove a Gowers inverse theorem for spherical sets. We start by recalling the definition of local Gowers norms. For any finite set A and function $f\colon A \to \mathbb{C}$, denote $\mathbb{E}_{n\in A} f(n) := \frac{1}{|A|}\sum_{n\in A} f(n)$.

Definition 1.1 (Local Gowers norms)**.** Let Ω be a subset of $\mathbb{F}_p^d$, $s \in \mathbb{N}_+$, and $f\colon \mathbb{F}_p^d \to \mathbb{C}$ be a function. The *s-th Ω-Gowers norm* of f is defined by the quantity

$$\|f\|_{U^s(\Omega)} := \Big|\mathbb{E}_{(n,h_1,\dots,h_s)\in\Box_s(\Omega)} \prod_{\epsilon=(\epsilon_1,\dots,\epsilon_s)\in\{0,1\}^s} C^{|\epsilon|} f(n + \epsilon_1 h_1 + \dots + \epsilon_s h_s)\Big|^{\frac{1}{2^s}},$$

where

$$\Box_s(\Omega) := \{(n, h_1, \dots, h_s) \in (\mathbb{F}_p^d)^{s+1} \colon n + \epsilon_1 h_1 + \dots + \epsilon_s h_s \in \Omega \text{ for all } (\epsilon_1, \dots, \epsilon_s) \in \{0,1\}^s\},$$

$|\epsilon| := \epsilon_1 + \dots, \epsilon_s$ and $C^{2n+1} f = \overline{f}$, $C^{2n} f = f$ for all $n \in \mathbb{Z}$.[1]

The inverse theorems for Gowers norms is the central part of higher order Fourier analysis, which has been studied extensively in the last two decades, and it has many

[1]One can show that $\|\cdot\|_{U^s(\Omega)}$ is indeed a norm when $s \geq 2$.

applications in Semerédi-type problems. See [1, 2, 3, 5, 6, 7, 8, 9, 11, 12, 13, 14, 22, 23, 24] for a (far from being complete) list of related works. For example, it was shown by Green, Tao and Ziegler [12] (see also [17]) that for $\Omega = \mathbb{F}_p$, for any function $f\colon \Omega \to \mathbb{C}$ bounded by 1 with $\|f\|_{U^{s+1}(\Omega)} \gg 1$, there exists an s-step nilsequence $\phi\colon \Omega \to \mathbb{C}$ of low complexity such that $|\mathbb{E}_{n\in\Omega} f(n)\phi(n)| \gg 1$. This result can be generalized to the case $\Omega = \mathbb{F}_p^d$ without much difficulty.

In this paper, we are particularly interested in the case when

$$\Omega = V(M) := \{n \in \mathbb{F}_p^d \colon M(n) = 0\}$$

is the set of zeros of some quadratic form $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ (see Section 1.3 for definitions), since such a norm is strongly connected to our study of the geometric Ramsey conjecture in [21]. A typical example of M is $M(n) = n \cdot n - r$ for some $r \in \mathbb{F}_p$, in which case $V(M)$ is the sphere of square radius r centered at $\mathbf{0}$. For convenience we call $\|\cdot\|_{U^s(V(M))}$ the *s-th spherical Gowers norm*. Let $\mathrm{Nil}_p^{s;C,1}(\mathbb{F}_p^d)$ denote the set of all scalar valued p-periodic nilsequences of degree at most s and complexity at most C (see Section 2.7 for the precise definition). Throughout this paper, for $s \in \mathbb{Z}$, denote

$$N(s) := (2s + 16)(15s + 453). \tag{1.1}$$

We proof the following spherical Gowers inverse theorem of step s (abbreviated as SGI(s)):

Theorem 1.2. [SGI(s)] For every $d \in \mathbb{N}_+$, $s \in \mathbb{N}$ with $d \geq N(s-1) = (2s+14)(15s+438)$ and $\epsilon > 0$, there exist $\delta := \delta(d, \epsilon), C := C(d, \epsilon) > 0$, and $p_0 := p_0(d, \epsilon) \in \mathbb{N}$ such that for every prime $p \geq p_0$, every non-degenerate quadratic form $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$, and every function $f\colon \mathbb{F}_p^d \to \mathbb{C}$ bounded in magnitude by 1, if $\|f\|_{U^{s+1}(V(M))} > \epsilon$, then there exists $\phi \in \mathrm{Nil}_p^{s;C,1}(\mathbb{F}_p^d)$ such that

$$\left|\mathbb{E}_{n\in V(M)} f(n) \cdot \phi(n)\right| > \delta.$$

Remark 1.3. By considering the cases $p \geq p_0$ and $p < p_0$ separately, one can show that the restriction $p \geq p_0$ in Theorem 1.2 is not necessary. We leave the details to the interested readers.

It is also natural to ask whether one can remove the dependence of δ and C on the dimension d in Theorem 1.2. We do not pursuit this imporvement in this paper since Theorem 1.2 is good enough for us for the study in [21].

We remark that the converse of SGI(s) also holds. See Proposition D.1.

1.2. **Outline of the proof.** The outline of our proof of SGI(s) is similar to the proof Theorem 1.3 of the work of Green, Tao and Ziegler [12]. However, there are also many significant differences between the two results. For convenience we informally say that two sequences $f, g\colon \Omega \to \mathbb{C}$ *correlate* (on Ω) if the average of fg over Ω is bounded away from zero. We say that f and g *s-correlate* (on Ω) if fg correlates (on Ω) with a nilsequence of step s with low complexity.

Step 1: some preliminary reductions. We first show that SGI(2) holds. The case $s = 2$ can be proved using pure Fourier analysis method. Our proof is similar to the method used by Lyall, Magyar and Parshall [16]. Now suppose that SGI(s) holds and we prove that SGI($s+1$) holds. If $\|f\|_{U^{s+2}(V(M))} \gg 1$, then for many $h \in \mathbb{F}_p^d$, we have that $\|\Delta_h f\|_{U^{s+1}(V(M)^h)} \gg 1$, where $V(M)^h$ is the set of $n \in \mathbb{F}_p^d$ with $M(n) = M(n+h) = 0$. So it follows from SGI(s)+ (and variation of SGI(s) which follows immediately from SGI(s), see Proposition 3.2) that $\Delta_h f$ correlates with an s-step nilsequence χ_h. Using the approximation properties developed in Appendix B, we may assume without loss of generality that χ_h are nilcharacters. Step 1 is conducted in Section 3.

Step 2: a Furstenburg-Weiss type argument. By using the cocycle identity $\Delta_{h+k} f(n) = \Delta_n f(n+k)\Delta_k f(n)$, we can subtract more information for χ_h. Roughly speaking, we can show that for many $(h_1, h_2, h_3, h_4) \in (\mathbb{F}_p^d)^4$ with $h_1 + h_2 = h_3 + h_4$, the function

$$n \mapsto \chi_{h_1}(n) \cdot \chi_{h_2}(n + h_1 - h_4) \cdot \overline{\chi}_{h_3}(n) \cdot \overline{\chi}_{h_4}(n + h_1 - h_4) \tag{1.2}$$

is correlated with an $(s-1)$-step nilsequence on $V(M)^{h_1,h_3,h_3-h_2}$, the set of $n \in \mathbb{F}_p^d$ with $M(n) = M(n+h_1) = M(n+h_3) = M(n+h_3-h_2) = 0$. This is done in Section 4. Our method is similar to the one used in Section 8 of [12]. However, a new Fubini-type theorem is needed (Theorem A.12) in order to complete this step.

To better understand the idea behind the proof, it is helpful to temporarily think of $\chi_h(n)$ as a polynomial function $\exp(P_h(n) + P'_h(n))$ for some homogeneous polynomial $P_h \colon \mathbb{F}_p^d \to \mathbb{F}_p$ of degree $s-1$ and some polynomial $P'_h \colon \mathbb{F}_p^d \to \mathbb{F}_p$ of degree at most $s-2$. In the setting of [12], roughly speaking, the Furstenburg-Weiss type argument implies that for many $h_1, \dots, h_4$ with $h_1 + h_2 = h_3 + h_4$, we have that $P_{h_1} + P_{h_2} = P_{h_3} + P_{h_4}$. One can then use the tools from additive combinatorics to find an explicit expression for P_h.

However, in our setting, there is a lost of information when applying the Furstenburg-Weiss type argument. As a result, we only obtain the weaker conclusion that $P_{h_1} + P_{h_2} \equiv P_{h_3} + P_{h_4} \mod J^M_{h_1,h_2,h_3}$ (see Appendix A.8 for the definition) for many $h_1, \dots, h_4$ with $h_1 + h_2 = h_3 + h_4$. To overcome this difficulty, we need to make essential use of the additive combinatorial tools for M-ideals developed in the second part of the series [20]. Using the linearization result for M-ideals (Theorem A.16), we are able to express $P_h \mod J^M_h$ as an almost linear Freiman homomorphism.

Step 3: building a nilobject. Using the information obtained in Step 2 on $\chi_h(n)$, we construct a nilobject and express $\chi_h(n)$ as $\chi(h, n)$ for some nilcharacter χ of multi-degree $(1, s)$. This is done in Sections 5, 6 and 7. The outline is similar to the one used in Sections 9-12 of [12]. But our method also have many differences from [12]. For example, in Section 5, we need to use a more sophisticated sun-flower lemma (Lemma 5.12, based on the work in the second part of the series [20]) compared with Lemma 10.10 of [12]. In Section 6, we need to use a new Ratner-type thoerem (Theorem 2.28) based on the factorization theorem developed in the first part of the series [19]. In Section 7, in the construction of χ, compared with the approach used in Section 12 of [12], we need to pay extra attention to ensure that χ is p-periodic.

Step 4: a symmetric argument. Using the symmetric identity $\Delta_h \Delta_{h'} f(n) = \Delta_{h'} \Delta_h f(n)$, we extract further information from $\chi_h(n) = \chi(n, h)$ and show that $\chi_h(n) = \Theta(n+h)\Theta'(n)$ for some $(s+1)$-step nilsequences Θ and Θ'. By some standard arguments, this implies that f correlates to Θ and thus completes the proof of SGI($s+1$). This step is done in

Sections 8 (for the case $s = 2$) and 9 (for the case $s \geq 3$), which is in analogous to Section 13 of [12].

However, unlike in [12], the symmetric argument in our setting causes a loss of information. Therefore, in Sections 8 and 9, we need to have an estimate finer than [12] in order to keep track of the information for nilcharacters and to minimize the loss of information (see Remark 8.1). Also as a result of the loss of information, we need to use new methods to improve the results of a significant part of Section 13 of [12] in order to carry out the symmetric argument. See the discussion at the beginning of Section 8.2 for more details.

Organization of the paper. We provide the background material for nilsequences in Section 2. Sections 3–9 are devoted to the proof the spherical Gowers inverse theorem, whose roles are explained in the above mentioned outline of the proof.

In order to prevent the readers from being distracted by technical details, we put the proofs of some results in the appendices. In Appendix A, we collect results from the previous two parts of the series [19, 20] which are used in this paper. In Appendix B, we prove some approximation properties for nilsequences. In Appendix C, we prove some properties on an equivalence property for nilsequences defined Definition 2.35. In Appendix D, we prove the converse of the spherical Gowers inverse theorem.

1.3. **Definitions and notations.**

Convention 1.4. Throughout this paper, we use $\tau\colon \mathbb{F}_p \to \{0,\dots,p-1\}$ to denote the natural bijective embedding, and use $\iota\colon \mathbb{Z} \to \mathbb{F}_p$ to denote the map given by $\iota(n) := \tau^{-1}(n \mod p\mathbb{Z})$. We also use τ to denote the map from $\mathbb{F}_p^k$ to $\mathbb{Z}^k$ given by $\tau(x_1,\dots,x_k) := (\tau(x_1),\dots,\tau(x_k))$, and ι to denote the map from $\mathbb{Z}^k$ to $\mathbb{F}_p^k$ given by $\iota(x_1,\dots,x_k) := (\iota(x_1),\dots,\iota(x_k))$.

We may also extend the domain of ι to all the rational numbers of the form x/y with $(x,y) = 1, x \in \mathbb{Z}, y \in \mathbb{Z}\backslash p\mathbb{Z}$ by setting $\iota(x/y) := \iota(xy^*)$, where y^* is any integer with $yy^* \equiv 1 \mod p\mathbb{Z}$.

Below are the notations we use in this paper:

- Let $\mathbb{N}, \mathbb{N}_+, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{R}+, \mathbb{C}$ denote the set of non-negative integers, positive integers, integers, rational numbers, real numbers, positive real numbers, and complex numbers, respectively. Denote $\mathbb{T} := \mathbb{R}/\mathbb{Z}$. Let $\mathbb{F}_p$ denote the finite field with p elements.
- Throughout this paper, d is a fixed positive integer and p is a prime number.
- Throughout this paper, unless otherwise stated, all vectors are assumed to be horizontal vectors.
- Let C be a collection of parameters and $A, B, c \in \mathbb{R}$. We write $A \gg_C B$ if $|A| \geq K|B|$ and $A = O_C(B)$ if $|A| \leq K|B|$ for some $K > 0$ depending only on the parameters in C. In the above definitions, we allow the set C to be empty. In this case K will be a universal constant.

- For $i = (i_1, \dots, i_k) \in \mathbb{Z}^k$, denote $|i| := |i_1| + \dots + |i_k|$. For $n = (n_1, \dots, n_k) \in \mathbb{Z}^k$ and $i = (i_1, \dots, i_k) \in \mathbb{N}^k$, denote $n^i := n_1^{i_1} \dots n_k^{i_k}$ and $i! := i_1! \dots i_k!$. For $n = (n_1, \dots, n_k) \in \mathbb{N}^k$ and $i = (i_1, \dots, i_k) \in \mathbb{N}^k$, denote $\binom{n}{i} := \binom{n_1}{i_1} \dots \binom{n_k}{i_k}$.
- We say that $\mathcal{F} \colon \mathbb{R}_+ \to \mathbb{R}_+$ is a *growth function* if $\mathcal{F}$ is strictly increasing and $\mathcal{F}(n) \geq n$ for all $n \in \mathbb{R}_+$.
- Let $\mathrm{HP}_d(s)$ denote the set of all the homogeneous polynomial in $\mathbb{F}_p[x_1, \dots, x_d]$ of degree s.
- For $D \in \mathbb{N}_+$, let $\mathbb{S}^D$ denote the set of $(z_1, \dots, z_D) \in \mathbb{C}^D$ with $|z_1|^2 + \dots + |z_D|^2 = 1$.
- For $D, D' \in \mathbb{N}_+$, $v = (v_1, \dots, v_D) \in \mathbb{C}^D$ and $w = (w_1, \dots, w_{D'}) \in \mathbb{C}^{D'}$. Let $v \otimes w := (v_1 w_1, \dots, v_D w_{D'}) \in \mathbb{C}^{DD'}$. We remark that if $v \in \mathbb{S}^D$ and $w \in \mathbb{S}^{D'}$, then $v \otimes w \in \mathbb{S}^{DD'}$. Similarly, if X is a set and $f \colon X \to \mathbb{C}^D$, $g \colon X \to \mathbb{C}^{D'}$ are functions, then we use $f \otimes g \colon X \to \mathbb{C}^{DD'}$ to denote the function $(f \otimes g)(x) := f(x) \otimes g(x)$.
- For $x \in \mathbb{R}$, let $\lfloor x \rfloor$ denote the largest integer which is not larger than x, and $\lceil x \rceil$ denote the smallest integer which is not smaller than x. Let $\{x\} := x - \lfloor x \rfloor$.
- For $D \in \mathbb{N}_+$, a set X and a function $f \colon X \to \mathbb{C}^D$, let $\overline{f} \colon X \to \mathbb{C}^D$ denote the function obtained by taking the conjugate of each coordinates of f.
- Let $D \in \mathbb{N}_+$, X be a finite set and $f \colon X \to \mathbb{C}^D$ be a function. Denote $\mathbb{E}_{x \in X} f(x) := \frac{1}{|X|} \sum_{x \in X} f(x)$, the average of f on X.
- For $F = \mathbb{Z}^k$ or $\mathbb{F}_p^k$, and $x = (x_1, \dots, x_k), y = (y_1, \dots, y_k) \in F$, let $x \cdot y \in \mathbb{Z}$ or $\mathbb{F}_p$ denote the dot product given by $x \cdot y := x_1 y_1 + \dots + x_k y_k$.
- Let $\exp \colon \mathbb{R} \to \mathbb{C}$ denote the function $\exp(x) := e^{2\pi i x}$.
- If G is a connected, simply connected Lie group, then we use $\log G$ to denote its Lie algebra. Let $\exp \colon \log G \to G$ be the exponential map, and $\log \colon G \to \log G$ be the logarithm map. For $t \in \mathbb{R}$ and $g \in G$, denote $g^t := \exp(t \log g)$.
- If $f \colon H \to G$ is a function from an abelian group $H = (H, +)$ to some group $(G, \cdot)$, denote $\Delta_h f(n) := f(n + h) \cdot f(n)^{-1}$ for all $n, h \in H$.
- If $f \colon H \to \mathbb{C}^D$ is a function on an abelian group $H = (H, +)$, denote $\Delta_h f(n) := f(n + h) \otimes \overline{f}(n)$ for all $n, h \in H$.
- We write affine subspaces of $\mathbb{F}_p^d$ as $V + c$, where V is a subspace of $\mathbb{F}_p^d$ passing through $\mathbf{0}$, and $c \in \mathbb{F}_p^d$.
- There is a natural correspondence between polynomials taking values in $\mathbb{F}_p$ and polynomials taking values in $\mathbb{Z}/p$. Let $F \in \mathrm{poly}(\mathbb{F}_p^d \to \mathbb{F}_p^{d'})$ and $f \in \mathrm{poly}(\mathbb{Z}^d \to (\mathbb{Z}/p)^{d'})$ be polynomials of degree at most s for some $s < p$. If $F = \iota \circ pf \circ \tau$, then we say that F is *induced* by f and f is a *lifting* of F.[2] We say that f is a *regular lifting* of F if in addition f has the same degree as F and f has $\{0, \frac{1}{p}, \dots, \frac{p-1}{p}\}$-coefficients.
- Let (X, d_X) be a metric space. The *Lipschitz norm* of a function $F \colon X \to \mathbb{C}^D$ is defined as

$$\|F\|_{\mathrm{Lip}(X)} := \sup_{x \in X} |F(x)| + \sup_{x, y \in X, x \neq y} \frac{|F(x) - F(y)|}{d_X(x, y)}.$$

[2]As is explained in [19], $\iota \circ pf \circ \tau$ is well defined.

For a vector valued function $F = (F_1, \dots, F_D)\colon X \to \mathbb{C}^D$, we denote $\|F\|_{\mathrm{Lip}(X)} := \max_{1\leq i\leq D} \|F_i\|_{\mathrm{Lip}(X)}$. When there is no confusion, we write $\|F\|_{\mathrm{Lip}} = \|F\|_{\mathrm{Lip}(X)}$ for short. For a subset A of $\mathbb{C}^D$, we say that $F\colon X \to A$ is a *Lipschitz function* if $\|F\|_{\mathrm{Lip}(X)} < \infty$. Let $\mathrm{Lip}(X \to A)$ denote the set of all Lipschitz functions $F\colon X \to A$.

Let $D, D' \in \mathbb{N}_+$ and $C > 0$. Here are some basic notions of complexities:

- **Real and complex numbers:** a number $r \in \mathbb{R}$ is of *complexity* at most C if $r = a/b$ for some $a, b \in \mathbb{Z}$ with $-C \leq a, b \leq C$. If $r \notin \mathbb{Q}$, then we say that the complexity of r is infinity. A complex number is of *complexity* at most C if both its real and imaginary parts are of complexity at most C.
- **Vectors and matrices:** a vector or matrix is of *complexity* at most C if all of its entries are of complexity at most C.
- **Subspaces:** a subspace of $\mathbb{R}^D$ is of *complexity* at most C if it is the null space of a matrix of complexity at most C.
- **Linear transformations:** let $L\colon \mathbb{C}^D \to \mathbb{C}^{D'}$ be a linear transformation. Then L is associated with an $D \times D'$ matrix A in $\mathbb{C}$. We say that L is of *complexity* at most C if A is of complexity at most C.
- **Lipschitz function:** the *complexity* of a Lipschitz function is defined to be its Lipschitz norm.

We also need to recall the notations regarding quadratic forms defined in [19].

Definition 1.5. We say that a function $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ is a *quadratic form* if

$$M(n) = (nA) \cdot n + n \cdot u + v$$

for some $d \times d$ symmetric matrix A in $\mathbb{F}_p$, some $u \in \mathbb{F}_p^d$ and $v \in \mathbb{F}_p$. We say that A is the matrix *associated to M*. We say that M is *pure* if $u = \mathbf{0}$. We say that M is *homogeneous* if $u = \mathbf{0}$ and $v = 0$. We say that M is *non-degenerate* if M is of rank d, or equivalently, $\det(A) \neq 0$.

We use $\mathrm{rank}(M) := \mathrm{rank}(A)$ to denote the *rank* of M. Let $V + c$ be an affine subspace of $\mathbb{F}_p^d$ of dimension r. There exists a (not necessarily unique) bijective linear transformation $\phi\colon \mathbb{F}_p^r \to V$. We define the *rank* $\mathrm{rank}(M|_{V+c})$ of M restricted to $V + c$ as the rank of the quadratic form $M(\phi(\cdot) + c)$. It was proved in [19] that $\mathrm{rank}(M|_{V+c})$ is independent of the choice of ϕ.

We also need to use the following notions.

- For a polynomial $P \in \mathrm{poly}(\mathbb{F}_p^k \to \mathbb{F}_p)$, let $V(P)$ denote the set of $n \in \mathbb{F}_p^k$ such that $P(n) = 0$.
- Let $r \in \mathbb{N}_+$, $h_1, \dots, h_r \in \mathbb{F}_p^d$ and $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form. Denote
$$V(M)^{h_1,\dots,h_r} := \cap_{i=1}^r (V(M(\cdot + h_i)) \cap V(M).$$
- Let Ω be a subset of $\mathbb{F}_p^d$ and $s \in \mathbb{N}$. Let $\square_s(\Omega)$ denote the set of $(n, h_1, \dots, h_s) \in (\mathbb{F}_p^d)^{s+1}$ such that $n + \epsilon_1 h_1 + \cdots + \epsilon_s h_s \in \Omega$ for all $(\epsilon_1, \dots, \epsilon_s) \in \{0, 1\}^s$. Here we

allow s to be 0, in which case $\square_0(\Omega) = \Omega$. We say that $\square_s(\Omega)$ is the *s-th Gowers set* of Ω.

Quadratic forms can also be defined in the $\mathbb{Z}/p$-setting.

Definition 1.6. We say that a function $M\colon \mathbb{Z}^d \to \mathbb{Z}/p$ is a *quadratic form* if

$$M(n) = \frac{1}{p}((nA)\cdot n + n \cdot u + v)$$

for some $d \times d$ symmetric matrix A in $\mathbb{Z}$, some $u \in \mathbb{Z}^d$ and $v \in \mathbb{Z}$. We say that A is the matrix *associated to M*.

By Lemma A.1 of [19], any quadratic form $\tilde{M}\colon \mathbb{Z}^d \to \mathbb{Z}/p$ associated with the matrix $\tilde{A}$ induces a quadratic form $M := \iota \circ p\tilde{M} \circ \tau\colon \mathbb{F}_p^d \to \mathbb{F}_p$ associated with the matrix $\iota(\tilde{A})$. Conversely, any quadratic form $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ associated with the matrix A admits a regular lifting $\tilde{M}\colon \mathbb{Z}^d \to \mathbb{Z}/p$, which is a quadratic form associated with the matrix $\tau(A)$.

For a quadratic form $\tilde{M}\colon \mathbb{Z}^d \to \mathbb{Z}/p$, we say that $\tilde{M}$ is *pure/homogeneous/p-non-degenerate* if the quadratic form $M := \iota \circ p\tilde{M} \circ \tau$ induced by $\tilde{M}$ is pure/homogeneous/non-degenerate. The *p-rank* of $\tilde{M}$, denoted by $\text{rank}_p(\tilde{M})$, is defined to be the rank of M.

We say that $h_1, \dots, h_k \in \mathbb{Z}^d$ are *p-linearly independent* if for all $c_1, \dots, c_k \in \mathbb{Z}/p$, $c_1h_1 + \dots + c_kh_k \in \mathbb{Z}$ implies that $c_1, \dots, c_k \in \mathbb{Z}$, or equivalently, if $\iota(h_1), \dots, \iota(h_k)$ are linearly independent.

We also need to use the following notions.

- For a polynomial $P \in \text{poly}(\mathbb{Z}^k \to \mathbb{R})$, let $V_p(P)$ denote the set of $n \in \mathbb{Z}^k$ such that $P(n + pm) \in \mathbb{Z}$ for all $m \in \mathbb{Z}^k$.
- For $\Omega \subseteq \mathbb{Z}^d$ and $s \in \mathbb{N}$, let $\square_{p,s}(\Omega)$ denote the set of $(n, h_1, \dots, h_s) \in (\mathbb{Z}^d)^{s+1}$ such that $n + \epsilon_1 h_1 + \dots + \epsilon_s h_s \in \Omega + p\mathbb{Z}^d$ for all $\epsilon_1, \dots, \epsilon_s \in \{0, 1\}$. We say that $\square_{p,s}(\Omega)$ is the *s-th p-Gowers set* of Ω.

For all other definitions and notations which are used but not mentioned in this paper, we refer the readers to Appendix A for details.

2. Background material for nilsequence

We start with some basic definitions on nilmanifolds. Many notations we use are similar to the ones used in Section 6 of [12]. Unlike in the first part of the series [19], in this paper, instead of just working with $\mathbb{N}$-filtered nilmanifold, we also need to work with multi-degree and degree-rank filtrations. As a result, there are overlappings between Sections 2.1-2.4 of this paper and Section 3 of [19] since we need to extend definitions and results in [19] to more general filtrations. On the other hand, the materials in Sections 2.5-2.7 are new and did not appear in [19].

2.1. Nilmanifolds, filtrations and polynomial sequences.

Definition 2.1 (Ordering)**.** An *ordering* $I = (I, \prec, +, 0)$ is a set I equipped with a partial ordering $\prec$, a binary operation $+\colon I \times I \to I$, and a distinguished element $0 \in I$ with the following properties:

(i) the operation + is commutative and associative, and has 0 as the identity element;
(ii) the partial ordering $\prec$ has 0 as the minimal element;
(iii) if $i, j \prec I$ are such that $i \prec j$, then $i + k \prec j + k$ for all $k \in I$;
(iv) for every $j \in I$, the initial segment $\{i \in I : i \prec d\}$ is finite.

A *finite downset* in I is a finite subset J of I with the property that $j \in J$ whenever $j \in I$ and $j \prec i$ for some $i \in J$.

In this paper, we will use the following special orderings:

(i) the *degree ordering*, where $I = \mathbb{N}$ with the usual ordering, addition and zero element;
(ii) the *multi-degree ordering*, where $I = \mathbb{N}^k$ with the usual addition and zero element, and with the product ordering, i.e. $(i'_1, \dots, i'_k) \leq (i_1, \dots, i_k)$ if $i'_j \leq i_j$ for all $1 \leq j \leq k$;
(iii) the *degree-rank ordering*, where I is the sector $\mathrm{DR} := \{[s, r] \in \mathbb{N}^2 : 0 \leq r \leq s\}$ with the usual addition and zero element, and the lexicographical ordering, i.e. $[s', r'] \prec [s, r]$ if $s' < s$ or if $s' = s$ and $r' < r$.[3]

It is not hard to verify that all these three concepts are indeed orderings. We remark that the degree and the degree-rank orderings are *total orderings*, meaning that for $i, j \in I$, either $i \leq j$ or $j \leq i$. However, the multi-degree ordering is not a total ordering.

Let I be the degree, multi-degree or degree-rank ordering. We say that the *dimension* of I (denoted as $\dim(I)$) is 1 if $I = \mathrm{DR}$, and is k if $I = \mathbb{N}^k$.

Let G be a group and $g, h \in G$. Denote $[g, h] := g^{-1}h^{-1}gh$. For subgroups H, H' of G, let $[H, H']$ denote the group generated by $[h, h']$ for all $h \in H$ and $h' \in H'$.

Definition 2.2 (Filtered group)**.** Let I be an ordering and G be a group. An *I-filtration* on G is a collection $G_I = (G_i)_{i \in I}$ of subgroups of G indexed by I such that the following holds:

(i) for all $i, j \in I$ with $i \prec j$, we have that $G_i \supseteq G_j$;
(ii) for all $i, j \in I$, we have $[G_i, G_j] \subseteq G_{i+j}$.

For $s \in I$, we say that G is an *(I-filtered) nilpotent group* of *degree* at most s (or of *degree* $\leq s$) with respect to some I-filtration $(G_i)_{i \in I}$ if G_i is trivial whenever $i \npreceq s$. For a downset J of I, we say that G is an *(I-filtered) nilpotent group* of *degree* $\subseteq J$ with respect to some I-filtration $(G_i)_{i \in I}$ if G_i is trivial whenever $i \notin J$.

In this paper, we will use the following filtrations.

(i) Let $(d_1, \dots, d_k) \in \mathbb{N}^k$. A *nilpotent Lie group of multi-degree* $\leq (d_1, \dots, d_k)$ (or a *nilpotent Lie group of degree* $\leq d_1$, if $k = 1$) is a nilpotent I-filtered Lie group of degree $\leq (d_1, \dots, d_k)$, where $I = \mathbb{N}^k$ is the multi-degree ordering. If J is a downset, we define nilpotent Lie group of multi-degree $\subseteq J$ in a similar way. For convenience we further require that $G_0 = G$.

[3]We write vectors in DR as $[s, r]$ instead of (s, r) in order to distinguish them from the $\mathbb{N}^2$ multi-degree ordering.

(ii) Let $[s,r] \in \mathrm{DR}$, A *nilpotent Lie group of degree-rank* $\leq [s,r]$ is a nilpotent DR-filtered Lie group of degree $\leq [s,r]$, with the additional assumption that $G_{[0,0]} = G$ and $G_{[i,0]} = G_{[i,1]}$ for all $i \geq 1$.

Definition 2.3 (Nilmanifold)**.** Let I be an ordering. Let Γ be a discrete and cocompact subgroup of a connected, simply-connected nilpotent Lie group G with filtration $G_I = (G_i)_{i\in I}$ such that $\Gamma_i := \Gamma\cap G_i$ is a cocompact subgroup of G_i[4] for all $i \in I$. Then we say that G/Γ is an *(I-filtered) nilmanifold*, and we use $(G/\Gamma)_I$ to denote the collection $(G_i/\Gamma_i)_{i\in I}$ (which is called the *I-filtration* of G/Γ). We say that G/Γ has degree $\leq s$ or $\subseteq J$ with respect to $(G/\Gamma)_I$ if G has degree $\leq s$ or $\subseteq J$ with respect to G_I.

Definition 2.4 (Sub-nilmanifold)**.** Let G/Γ be an I-filtered nilmanifold of degree $\subseteq J$ with filtration G_I and H be a rational subgroup of G. Then $H/(H\cap\Gamma)$ is also an I-filtered nilmanifold of degree $\subseteq J$ with the filtration H_I given by $H_i := G_i \cap H$ for all $i \in I$ (see Example 6.14 of [12]). We say that $H/(H\cap\Gamma)$ is a *sub-nilmanifold* of G/Γ, H_I (or $(H/(H\cap\Gamma))_I$) is the filtration *induced by* G_I (or $(G/\Gamma)_I$).

Definition 2.5 (Quotient nilmanifold)**.** Let G/Γ be an I-filtered nilmanifold of degree $\subseteq J$ with filtration G_I and H be a normal subgroup of G. Then $(G/H)/(\Gamma/(\Gamma\cap H))$ is also an I-filtered nilmanifold of most $\subseteq J$ with the filtration $(G/H)_I$ given by $(G/H)_i := G_i/(H\cap G_i)$ for all $i \in I$. We say that $(G/H)/(\Gamma/(\Gamma\cap H))$ is the *quotient nilmanifold* of G/Γ by H and that $(G/H)_I$ is the filtration *induced by* G_I.

Definition 2.6 (Product nilmanifold)**.** Let G/Γ and G'/Γ' be I-filtered nilmanifolds of degree $\subseteq J$ with filtration G_I and G'_I. Then $G\times G'/\Gamma\times\Gamma'$ is also an I-filtered nilmanifold of most $\subseteq J$ with the filtration $(G\times G'/\Gamma\times\Gamma')_I$ given by $(G\times G'/\Gamma\times\Gamma')_i := G_i\times G'_i/\Gamma_i\times\Gamma'_i$ for all $i \in I$. We say that $G\times G'/\Gamma\times\Gamma'$ is the *product nilmanifold* of G/Γ and G'/Γ' and that $(G\times G'/\Gamma\times\Gamma')_I$ is the filtration *induced by* G_I and G'_I.

The following definition is based on from Example 6.11 of [12], which explains the connections between different filtrations.

Definition 2.7 (Inductions of filtrations)**.** Let $s \in \mathbb{N}_+$ and G be an $\mathbb{N}$-filtered nilmanifold of degree $\leq s$ with filtration $G_{\mathbb{N}}$. Then for any $k \in \mathbb{N}_+$, $G_{\mathbb{N}}$ *induces* a multi-degree filtration $G'_{\mathbb{N}^k}$ of degree $\subseteq \{m \in \mathbb{N}^k \colon |m| \leq s\}$ given by $G'_m := G_{|m|}$ for all $m \in \mathbb{N}^k$. Conversely, let $k \in \mathbb{N}_+, J \subseteq \mathbb{N}^k$, and G be an $\mathbb{N}^k$-filtered nilmanifold of degree $\leq J$ with filtration $G_{\mathbb{N}^k}$. Then $G_{\mathbb{N}^k}$ *induces* a degree filtration $G'_{\mathbb{N}}$ of degree $\max\{|m| \colon m \in \mathbb{N}^k\}$ given by $G'_i := \bigvee_{m\in\mathbb{N}^k, |m|=i} G_m$ for all $i \in \mathbb{N}$, where $\vee_{a\in A} G_a$ is the group generated by $\cup_{a\in A} G_a$.

Let $s \in \mathbb{N}_+$ and G be an $\mathbb{N}$-filtered nilmanifold of degree $\leq s$ with filtration $G_{\mathbb{N}}$. Since $G = G_0$, $G_{\mathbb{N}}$ *induces* a degree-rank filtration G'_{DR} of degree $\leq [s,s]$ given by setting $G'_{[t,r]}$ to be the group generated by all the iterated commutators of $g_{i_1},\dots,g_{i_m}$ with $g_{i_j} \in G_{i_j}$ for $1 \leq j \leq m$ such that either $i_1+\cdots+i_m > t$ or $i_1+\cdots+i_m = t$ and $m \geq \max\{r,1\}$. Conversely, let $s,r \in \mathbb{N}_+, r \leq s$ and G be an DR-filtered nilmanifold of degree $\leq [s,r]$ with filtration G_{DR}. Then G_{DR} *induces* a degree filtration $G'_{\mathbb{N}}$ of degree $\leq s$ given by setting $G'_i := G_{[i,0]}$ for all $i \in \mathbb{N}$.

[4] In some papers, such Γ_i is called a *rational* subgroup of G.

A multi-degree filtration $G_{\mathbb{N}^k}$ *induces* a degree-rank filtration G'_{DR} by first induces a degree filtration $G''_{\mathbb{N}}$ and then induces from $G''_{\mathbb{N}}$ a degree-rank filtration G'_{DR} in the above mentioned senses. Similarly, a degree-rank filtration G_{DR} *induces* a multi-degree filtration $G'_{\mathbb{N}^k}$ by first induces a degree filtration $G''_{\mathbb{N}}$ and then induces from $G''_{\mathbb{N}}$ a multi-degree filtration $G'_{\mathbb{N}^k}$ in the above mentioned senses.

Let I, I' be the degree, multi-degree or degree-rank filtrations. We say that the I-filtration of a nilmanifold $(G/\Gamma)_I$ *induces* the I'-filtration of a nilmanifold $(G'/\Gamma')_{I'}$ if G_I induces $G'_{I'}$.

We remark that although we used the same terminology "induce" in Definitions 2.4, 2.5, 2.6 and 2.7, this will not cause any confusion in the paper as the meaning of "induce" will be clear from the context.

Definition 2.8 (Filtered homomorphism)**.** An *(I-filtered) homomorphism* $\phi\colon G/\Gamma \to G'/\Gamma'$ between two I-filtered nilmanifolds is a group homomorphism $\phi\colon G \to G'$ which maps Γ to Γ' and maps G_i to G'_i for all $i \in I$.

Every nilmanifold has an explicit algebraic description by using the Mal'cev basis:

Definition 2.9 (Mal'cev basis)**.** Let $s \in \mathbb{N}_+$, G/Γ be a nilmanifold of step at most s with the $\mathbb{N}$-filtration $(G_i)_{i\in\mathbb{N}}$. Let $\dim(G) = m$ and $\dim(G_i) = m_i$ for all $0 \leq i \leq s$. A basis $\mathcal{X} := \{X_1, \dots, X_m\}$ for the Lie algebra $\log G$ of G (over $\mathbb{R}$) is a *Mal'cev basis* for G/Γ adapted to the filtration $G_{\mathbb{N}}$ if

- for all $0 \leq j \leq m-1$, $\log H_j := \mathrm{Span}_{\mathbb{R}}\{\xi_{j+1}, \dots, \xi_m\}$ is a Lie algebra ideal of $\log G$ and so $H_j := \exp(\log H_j)$ is a normal Lie subgroup of G;
- $G_i = H_{m-m_i}$ for all $0 \leq i \leq s$;
- the map $\psi^{-1}\colon \mathbb{R}^m \to G$ given by
$$\psi^{-1}(t_1, \dots, t_m) = \exp(t_1X_1)\dots\exp(t_mX_m)$$
is a bijection;
- $\Gamma = \psi^{-1}(\mathbb{Z}^m)$.

We call ψ the *Mal'cev coordinate map* with respect to the Mal'cev basis $\mathcal{X}$. If $g = \psi^{-1}(t_1, \dots, t_m)$, we say that $(t_1, \dots, t_m)$ are the *Mal'cev coordinates* of g with respect to $\mathcal{X}$.

We say that the Mal'cev basis $\mathcal{X}$ is *C-rational* (or of *complexity* at most C) if all the structure constants $c_{i,j,k}$ in the relations
$$[X_i, X_j] = \sum_k c_{i,j,k}X_k$$
are rational with complexity at most C.

It is known that for every $\mathbb{N}$-filtration $G_{\mathbb{N}}$ which is rational for Γ, there exists a Mal'cev basis adapted to it. See for example the discussion on pages 11–12 of [10].

If G/Γ is a nilmanifold with a multi-degree filtration $(G_i)_{i\in\mathbb{N}^k}$. Then there is a Mal'cev basis $\mathcal{X}$ of G/Γ adapted to the degree filtration $(G'_i)_{i\in\mathbb{N}}$, where G'_i is the group generated by

$G_{(i_1,\dots,i_k)}$, $i_1+\dots+i_k=i$ for all $i\in\mathbb{N}$. We say that $\mathcal{X}$ is the *Mal'cev basis* of G/Γ adapted to $(G_i)_{i\in\mathbb{N}^k}$.

If G/Γ is a nilmanifold with a degree-rank filtration $(G_{[s,r]})_{[s,r]\in\mathrm{DR}}$. Then there is a Mal'cev basis $\mathcal{X}$ of G/Γ adapted to the degree filtration $(G'_i)_{i\in\mathbb{N}}$, where $G'_i=G_{[i,0]}$ for all $i\in\mathbb{N}$. We say that $\mathcal{X}$ is the *Mal'cev basis* of G/Γ adapted to $(G_{[s,r]})_{[s,r]\in\mathrm{DR}}$.

We use the following quantities to describe the complexities of the objected defined above.

Definition 2.10 (Notions of complexities for nilmanifolds)**.** Let G/Γ be a nilmanifold with filtration G_I and a Mal'cev basis $\mathcal{X}=\{X_1,\dots,X_D\}$ adapted to it. We say that G/Γ is of *complexity* at most C if the Mal'cev basis $\mathcal{X}$ is C-rational and $\dim(G)\leq C$.

An element $g\in G$ is of *complexity* at most C (with respect to the Mal'cev coordinate map $\psi\colon G/\Gamma\to\mathbb{R}^m$) if $\psi(g)\in[-C,C]^m$.

Let G'/Γ' be a nilmanifold endowed with the Mal'cev basis $\mathcal{X}'=\{X'_1,\dots,X'_{D'}\}$ respectively. Let $\phi\colon G/\Gamma\to G'/\Gamma'$ be a filtered homomorphism, we say that ϕ is of *complexity* at most C if the map $X_i\to\sum_j a_{i,j}X'_j$ induced by ϕ is such that all $a_{i,j}$ are of complexity at most C.

Let $G'\subseteq G$ be a closed connected subgroup. We say that G' is *C-rational* (or of *complexity* at most C) relative to $\mathcal{X}$ if the Lie algebra $\log G$ has a basis consisting of linear combinations $\sum_i a_iX_i$ such that a_i are rational numbers of complexity at most C.

Convention 2.11. In the rest of the paper, all nilmanifolds are assumed to have a fixed filtration, Mal'cev basis and a smooth Riemannian metric induced by the Mal'cev basis. Therefore, we will simply say that a nilmanifold, Lipschitz function, sub-nilmanifold etc. is of complexity C without mentioning the reference filtration and Mal'cev basis.

Definition 2.12 (Polynomial sequences)**.** Let $d,k\in\mathbb{N}_+$ and G be a connected simply-connected nilpotent Lie group.

(i) Let $H=\mathbb{Z}^d$ or $\mathbb{R}^d$[5] and $(G_i)_{i\in\mathbb{N}}$ be a degree filtration of G. A map $g\colon H\to G$ is a *($\mathbb{N}$-filtered) d-integral polynomial sequence* if
$$\Delta_{h_m}\dots\Delta_{h_1}g(n)\in G_m$$
for all $m\in\mathbb{N}$ and $n,h_1,\dots,h_m\in H$.[6]

(ii) Let $H=(\mathbb{Z}^d)^k$ or $(\mathbb{R}^d)^k$ and $(G_i)_{i\in\mathbb{N}^k}$ be a multi-degree filtration of G. Let $e_{i,j}$, $1\leq i\leq k, 1\leq j\leq d$ be the standard unit vectors in $(\mathbb{Z}^d)^k$. A map $g\colon H\to G$ is a *($\mathbb{N}^k$-filtered) d-integral polynomial sequence* if
$$\Delta_{e_{i_m,j_m}}\dots\Delta_{e_{i_1,j_1}}g(n)\in G_t$$
for all $m\in\mathbb{N}$, $1\leq i_1,\dots,i_m\leq k, 1\leq j_1,\dots,j_m\leq d$, and $n\in H$, where $t=(t_1,\dots,t_k)\in\mathbb{N}^k$ is such that t_ℓ equals to the number of $i_1,\dots,i_m$ which equals to ℓ for all $1\leq\ell\leq k$.

[5]In this paper, we mainly consider polynomials sequences from $\mathbb{Z}^d$ to G. The only place we use polynomials sequences from $\mathbb{R}^d$ to G is Theorem B.2.

[6]Recall that $\Delta_h g(n):=g(n+h)g(n)^{-1}$ for all $n,h\in H$.

(iii) Let $H = \mathbb{Z}^d$ or $\mathbb{R}^d$ and $(G_{[s,r]})_{[s,r]\in\mathrm{DR}}$ be a degree-rank filtration of G. A map $g\colon H \to G$ is a *(DR-filtered) d-integral polynomial sequence* if

$$\Delta_{h_m} \dots \Delta_{h_1} g(n) \in G_{[m,0]}$$

for all $m \in \mathbb{N}$ and $n, h_1, \dots, h_m \in H$.

The sets of all $\mathbb{N}$-, $\mathbb{N}^k$-, and DR-filtered d-integral polynomial sequences are denoted by $\mathrm{poly}(H \to G_{\mathbb{N}})$, $\mathrm{poly}(H^k \to G_{\mathbb{N}^k})$ and $\mathrm{poly}(H \to G_{\mathrm{DR}})$, respectively.[7]

Remark 2.13. Our definition of polynomial sequences coincides with the one in [12] when $d = 1$. When $d > 1$, the set $\mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$ defined in this paper is the set $\mathrm{poly}(\mathbb{Z}^d \to G'_{\mathbb{N}^d})$ defined in [12], where $G'_{(i_1,\dots,i_d)} = G_{i_1+\dots+i_d}$ for all $i_1, \dots, i_d \in \mathbb{N}$, and the set $\mathrm{poly}((\mathbb{Z}^d)^k \to G_{\mathbb{N}})$ defined in this paper is the set $\mathrm{poly}((\mathbb{Z}^d)^k \to G'_{\mathbb{N}^{dk}})$ defined in [12], where $G'_{(i_{1,1},\dots,i_{k,d})} = G_{(i_{1,1}+\dots+i_{1,d},\dots,i_{k,1}+\dots+i_{k,d})}$ for all $i_{1,1}, \dots, i_{k,d} \in \mathbb{N}$. The reason we adopt a different notation from [12] is that the new notation is more convenient to use since we frequently view vectors in $\mathbb{Z}^{dk}$ as k-tuples of elements in $\mathbb{Z}^d$ in this paper.

Convention 2.14. In the rest of the paper, we will omit the phrase "d-integral" when mentioning a polynomial sequence, since the quantity d will always be clear from the context. For example, let I be the degree, multi-degree or degree-rank ordering, $k \in \mathbb{N}_+$ with $\dim(I)|k$, and G/Γ be an I-filtered nilmanifold. Denote $d = k/\dim(I)$. When writing $\mathrm{poly}(\Omega \to G_I)$, we regard $\mathbb{Z}^k$ as $(\mathbb{Z}^d)^{\dim(I)}$, and $\mathrm{poly}(\mathbb{Z}^k \to G_I)$ is understood as $\mathrm{poly}((\mathbb{Z}^d)^{\dim(I)} \to G_I)$, the set of d-integral polynomial sequences (where d is uniquely determined by I and k).

By Corollary B.4 of [12], all of $\mathrm{poly}(\mathbb{Z}^k \to G_{\mathbb{N}})$, $\mathrm{poly}((\mathbb{Z}^k)^{k'} \to G_{\mathbb{N}^{k'}})$, $\mathrm{poly}(\mathbb{Z}^k \to G_{\mathrm{DR}})$, and $\mathrm{poly}((\mathbb{Z}^k)^{k'} \to G_{\mathbb{N}})$ are groups with respect to the pointwise multiplicative operation. We refer the readers to Appendix B of [12] for more properties for polynomial sequences.

We conclude this section with a lemma regarding the expression of a polynomial sequence using Mal'cev basis.

Lemma 2.15. Let $s \in \mathbb{N}_+$ and G be an s-step connected, simply-connected nilpotent Lie group with the degree filtration $G_{\mathbb{N}}$ and Mal'cev basis $\mathcal{X} = \{X_1, \dots, X_m\}$, where $m = \dim(G)$. Denote $m_i := \dim(G_i)$. Suppose that G_i is generated by $X_{m-m_i+1}, \dots, X_m$. Let $g\colon \mathbb{Z}^d \to G$ be a map given by

$$g(n) = \prod_{j=1}^{m} \exp(P_j(n)X_j)$$

for some polynomials $P_j\colon \mathbb{Z}^d \to \mathbb{R}$. Then $g \in \mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$ if and only if $\deg(P_j) \leq i$ for all $m - m_i + 1 \leq j \leq m - m_{i+1}$.

Proof. If $\deg(P_j) \leq i$ for all $m - m_i + 1 \leq j \leq m - m_{i+1}$, then since $\exp((a + b)X_j) = \exp(aX_j)\exp(bX_j)$ for all $a, b \in \mathbb{R}$, it is not hard to see that $n \mapsto \exp(P_j(n)X_j)$ is a map in

[7] Polynomial sequences can be defined for more general groups H and for more general filtrations (see for example [12]). But we do not need them in this paper.

poly($\mathbb{Z}^d \to G_{\mathbb{N}}$) for all $1 \leq j \leq m$. By Corollary B.4 of [12], we have that $g \in \text{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$.

Conversely, suppose that $g \in \text{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$. We say that *Property-i* holds if $\deg(P_j) \leq i'$ for all $m - m_{i'} + 1 \leq j \leq m - m_{i'+1}$ with $i' \leq i$. Clearly Property-0 holds. Suppose that Property-$(i-1)$ holds for some $1 \leq i \leq s$. Then by the converse direction and Corollary B.4 of [12],

$$\Big(\prod_{j=1}^{m-m_i} \exp(P_j(n)X_j)\Big)^{-1} g(n) = \prod_{j=m-m_i+1}^{m} \exp(P_j(n)X_j)$$

belongs to $g \in \text{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$. So $\prod_{j=m-m_i+1}^{m-m_{i+1}} \exp(P_j(n)X_j) \mod G_{i+1}$ belongs to poly($\mathbb{Z}^d \to (G/G_{i+1})_{\mathbb{N}}$). Since G_i/G_{i+1} is abelian, by the definition of poly($\mathbb{Z}^d \to (G/G_{i+1})_{\mathbb{N}}$),

$$\Delta_{h_{i+1}} \dots \Delta_{h_1} \prod_{j=m-m_i+1}^{m-m_{i+1}} \exp(P_j(n)X_j) \equiv \prod_{j=m-m_i+1}^{m-m_{i+1}} \exp(\Delta_{h_{i+1}} \dots \Delta_{h_1} P_j(n)X_j) \equiv 0 \mod G_{i+1}$$

for all $n, h_1, \dots, h_{i+1} \in \mathbb{Z}^d$. This means that $\deg(P_j) \leq i$ for all $m - m_i + 1 \leq j \leq m - m_{i+1}$. So Property-$i$ holds.

Inductively we have that Property-s holds and we are done. □

2.2. **Partially periodic polynomial sequences.**

We may extend the concept of partially periodic polynomial sequences defined in [19] naturally for more general filtrations.

Definition 2.16 (Partially periodic polynomial sequences)**.** Let I be the degree, multi-degree or degree-rank ordering, $k \in \mathbb{N}_+$ with $\dim(I)|k$, G/Γ be an I-filtered nilmanifold, p be a prime, and Ω be a subset of $\mathbb{Z}^k$. Let poly($\Omega \to G_I|\Gamma$) denote the set of all $g \in \text{poly}(\mathbb{Z}^k \to G_I)$ such that $g(n) \in \Gamma$ for all $n \in \Omega + p\mathbb{Z}^k$, and $\text{poly}_p(\Omega \to G_I|\Gamma)$ denote the set of all $g \in \text{poly}(\mathbb{Z}^k \to G_I)$ such that $g(n+pm)^{-1}g(n) \in \Gamma$ for all $n \in \Omega + p\mathbb{Z}^k$ and $m \in \mathbb{Z}^k$.

Let poly($\mathbb{F}_p^k \to G_I$) denote the set of all functions of the form $f \circ \tau$ for some $f \in \text{poly}(\mathbb{Z}^k \to G_I)$. For a subset Ω of $\mathbb{F}_p^k$, let $\text{poly}_p(\Omega \to G_I|\Gamma)$ denote the set of all functions of the form $f \circ \tau$ for some $f \in \text{poly}_p(\iota^{-1}(\Omega) \to G_I|\Gamma) = \text{poly}_p(\tau(\Omega) \to G_I|\Gamma)$ (note that $\iota^{-1}(\Omega) = \tau(\Omega) + p\mathbb{Z}^k$). We define poly($\Omega \to G_I|\Gamma$) similarly.

For $\Omega \subseteq \mathbb{Z}^k$ or $\mathbb{F}_p^k$, we call functions in $\text{poly}_p(\Omega \to G_I|\Gamma)$ *partially p-periodic polynomial sequences on* Ω.

As is explain in Example 3.16 of [19], $\text{poly}_p(\Omega \to G_I|\Gamma)$ is not closed under multiplication. However, we have the following:

Lemma 2.17. Let I be the degree, multi-degree or degree-rank ordering, $k \in \mathbb{N}_+$ with $\dim(I)|k$, G/Γ be an I-filtered nilmanifold, p be a prime, and Ω be a subset of $\mathbb{Z}^k$. For all $f \in \text{poly}_p(\Omega \to G_I|\Gamma)$ and $g \in \text{poly}(\Omega \to G_I|\Gamma)$, we have that $fg \in \text{poly}_p(\Omega \to G_I|\Gamma)$.

The proof of Lemma 2.17 is almost identical to that of Lemma 3.17 in [19]. So we omit the details.

As is explained in Example 3.18 of [19], partially p-periodic polynomial sequences are not closed under translations and transformations. However, the following proposition asserts that partially p-periodic are closed under translations and transformations modulo Γ.

Proposition 2.18. Let I be the degree, multi-degree or degree-rank ordering, $k \in \mathbb{N}_+$ with $\dim(I)|k$, G/Γ be an I-filtered nilmanifold, p be a prime, Ω be a subset of $\mathbb{F}_p^k$, and $g \in \text{poly}_p(\Omega \to G_I|\Gamma)$.

(i) For any $h \in \mathbb{F}_p^k$, there exists $g' \in \text{poly}_p((\Omega - h) \to G_I|\Gamma)$ such that $g(n+h)\Gamma = g'(n)\Gamma$ for all $n \in \Omega - h$.
(ii) If I is the degree filtration, then for any $k' \in \mathbb{N}_+$ and linear transformation $L\colon \mathbb{F}_p^{k'} \to \mathbb{F}_p^k$, there exists $g'' \in \text{poly}_p(L^{-1}(\Omega) \to G_I|\Gamma)$ such that $g(L(n))\Gamma = g''(n)\Gamma$ for all $n \in L^{-1}(\Omega)$.

The proof of Proposition 2.18 is almost identical to that of Lemma 3.19 in [19]. So we omit the details.

2.3. The Baker-Campbell-Hausdorff formula.

The material of this section comes from Appendix C of [8]. We write it down for completeness.

Let G be a group, $t \in \mathbb{N}_+$ and $g_1, \dots, g_t \in G$. The *iterated commutator* of g_1 is defined to be g_1 itself. Iteratively, we define an *iterated commutator* of $g_1, \dots, g_t$ to be an element of the form $[w, w']$, where w is an iterated commutator of $g_{i_1}, \dots, g_{i_r}$, w' is an iterated commutator of $g_{i'_1}, \dots, g_{i'_{r'}}$ for some $1 \leq r, r' \leq t-1$ with $r + r' = t$ and $\{i_1, \dots, i_r\} \cup \{i'_1, \dots, i'_{r'}\} = \{1, \dots, t\}$.

Similarly, let $X_1, \dots, X_t$ be elements of a Lie algebra. The *iterated Lie bracket* of X_1 is defined to be X_1 itself. Iteratively, we define an *iterated Lie bracket* of $X_1, \dots, X_t$ to be an element of the form $[w, w']$, where w is an iterated Lie bracket of $X_{i_1}, \dots, X_{i_r}$, w' is an iterated Lie bracket of $X_{i'_1}, \dots, X_{i'_{r'}}$ for some $1 \leq r, r' \leq t-1$ with $r + r' = t$ and $\{i_1, \dots, i_r\} \cup \{i'_1, \dots, i'_{r'}\} = \{1, \dots, t\}$.

Let G be a connected and simply connected nilpotent Lie group. The *Baker-Campbell-Hausdorff formula* asserts that for all $X_1, X_2 \in \log G$, we have

$$\exp(X_1)\exp(X_2) = \exp\Big(X_1 + X_2 + \frac{1}{2}[X_1, X_2] + \prod_{\alpha} c_\alpha X_\alpha\Big),$$

where α is a finite set of labels, c_α are real constants, and X_α are iterated Lie brackets with $k_{1,\alpha}$ copies of X_1 and $k_{2,\alpha}$ copies of X_2 for some $k_{1,\alpha}, k_{2,\alpha} \geq 1$ and $k_{1,\alpha} + k_{2,\alpha} \geq 3$. One may use this formula to show that for all $g_1, g_2 \in G$ and $x \in \mathbb{R}$, we have that

$$(g_1 g_2)^x = g_1^x g_2^x \prod_{\alpha} g_\alpha^{Q_\alpha(x)},$$

where α is a finite set of labels, g_α are iterated commutators with $k_{1,\alpha}$ copies of g_1 and $k_{2,\alpha}$ copies of X_2 for some $k_{1,\alpha}, k_{2,\alpha} \geq 1$, and $Q_\alpha\colon \mathbb{R} \to \mathbb{R}$ are polynomials of degrees at most $k_{1,\alpha} + k_{2,\alpha}$ without constant terms.

Similarly, one can show that for any $g_1, g_2 \in G$ and $x_1, x_2 \in \mathbb{R}$, we have that

$$[g_1^{x_1}, g_2^{x_2}] = [g_1, g_2]^{x_1 x_2} \prod_{\alpha} g_\alpha^{P_\alpha(x_1, x_2)},$$

where α is a finite set of labels, g_α are iterated commutators with $k_{1,\alpha}$ copies of g_1 and $g_{2,\alpha}$ copies of X_2 for some $k_{1,\alpha}, k_{2,\alpha} \geq 1$, $k_{1,\alpha} + k_{2,\alpha} \geq 3$, and $P_\alpha \colon \mathbb{R}^2 \to \mathbb{R}$ are polynomials of degrees at most $k_{1,\alpha}$ in x_1 and at most $k_{2,\alpha}$ in x_2 which vanishes when $x_1 x_2 = 0$.

2.4. **Type-I horizontal torus and character.** In this section, we recall the definitions of type-I horizontal torus and character defined in [19].

Definition 2.19 (Type-I horizontal torus and character)**.** Let G/Γ be a nilmanifold endowed with a Mal'cev basis $\mathcal{X}$. The *type-I horizontal torus* of G/Γ is $G/[G,G]\Gamma$. A *type-I horizontal character* is a continuous homomorphism $\eta \colon G \to \mathbb{R}$ such that $\eta(\Gamma) \subseteq \mathbb{Z}$. When written in the coordinates relative to $\mathcal{X}$, we may write $\eta(g) = k \cdot \psi(g)$ for some unique $k = (k_1, \dots, k_m) \in \mathbb{Z}^m$, where $\psi \colon G \to \mathbb{R}^m$ is the coordinate map with respect to the Mal'cev basis $\mathcal{X}$. We call the quantity $\|\eta\| := |k| = |k_1| + \cdots + |k_m|$ the *complexity* of η (with respect to $\mathcal{X}$).

It is not hard to see that any type-I horizontal character mod $\mathbb{Z}$ vanishes on $[G,G]\Gamma$ and thus descent to a continuous homomorphism between the type-I horizontal torus $G/[G,G]\Gamma$ and $\mathbb{R}/\mathbb{Z}$. Moreover, $\eta \mod \mathbb{Z}$ is a well defined map from G/Γ to $\mathbb{R}/\mathbb{Z}$.

Type-I horizontal torus and character are used to characterize whether a nisequence is equidistributed on a nilmanifold [10, 15]. The following lemma is a generalization of Lemma 3.21 of [19] to multi-degree filtrations. Its proof is similar to that of Lemma B.9 of [12] and Lemma 2.8 of [2]. We omit the details.

Lemma 2.20. Let $d, k \in \mathbb{N}_+$, $H = (\mathbb{Z}^d)^k$ or $(\mathbb{R}^d)^k$, and G be an $\mathbb{N}^k$-filtered group of degree $\leq J$ for some finite downset J. We complete the partial order on J to a total ordering in an arbitrary fashion. A function $g \colon H \to G$ belongs to poly$(H \to G_{\mathbb{N}^k})$ if and only if for all $m = (m_1, \dots, m_k) \in (\mathbb{N}^d)^k$ with $(|m_1|, \dots, |m_k|) \in J$, there exists $X_m \in \log(G_{(|m_1|,\dots,|m_k|)})$ such that

$$g(n) = \prod_{m=(m_1,\dots,m_k)\in(\mathbb{N}^d)^k, (|m_1|,\dots,|m_k|)\in J} \exp\Big(\binom{n}{m} X_m\Big).$$

Moreover, if $g \in \text{poly}(H \to G_{\mathbb{N}^k})$, then the choice of X_m are unique.

Furthermore, if G' is a subgroup of G and g takes values in G', then $X_m \in \log(G')$ for all m.

An equivalent way of saying that a function $g \colon H \to G$ belongs to poly$(\mathbb{Z}^k \to G_{\mathbb{N}^k})$ is that

$$g(n) = \prod_{m=(m_1,\dots,m_k)\in(\mathbb{N}^d)^k, (|m_1|,\dots,|m_k|)\in J} g_m^{\binom{n}{m}} \tag{2.1}$$

for some $g_m \in G_{(|m_1|,\dots,|m_k|)}$ (with any fixed order in $\mathbb{N}^k$). We call g_m the *(m-th) type-I Taylor coefficient* of g, and (2.1) the *type-I Taylor expansion* of g.

2.5. **Type-II horizontal torus and character.** In this paper, in addition the type-I horizontal torus and character, we also need to use the type-II horizontal torus and character, which are defined in Definition 9.6 of [12] and can be used to characterize Ratner-type theorem for nilmanifolds with degree-rank filtrations.

Definition 2.21 (Type-II Taylor coefficients)**.** Let G be a degree-rank-filtered connected, simply-connected nilpotent Lie group with filtration $(G_{[s,r]})_{[s,r]\in \mathrm{DR}}$. For every $i \in \mathbb{N}_+$, define the *i-th type-II horizontal space* $\mathrm{Horiz}_i(G)$ to be the abelian group

$$\mathrm{Horiz}_i(G) := G_{[i,1]}/G_{[i,2]},$$

with the convention that $G_{[s,r]} := G_{[s+1,0]}$ if $r > s$ (in particular, $G_{[1,2]} = G_{[2,0]}$). If Γ is a subgroup of G, then define

$$\mathrm{Horiz}_i(G/\Gamma) := \mathrm{Horiz}_i(G)/\mathrm{Horiz}_i(\Gamma).$$

It is easy to see that the type-II horizontal spaces $\mathrm{Horiz}_i(G)$ are abelian Lip groups and that $\mathrm{Horiz}_i(\Gamma)$ is a sublattice of $\mathrm{Horiz}_i(G)$. So $\mathrm{Horiz}_i(G/\Gamma)$ is a torus, which we call the *i-th type-II horizontal torus* of G/Γ.

For any $d \in \mathbb{N}_+$, $g \in \mathrm{poly}(\mathbb{Z}^d \to G_{\mathrm{DR}})$ and $m = (m_1, \dots, m_d) \in \mathbb{N}^d$ with $|m| = i$, we define the *m-th type-II Taylor coefficient* $\mathrm{Taylor}_m(g)$ to be the quantity

$$\mathrm{Taylor}_m(g) := \Delta_{e_1} \dots \Delta_{e_1} \dots \Delta_{e_d} \dots \Delta_{e_d} g(n) \mod G_{[i,2]}$$

for any $n \in \mathbb{Z}^d$, where Δ_{e_j} occurs m_j times for all $1 \leq j \leq d$ and $e_j \in \mathbb{Z}^d$ is the j-th standard unit vector. Note that this map is well defined since $\mathrm{Taylor}_m(g)$ takes values in $G_{[i,1]}$ and has first derivatives in $G_{[i+1,1]} \subseteq G_{[i,2]}$ (and thus is independent of the choice of n). Denote

$$\mathrm{Taylor}_i(g)(h_1, \dots, h_i) := \Delta_{h_i} \dots \Delta_{h_1} g(n) \mod G_{[i,2]}$$

for $h_1, \dots, h_i \in \mathbb{Z}^d$ (again this quantity is independent of the choice of n). By Corollary B.7 of [12], it is not hard to see that $\mathrm{Taylor}_i(g)(h_1, \dots, h_i)$ is a homogeneous polynomial of multi-degree $(1, \dots, 1)$ in the variables $h_1, \dots, h_i$, and is symmetric in the variables $h_1, \dots, h_i$.

If G is a connected, simply-connected nilpotent Lie group with a degree filtration $G_{\mathbb{N}}$, then we may define the type-II Taylor coefficient and type-II horizontal toruses similarly with $(G_{[s,r]})_{[s,r]\in \mathrm{DR}}$ being the natural degree-rank filtration induced by $G_{\mathbb{N}}$.

By Corollary B.7 of [12] (see also the discussion on page 1282 of [12]), we have:

Lemma 2.22. Let $d \in \mathbb{N}_+$ and G/Γ be a nilmanifold of degree at most s with a filtration $G_{\mathbb{N}}$. For all $g, g' \in \mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$ and $m \in \mathbb{N}^d$, we have that

$$\mathrm{Taylor}_m(gg')(h_1, \dots, h_i) = \mathrm{Taylor}_m(g)(h_1, \dots, h_i) \cdot \mathrm{Taylor}_m(g')(h_1, \dots, h_i) \mod G_{[i,2]}.$$

In other words, the map $g \mapsto \mathrm{Taylor}_m(g)$ is a homomorphism.

Definition 2.23 (i-th type-II horizontal character)**.** For $i \in \mathbb{N}_+$ and an $\mathbb{N}$-filtered or DR-filtered nilmanifold G/Γ. An *i-th type-II horizontal character* is a continuous homomorphisms $\xi_i \colon \mathrm{Horiz}_i(G) \to \mathbb{R}$ with $\xi_i(\mathrm{Horiz}_i(\Gamma)) \subseteq \mathbb{Z}$. We use $\mathfrak{R}_i(G/\Gamma)$ to denote the group of all i-th type-II horizontal characters.

The Mal'cev basis $\mathcal{X}$ induces a natural isomorphism $\psi \colon G_{[i,1]}/G_{[i,2]} \to \mathbb{R}^k$, and thus we may write $\eta_i(g_i) := (m_1, \dots, m_k) \cdot \psi(g_i)$ for some $m_1, \dots, m_k \in \mathbb{Z}$, where $k = \dim(G_{[i,1]}) - \dim(G_{[i,2]})$. We call the quantity $\|\eta_i\| := |m_1| + \dots + |m_k|$ the *complexity* of η_i (with respect to $\mathcal{X}$).

We provide a formula for further uses.

Lemma 2.24. Let $d, i \in \mathbb{N}_+$, G/Γ be an $\mathbb{N}$-filtered nilmanifold, $g(n) \in \text{poly}(\mathbb{Z}^d \to G_\mathbb{N})$, and $\xi_i \in \mathfrak{R}_i(G/\Gamma)$. Then

$$\xi_i(\text{Taylor}_i(g)(h_1, \dots, h_i)) = \Delta_{h_i} \dots \Delta_{h_1}\Big(\sum_{m \in \mathbb{N}^d, |m|=i} \xi_i(\text{Taylor}_m(g))\binom{n}{m}\Big). \tag{2.2}$$

Proof. By Corollary B.7 of [12], it is not hard to see that both sides of (2.2) are symmetric and multi-linear in $h_1, \dots, h_i$, and are independent of n (see also the remark on page 1281 of [12]). On the other hand, for any $m = (m_1, \dots, m_d) \in \mathbb{N}^d$ with $|m| = i$, if we take m_t many $h_1, \dots, h_i$ to be equal to e_t for all $1 \leq t \leq d$, then both sides of (2.2) equals to $\xi_i(\text{Taylor}_m(g))$. So by multi-linearity, we have that (2.2) holds. □

For p-periodic polynomial sequences, we have some p-periodicity properties for their type-II Taylor coefficients.

Lemma 2.25. Let $d, i \in \mathbb{N}_+$, p be a prime, G/Γ be an $\mathbb{N}$-filtered nilmanifold, $g \in \text{poly}_p(\mathbb{Z}^d \to G_\mathbb{N}|\Gamma)$, and ξ_i be an i-th type-II horizontal character of G/Γ. Then the map $(h_1, \dots, h_i) \mapsto \xi_i(\text{Taylor}_i(g)(h_1, \dots, h_i))$ is a homogeneous p-periodic symmetric polynomial of multi-degree $(1, \dots, 1)$. Moreover, $\text{Taylor}_m(g)^p \in G_{[i,2]}\Gamma$ and $\xi_i(\text{Taylor}_m(g)) \in \mathbb{Z}/p$ for all $m \in \mathbb{N}^d$ with $|m| = i$.

Proof. It follows from Corollary B.7 of [12] that $\xi_i(\text{Taylor}_i(g)(h_1, \dots, h_i))$ a homogeneous symmetric polynomial of multi-degree $(1, \dots, 1)$ (see also the remark on page 1281 of [12]). We now show that it is p-periodic. Note that if $h_1 \in p\mathbb{Z}^d$, then by the Baker-Campbell-Hausdorff formula,

$$\Delta_{h_i} \dots \Delta_{h_1} g(n) \equiv \Delta_{h_i} \dots \Delta_{h_2}(g_{h_1}(n)) \mod G_{[2,2]},$$

where $g_{h_1}(n) := g(n)^{-1}g(n + h_1)$ takes values in Γ. So $\Delta_{h_i} \dots \Delta_{h_1} g(n) \in G_{[2,2]}\Gamma$ and thus

$$\xi_i(\text{Taylor}_i(g)(h_1, \dots, h_i)) \equiv \xi_i(\Delta_{h_i} \dots \Delta_{h_1} g(n) \mod G_{[i,2]}) \in \mathbb{Z} \text{ if } h_1 \in p\mathbb{Z}^d. \tag{2.3}$$

Since $\xi_i(\text{Taylor}_i(g)(h_1, \dots, h_i))$ is multi-linear and symmetric, it is not hard to deduce from (2.3) that $\xi_i(\text{Taylor}_i(g)(h_1, \dots, h_i))$ is p-periodic.

So by Lemma 2.24, $\xi_i(\text{Taylor}_m(g)) \in \mathbb{Z}/p$ for all $m \in \mathbb{N}^d$ with $|m| = i$. Since ξ_i is an arbitrary i-th type-II horizontal character of G/Γ, we have that $\text{Taylor}_m(g)^p$ vanishes at every i-th type-II horizontal character of G/Γ. So $\text{Taylor}_m(g)^p \in G_{[i,2]}\Gamma$. □

2.6. **Ratner-type theorem.** Following the idea of [12], we use the factorization theorem to obtain a Ratner-type theorem. For an $\mathbb{N}$-filtered nilmanifold G/Γ and $i \in \mathbb{N}_+$, recall that $\mathfrak{R}_i(G/\Gamma)$ is the group of all i-th type-II horizontal characters.

Definition 2.26. Let $g(n) \in \text{poly}_p(\mathbb{Z}^d \to G_\mathbb{N}|\Gamma)$ and Ω be a subset of $\mathbb{Z}^d$. Let $\Xi_{\Omega,i}(g)$ denote the group of all $\xi_i \in \mathfrak{R}_i(G/\Gamma)$ such that

$$\xi_i(\text{Taylor}_i(g)(h_1, \dots, h_i)) = \xi_i(\Delta_{h_i} \dots \Delta_{h_1} g(n) \mod G_{[i,2]}) \in \mathbb{Z}$$

for all $(n, h_1, \dots, h_i) \in \square_{p,i}(\Omega)$.[8] Let

$$\Xi^{\perp}_{\Omega,i}(g) := \{x \in \mathrm{Horiz}_i(G) \colon \xi_i(x) \in \mathbb{Z} \text{ for all } \xi_i \in \Xi_{\Omega,i}(g)\}.$$

If Ω is a subset of $\mathbb{F}_p^d$ and $g = g' \circ \tau \in \mathrm{poly}_p(\mathbb{F}_p^d \to G_{\mathbb{N}}|\Gamma)$ for some $g' \in \mathrm{poly}_p(\mathbb{Z}^d \to G_{\mathbb{N}}|\Gamma)$, then denote $\Xi_{\Omega,i}(g) := \Xi_{\tau(\Omega),i}(g')$ and $\Xi^{\perp}_{\Omega,i}(g) := \Xi^{\perp}_{\tau(\Omega),i}(g')$.

Lemma 2.27. Suppose that $g = g' \circ \tau = g'' \circ \tau$ for some $g', g'' \in \mathrm{poly}_p(\tau(\Omega) \to G_{\mathbb{N}}|\Gamma)$. Then $\Xi_{\tau(\Omega),i}(g') = \Xi_{\tau(\Omega),i}(g'')$ and $\Xi^{\perp}_{\tau(\Omega),i}(g') = \Xi^{\perp}_{\tau(\Omega),i}(g'')$. In particular, $\Xi_{\Omega,i}(g)$ and $\Xi^{\perp}_{\Omega,i}(g)$ are well defined and are independent of the choice of g'.

Proof. It suffices to show that for all $\xi_i \in \mathfrak{R}_i(G/\Gamma)$ and $(n, h_1, \dots, h_i) \in \square_{p,i}(\tau(\Omega))$, we have that

$$\xi_i(\mathrm{Taylor}_i(g')(h_1, \dots, h_i)) = \xi_i(\mathrm{Taylor}_i(g'')(h_1, \dots, h_i)). \tag{2.4}$$

Note that $g'(n) = g''(n)$ for all $n \in \tau(\Omega)$. So for all $n \in \tau(\Omega)$ and $m, m' \in (\mathbb{Z}^d)^k$,

$$g'(n + pm)^{-1} g''(n + pm') = (g'(n + pm)^{-1} g'(n))(g''(n)^{-1} g''(n + pm')) \in \Gamma. \tag{2.5}$$

For all $(n, h_1, \dots, h_i) \in \square_{p,i}(\tau(\Omega))$, since $n + \epsilon_1 h_1 + \cdots + \epsilon_i h_i \in \tau(\Omega) + p(\mathbb{Z}^d)^k$ for all $\epsilon_1, \dots, \epsilon_i \in \{0, 1\}$, it follows from (2.5) that $\Delta_{h_i} \dots \Delta_{h_1}(g'^{-1} g'')(n) \in G_{[i,2]}\Gamma$. Therefore, we have that

$$\xi_i(\mathrm{Taylor}_i(g'^{-1} g'')(h_1, \dots, h_i)) = \xi_i(\Delta_{h_i} \dots \Delta_{h_1}(g'^{-1} g'')(n) \mod G_{[i,2]}) \in \mathbb{Z}$$

for all $(n, h_1, \dots, h_i) \in \square_{p,i}(\tau(\Omega))$. By Lemma 2.22, we get (2.4). □

We are now ready to state the Ratner-type theorem, in the proof of which we make essential use of the factorization theorem proved in the first part of the series [19]. For $i \in \mathbb{N}$, let $\pi_{\mathrm{Horiz}_i(G)} \colon G_i \to \mathrm{Horiz}_i(G)$ be the projection map.

Theorem 2.28 (Ratner-type theorem)**.** Let $0 < \delta < 1/2, C > 0$, $d, D \in \mathbb{N}_+$, $r, s \in \mathbb{N}$ and $p \gg_{C,\delta,d,D} 1$ be a prime. Let $M \colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a non-degenerate quadratic form and $V + c$ be an affine subspace of $\mathbb{F}_p^d$ of co-dimension r with $\mathrm{rank}(M|_{V+c}) = d - r$. Let G/Γ be an s-step $\mathbb{N}$-filtered nilmanifold of complexity at most C, equipped with an C-rational Mal'cev basis $\mathcal{X}$, and that $g \in \mathrm{poly}_p(V(M) \cap (V + c) \to G_{\mathbb{N}}|\Gamma)$. Let $F \in \mathrm{Lip}(G/\Gamma \to \mathbb{C}^D)$ be a function with $\|F\|_{\mathrm{Lip}(G/\Gamma \to \mathbb{C}^D)} \leq C$ such that

$$|\mathbb{E}_{n \in V(M) \cap (V+c)} F(g(n)\Gamma)| \geq \delta.$$

If $d \geq s + r + 13$ and $p \gg_{C,\delta,d,D} 1$, then there exist $\epsilon \in G$ of complexity $O_{C,\delta,D,d}(1)$ and a rational subgroup G_P of G which is $O_{C,\delta,d,D}(1)$-rational relative to $\mathcal{X}$ such that for all $i \in \mathbb{N}_+$, $\pi_{\mathrm{Horiz}_i(G)}(G_P \cap G_i) \geq \Xi^{\perp}_{V(M) \cap (V+c),i}(g)$ and

$$\Big| \int_{G_P/\Gamma_P} F(\epsilon x) \, dm_{G_P/\Gamma_P}(x) \Big| \geq \delta/2,$$

where m_{G_P/Γ_P} is the Haar measure of G_P/Γ_P.

[8] Note that $\xi_i(\Delta_{h_i} \dots \Delta_{h_1} g(n) \mod G_{[i,2]})$ is well defined since $\Delta_{h_i} \dots \Delta_{h_1} g(n) \in G_i$.

Proof. By taking components we may assume that F is scalar-valued. Let $\mathcal{F}: \mathbb{R}_+ \to \mathbb{R}_+$ be a growth function to be chosen later depending only on C, δ, d, D. By Theorem A.14, if $p \gg_{C,\delta,d-r,\mathcal{F}} 1$, then we may write

$$g = \epsilon g' \gamma,$$

where $\epsilon \in G$ is of complexity $O_{C'}(1)$, $g' \in \text{poly}_p(V(M) \cap (V+c) \to (G_P)_{\mathbb{N}}|\Gamma)$ with $(g'(n)\Gamma)_{n \in V(M)\cap(V+c)}$ being $\mathcal{F}(C')^{-1}$-equidistributed (see Appendix A.7 for definition) on some sub-nilmanifold $G_P/(G_P \cap \Gamma)$ of G/Γ for some subgroup G_P of G which is C'-rational relative to $\mathcal{X}$ for some $C \leq C' \leq O_{C,d-r,\mathcal{F}}(1)$, and $\gamma \in \text{poly}(V(M) \cap (V+c) \to G_{\mathbb{N}}|G_P \cap \Gamma)$. So

$$|\mathbb{E}_{n \in V(M)\cap(V+c)} F(\epsilon g'(n)\Gamma)| = |\mathbb{E}_{n \in V(M)\cap(V+c)} F(g(n)\Gamma)| \geq \delta. \tag{2.6}$$

Since $(g'(n)\Gamma)_{n \in V(M)\cap(V+c)}$ is $\mathcal{F}(C')^{-1}$-equidistributed on $G_P/(G_P \cap \Gamma)$, we have

$$\left|\mathbb{E}_{n \in V(M)\cap(V+c)} F(\epsilon g'(n)\Gamma) - \int_{G_P/\Gamma_P} F(\epsilon x)\, dm_{G_P/\Gamma_P}(x)\right| \leq \mathcal{F}(C')^{-1} \|F(\epsilon\cdot)\|_{\text{Lip}(G_P/\Gamma_P)}.$$

Since $\|F\|_{\text{Lip}(G/\Gamma)} \leq C$, we have that $\|F\|_{\text{Lip}(G_P/\Gamma_P)} \leq O_{C,C',\delta,s}(1)$ and so $\|F(\epsilon\cdot)\|_{\text{Lip}(G_P/\Gamma_P)} \leq O_{C,C',\delta,s}(1)$ by Lemma A.5 of [10]. Therefore, if $\mathcal{F}$ grows sufficiently fast depending on C, δ, d, D, then we have that

$$\left|\mathbb{E}_{n \in V(M)\cap(V+c)} F(\epsilon g'(n)\Gamma) - \int_{G_P/\Gamma_P} F(\epsilon x)\, dm_{G_P/\Gamma_P}(x)\right| < \delta/2$$

and so it follows from (2.6) that

$$\left|\int_{G_P/\Gamma_P} F(\epsilon x)\, dm_{G_P/\Gamma_P}(x)\right| \geq \delta/2.$$

It remains to show that

$$\pi_{\text{Horiz}_i(G)}(G_P \cap G_i) \geq \Xi^{\perp}_{V(M)\cap(V+c),i}(g)$$

for all $i \in \mathbb{N}_+$. If not, then by duality and the rational nature of G_P, there exists $\xi_i \in \mathfrak{N}_i(G/\Gamma)$ not belonging to $\Xi_{V(M)\cap(V+c),i}(g)$ which annihilates $\pi_{\text{Horiz}_i(G)}(G_P \cap G_i)$.

Write $g' = \tilde{g} \circ \tau$ and $\gamma = \tilde{\gamma} \circ \tau$ for some $\tilde{g} \in \text{poly}_p(\tau(V(M) \cap (V+c)) + p\mathbb{Z}^d \to (G_P)_{\mathbb{N}}|G_P \cap \Gamma)$ and $\tilde{\gamma} \in \text{poly}(\tau(V(M) \cap (V+c)) + p\mathbb{Z}^d \to G_{\mathbb{N}}|\Gamma)$. Then $g = (\epsilon\tilde{g}\tilde{\gamma}) \circ \tau$ and it follows from Lemma 2.17 that $\epsilon\tilde{g}\tilde{\gamma} \in \text{poly}_p(\tau(V(M) \cap (V+c)) + p\mathbb{Z}^d \to G_{\mathbb{N}}|\Gamma)$. By Lemma 2.22,

(2.7)

$$\begin{aligned} &\xi_i(\Delta_{h_i} \dots \Delta_{h_1} \epsilon\tilde{g}\tilde{\gamma}(n) \mod G_{[i,2]}) \\ =\ &\xi_i(\Delta_{h_i} \dots \Delta_{h_1} \epsilon \mod G_{[i,2]})\xi_i(\Delta_{h_i} \dots \Delta_{h_1} \tilde{g}(n) \mod G_{[i,2]})\xi_i(\Delta_{h_i} \dots \Delta_{h_1} \tilde{\gamma}(n) \mod G_{[i,2]}) \\ \equiv\ &\xi_i(\Delta_{h_i} \dots \Delta_{h_1} \tilde{\gamma}(n) \mod G_{[i,2]}) \mod \mathbb{Z}. \end{aligned}$$

Since $\tilde{\gamma}(n)$ takes values in Γ when $n \in \tau(V(M) \cap (V+c)) + p\mathbb{Z}^d$, we have that

$$\xi_i(\Delta_{h_i} \dots \Delta_{h_1} \tilde{\gamma}(n) \mod G_{[i,2]}) \in \mathbb{Z}$$

for all $(n, h_1, \dots, h_i) \in \square_{p,i}(\tau(V(M) \cap (V+c)))$. By (2.7), we have that

$$\xi_i(\Delta_{h_i} \dots \Delta_{h_1} \epsilon\tilde{g}\tilde{\gamma}(n) \mod G_{[i,2]}) \in \mathbb{Z}$$

for all $(n, h_1, \dots, h_i) \in \square_{p,i}(\tau(V(M) \cap (V + c)))$. So $\xi_i \in \Xi_{\tau(V(M)\cap(V+c)),i}(\epsilon\tilde{g}\tilde{\gamma})$. By definition, $\xi_i \in \Xi_{V(M)\cap(V+c),i}(g)$, a contradiction. This finishes the proof. □

2.7. **Nilsequences and nilcharacters.** We start with the definition of vertical torus and character, which plays an important role in the equiditribution properties on nilmanifolds.

Definition 2.29 (Vertical torus and character)**.** Let I be the degree, multi-degree, or degree-rank ordering, and J be a finite down set of I having a maximum element s. Let G/Γ be a nilmanifold with an I-filtration $(G_i)_{i\in I}$ with a Mal'cev basis $\mathcal{X}$ adapted to it. Then G_s lies in the center of G. The *vertical torus* of G/Γ is the set $G_s/(\Gamma \cap G_s)$. A *vertical character* of G/Γ (with respect to the filtration $(G_i)_{i\in I}$) is a continuous homomorphism $\xi\colon G_s \to \mathbb{R}$ such that $\xi(\Gamma \cap G_s) \subseteq \mathbb{Z}$ (in particular, ξ descents to a continuous homomorphism from the vertical torus $G_s/(\Gamma \cap G_s)$ to $\mathbb{T}$). Let $\psi\colon G \to \mathbb{R}^m$ be the Mal'cev coordinate map and denote $m_s = \dim(G_s)$. Then there exists $k \in \mathbb{Z}^{m_s}$ such that $\xi(g_s) = (\mathbf{0}, k) \cdot \psi(g_s)$ for all $g_s \in G_s$ (note that the first $\dim(G) - m_s$ coordinates of $\psi(g_s)$ are all zero). We call the quantity $\|\xi\| := |k|$ the *complexity* of ξ (with respect to $\mathcal{X}$).

We now define nilsequences and nilcharacters. In addition to the conventional definitions, we also define periodic and partially periodic nilsequences and nilcharacters. Recall that for $D \in \mathbb{N}_+$, $\mathbb{S}^D$ denotes the set of $(z_1, \dots, z_D) \in \mathbb{C}^D$ with $|z_1|^2 + \dots + |z_D|^2 = 1$.

Definition 2.30 (Nilsequences and nilcharacters)**.** Let $D, k \in \mathbb{N}_+$, I be the degree, multi-degree, or degree-rank ordering such that $\dim(I)|k$, and J be a finite down set of I. Let Ω be a subset $\mathbb{Z}^k$. We say that a function $\phi\colon \Omega \to \mathbb{C}^D$ is an *(I-filtered) nilsequence* on Ω of degree $\subseteq J$ if

$$\phi(n) = F(g(n)\Gamma) \text{ for all } n \in \Omega \tag{2.8}$$

for some nilmanifold G/Γ with an I-filtration $(G_i)_{i\in I}$ of degree $\subseteq J$ equipped with some smooth Riemannian metric $d_{G/\Gamma}$, some $g \in \text{poly}(\mathbb{Z}^k \to G_I)$, and some Lipschitz function $F\colon G/\Gamma \to \mathbb{C}^D$. We say that D is the *dimension* of ϕ. If in addition,

- J has a maximal element s;
- $F(g_s x) = \exp(\eta(g_s))F(x)$ for all $x \in G/\Gamma$ and $g_s \in G_s$ for some vertical character $\eta\colon G_s \to \mathbb{R}$ (we say that η is the *vertical frequency* of F);
- F takes values in $\mathbb{S}^D$,

then we say that ϕ is a *k'-integral (I-filtered) nilcharacter* on Ω of degree $\subseteq J$, where $k' = \frac{k}{\dim(I)}$. Following Convention 2.14, we omit the phrase k'-integral throughout as the value of k' is clear given I and Ω. A similar remark applies to the degree and the degree-rank filtrations.

For any $s \in I$, we say that a function $\phi\colon \Omega \to \mathbb{C}^D$ is a (I-filtered) nilsequence/nilcharacter on Ω of degree $\leq s$ if it is a (I-filtered) nilsequence/nilcharacter on Ω of degree $\subseteq \{i \in I\colon i \leq s\}$.

If in addition, the polynomial sequence g in (2.8) belongs to $\text{poly}_p(\Omega \to G_I|\Gamma)$, then we say that ϕ is a *partially p-periodic* (I-filtered) nilsequence/nilcharacter on Ω.

If in addition, the polynomial sequence g in (2.8) belongs to $\text{poly}_p(\mathbb{Z}^k \to G_I|\Gamma)$, then we say that ϕ is a *p-periodic* (I-filtered) nilsequence/nilcharacter on Ω.[9]

Let Ω be a subset of $\mathbb{F}_p^k$. Denote $\Omega' = \iota^{-1}(\Omega) = \tau(\Omega) + p\mathbb{Z}^k$. A function $\phi\colon \Omega \to \mathbb{C}^D$ is an *(I-filtered) nilsequence/nilcharacter* on Ω of degree $\leq s$ (or $\subseteq J$) if $\phi(n) = \phi' \circ \tau(n)$ for all $n \in \Omega$ for some (I-filtered) nilsequence/nilcharacter $\phi'\colon \Omega' \to \mathbb{C}^D$ on Ω of degree $\leq s$ (or $\subseteq J$). We say that ϕ is *partially p-periodic/p-periodic* (or simply *partially perodic/periodic*) on Ω if we may require ϕ' to be partially p-perodic/p-periodic on Ω'.

Note that we allow nilsequences and nilcharacters to be vector valued (i.e. $D \geq 2$) in order to avoid certain topological obstructions in the constructions of nilsequences. See [12] pages 1252–1255 for further discussions.

We next define the spaces of nilsequences and nilcharacters that we will work with in this paper.

Definition 2.31 (Spaces of nilsequences and nilcharacters)**.** Let I be the degree, multi-degree, or degree-rank ordering, $k \in \mathbb{N}_+$ with $\dim(I)|k$, J be a finite down set of I, p be a prime, and Ω be a subset of $\mathbb{Z}^k$ or $\mathbb{F}_p^k$.

- Let $\text{Nil}^J(\Omega)$ denote the set of all nilsequences on Ω of degree $\subseteq J$.
- Let $\text{Nil}_p^J(\Omega)$ denote the set of all p-periodic nilsequences on Ω of degree $\subseteq J$.
- Let $\Xi_p^J(\Omega)$ denote the set of p-periodic nilcharacters on Ω, if J has a maximum element.
- We define $\text{Nil}^{<J}(\Omega) := \text{Nil}^{J'}(\Omega)$ and $\text{Nil}_p^{<J}(\Omega) := \text{Nil}_p^{J'}(\Omega)$, where J' is the largest down set of I which is contained in but not equals to J. Since I is the degree, multi-degree or the degree-rank ordering, it is not hard to see that such J' exists and is unique.
- For $s \in I$, we write $\text{Nil}^s(\Omega) := \text{Nil}^J(\Omega)$ and $\text{Nil}^{<s}(\Omega) := \text{Nil}^{J'}(\Omega)$, where J is the down set consisting of all $i \in I$ which does not exceed s and J' is the down set consisting of all $i \in I$ which does not exceed s', where s' is the largest element which is smaller than s (if such s' exists). We use the similar notations for $\text{Nil}_p^s(\Omega)$, $\text{Nil}_p^{<s}(\Omega)$, $\Xi_p^s(\Omega)$ and $\Xi_p^{<s}(\Omega)$.

For convenience, we introduce the following definition.

Definition 2.32 (Representations of nilsequences and nilcharacters)**.** Let I be the degree, multi-degree, or degree-rank ordering, $k \in \mathbb{N}_+$ with $\dim(I)|k$, J be a finite down set of I, p be a prime, and Ω be a subset of $\mathbb{Z}^k$.

- For any $\phi \in \text{Nil}^J(\Omega)$, we may write $\phi(n) = F(g(n)\Gamma), n \in \Omega$ for some degree-J nilmanifold $(G/\Gamma)_I$, $g \in \text{poly}(\mathbb{Z}^k \to G_I)$, and Lipschitz function $F\colon G/\Gamma \to \mathbb{C}^D$. We say that $((G/\Gamma)_I, g, F)$ is a *$\text{Nil}^J(\Omega)$-representation* of ϕ.

[9]Note that we only define p-periodic nilsequence/nilcharacter on the designated set Ω. However, one can easily extend the definition of such a sequence to the whole space $\mathbb{Z}^k$ by requiring (2.8) to hold for all $n \in \mathbb{Z}^k$.

- For any $\phi \in \mathrm{Nil}^J_p(\Omega)$, we may write $\phi(n) = F(g(n)\Gamma), n \in \Omega$ for some degree-J nilmanifold $(G/\Gamma)_I$, $g \in \mathrm{poly}_p(\mathbb{Z}^k \to G_I|\Gamma)$, and Lipschitz function $F\colon G/\Gamma \to \mathbb{C}^D$. We say that $((G/\Gamma)_I, g, F)$ is a *$Nil^J_p(\Omega)$-representation* of ϕ.
- For any $\xi \in \Xi^J_p(\Omega)$, we may write $\chi(n) = F(g(n)\Gamma), n \in \Omega$ for some degree-J nilmanifold $(G/\Gamma)_I$, $g \in \mathrm{poly}(\mathbb{Z}^k \to G_I|\Gamma)$, and Lipschitz function $F\colon G/\Gamma \to \mathbb{C}^D$ having a vertical frequency η. We say that $((G/\Gamma)_I, g, F, \eta)$ is a *$\Xi^J_p(\Omega)$-representation* of χ.
- The complexity of a $\mathrm{Nil}^J(\Omega)/\mathrm{Nil}^J_p(\Omega)/\Xi^J_p(\Omega)$-representation is the the maximum of the complexities of G/Γ, F and η (if η appears in the corresponding definitions). We say that ϕ is of *complexity* at most C as an element of $\mathrm{Nil}^J(\Omega)/\mathrm{Nil}^J_p(\Omega)/\Xi^J_p(\Omega)$ if ϕ admits a $\mathrm{Nil}^J(\Omega)/\mathrm{Nil}^J_p(\Omega)/\Xi^J_p(\Omega)$-representation of complexity at most C.
- Let $\Omega' \subseteq \mathbb{F}_p^k$ and $\phi \in \mathrm{Nil}^J(\Omega')$. We say that $((G/\Gamma)_I, g, F)$ is a *$Nil^J(\Omega')$-representation* of ϕ if it is a $\mathrm{Nil}^J(\tau(\Omega') + p\mathbb{Z}^k)$-representation of some $\phi' \in \mathrm{Nil}^J(\tau(\Omega') + p\mathbb{Z}^k)$ such that $\phi(n) = \phi' \circ \tau(n)$ for all $n \in \Omega'$. We say that ϕ is of *complexity* at most C as an element of $\mathrm{Nil}^J(\Omega')$ if ϕ admits a $\mathrm{Nil}^J(\Omega')$-representation of complexity at most C. We adopt similar notations for $\mathrm{Nil}^J_p(\Omega')$- and $\Xi^J_p(\Omega')$-representations and complexities.
- We define similar notations for $\mathrm{Nil}^{<J}(\Omega)$, $\mathrm{Nil}^{<s}(\Omega)$, ect.

Remark 2.33. In certain situations, a sequence can belong to multiple sets defined above, and have different complexities in different sets. For example, a sequence ϕ can simultaneously belong to $\Xi^{s+1}_p(\Omega)$ and $\mathrm{Nil}^s(\Omega)$. In this case, the complexity of ϕ as an element in $\Xi^{s+1}_p(\Omega)$ and the complexity of ϕ as an element in $\mathrm{Nil}^s(\Omega)$ are different. We will specify which complexity we are referring to when such ambiguity happens.

Definition 2.34 (Complexity of nilsequences and nilcharacters)**.** Let $\mathrm{Nil}^{J;C}(\Omega)$ denote the set of all nilsequences in $\mathrm{Nil}^J(\Omega)$ of complexity at most C (as an element in $\mathrm{Nil}^{J;C}(\Omega)$), and $\mathrm{Nil}^{J;C,D}(\Omega)$ denote the set of all nilsequences in $\mathrm{Nil}^{J;C}(\Omega)$ of complexity at most C and dimension at most D. We adopt similar notations for $\mathrm{Nil}^J_p(\Omega)$, $\Xi^J_p(\Omega)$, $\mathrm{Nil}^{<J}(\Omega)$, $\mathrm{Nil}^s(\Omega)$, ect.

We refer the readers to Appendix B for some approximation properties for nilsequeneces which will be used in this paper.

Sometimes it is convenient to identify two nilcharacters when their difference is a nilsequence of lower degree. Therefore, we introduce the following notion:

Definition 2.35 (An equivalence relation for nilcharacters)**.** Let $k \in \mathbb{N}_+$, $C > 0$, p be a prime, Ω be a subset of $\mathbb{F}_p^k$, I be the degree, multi-degree, or degree-rank ordering with $\dim(I)|k$ and let $s \in I$. For $\chi, \chi' \in \Xi^s_p(\Omega)$, we write $\chi \sim_C \chi' \mod \Xi^s_p(\Omega)$ if $\chi \otimes \overline{\chi} \in \mathrm{Nil}^{<s;C}(\Omega)$.[10] We write $\chi \sim \chi' \mod \Xi^s_p(\Omega)$ if $\chi \sim_C \chi' \mod \Xi^s_p(\Omega)$ for some $C > 0$.

[10]Note that if χ and χ' to are p-periodic s-step nilcharacters, then so is $\chi \otimes \overline{\chi}$. However, if $\chi \otimes \overline{\chi}$ coincides with a $\prec s$-step nilsequence, it is unclear whether $\chi \otimes \overline{\chi}$ is p-periodic as a $\prec s$-step nilsequence. In our definition, if $\chi \sim_C \chi' \mod \Xi^s_p(\Omega)$, then we do not require $\chi \otimes \overline{\chi}$ to be p-periodic as a $\prec s$-step nilsequence.

We summarized some properties for the equivalence relation $\sim$ in Appendix C for later uses.

3. Some preliminary results

In this section, we provide some preliminary results and setup the initial steps for the proof of SGI(s).

3.1. **The proof of SGI**(1)**.** Note that SGI(0) holds trivially by taking $\delta = \epsilon$ and $\phi \equiv 1$. We now prove SGI(1). It turns out that the second local Gowers norm is connected to the Fourier coefficients. For a function $f\colon \mathbb{F}_p^d \to \mathbb{C}$, let $\widehat{f}\colon \mathbb{F}_p^d \to \mathbb{C}$ denote the Fourier transform of f given by $\widehat{f}(\xi) := \mathbb{E}_{x\in\mathbb{F}_p^d} f(x)\exp(-\frac{1}{p}\tau(\xi\cdot x))$. We have

Lemma 3.1. Let $d \in \mathbb{N}$ with $d \geq 9$, p be a prime, $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a non-degenerate quadratic form, and $f\colon \mathbb{F}_p^d \to \mathbb{C}$ with $|f| \leq 1$. Then

$$\|f\|^4_{U^2(V(M))} \leq \sup_{\xi\in\mathbb{F}_p^d} p^{1-d}|\widehat{\mathbf{1}_{V(M)}f}(\xi)|(1 + O(p^{-\frac{1}{8}})).$$

Proof. The approach we use is similar to the proof of Theorem 10 of [16]. By Example 9.5 of [19], $\Box_2(V(M))$ is an M-set of total co-dimension 4. By Lemma A.7, since $d \geq 9$, we have that

$$|\Box_2(V(M))| = p^{3d-4}(1 + O(p^{-1/2})).$$

For convenience denote $\delta := \mathbf{1}_{V(M)}$. Then

$$\begin{aligned}
\|f\|^4_{U^2(V(M))} &= \mathbb{E}_{(n,h_1,h_2)\in\Box_2(V(M))} \prod_{\epsilon=(\epsilon_1,\epsilon_2)\in\{0,1\}^2} C^{|\epsilon|} f(n + \epsilon_1 h_1 + \epsilon_2 h_2) \\
&= \frac{p^{3d}}{|\Box_2(V(M))|} \cdot \mathbb{E}_{(n,h_1,h_2)\in(\mathbb{F}_p^d)^3} \prod_{\epsilon=(\epsilon_1,\epsilon_2)\in\{0,1\}^2} C^{|\epsilon|}(\delta f)(n + \epsilon_1 h_1 + \epsilon_2 h_2) \\
&\leq p^4 \mathbb{E}_{n\in\mathbb{F}_p^d}\delta(n)\Big|\mathbb{E}_{y,z,w\in\mathbb{F}_p^d}\delta\overline{f}(y)\delta\overline{f}(z)\delta f(w)\mathbf{1}_{n=y+z-w}\Big|\Big(1 + O(p^{-\frac{1}{2}})).
\end{aligned}$$

By the Cauchy-Schwartz inequality,

$$\|f\|^8_{U^2(V(M))} \leq p^8\Big(\mathbb{E}_{n\in\mathbb{F}_p^d}\delta(n)\Big)\Big(\mathbb{E}_{n\in\mathbb{F}_p^d}\delta(n)\Big|\mathbb{E}_{y,z,w\in\mathbb{F}_p^d}\delta\overline{f}(y)\delta\overline{f}(z)\delta f(w)\mathbf{1}_{n=y+z-w}\Big|^2\Big)\Big(1 + O(p^{-1})\Big).$$

By Lemma A.7,

$$\begin{aligned}
&\|f\|^8_{U^2(V(M))}\\
\leq\ & p^7\mathbb{E}_{n\in\mathbb{F}_p^d}\delta(n)\mathbb{E}_{y_1,z_1,w_1,y_2,z_2,w_2\in\mathbb{F}_p^d}\prod_{j=1,2}C^j\Big(\delta\overline{f}(y_j)\delta\overline{f}(z_j)\delta f(w_j)\mathbf{1}_{n=y_j+z_j-w_j}\Big)\Big(1+O(p^{-\frac{1}{2}})\Big)\\
=\ & p^{7-d}\mathbb{E}_{y_1,z_1,w_1,y_2,z_2,w_2\in\mathbb{F}_p^d}\\
&\qquad\prod_{j=1,2}C^j\Big(\delta\overline{f}(y_j)\delta\overline{f}(z_j)\delta f(w_j)\Big)\delta(y_1+z_1-w_1)\mathbf{1}_{y_1+z_1-w_1=y_2+z_2-w_2}\Big(1+O(p^{-\frac{1}{2}})\Big)\\
\leq\ & p^{7-d}\mathbb{E}_{w_1,w_2\in\mathbb{F}_p^d}\delta(w_1)\delta(w_2)\\
&\qquad\cdot\Big|\mathbb{E}_{y_1,z_1,y_2,z_2\in\mathbb{F}_p^d}\delta(y_1+z_1-w_1)\mathbf{1}_{y_1+z_1-w_1=y_2+z_2-w_2}\prod_{j=1,2}C^j(\delta f(y_j)\delta\overline{f}(z_j))\Big|\Big(1+O(p^{-\frac{1}{2}})\Big).
\end{aligned}$$

Again by a similar argument using the Cauchy-Schwartz inequality and Lemma A.7, we may bound $\|f\|^{16}_{U^2(V(M))}$ by

$$p^{12-3d}\mathbb{E}_{y_j,z_j\in\mathbb{F}_p^d,1\leq j\leq 4}\prod_{j=1}^4 C^j(\delta f(y_j)\delta\overline{f}(z_j))\mathbf{1}_{y_2-y_1+z_2-z_1=y_4-y_3+z_4-z_3}W_{y_1,\dots,z_4}\Big(1+O(p^{-\frac{1}{2}})\Big),$$

where $W_{y_1,\dots,z_4}$ is the qunatitiy

$$\mathbb{E}_{w\in\mathbb{F}_p^d}\delta(w)\delta(y_2-y_1+z_2-z_1+w)\delta(y_1+z_1-w)\delta(y_3+z_3-w)=\frac{|V(M)\cap V_{y_1,\dots,z_4}|}{p^d},$$

with $V_{y_1,\dots,z_4}$ being the set of $w\in\mathbb{F}_p^d$ such that $M(w)=M(y_2-y_1+z_2-z_1+w)=M(y_1+z_1-w)=M(y_3+z_3-w)$.

If $y_2-y_1+z_2-z_1, y_1+z_1, y_3+z_3$ are linearly independent, then $V_{y_1,\dots,z_4}$ is an affine subspace of $\mathbb{F}_p^d$ of co-dimension 3. Since $d\geq 9$, by Lemma A.7, we have that $W_{y_1,\dots,z_4}=p^{-4}(1+O(p^{-\frac{1}{2}}))$. On the other hand, by Lemma A.5, it is not hard to compute that the number of tuples $(y_1,y_2,y_3,z_1,z_2,z_3)\in(\mathbb{F}_p^d)^6$ such that $y_2-y_1+z_2-z_1, y_1+z_1, y_3+z_3, V$ are not linearly independent is at most $3p^{5d+2}$. So $\|f\|^{16}_{U^2(V(M))}$ is bounded by

$$p^{8-3d}\mathbb{E}_{y_j,z_j\in\mathbb{F}_p^d,1\leq j\leq 4}\prod_{j=1}^4 C^j(\delta f(y_j)\delta\overline{f}(z_j))\mathbf{1}_{y_2-y_1+z_2-z_1=y_4-y_3+z_4-z_3}\Big(1+O(p^{-\frac{1}{2}})\Big). \tag{3.1}$$

Since

$$\mathbf{1}_{y_2-y_1+z_2-z_1=y_4-y_3+z_4-z_3}=p^{-d}\sum_{\xi\in\mathbb{F}_p^d}\exp\Big(-\frac{1}{p}\tau(\xi\cdot(y_4-y_3-y_2+y_1+z_4-z_3-z_2+z_1))\Big),$$

we may rewrite (3.1) as $p^{8-4d}\sum_{\xi\in\mathbb{F}_p^d}|\widehat{\delta f}(\xi)|^8(1+O(p^{-\frac{1}{2}}))$. So

$$
\begin{aligned}
\|f\|_{U^2(V(M))}^{16} &\leq p^{8-4d}\sum_{\xi\in\mathbb{F}_p^d}|\widehat{\delta f}(\xi)|^8(1+O(p^{-\frac{1}{2}}))\\
&\leq p^{8-4d}\sup_{\xi\in\mathbb{F}_p^d}|\widehat{\delta f}(\xi)|^4\sum_{\xi\in\mathbb{F}_p^d}|\widehat{\delta f}(\xi)|^4(1+O(p^{-\frac{1}{2}}))\\
&= p^{8-4d}\sup_{\xi\in\mathbb{F}_p^d}|\widehat{\delta f}(\xi)|^4\|\delta f\|_{U^2(\mathbb{F}_p^d)}^4(1+O(p^{-\frac{1}{2}}))\text{ (see page 14 of [16])}\\
&= p^{4-4d}\sup_{\xi\in\mathbb{F}_p^d}|\widehat{\delta f}(\xi)|^4\|f\|_{U^2(V(M))}^4(1+O(p^{-\frac{1}{2}}))\\
&\leq p^{4-4d}\sup_{\xi\in\mathbb{F}_p^d}|\widehat{\delta f}(\xi)|^4(1+O(p^{-\frac{1}{2}})).
\end{aligned}
$$

This finishes the proof. □

We are now ready to prove SGI(1). Suppose that $\|f\|_{U^2(V(M))}>\epsilon$. By Lemma 3.1, if $d\geq 9$ and $p\gg_\epsilon 1$, then there exists $\xi\in\mathbb{F}_p^d$ such that $p^{1-d}\left|\widehat{\mathbf{1}_{V(M)}f}(\xi)\right|\geq(\epsilon/2)^4$. By Lemma A.7,

$$
\begin{aligned}
(\epsilon/2)^4 \leq p^{1-d}\left|\widehat{\mathbf{1}_{V(M)}f}(\xi)\right| &= p^{1-d}\left|\mathbb{E}_{n\in\mathbb{F}_p^d}\mathbf{1}_{V(M)}f(n)\cdot\exp(-\frac{\tau(\xi\cdot n)}{p})\right|\\
=\left|\mathbb{E}_{n\in V(M)}f(n)\cdot\exp(-\frac{\tau(\xi\cdot n)}{p})\right|+O(p^{-\frac{1}{2}}).
\end{aligned}
$$

It is clear that the map $n\mapsto\exp(-\frac{\tau(\xi\cdot n)}{p})$ is a 1-step p-periodic nilsequence of complexity $O(1)$. This completes the proof of SGI(1).

3.2. **A preliminary reduction for SGI(s).** For our applications, we also need to use the following improved version of the inverse theorem, which we denote as SGI(s)+:

Proposition 3.2. [SGI(s)+] Let $d\in\mathbb{N}_+, r,s\in\mathbb{N}$, with $d-r\geq N(s-1)$ and $\epsilon>0$. There exist $\delta:=\delta(d-r,\epsilon), C:=C(d-r,\epsilon)>0$, and $p_0:=p_0(d-r,\epsilon)\in\mathbb{N}$ such that for every prime $p\geq p_0$, every non-degenerate quadratic form $M\colon\mathbb{F}_p^d\to\mathbb{F}_p$, every affine subspace $V+c$ of $\mathbb{F}_p^d$ of co-dimension r, and every function $f\colon\mathbb{F}_p^d\to\mathbb{C}$ bounded in magnitude by 1, if $\text{rank}(M|_{V+c})=d-r$, and $\|f\|_{U^{s+1}(V(M)\cap(V+c))}>\epsilon$, then there exists $\phi\in\text{Nil}_p^{s;C,1}(\mathbb{F}_p^d)$ such that

$$\left|\mathbb{E}_{n\in V(M)\cap(V+c)}f(n)\cdot\phi(n)\right|>\delta.$$

Lemma 3.3. We have that SGI(s) ⇒ SGI(s)+.

Proof. Let $L\colon\mathbb{F}_p^{d-r}\to V$ be a bijective linear transformation and $\tilde{M}\colon\mathbb{F}_p^{d-r}\to\mathbb{F}_p$ be the quadratic form given by $\tilde{M}(m):=M(L(m)+c)$. It is not hard to see that

$$\|f\|_{U^{s+1}(V(M)\cap(V+c))}=\|f(L(\cdot)+c)\|_{U^{s+1}(V(\tilde{M}))}.$$

Since $\mathrm{rank}(M|_{V+c}) = d - r$, $\tilde{M}$ is non-degenerate. By SGI(s), there exist $\delta := \delta(d - r, \epsilon), C := C(d-r, \epsilon) > 0$, $p_0 := p_0(d-r, \epsilon) \in \mathbb{N}$, and $\phi \in \mathrm{Nil}_p^{s;C,1}(\mathbb{F}_p^{d-r})$ such that if $p \geq p_0$, then

$$\left|\mathbb{E}_{m\in V(\tilde{M})} f(L(m)+c)\cdot\phi(m)\right| = \left|\mathbb{E}_{n\in V(M)\cap(V+c)} f(n)\cdot\phi(L^{-1}(n-c))\right| > \delta.$$

Assume that $\phi(m) = F(g(m)\Gamma)$ for some s-step nilmanifold G/Γ of complexity at most C, Lipschitz function $F\colon G/\Gamma \to \mathbb{C}$ of complexity at most C, and some $g \in \mathrm{poly}_p(\mathbb{F}_p^d \to G_{\mathbb{N}}|\Gamma)$. By Proposition 2.18, there exists $g' \in \mathrm{poly}_p(\mathbb{F}_p^d \to G_{\mathbb{N}}|\Gamma)$ such that $g'(n)\Gamma = g(L^{-1}(n-c))\Gamma$ for all $n \in \mathbb{F}_p^d$. So

$$\left|\mathbb{E}_{n\in V(M)\cap(V+c)} f(n)\cdot F(g'(n)\Gamma)\right| = \left|\mathbb{E}_{n\in V(M)\cap(V+c)} f(n)\cdot\phi(L^{-1}(n-c))\right| > \delta.$$

□

Now suppose that SGI(s) holds for some $s \geq 1$ and our goal is to show SGI($s+1$). By Lemma 3.3, SGI(s)+ holds. Suppose that $\|f\|_{U^{s+2}(V(M))} > \epsilon$. By Example 9.5 of [19], $\Box_{s+2}(V(M))$ is an M-set of total co-dimension $(s^2+5s+8)/2$. Note that for all $h_{s+2} \in \mathbb{F}_p^d$, $(n, h_1, \dots, h_{s+2}) \in \Box_{s+2}(V(M))$ if and only if $(n, h_1, \dots, h_{s+1}) \in \Box_{s+1}(V(M)^{h_{s+2}})$ (recall Section 1.3 for definitions). Since $d \geq s^2 + 5s + 9$, by Theorem A.12, we have that

$$\|f\|_{U^{s+2}(V(M))}^{2^{s+2}} = \mathbb{E}_{(n,h_1,\dots,h_{s+2})\in\Box_{s+2}(V(M))} \prod_{\epsilon=(\epsilon_1,\dots,\epsilon_{s+2})\in\{0,1\}^{s+2}} C^{|\epsilon|} f(n+\epsilon_1 h_1+\dots+\epsilon_{s+2}h_{s+2})$$

$$= \mathbb{E}_{h_{s+2}\in\mathbb{F}_p^d}\mathbb{E}_{(n,h_1,\dots,h_{s+1})\in\Box_{s+1}(V(M)^{h_{s+2}})} \prod_{\epsilon=(\epsilon_1,\dots,\epsilon_{s+1})\in\{0,1\}^{s+1}} C^{|\epsilon|}\overline{\Delta_h f}(n+\epsilon_1 h_1+\dots+\epsilon_{s+1}h_{s+1}) + O_s(p^{-\frac{1}{2}})$$

$$= \mathbb{E}_{h_{s+2}\in\mathbb{F}_p^d}\|\Delta_{h_{s+2}} f\|_{U^{s+1}(V(M)^{h_{s+2}})}^{2^{s+1}} + O_s(p^{-\frac{1}{2}}).$$

By the Pigeonhole Principle, if $p \gg_{\epsilon,s} 1$, then there exists a subset H of $\mathbb{F}_p^d$ of cardinality $\gg_{d,\epsilon} p^d$ such that for all $h \in H$, we have that

$$\|\Delta_h f\|_{U^{s+1}(V(M)^h)}^{2^{s+1}} \gg_{d,\epsilon} 1.$$

Since $d \geq 3$, by Lemma A.7 and passing to a subset if necessary, we may further assume that $(hA)\cdot h \neq 0$ for all $h \in H$, where A is the matrix associated with M.

Fix $h \in H$. Note that $V(M)^h$ is the intersection of $V(M)$ with an affine subspace W_h of $\mathbb{F}_p^d$ of co-dimension 1. Since $(hA)\cdot h \neq 0$, $\mathrm{span}_{\mathbb{F}_p}\{h\} \cap \mathrm{span}_{\mathbb{F}_p}\{h\}^{\perp_M} = \{\mathbf{0}\}$. By Proposition A.4, $\mathrm{rank}(M|_{W_h}) = d-1$. By SGI($s$)+, Corollary B.6 and the Pigeonhole Principle, for all $h \in H$, there exists an s-step nilcharacter $\chi_h \in \Xi_p^{s;O_{d,\epsilon}(1),O_{d,\epsilon}(1)}(\mathbb{F}_p^d)$ such that

$$\left|\mathbb{E}_{n\in V(M)^h} f(n+h)\overline{f}(n)\otimes\chi_h(n)\right| \gg_{d,\epsilon} 1. \tag{3.2}$$

This completes Step 1 described in Section 1.2.

4. A Furstenberg-Weiss type argument

In this section, we conduct Step 2 described in Section 1.2 by showing the following intermediate result:

Proposition 4.1. Let $d, D, s \in \mathbb{N}_+, d \geq 9, C, \epsilon > 0$, $p \gg_{C,d,D,\epsilon,s} 1$ be a prime, $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a non-degenerate quadratic form, and $f\colon \mathbb{F}_p^d \to \mathbb{C}$ be bounded by 1. Suppose that for some $1 \leq r_* \leq s$, there exist a subset $H \subseteq \mathbb{F}_p^d$ with $|H| \geq \epsilon p^d$, some $\chi_0 \in \Xi_p^{(1,s);C,D}((\mathbb{F}_p^d)^2)$, some $\chi_h \in \Xi_p^{[s,r_*];C,D}(\mathbb{F}_p^d)$, and some $\psi_h \in \mathrm{Nil}_p^{s-1;C,1}(\mathbb{F}_p^d)$ for all $h \in H$ such that

$$\left|\mathbb{E}_{n\in V(M)^h} f(n+h)\overline{f}(n)\chi_0(h,n)\otimes\chi_h(n)\psi_h(n)\right| > \epsilon \tag{4.1}$$

for all $h \in H$. Then there exist a subset $U \subseteq H^3$ with $|U| \gg_{C,d,D,\epsilon,s} p^{3d}$ such that writing

$$\chi_{h_1,h_2,h_3}(n) := \chi_{h_1}(n)\otimes\chi_{h_2}(n+h_3-h_2)\otimes\overline{\chi}_{h_3}(n)\otimes\overline{\chi}_{h_1+h_2-h_3}(n+h_3-h_2),$$

we have that

$$\left|\mathbb{E}_{n\in V(M)^{h_1,h_3,h_3-h_2}}\chi_{h_1,h_2,h_3}(n)\psi_{h_1,h_2,h_3}(n)\right| \gg_{C,d,D,\epsilon,s} 1 \tag{4.2}$$

for some $\psi_{h_1,h_2,h_3} \in \mathrm{Nil}_p^{s-1;O_{C,d,D,\epsilon,s}(1),1}(\mathbb{F}_p^d)$ for all $(h_1,h_2,h_3) \in U$. Moreover, for all (h_1,h_2,h_3) in U, h_1,h_2,h_3 are linearly independent and M-non-isotropic (see Appendix A.2 for the definition).

Proof. Our strategy is similar to the one used in [4, 12]. An additional difficulty is that the Fubini's theorem in our setting (Theorem A.12) is more delicate, which makes the computations more complicated.

By (4.1), Proposition B.7 and the Pigeonhole Principle, we have that

$$\left|\mathbb{E}_{n\in V(M)^h}\Delta_h f(n)c(n)^{\tau(h)}\psi(n)\otimes\chi_h(n)\psi_h(n)\right| \gg_{C,d,D,\epsilon,s} 1$$

for some function $\psi\colon \mathbb{F}_p^d \to \mathbb{C}$ bounded by 1 and some function $c = (c_1,\dots,c_d)$ with $c_1,\dots,c_d\colon \mathbb{F}_p^d \to \mathbb{S}$ such that $c_i(n)^{pm} = 1$ for all $n \in \mathbb{F}_p^d$ and $m \in \mathbb{Z}^d$, and that the map $n \mapsto c(n-\ell)c(n)^{\tau(h)}$ belongs to $\mathrm{Nil}_p^{s-1;O_{C,d,D,\epsilon,s}(1)}(\mathbb{F}_p^d)$, where $c(n)^{\tau(h)} := c_1(n)^{\tau(h_1)}\dots c_d(n)^{\tau(h_d)}$. Taking an average over H, we have that

$$\mathbb{E}_{h\in\mathbb{F}_p^d}\left|\mathbb{E}_{n\in V(M)^h}\Delta_h f(n)c(n)^{\tau(h)}\psi(n)\otimes\chi_h(n)\psi_h(n)\right| \gg_{C,d,D,\epsilon,s} 1$$

(for those h such that χ_h and ψ_h are not yet defined, set $\chi_h = \psi_h \equiv 0$[11]). By switching $\psi_h(n)$ to $a_h\psi_h(n)$ for some $a_h \in \mathbb{S}$ if necessary, we may assume without loss of generality that

$$|\mathbb{E}_{h\in\mathbb{F}_p^d}\mathbb{E}_{n\in V(M)^h}\Delta_h f(n)c(n)^{\tau(h)}\psi(n)\otimes\chi_h(n)\psi_h(n)| \gg_{C,d,D,\epsilon,s} 1. \tag{4.3}$$

[11]Note that the function χ_h defined this way is not a nilcharacter as its modulo is not 1. However, this does not affect our proof.

Since $\Box_1(V(M))$ is an M-set of total co-dimension 2, $d \geq 5$, and since $c_i(n)^{pm} = 1$ for all $n \in \mathbb{F}_p^d$ and $m \in \mathbb{Z}^d$, it follows from Theorem A.12 that

$$\begin{aligned}
&\left|\mathbb{E}_{h\in\mathbb{F}_p^d}\mathbb{E}_{n\in V(M)^h}\Delta_h f(n)c(n)^{\tau(h)}\psi(n)\otimes\chi_h(n)\psi_n(n)\right|^2\\
=&\left|\mathbb{E}_{(n,h)\in\Box_1(V(M))}\Delta_h f(n)c(n)^{\tau(h)}\psi(n)\otimes\chi_h(n)\psi_h(n)\right|^2+O(p^{-1/2})\\
=&\left|\mathbb{E}_{y\in V(M)}f(y)\mathbb{E}_{n\in V(M)}f'(n)\cdot c(n)^{\tau(y-n)}\otimes\chi_{y-n}(n)\psi_{y-n}(n)\right|^2+O(p^{-1/2})\\
=&\left|\mathbb{E}_{y\in V(M)}f(y)\mathbb{E}_{h\in y-V(M)}f'(y-h)\cdot c(y-h)^{\tau(h)}\otimes\chi_h(y-h)\psi_h(y-h)\right|^2+O(p^{-1/2}) \qquad (4.4)\\
\leq&\ \mathbb{E}_{y\in V(M)}\left|\mathbb{E}_{h\in y-V(M)}f'(y-h)\cdot c(y-h)^{\tau(h)}\otimes\chi_h(y-h)\psi_h(y-h)\right|^2+O(p^{-1/2})\\
\ll_D&\left|\mathbb{E}_{y\in V(M)}\mathbb{E}_{h,h'\in y-V(M)}c(y-h)^{\tau(h)}c(y-h')^{-\tau(h')}f'(y-h)\otimes\overline{f'}(y-h')\right.\\
&\left.\otimes\chi_h(y-h)\otimes\overline{\chi}_{h'}(y-h')\psi_h(y-h)\overline{\psi}_{h'}(y-h')\right|+O(p^{-1/2}),
\end{aligned}$$

where $f' := \overline{f}\cdot\psi$ and we used the vector-valued Cauchy-Schwartz inequality in the last step. Define

$$c_2(\ell,n) := c(n-\ell)^{-\tau(\ell)}f'(n)\otimes\overline{f'}(n-\ell), \psi_{h,\ell}(n) := c(n)^{\tau(h)}c(n-\ell)^{-\tau(h)}\psi_h(n)\overline{\psi}_{h+\ell}(n-\ell).$$

Then c_2 is bounded by $O_{C,d,D,\epsilon,s}(1)$. Since $\psi_h \in \mathrm{Nil}_p^{s-1;C,1}(\mathbb{F}_p^d)$ and $n \mapsto c(n)^{\tau(h)}c(n-\ell)^{-\tau(h)}$ belongs to $\mathrm{Nil}_p^{s-1;O_{C,d,D,\epsilon,s}(1)}(\mathbb{F}_p^d)$, we have that $\psi_{h,\ell}$ belongs to $\mathrm{Nil}_p^{s-1;O_{C,d,D,\epsilon,s}(1),1}(\mathbb{F}_p^d)$ (note that the product of two p-periodic nilsequence is also a p-periodic nilsequence).

Setting $\ell = h'-h$ and $y = n+h$, we have that the square of right hand side of (4.4) is at most

$$\begin{aligned}
&\left|\mathbb{E}_{(n,\ell,h)\colon n,n+h,n-\ell\in V(M)}c_2(\ell,n)\otimes\chi_h(n)\otimes\overline{\chi}_{h+\ell}(n-\ell)\psi_{h,\ell}(n)\right|^2+O(p^{-1/2})\\
=&\left|\mathbb{E}_{(n,\ell)\in\Box_1(V(M))}c_2(-\ell,n)\otimes\mathbb{E}_{h\in V(M)-n}\chi_h(n)\otimes\overline{\chi}_{h-\ell}(n+\ell)\psi_{h,-\ell}(n)\right|^2+O(p^{-1/2})\\
\ll_{C,d,D,\epsilon,s}&\ \mathbb{E}_{(n,\ell)\in\Box_1(V(M))}\left|\mathbb{E}_{h\in V(M)-n}\chi_h(n)\otimes\overline{\chi}_{h-\ell}(n+\ell)\psi_{h,-\ell}(n)\right|^2+O(p^{-1/2})\\
\ll_{C,d,D,\epsilon,s}&\left|\mathbb{E}_{(n,\ell)\in\Box_1(V(M))}\mathbb{E}_{h,h'\in V(M)-n}\right.\\
&\left.\chi_h(n)\otimes\overline{\chi}_{h-\ell}(n+\ell)\otimes\overline{\chi}_{h'}(n)\otimes\chi_{h'-\ell}(n+\ell)\psi_{h,-\ell}(n)\overline{\psi}_{h',-\ell}(n)\right|+O(p^{-1/2})\\
\ll_{C,d,D,\epsilon,s}&\left|\mathbb{E}_{(n,\ell,h,h')\colon n,n+\ell,n+h,n+h'\in V(M)}\right.\\
&\left.\chi_h(n)\otimes\overline{\chi}_{h-\ell}(n+\ell)\otimes\overline{\chi}_{h'}(n)\otimes\chi_{h'-\ell}(n+\ell)\psi_{h,-\ell}(n)\overline{\psi}_{h',-\ell}(n)\right|+O(p^{-1/2}),
\end{aligned}$$

where we used the fact that the cardinality of the set $V(M)+c$ is independent of c. Since $d \geq 9$, Setting $h_1 = h$, $h_2 = h'-\ell$, $h_3 = h'$ and $h_4 = h-\ell$, we have that $h_1+h_2 = h_3+h_4$

and that

$$
\begin{aligned}
&\Big|\mathbb{E}_{(n,\ell,h,h')\colon n,n+\ell,n+h,n+h'\in V(M)}\chi_h(n)\otimes\overline{\chi}_{h-\ell}(n+\ell)\otimes\overline{\chi}_{h'}(n)\otimes\chi_{h'-\ell}(n+\ell)\psi_{h,-\ell}(n)\overline{\psi}_{h',-\ell}(n)\Big|\\
=&\Big|\mathbb{E}_{(n,h_1,h_2,h_3)\in(\mathbb{F}_p^d)^4\colon n\in V(M)^{h_1,h_3,h_3-h_2}}\\
&\qquad\chi_{h_1}(n)\otimes\chi_{h_2}(n+h_3-h_2)\otimes\overline{\chi}_{h_3}(n)\otimes\overline{\chi}_{h_1+h_2-h_3}(n+h_3-h_2)\psi_{h_1,h_2,h_3}(n)\Big|+O(p^{-1/2})\\
=&\Big|\mathbb{E}_{h_1,h_2,h_3\in\mathbb{F}_p^d}\mathbb{E}_{n\in V(M)^{h_1,h_3,h_3-h_2}}\\
&\qquad\chi_{h_1}(n)\otimes\chi_{h_2}(n+h_3-h_2)\otimes\overline{\chi}_{h_3}(n)\otimes\overline{\chi}_{h_1+h_2-h_3}(n+h_3-h_2)\psi_{h_1,h_2,h_3}(n)\Big|+O(p^{-1/2}),
\end{aligned}
$$

where $\psi_{h_1,h_2,h_3} := \psi_{h_1,h_2-h_3}\overline{\psi}_{h_3,h_2-h_3}$ belongs to $\mathrm{Nil}_p^{s-1;O_{C,d,D,\epsilon,s}(1),1}(\mathbb{F}_p^d)$, and we used Theorem A.12 for the set

$$\{(n,h_1,h_2,h_3)\in(\mathbb{F}_p^d)^4\colon n\in V(M)^{h_1,h_3,h_3-h_2}\},$$

which one can check easily is an M-set of total co-dimension 4. From (4.3) we conclude that there exists a subset $U\subseteq H^3$ with $|U|\gg_{C,d,D,\epsilon,s} p^{3d}$ such that (4.2) holds for all $(h_1,h_2,h_3)\in U$ (note that (4.2) fails if $(h_1,h_2,h_3)\notin H^3$ by our construction). Finally, since the number of $(h_1,h_2,h_3)\in U$ with h_1,h_2,h_3 being linearly dependent or M-isotropic is at most $3p^{2d+2}+O_{d,k}(p^{3d-1})$ by Lemma A.5. Since $d\geq 3$, we are done by removing from U the tuples (h_1,h_2,h_3) with h_1,h_2,h_3 being linearly dependent or M-isotropic. □

We also have the following simpler version of Proposition 4.1.

Proposition 4.2. Let $d,D,s\in\mathbb{N}_+$ with $d\geq 9$, $C,\epsilon>0$, $p\gg_{C,d,D,\epsilon,s} 1$ be a prime, $M\colon\mathbb{F}_p^d\to\mathbb{F}_p$ be a non-degenerate quadratic form, and $f\colon\mathbb{F}_p^d\to\mathbb{C}$ be bounded by 1. Suppose that there exist a subset $H\subseteq\mathbb{F}_p^d$ with $|H|\geq\epsilon p^d$, and some functions $g_h\colon\mathbb{F}_p^d\to\mathbb{C}^D$ bounded by C for all $h\in H$ such that

$$\Big|\mathbb{E}_{n\in V(M)^h}f(n+h)\overline{f}(n)g_h(n)\Big|>\epsilon$$

for all $h\in H$. Denote $g_h\equiv 0$ if $h\notin H$. Then for all $h_1,h_2,h_3\in H$, writing

$$g_{h_1,h_2,h_3}(n):=g_{h_1}(n)\otimes g_{h_2}(n+h_3-h_2)\otimes\overline{g}_{h_3}(n)\otimes\overline{g}_{h_1+h_2-h_3}(n+h_3-h_2),$$

we have that

$$\Big|\mathbb{E}_{h_1,h_2,h_3\in\mathbb{F}_p^d}\mathbb{E}_{n\in V(M)^{h_1,h_3,h_3-h_2}}g_{h_1,h_2,h_3}(n)\Big|\gg_{C,d,D,\epsilon,s} 1.$$

In particular, there exist a subset $U\subseteq H^3$ with $|U|\gg_{C,d,D,\epsilon,s} p^{3d}$ such that for all $(h_1,h_2,h_3)\in U$, h_1,h_2,h_3 are linearly independent and M-non-isotropic, and that

$$\Big|\mathbb{E}_{n\in V(M)^{h_1,h_3,h_3-h_2}}g_{h_1,h_2,h_3}(n)\Big|\gg_{C,d,D,\epsilon,s} 1.$$

The proof of Proposition 4.2 is almost the same as that of Proposition 4.1. We leave its proof to the interested readers.

5. The sunflower lemma

We will use Proposition 4.1 to analysis the "frequency" of nilcharacters. It is helpful for the readers to imagine each nilcharacter χ_h as a function of the form $\exp(P_h(n))$ for some polynomial $P_h\colon \mathbb{F}_p^d \to \mathbb{F}_p$, and to imagine its "frequency" as the leading terms of P_h. In Section 10 of [12], Green, Tao and Ziegler proved a sunflower lemma (Lemma 10.10 of [12]) explaining how to decompose a sequence of real number as linear combinations of generators with some "sunflower-type" independence properties. In our setting, the "frequency" of a nilcharacter is a homogeneous polynomial over d-variables (which can be viewed as a "string" of real numbers) instead of a real number. Therefore, we need to extend many concepts in [12] from real numbers to strings of real numbers. For convenience, throughout this section, we use m to denote elements in $\mathbb{N}^d$.

Definition 5.1 (Strings)**.** Let $d, p, s \in \mathbb{N}_+$, $\vec{D} = (D_j)_{1\leq j\leq s} \in \mathbb{N}^s$ and $\mathcal{T} = (\xi_{m,k})_{1\leq |m|\leq s, 1\leq k\leq D_{|m|}}$ be a tuple of elements of the form $\xi_{m,k} \in \mathbb{Z}/p$ (here $m \in \mathbb{N}^d$). We say that $(\vec{D}, \mathcal{T})$ (or $\mathcal{T}$) is a (d, p)-*tower* (or simply a *tower* when there is no confusion). We can also write $\mathcal{T} = (\mathcal{F}_j)_{1\leq j\leq s}$, where $\mathcal{F}_j := (\xi_{m,k})_{|m|=j, 1\leq k\leq D_j}$ is the j-*th floor* of $\mathcal{T}$. We can further write $\mathcal{F}_j = (\vec{\xi}_{j,k})_{1\leq k\leq D_j}$, where $\vec{\xi}_{j,k} := (\xi_{m,k})_{|m|=j}$ is a (j, d, p)-*string* (when there is no confusion, we call it to be a (j, d)-*string*, j-*string* or simply a *string*). We may define the addition and scalar multiplication operators for strings, floors and towers by treating them as vectors.

It is helpful to think of a string as a homogeneous polynomial. Indeed, for any $a \in \mathbb{Z}\backslash\{0\}$ with $|a| < p$, let a^* denote the unique integer in $\{1, \dots, p-1\}$ such that $aa^* \equiv 1 \mod p\mathbb{Z}$. We may *associate* every j-string $\vec{\xi} = (\xi_m)_{|m|=j}$ with a $\mathbb{Z}/p$-coefficient polynomial $f \in \text{poly}_p(\mathbb{Z}^d \to \mathbb{Z}/p|\mathbb{Z})$ given by $f(n) = \sum_{|m|=j}(m!)^*\xi_m n^m$, and conversely *associate* every polynomial $f \in \text{poly}_p(\mathbb{Z}^d \to \mathbb{Z}/p|\mathbb{Z})$ given by $f(n) = \sum_{|m|=j} a_m n^m, a_m \in \frac{1}{m!p}\mathbb{Z}$ with a j-string $\vec{\xi} = (m! a_m)_{|m|=j}$.

In our setting, it is convenient to identify two polynomials if their difference vanishes on a subset of $\mathbb{Z}^d$ of our interest. This leads to the following definition:

Definition 5.2 (Reducible strings)**.** Let Ω be a subset of $\mathbb{Z}^d$. We say that a polynomial $f\colon \mathbb{Z}^d \to \mathbb{R}$ of degree j is (Ω, p)-*reducible* if $\Delta_{h_j}\dots\Delta_{h_1} f(n) \in \mathbb{Z}$ for all $(n, h_1, \dots, h_j) \in \Box_{p,j}(\Omega)$ (recall the definition in Section 1.3). A j-string $\vec{\xi} = (\xi_m)_{|m|=j}$ is (Ω, p)-*reducible* if the polynomial associated to it is (Ω, p)-reducible. We say that a tower is (Ω, p)-*reducible* if every string of the tower is (Ω, p)-reducible.

We have the following characterization for $(\tau(V(M)^{h_1,\dots,h_s}), p)$-reducible strings, using the language developed in [20] (see Appendix A.8 for definitions).

Lemma 5.3. Let $d, j, s \in \mathbb{N}_+, d \geq s + j + 3$ $p \gg_d 1$ be a prime, $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a non-degenerate quadratic form, and $h_1, \dots, h_s \in \mathbb{F}_p^d$ be linearly independent and M-non-isotropic vectors. Let $\vec{\xi} = (\xi_m)_{|m|=j}$ be a (j, d, p)-string. Let $f\colon \mathbb{Z}^d \to \mathbb{Z}/p$ be the polynomial associated to $\vec{\xi}$ and $F\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be the map induced by f. Then $\vec{\xi}$ (or f) is $(\tau(V(M)^{h_1,\dots,h_s}), p)$-reducible if and only if $F \in J^M_{h_1,\dots,h_s}$.

Proof. Since f is homogeneous, so is F. By definition, $\vec{\xi}$ is $(\tau(V(M)^{h_1,\dots,h_s}),p)$-reducible if and only if $\Delta_{n_j}\dots\Delta_{n_1}f(n)\in\mathbb{Z}$ for all $(n,n_1,\dots,n_j)\in\Box_{p,j}(\tau(V(M)^{h_1,\dots,h_s}))$. By Proposition A.13 (the part that (i) $\Leftrightarrow$ (iii)) and the fact that f is homogeneous, this is equivalent of saying that $F\in J^M_{h_1,\dots,h_s}$. □

We may define linear independent/combination for strings in the natural way (modulo reducible strings):

Definition 5.4 (Linear independent/combination for strings)**.** Let Ω be a subset of $\mathbb{Z}^d$ and $c\in\mathbb{N}$. We say that j-strings $\vec{\xi_{(1)}},\dots,\vec{\xi_{(\ell)}}$ are (Ω,c,p)-*independent* if for every $1\leq j\leq s$ and $a_1,\dots,a_\ell\in\mathbb{Z}$ with $-c\leq a_k\leq c$, if $\sum_{k=1}^{\ell}a_k\vec{\xi_{(k)}}$ is (Ω,p)-reducible, then $a_1=\dots=a_k=0$. A string which can be written in the form $\vec{\xi}+\sum_{k=1}^{\ell}a_k\vec{\xi_{(k)}}$ for some $-c\leq a_k\leq c$ and some (Ω,p)-reducible string $\vec{\xi}$ is called an (Ω,c,p)-*linear combination* of $\vec{\xi_{(1)}},\dots,\vec{\xi_{(\ell)}}$.

Let $\mathcal{T}=(\vec{\xi}_{j,k})_{1\leq j\leq s,1\leq k\leq D_j}$ be a (d,p)-tower. We say that $\mathcal{T}$ is (Ω,c,p)-*independent* if for every $1\leq j\leq s$ and the j-strings $\vec{\xi}_{j,1},\dots,\vec{\xi}_{j,D_j}$ are (Ω,c,p)-independent.

We may also extend the definition of p-almost linear function/Freiman homomorphism defined in [20] to strings (see Appendix A.8 for definitions):

Definition 5.5 (p-almost linear function/Freiman homomorphism for strings)**.** Let $d,j,K\in\mathbb{N}_+$, p be a prime, and H be a subset of $\mathbb{F}_p^d$. We say that a map $h\mapsto\vec{\xi}_h=(\xi_{h,m})_{m\in\mathbb{N}^d,|m|=j}$ from H to the set of (j,d,p)-strings is a *p-almost linear function/Freiman homomorphism* (of complexity at most K) if the map $h\mapsto\xi_{h,m}$ is a p-almost linear function/Freiman homomorphism for all $m\in\mathbb{N}^d,|m|=j$.

We are now ready to prove the main result for this section, which says that any sequences of strings can be written as linear combinations of generators with some sunflower-type independence properties. This is where we make essential use of the additive combinatorics result obtained in the previous part of the series [20] (i.e. Theorem A.16 of this paper).

Lemma 5.6 (Sunflower lemma for frequencies)**.** Let $d,j,\ell,L\in\mathbb{N}_+$, $\delta,\epsilon>0$, $p\gg_{\delta,d,\epsilon,\ell,L}1$ be a prime, $M\colon\mathbb{F}_p^d\to\mathbb{F}_p$ be a non-degenerate quadratic form, and H be a subset of $\mathbb{F}_p^d$ with $|H|>\epsilon p^d$. For all $h\in H$, let $\vec{\xi}_{h,1},\dots,\vec{\xi}_{h,\ell}$ be (j,d)-strings. If $d\geq N(j)$, then there exist $a\in\mathbb{N}_+$, $a=O_\ell(1)$, a subset H' of H with $|H'|\gg_{\delta,d,\epsilon,\ell,L}p^d$, integers $\ell_{\text{cor}},\ell_{\text{ind}},\ell_{\text{lin}}=O_\ell(1)$, a collection of "core" (j,d)-strings

$$\mathcal{S}^{\text{cor}}=\{\vec{\xi}^{\text{cor}}_1,\dots,\vec{\xi}^{\text{cor}}_{\ell_{\text{cor}}}\}$$

collections of "petal" (j,d)-strings

$$\mathcal{S}^{\text{ind}}_h=\{\vec{\xi}^{\text{ind}}_{h,1},\dots,\vec{\xi}^{\text{ind}}_{h,\ell_{\text{ind}}}\},\mathcal{S}^{\text{lin}}_h=\{\vec{\xi}^{\text{lin}}_{h,1},\dots,\vec{\xi}^{\text{lin}}_{h,\ell_{\text{lin}}}\}$$

for all $h\in H'$ such that the following holds:

(i) For all $h\in H'$ and $1\leq k\leq\ell$, $\vec{\xi}_{h,k}$ is a $(\tau(V(M)^h),L^a,p)$-linear combination of
$$\mathcal{S}^{\text{cor}}\cup\mathcal{S}^{\text{ind}}_h\cup\mathcal{S}^{\text{lin}}_h.$$

(ii) There exists a subset Γ' of $\Gamma := \{(h_1,\dots,h_4) \in H'^4 : h_1 + h_2 = h_3 + h_4\}$ with $|\Gamma'| > (1-\delta)|\Gamma|$ such that for all $(h_1,\dots,h_4) \in \Gamma'$, the (j,d)-strings in
$$\mathcal{S}^{\text{cor}} \cup \cup_{i=1}^{4} \mathcal{S}_{h_i}^{\text{ind}} \cup \cup_{i=1}^{3} \mathcal{S}_{h_i}^{\text{lin}}$$
are $(\tau(V(M)^{h_1,h_2,h_3}), L, p)$-independent.

(iii) For all $1 \leq k \leq \ell_{\text{lin}}$, the map $h \mapsto \vec{\xi}_{h,k}^{\text{lin}}, h \in H'$ is a p-almost linear Freiman homomorphism of complexity $O_{\delta,d,\epsilon,\ell,L}(1)$.

Proof. The outline of the proof of Lemma 5.6 is similar to that of Lemma 10.5 of [12]. We say that a choice of $H', \ell_{\text{ind}}, \ell_{\text{lin}}, \ell_{\text{cor}}, a, \mathcal{S}_h^{\text{ind}}, \mathcal{S}_h^{\text{lin}}, \mathcal{S}^{\text{cor}}, h \in H'$ is a *partial solution with weight* $(\ell_{\text{ind}}, \ell_{\text{lin}}, \ell_{\text{cor}})$ if all the stated conditions are satisfied expect Condition (ii). Define the lexicographical ordering on the weights by setting $(\ell_{\text{ind}}, \ell_{\text{lin}}, \ell_{\text{cor}}) > (\ell'_{\text{ind}}, \ell'_{\text{lin}}, \ell'_{\text{cor}})$ if $\ell_{\text{ind}} > \ell'_{\text{ind}}$; or $\ell_{\text{ind}} = \ell'_{\text{ind}}, \ell_{\text{lin}} > \ell'_{\text{lin}}$; or $\ell_{\text{ind}} = \ell'_{\text{ind}}, \ell_{\text{lin}} = \ell'_{\text{lin}}, \ell_{\text{cor}} > \ell'_{\text{cor}}$.

By setting $H' = H, \ell_{\text{ind}} = \ell, \ell_{\text{lin}} = \ell_{\text{cor}} = 0, a = 1$ and $\vec{\xi}_{h,k}^{\text{ind}} := \vec{\xi}_{h,k}$ for all $h \in H, 1 \leq k \leq \ell$, we get a partial solution with weight $(\ell, 0, 0)$.

Claim: If
$$H', \ell_{\text{ind}}, \ell_{\text{lin}}, \ell_{\text{cor}}, a, \mathcal{S}_h^{\text{ind}}, \mathcal{S}_h^{\text{lin}}, \mathcal{S}^{\text{cor}}, h \in H'$$
is a partial solution with $|H'| \gg_{\delta,d,\epsilon,\ell,L} p^d$ such that Condition (ii) is not satisfied and that $2\ell_{\text{ind}} + \ell_{\text{lin}} + \ell_{\text{cor}} \leq 2\ell$, then there exists a partial solution
$$H'', \ell'_{\text{ind}}, \ell'_{\text{lin}}, \ell'_{\text{cor}}, a+1, \mathcal{S'}_h^{\text{ind}}, \mathcal{S'}_h^{\text{lin}}, \mathcal{S'}^{\text{cor}}, h \in H''$$
with
$$(\ell'_{\text{ind}}, \ell'_{\text{lin}}, \ell'_{\text{cor}}) < (\ell_{\text{ind}}, \ell_{\text{lin}}, \ell_{\text{cor}}), \text{ and } 2\ell'_{\text{ind}} + \ell'_{\text{lin}} + \ell'_{\text{cor}} \leq 2\ell_{\text{ind}} + \ell_{\text{lin}} + \ell_{\text{cor}}, \tag{5.1}$$
and that $|H''| \gg_{\delta,d,\epsilon,\ell,L} p^d$.

By Lemma A.5, we may assume without loss of generality that all the vectors in H' are M-non-isotropic. We may write
$$\mathcal{S}_h^{\text{ind}} = \{\vec{\xi}_{h,k}^{\text{ind}} : 1 \leq k \leq \ell_{\text{ind}}\}, \mathcal{S}_h^{\text{lin}} = \{\vec{\xi}_{h,k}^{\text{lin}} : 1 \leq k \leq \ell_{\text{lin}}\}, \mathcal{S}^{\text{cor}} = \{\vec{\xi}_k^{\text{cor}} : 1 \leq k \leq \ell_{\text{cor}}\}$$
for all $h \in H'$. For $h \in \mathbb{F}_p^d$, let H'_h denote the set of $(h_1,h_2) \in H'^2$ such that $h_1 + h_2 = h$. Then
$$|\{(h_1,h_2,h_3,h_4) \in H'^4 : h_1 + h_2 = h_3 + h_4\}| = \sum_{h \in \mathbb{F}_p^d} |H'_h|^2 \geq \frac{1}{p^d}\Big(\sum_{h \in \mathbb{F}_p^d} |H'_h|\Big)^2 = \frac{|H'|^4}{p^d}. \tag{5.2}$$

Since $|H'| \gg_{\delta,d,\epsilon,\ell,L} p^d$, by assumption, (5.2) and the Pigeonhole Principle, there exist $-L \leq a_k^{\text{cor}}, a_{i,k}^{\text{ind}}, a_{i,k}^{\text{lin}} \leq L$ not all equal to zero and a subset W of $\{(h_1,h_2,h_3,h_4) \in H'^4 : h_1 + h_2 = h_3 + h_4\}$ with $|W| \gg_{\delta,d,\epsilon,\ell,L} p^{3d}$ such that for all $(h_1,h_2,h_3,h_4) \in W$, we have that
$$\sum_{k=1}^{\ell_{\text{cor}}} a_k^{\text{cor}} \vec{\xi}_k^{\text{cor}} + \sum_{i=1}^{4} \sum_{k=1}^{\ell_{\text{ind}}} a_{i,k}^{\text{ind}} \vec{\xi}_{h_i,k}^{\text{ind}} + \sum_{i=1}^{3} \sum_{k=1}^{\ell_{\text{lin}}} a_{i,k}^{\text{lin}} \vec{\xi}_{h_i,k}^{\text{lin}} \text{ is } (\tau(V(M)^{h_1,h_2,h_3}), p)\text{-reducible.} \tag{5.3}$$

Let Z be the set of $h_1,h_2,h_3,h_4 \in \mathbb{F}_p^d$ such that $h_1 + h_2 = h_3 + h_4$ and that h_1,h_2,h_3 are either linearly dependent or M-non-isotropic. Let $W' := W \backslash Z$. Since $d \geq 3$, by Lemma A.5, we have that $|W'| \gg_{\delta,d,\epsilon,\ell,L} p^{3d}$. We consider three different cases.

Case 1: all of $a^{\text{ind}}_{i,k}, a^{\text{lin}}_{i,k}$ are zero. We may assume without loss of generality that $a^{\text{cor}}_1 \neq 0$. Let $\vec{\xi'}^{\text{cor}}_k := (a^{\text{cor}}_1)^* \vec{\xi}^{\text{cor}}_k$, $\vec{\xi'}^{\text{ind}}_{h,k} := (a^{\text{cor}}_1)^* \vec{\xi}^{\text{ind}}_{h,k}$ and $\vec{\xi'}^{\text{lin}}_{h,k} := (a^{\text{cor}}_1)^* \vec{\xi}^{\text{ind}}_{h,k}$. Note that (5.3) implies that $\vec{\xi}^{\text{cor}}_1 + \sum_{k=2}^{\ell_{\text{cor}}} a^{\text{cor}}_k \vec{\xi'}^{\text{cor}}_k$ is $(\tau(V(M)^{h_1,h_2,h_3}), p)$-reducible for all $(h_1, h_2, h_3, h_4) \in W'$. We show that $\vec{\xi}^{\text{cor}}_1 + \sum_{k=2}^{\ell_{\text{cor}}} a^{\text{cor}}_k \vec{\xi'}^{\text{cor}}_k$ is $(\tau(V(M)), p)$-reducible by the intersection method introduced in [20].

Let F be the polynomial induced by the polynomial associated to the j-string

$$\vec{\xi}^{\text{cor}}_1 + \sum_{k=2}^{\ell_{\text{cor}}} a^{\text{cor}}_k \vec{\xi'}^{\text{cor}}_k.$$

Since $|W'| \gg_{\delta,d,\epsilon,\ell,L} p^{3d}$ and $d \geq N(j)$, there exist $(h_{i,1}, h_{i,2}, h_{i,3}, h_{i,4}) \in W''$, $1 \leq i \leq j+1$ such that $h_{i,1}, h_{i,2}, h_{i,3}$, $1 \leq i \leq j+1$ are linearly independent. So for all $1 \leq i \leq j+1$, the j-string in $\vec{\xi}^{\text{cor}}_1 + \sum_{k=2}^{\ell_{\text{cor}}} a^{\text{cor}}_k \vec{\xi'}^{\text{cor}}_k$ is $(\tau(V(M)^{h_{i,1},h_{i,2},h_{i,3}}), p)$-reducible. Since $h_{i,1}, h_{i,2}, h_{i,3}$ are not M-isotropic, by Lemma 5.3, F belongs to $J^M_{h_{i,1},h_{i,2},h_{i,3}}$. By Proposition A.17 (setting $m = 0, r = 3$ and $L = j$), we have that F belongs to J^M. So by Lemma 5.3, $\vec{\xi}^{\text{cor}}_1 + \sum_{k=2}^{\ell_{\text{cor}}} a^{\text{cor}}_k \vec{\xi'}^{\text{cor}}_k$ is $(\tau(V(M)), p)$-reducible.

This means that $\vec{\xi}^{\text{cor}}_1$ is a $(\tau(V(M)), L, p)$-linear combination of $\vec{\xi'}^{\text{cor}}_2, \dots, \vec{\xi'}^{\text{cor}}_{\ell_{\text{cor}}}$. Since $\vec{\xi}_{h,k}$ is a $(\tau(V(M)), L^a, p)$-linear combination of the strings in $\mathcal{S}^{\text{cor}} \cup \mathcal{S}^{\text{ind}}_h \cup \mathcal{S}^{\text{lin}}_h$, it is also a $(\tau(V(M)), L^{a+1}, p)$-linear combination of the strings in $\mathcal{S'}^{\text{ind}}_h, \mathcal{S'}^{\text{lin}}_h, \mathcal{S'}^{\text{cor}}$, where

$$\mathcal{S'}^{\text{ind}}_h = \{\vec{\xi'}^{\text{ind}}_{h,k} \colon 1 \leq k \leq \ell_{\text{ind}}\}, \mathcal{S'}^{\text{lin}}_h = \{\vec{\xi'}^{\text{lin}}_{h,k} \colon 1 \leq k \leq \ell_{\text{lin}}\}, \mathcal{S'}^{\text{cor}} = \{\vec{\xi'}^{\text{cor}}_k \colon 2 \leq k \leq \ell_{\text{cor}}\}.$$

This means that

$$H', \ell_{\text{ind}}, \ell_{\text{lin}}, \ell_{\text{cor}} - 1, a+1, \mathcal{S'}^{\text{ind}}_h, \mathcal{S'}^{\text{lin}}_h, \mathcal{S'}^{\text{cor}}, h \in H'$$

is a partial solution satisfying (5.1).

Case 2: all of $a^{\text{ind}}_{i,k}$ are zero, but not all of $a^{\text{lin}}_{i,k}$ are zero. We may assume without loss of generality that $a^{\text{lin}}_{1,1} \neq 0$. For $h \in H'$, let U_h denote the set of $(h_2, h_3) \in H'^2$ such that $(h, h_2, h_3, h + h_2 - h_3) \in W'$ for all $1 \leq i \leq j+3$. and U'_h denote the set of $(h_{1,2}, h_{1,3}, \dots, h_{j+3,2}, h_{j+3,3}) \in H'^{2(j+3)}$ such that $(h, h_{i,2}, h_{i,3}, h + h_{i,2} - h_{i,3}) \in W'$ for all $1 \leq i \leq j+3$. For $\mathbf{h} = (h_{1,2}, h_{1,3}, \dots, h_{j+3,2}, h_{j+3,3}) \in H'^{2(j+3)}$, let $U''_{\mathbf{h}}$ denote the set of $h \in H'$ such that $\mathbf{h} \in U'_h$. Then

$$\sum_{\mathbf{h} \in H'^{2(j+3)}} |U''_{\mathbf{h}}| = \sum_{h \in H'} |U'_h| = \sum_{h \in H'} |U_h|^{j+3} \geq \frac{1}{|H'|^{j+2}} \Big(\sum_{h \in H'} |U_h|\Big)^{j+3} = \frac{|W'|^{j+3}}{|H'|^{j+2}} \gg_{\delta,d,\epsilon,\ell,L} p^{(2j+7)d}.$$

By the Pigeonhole Principle, there exists $\tilde{H} \subseteq H'^{2(j+3)}$ with $|\tilde{H}| \gg_{\delta,d,\epsilon,\ell,L} p^{2(j+3)d}$ such that $|U''_{\mathbf{h}}| \gg_{\delta,d,\epsilon,\ell,L} p^d$ for all $\mathbf{h} \in \tilde{H}$. By Lemma A.5, we may pick some

$$\mathbf{h} = (h_{1,2}, h_{1,3}, \dots, h_{j+3,2}, h_{j+3,3}) \in \tilde{H}$$

with $h_{1,2}, h_{1,3}, \dots, h_{j+3,2}, h_{j+3,3}$ being linearly independent.

Fix any such $\mathbf{h}$ and set $H'' = U''_{\mathbf{h}}$. For all $h \in H''$, let $\vec{\xi'}^{\text{cor}}_k := (a^{\text{lin}}_{1,1})^*\vec{\xi}^{\text{cor}}_k$, $\vec{\xi'}^{\text{ind}}_{h,k} := (a^{\text{lin}}_{1,1})^*\vec{\xi}^{\text{ind}}_{h,k}$ and $\vec{\xi'}^{\text{lin}}_{h,k} := (a^{\text{lin}}_{1,1})^*\vec{\xi}^{\text{ind}}_{h,k}$. By (5.3) and the construction of H'', for all $h \in H''$, we have that

$$\sum_{k=1}^{\ell_{\text{cor}}} a^{\text{cor}}_k \vec{\xi'}^{\text{cor}}_k + \sum_{j=2}^{3}\sum_{k=1}^{\ell_{\text{lin}}} a^{\text{lin}}_{j,k}\vec{\xi'}^{\text{lin}}_{h_{i,j},k} + \sum_{k=2}^{\ell_{\text{lin}}} a^{\text{lin}}_{1,k}\vec{\xi'}^{\text{lin}}_{h,k} + \vec{\xi}^{\text{lin}}_{h,1} \text{ is } (\tau(V(M)^{h,h_{i,2},h_{i,3}}), p)\text{-reducible.} \tag{5.4}$$

Let F_i be the polynomial induced by the polynomial associated to the j-string

$$\sum_{k=1}^{\ell_{\text{cor}}} a^{\text{cor}}_k \vec{\xi'}^{\text{cor}}_k + \sum_{j=2}^{3}\sum_{k=1}^{\ell_{\text{lin}}} a^{\text{lin}}_{j,k}\vec{\xi'}^{\text{lin}}_{h_{i,j},k},$$

and G_h be the polynomial induced by the polynomial associated to the j-string

$$\sum_{k=2}^{\ell_{\text{lin}}} a^{\text{lin}}_{1,k}\vec{\xi'}^{\text{lin}}_{h,k} + \vec{\xi}^{\text{lin}}_{h,1}.$$

Since $h, h_{i,2}, h_{i,3}$ are linearly independent and M-non-isotropic, by Lemma 5.3 and (5.4), $F_i - G_h$ belongs to $J^M_{h,h_{i,2},h_{i,3}}$ for all $1 \leq i \leq j+3$ and $h \in H''$. So for all $h, h' \in H''$ and $1 \leq i \leq j+3$, we have that

$$G_h - G_{h'} \equiv (G_h - F_i) - (G_{h'} - F_i) \equiv 0 \mod J^M_{h,h',h_{i,2},h_{i,3}}.$$

Since $h_{1,2}, h_{1,3}, \dots, h_{j+3,2}, h_{j+3,3}$ are linearly independent, by Proposition A.17 (setting $m \leq 2$, $s = j, r = 2$), we have that

$$G_h \equiv G_{h'} \mod J^M_{h,h'}$$

for all $h, h' \in H''$. By Proposition A.18, there exists $G \in \text{HP}_d(j)$ (recall the definition in Section 1.3) such that

$$G_h \equiv G \mod J^M_h$$

for all $h \in H''$. In other words, letting $\vec{\xi}_0$ denote the j-string associated to any regular lifting of G, by Lemma 5.3, we have that

$$-\vec{\xi}_0 + \sum_{k=2}^{\ell_{\text{lin}}} a^{\text{lin}}_{1,k}\vec{\xi'}^{\text{lin}}_{h,k} + \vec{\xi}^{\text{lin}}_{h,1}$$

is $(\tau(V(M)^h), p)$-reducible for all $h \in H''$.

This mean that $\vec{\xi}^{\text{lin}}_{h,1}$ is a $(\tau(V(M)^h), L, p)$-linear combination of $\vec{\xi}_0, \vec{\xi'}^{\text{lin}}_{h,k}, 2 \leq k \leq \ell_{\text{lin}}$. Since $\vec{\xi}_{h,k}$ is a $(\tau(V(M)^h), L^a, p)$-linear combination of the strings in $\mathcal{S}^{\text{cor}} \cup \mathcal{S}^{\text{ind}}_h \cup \mathcal{S}^{\text{lin}}_h$, it is also a $(\tau(V(M)^h), L^{a+1}, p)$-linear combination of the strings in $\mathcal{S'}^{\text{ind}}_h, \mathcal{S'}^{\text{lin}}_h, \mathcal{S'}^{\text{cor}}$, where

$$\mathcal{S'}^{\text{ind}}_h = \{\vec{\xi'}^{\text{ind}}_{h,k} \colon 1 \leq k \leq \ell_{\text{ind}}\}, \mathcal{S'}^{\text{lin}}_h = \{\vec{\xi'}^{\text{lin}}_{h,k} \colon 2 \leq k \leq \ell_{\text{lin}}\}, \mathcal{S'}^{\text{cor}} = \{\vec{\xi}_0, \vec{\xi'}^{\text{cor}}_k \colon 1 \leq k \leq \ell_{\text{cor}}\}.$$

This means that

$$H'', \ell_{\text{ind}}, \ell_{\text{lin}} - 1, \ell_{\text{cor}} + 1, a + 1, \mathcal{S'}^{\text{ind}}_h, \mathcal{S'}^{\text{lin}}_h, \mathcal{S'}^{\text{cor}}, h \in H''$$

is a partial solution satisfying (5.1).

Case 3: not all of $a_{i,k}^{\text{ind}}$ are zero. We may assume without loss of generality that $a_{4,1}^{\text{lin}} \neq 0$. Let $\Xi_1, \dots, \Xi_4 \colon \mathbb{F}_p^d \to (\mathbb{Z}/p)^{\binom{j+d-1}{d-1}}$ be functions given by

$$\Xi_i(h) := \sum_{k=1}^{\ell_{\text{ind}}} a_{i,k}^{\text{ind}} \vec{\xi}_{h,k}^{\text{ind}} + \sum_{k=1}^{\ell_{\text{lin}}} a_{i,k}^{\text{lin}} \vec{\xi}_{h,k}^{\text{lin}}$$

for $i = 1, 2, 3$, and

$$\Xi_4(h) := \sum_{k=1}^{\ell_{\text{cor}}} a_k^{\text{cor}} \vec{\xi}_k^{\text{cor}} + \sum_{k=1}^{\ell_{\text{ind}}} a_{4,k}^{\text{ind}} \vec{\xi}_{h,k}^{\text{ind}}.$$

Then (5.3) implies that

$$\Xi_1(h_1) + \Xi_2(h_2) + \Xi_3(h_3) + \Xi_4(h_4) \text{ is } (\tau(V(M)^{h_1,h_2,h_3}), p)\text{-reducible} \tag{5.5}$$

for all $(h_1, h_2, h_3, h_1 + h_2 - h_3) \in W'$. For any (j, d)-string $\vec{\xi} := (\xi_m)_{|m|=j}$, let $\omega(\vec{\xi})$ the polynomial in $\text{HP}_d(j)$ given by

$$\omega(\vec{\xi})(n) := \sum_{m \in \mathbb{N}^d, |m|=j} \iota(p\xi_m)(m!)^* n^m, n \in \mathbb{F}_p^d,$$

i.e. $\omega(\vec{\xi})$ is the polynomial induced by the polynomial associated to $\vec{\xi}$. For $1 \leq i \leq 4$, let $\omega_i \colon H' \to \text{HP}_d(j)$ be the map given by $\omega_i(h) := \omega(\Xi_i(h))$. Since $d \geq N(j)$ and h_1, h_2, h_3 are linearly independent and M-non-isotropic, by (5.5) and Proposition A.13, we have that

$$\omega_1(h_1) + \omega_2(h_2) + \omega_3(h_3) + \omega_4(h_4) \in J_{h_1,h_2,h_3}^M.$$

Since $d \geq N(j)$, by Theorem A.16, there exists a subset $H'' \subseteq H'$ with $|H''| \gg_{\delta,d,\epsilon,\ell,L} p^d$, some $g \in \text{HP}_d(j)$ and an almost linear Freiman homomorphism $T \colon H \to \text{HP}_d(j)$ of complexity $O_{\delta,d,\epsilon,\ell,L}(1)$ such that

$$\omega_4(h) - T(h) - g \in J_h^M$$

for all $h \in H''$. By Lemma A.5, we may assume without loss of generality that all elements in H'' are non-zero and M-non-isotropic. By definition and Lemma 5.3, there exist a p-almost linear Freiman homomorphism $h \mapsto \vec{\xi}_h$ of complexity $O_{\delta,d,\epsilon,\ell,L}(1)$ and some j-string $\vec{\xi}_0$ such that

$$\Xi_4(h) - \vec{\xi}_h - \vec{\xi}_0 = a_{4,1}^{\text{ind}} \vec{\xi}_{h,1}^{\text{ind}} - \vec{\xi}_h - \vec{\xi}_0 + \sum_{k=1}^{\ell_{\text{cor}}} a_k^{\text{cor}} \vec{\xi}_k^{\text{cor}} + \sum_{k=2}^{\ell_{\text{ind}}} a_{4,k}^{\text{ind}} \vec{\xi}_{h,k}^{\text{ind}}$$

is $\tau(V(M)^h)$-reducible for all $h \in H''$. Let $\vec{\xi'}_0 := (a_{4,1}^{\text{ind}})^* \vec{\xi}_0$, $\vec{\xi'}_k^{\text{cor}} := (a_{4,1}^{\text{ind}})^* \vec{\xi}_k^{\text{cor}}$, $\vec{\xi'}_{h,k}^{\text{ind}} := (a_{4,1}^{\text{ind}})^* \vec{\xi}_{h,k}^{\text{ind}}$, and $\vec{\xi'}_h := (a_{4,1}^{\text{ind}})^* \vec{\xi}_h$.

This mean that $\vec{\xi}_{h,1}^{\text{ind}}$ is a $(\tau(V(M)^h), L, p)$-linear combination of $\vec{\xi'}_0, \vec{\xi'}_h, \vec{\xi'}_k^{\text{cor}}, \vec{\xi'}_{h,k}^{\text{ind}}$. Since $\vec{\xi}_{h,k}$ is a $(\tau(V(M)^h), L^a, p)$-linear combination of the strings in $\mathcal{S}^{\text{cor}} \cup \mathcal{S}_h^{\text{ind}} \cup \mathcal{S}_h^{\text{lin}}$, it is also a $(\tau(V(M)^h), L^{a+1}, p)$-linear combination of the strings in $\mathcal{S'}_h^{\text{ind}}, \mathcal{S'}_h^{\text{lin}}, \mathcal{S'}^{\text{cor}}$, where

$$\mathcal{S'}_h^{\text{ind}} = \{\vec{\xi'}_{h,k}^{\text{ind}} \colon 2 \leq k \leq \ell_{\text{ind}}\}, \mathcal{S'}_h^{\text{lin}} = \{\vec{\xi'}_h, \vec{\xi'}_{h,k}^{\text{lin}} \colon 1 \leq k \leq \ell_{\text{lin}}\}, \mathcal{S'}^{\text{cor}} = \{\vec{\xi'}_0, \vec{\xi'}_k^{\text{cor}} \colon 1 \leq k \leq \ell_{\text{cor}}\}.$$

This means that

$$H'', \ell_{\text{ind}} - 1, \ell_{\text{lin}} + 1, \ell_{\text{cor}} + 1, a + 1, \mathcal{S}'^{\text{ind}}_h, \mathcal{S}'^{\text{lin}}_h, \mathcal{S}'^{\text{cor}}, h \in H''$$

is a partial solution satisfying (5.1). This finishes the proof of the claim.

From the claim, we can find a series partial solutions with decreasing weights. Moreover, by (5.1), it is not hard to see that this process terminates after at most $O_\ell(1)$ steps. By the claim, this partial solution satisfies Condition (ii), and we are done. □

Next we use the decomposition of strings in Lemma 5.6 to obtain a decomposition of nilcharacters. We need to recall some definitions introduced in [12]:

Definition 5.7 (Universal nilmanifold)**.** A *dimension vector* is a tuple

$$\vec{D} = (D_1, \dots, D_s) \in \mathbb{N}^s.$$

Given a dimension vector, we defined the *universal nilpotent group* $G^{\vec{D},d}$ of degree-rank $[s, r_*]$ to be the Lip group generated by formal generators $e_{m,j}$ for $m \in \mathbb{N}^d, 1 \leq |m| \leq s$ and $1 \leq j \leq D_i$ such that any iterated commutator of $e_{m_1,j_1}, \dots, e_{m_\ell,j_\ell}$ is trivial if either $|m_1| + \dots + |m_\ell| > s$ or $|m_1| + \dots + |m_\ell| = s$ and $\ell \geq r + 1$.

We endow this group a degree-rank filtration $(G^{\vec{D},d}_{[s,r]})_{[s,r]\in\text{DR}}$ by letting $G^{\vec{D},d}_{[s,r]}$ be the Lie group generated by all the elements of the iterated commutators of $e_{m_1,j_1}, \dots, e_{m_\ell,j_\ell}$ with $1 \leq |m_1|, \dots, |m_\ell| \leq s$ and $1 \leq j_i \leq D_{|m_i|}, 1 \leq i \leq \ell$ such that either $|m_1| + \dots + |m_\ell| > s$ or $|m_1| + \dots + |m_\ell| = s$ and $\ell \geq r$. It is not hard to verify that this is indeed a filtration of degree-rank $\leq [s, r_*]$. We then let $\Gamma^{\vec{D},d}$ be the discrete group generated by $e_{m,j}, 1 \leq |m| \leq s, 1 \leq j \leq D_{|m|}$, and refer to $G^{\vec{D},d}/\Gamma^{\vec{D},d}$ as the *universal nilmanifold* with dimension vector $(\vec{D}, d)$ of degree-rank $[s, r_*]$.

A *universal vertical frequency* on the universal nilmanifold $G^{\vec{D},d}/\Gamma^{\vec{D},d}$ of degree-rank $[s, r_*]$ is a continuous homomorphism $\eta\colon G^{\vec{D},d}_{[s,r_*]} \to \mathbb{R}$ which sends $\Gamma^{\vec{D},d}_{[s,r_*]}$ to the integers. The *complexity* of a universal vertical frequency can be defined similar to Definition 2.29.

Let $(g_m)_{|m|=j}$ be a sequence of elements with $g_m \in G_{(j,1)}/G_{(j,2)}$ (for convenience we also call $(g_m)_{|m|=j}$ a *j-string*), and Ω be a subset of $\mathbb{Z}^d$. We say that $(g_m)_{|m|=j}$ is (Ω, p)*-reducible* if $(\xi_j(g_m))_{|m|=j}$ is (Ω, p)-reducible for all $\xi_j \in \mathfrak{R}_j(G/\Gamma)$.

Definition 5.8 (Universal representation)**.** Let $d, D \in \mathbb{N}_+, C > 0$, $[s, r_*] \in \text{DR}$, p be a prime, Ω_1, Ω_2 be subsets of $\mathbb{F}_p^d$, and $\chi \in \Xi_p^{[s,r_*]}(\Omega_1)$ be a nilcharacter of degree-rank $\leq [s, r_*]$ of dimension D. A *universal representation* of χ is a collection of the following data:

- a filtered nilmanifold G/Γ of degree-rank $\leq [s, r_*]$ and complexity at most C;
- a filtered nilmanifold G_0/Γ_0 of degree-rank $< [s, r_*]$ and complexity at most C;
- a function $F \in \text{Lip}((G/\Gamma \times G_0/\Gamma_0) \to \mathbb{S}^D)$ of Lipschitz norm at most C;
- p-periodic polynomial sequences $g \in \text{poly}_p(\mathbb{Z}^d \to G_\mathbb{N}|\Gamma)$ and $g_0 \in \text{poly}_p(\mathbb{Z}^d \to (G_0)_\mathbb{N}|\Gamma_0)$;
- a dimension vector $\vec{D} = (D_1, \dots, D_s) \in \mathbb{N}^s$ with $|\vec{D}| \leq C$;

- a universal vertical frequency $\eta\colon G^{\vec{D},d}_{[s,r_*]} \to \mathbb{R}$ of complexity at most C on the universal nilmanifold $G^{\vec{D},d}/\Gamma^{\vec{D},d}$ of degree-rank $[s,r_*]$;
- a filtered homomorphism $\phi\colon G^{\vec{D},d}/\Gamma^{\vec{D},d} \to G/\Gamma$ of complexity at most C;
- a frequency tower $\mathcal{T} = (\xi_{m,k})_{1\leq |m|\leq s, 1\leq k\leq D_{|m|}}$ with $\xi_{m,k} \in \mathbb{Z}/p$;

such that the followings hold:

- For all $n \in \Omega_1$, we have that

$$\chi(n) = F(g \circ \tau(n)\Gamma, g_0 \circ \tau(n)\Gamma_0). \tag{5.6}$$

- For all $t \in G^{\vec{D},d}_{[s,r_*]}$ and $(x, x_0) \in G/\Gamma \times G_0/\Gamma_0$, we have that

$$F(\phi(t)x, x_0) = \exp(\eta(t))F(x, x_0). \tag{5.7}$$

- For all $1 \leq j \leq s$, the j-string

$$\Big(\mathrm{Taylor}_m(g) \cdot \Big(\pi_{\mathrm{Horiz}_j(G/\Gamma)} \circ \phi(\prod_{k=1}^{D_j} e_{m,k}^{\xi_{m,k}})\Big)^{-1}\Big)_{|m|=j}$$

  is $(\tau(\Omega_2), p)$-reducible, where $\pi_{\mathrm{Horiz}_j(G/\Gamma)}\colon G_j \to \mathrm{Horiz}_i(G/\Gamma)$ is the projection map.

We say that the tuple $(\vec{D}, \mathcal{T}, \eta)$, or $(\vec{D}, \mathcal{T}, \eta, G/\Gamma \times G_0/\Gamma_0)$ or $(\vec{D}, \mathcal{T}, \eta, F, \phi, G/\Gamma \times G_0/\Gamma_0)$ is an (Ω_1, Ω_2, C)-*universal representation* of χ.

Remark 5.9. Note that if $(\vec{D}, \mathcal{T}, \eta, F, \phi, G/\Gamma \times G_0/\Gamma_0)$ is a universal representation of χ such that (5.6) holds, then by lifting g to a polynomial sequence on the universal nilmanifold $G^{\vec{D},d}$, we may write χ as

$$\chi(n) = F(\phi \circ \tilde{g} \circ \tau(n)\Gamma^{\vec{D},d}, g_0 \circ \tau(n)\Gamma_0)$$

for some $\tilde{g} \in \mathrm{poly}(\mathbb{Z}^d \to G^{\vec{D},d}_{\mathbb{N}})$. However, we caution the readers that $\tilde{g}$ is not necessarily p-periodic.

We collect some basic properties on the universal representation of nilcharcters. The following lemma says that every nilcharcter admits a universal representation, which is an extension of Lemma 9.12 of [12]:

Lemma 5.10 (Existence of universal representation)**.** Let Ω_1, Ω_2 be subsets of $\mathbb{F}_p^d$, $C > 0$ and $[s, r_*] \in \mathrm{DR}$. Every $\chi \in \Xi_p^{[s,r_*];C}(\Omega_1)$ has an $(\Omega_1, \Omega_2, O_{C,d,s}(1))$-universal representation $(\vec{D}, \mathcal{T}, \eta, G/\Gamma)$ if $p \gg_{C,d,s} 1$.[12]

Proof. Suppose that $\chi(n) = F(g \circ \tau(n)\Gamma), n \in \Omega_1$ for some nilmanifold G/Γ of degree-rank $\leq [s, r_*]$ and complexity at most C, some $g \in \mathrm{poly}_p(\mathbb{Z}^d \to G_{\mathbb{N}^d}|\Gamma)$ and some $F \in \mathrm{Lip}(G/\Gamma \to \mathbb{S}^D)$ of Lipschitz norm at most C with a vertical frequency of complexity at most C. For each $1 \leq i \leq s$, let $f_{i,1}, \dots, f_{i,D_i}$ be a basis of generators of Γ_i. Set

[12] We remark that G_0/Γ_0 can be taken to be the trivial system, and G/Γ can be taken to be the nilmanifold on which χ is defined.

$\vec{D} := (D_1, \dots, D_s)$. Then we have a filtered homomorphism $\phi\colon G^{\vec{D},d} \to G$ which maps $e_{m,k}$ to $f_{|m|,k}$. Since G/Γ is of complexity at most C, $G^{\vec{D},d}/\Gamma^{\vec{D},d}$ is of complexity at most $O_{C,d,s}(1)$ and ϕ is $O_{C,d,s}(1)$-rational.

Fix $m \in \mathbb{N}^d$ with $1 \leq |m| \leq s$. Since $g \in \mathrm{poly}_p(\mathbb{Z}^d \to G_{\mathbb{N}^d}|\Gamma)$, by Lemma 2.25, $\mathrm{Taylor}_m(g)^p \in G_{[i,2]}\Gamma$. Since $G_{[i,1]}/G_{[i,2]}$ is abelian and $f_{i,1}, \dots, f_{i,D_i}$ is a basis of generators of Γ_i, we have that

$$\mathrm{Taylor}_m(g) \equiv \prod_{j=1}^{D_{|m|}} f_{|m|,k}^{\xi_{m,k}} \equiv \Big(\pi_{\mathrm{Horiz}_{|m|}(G/\Gamma)}(\phi(\prod_{j=1}^{D_{|m|}} e_{m,k}^{\xi_{m,k}}))\Big) \mod G_{[i,2]}\Gamma$$

for some $\xi_{m,k} \in \mathbb{Z}/p$. Therefore,

$$\Big(\mathrm{Taylor}_m(g) \cdot \Big(\pi_{\mathrm{Horiz}_j(G/\Gamma)} \circ \phi(\prod_{k=1}^{D_j} e_{m,k}^{\xi_{m,k}})\Big)^{-1}\Big)_{|m|=j}$$

is a $G_{[i,2]}\Gamma$-valued string and thus is $(\tau(\Omega_2), p)$-reducible.

Since ϕ is $O_{C,d,s}(1)$-rational, condition (5.7) can be derived by pulling back the vertical frequency of F by ϕ (and thus η is of complexity at most $O_{C,d,s}(1)$). This finishes the proof by setting G_0/Γ_0 to be trivial. □

Let Ω be a subset of $\mathbb{Z}^d$ and $c \in \mathbb{N}$. Let $\mathcal{T} = (\vec{\xi}_{j,k})_{1\leq j\leq s, 1\leq k\leq D_j}$ and $\mathcal{T}' = (\vec{\xi}'_{j,k})_{1\leq j\leq s, 1\leq k\leq D'_j}$ be two (d, p)-towers. We say that $\mathcal{T}$ is (Ω, c, p)-*represented* by $\mathcal{T}'$ if for all $1 \leq j \leq s$ and $1 \leq k \leq D_j$, $\vec{\xi}_{j,k}$ is an (Ω, c, p)-linear combination of $\vec{\xi}'_{j,1}, \dots, \vec{\xi}'_{j,D'_j}$. The following lemma explain how different "basis" affects the representation of a nilcharacter, which is similar to Lemma 10.8 of [12]:

Lemma 5.11 (Change of basis)**.** Let Ω_1, Ω_2 be subsets of $\mathbb{F}_p^d$ and $[s, r_*] \in \mathrm{DR}$. Let $C > 0$ and $\chi \in \Xi_p^{[s,r_*];C}(\Omega_1)$ be a degree-rank $[s, r_*]$ nilcharacter with an (Ω_1, Ω_2, C)-universal representation $(\vec{D}, \mathcal{T}, \eta, G/\Gamma \times G_0/\Gamma_0)$. Suppose that the frequency tower $(\vec{D}, \mathcal{T})$ is $(\tau(\Omega_2), C, p)$-represented by another frequency tower $(\vec{D}', \mathcal{T}')$. Then there exists a vertical frequency $\eta'\colon G_{[s,r_*]}^{\vec{D}',d} \to \mathbb{R}$ such that χ has an $(\Omega_1, \Omega_2, O_{C,d,|\vec{D}'|}(1))$-universal representation $(\vec{D}', \mathcal{T}', \eta', G/\Gamma \times G_0/\Gamma_0)$.

Proof. Let $(\vec{D}, \mathcal{T}, \eta, F, \phi, G/\Gamma \times G_0/\Gamma_0)$ be an (Ω_1, Ω_2, C)-universal representation of χ. Suppose that $\vec{D} = (D_1, \dots, D_s)$, $\vec{D}' = (D'_1, \dots, D'_s)$, $\mathcal{T} = (\vec{\xi}_{j,k})_{1\leq j\leq s, 1\leq k\leq D_j}$, and $\mathcal{T}' = (\vec{\xi}'_{j,k})_{1\leq j\leq s, 1\leq k\leq D'_j}$. By assumption, each j-string $\vec{\xi}_{j,k}$ of $\mathcal{T}$ can be written as

$$\vec{\xi}_{j,k} = \vec{\zeta}_{j,k} + \sum_{k'=1}^{D'_j} a_{j,k,k'} \vec{\xi}'_{j,k'} \tag{5.8}$$

for some $(\tau(\Omega_2), p)$-reducible j-string $\vec{\zeta}_{j,k}$ and $-C \leq a_{j,k,k'} \leq C$. Let $\psi\colon G^{\vec{D}',d} \to G^{\vec{D},d}$ be the unique filtered homomorphism that maps $e'_{m,k'}$ to $\prod_{k=1}^{D_j} e_{m,k}^{a_{j,k,k'}}$. Write $\vec{\zeta}_{j,k} = (\zeta_{m,k})_{|m|=j}$,

$\vec{\xi}_{j,k}=(\xi_{m,k})_{|m|=j}$ and $\vec{\xi}'_{j,k}=(\xi'_{m,k})_{|m|=j}$. By assumption, χ can be written as $\chi(n)=F(g\circ\tau(n)\Gamma, g_0\circ\tau(n)\Gamma_0), n\in\Omega_1$ with

$$\Big(\text{Taylor}_m(g)\cdot\Big(\pi_{\text{Horiz}_j(G/\Gamma)}\circ\phi(\prod_{k=1}^{D_j}e_{m,k}^{\xi_{m,k}})\Big)^{-1}\Big)_{|m|=j}$$

being $(\tau(\Omega_2),p)$-reducible for some filtered homomorphism $\phi\colon G^{\vec{D},d}\to G$ and some tower $(\xi_{m,k})_{1\leq|m|\leq s,1\leq k\leq D_{|m|}}$ for all $1\leq j\leq s$. Then for all $|m|=j$, by (5.8),

$$\begin{aligned}&\text{Taylor}_m(g)\cdot\Big(\pi_{\text{Horiz}_j(G/\Gamma)}\circ\phi\circ\psi(\prod_{k'=1}^{D'_j}e'^{\xi_{m,k'}}_{m,k'})\Big)^{-1}\\
&=\text{Taylor}_m(g)\cdot\Big(\pi_{\text{Horiz}_j(G/\Gamma)}\circ\phi(\prod_{k=1}^{D_j}e_{m,k}^{\sum_{k'=1}^{D'_j}a_{j,k,k'}\xi_{m,k'}})\Big)^{-1}\\
&=\text{Taylor}_m(g)\cdot\Big(\pi_{\text{Horiz}_j(G/\Gamma)}\circ\phi(\prod_{k=1}^{D_j}e_{m,k}^{\xi_{m,k}})\Big)^{-1}\cdot\prod_{k=1}^{D_j}\Big(\pi_{\text{Horiz}_j(G/\Gamma)}\circ\phi(e_{m,k}^{\zeta_{m,k}})\Big).\end{aligned}\tag{5.9}$$

Since both $\Big(\text{Taylor}_m(g)\cdot\Big(\pi_{\text{Horiz}_j(G/\Gamma)}\circ\phi(\prod_{k=1}^{D_j}e_{m,k}^{\xi_{m,k}})\Big)^{-1}\Big)_{|m|=j}$ and $\vec{\zeta}_{j,1},\dots,\vec{\zeta}_{j,D_j}$ are $(\tau(\Omega_2),p)$-reducible, the left hand side of (5.9) is also $(\tau(\Omega_2),p)$-reducible.

On the other hand, $\phi\circ\psi\colon G^{\vec{D}',d}\to G$ is a filtered homomorphism and $\eta\circ\psi\colon G^{\vec{D}',d}_{[s,r_*]}\to\mathbb{R}$ is a vertical frequency. Since η,ϕ and G/Γ are of complexities at most C and $-C\leq a_{j,k,k'}\leq C$, it is not hard to see that $\phi\circ\psi$ is $O_{C,d,|\vec{D}'|}(1)$-rational and that $\eta\circ\psi$ is of complexity at most $O_{C,d,|\vec{D}'|}(1)$. This finishes the proof. □

We refer the readers to Section 9 of [12] for further properties of universal representations of nilcharacters.

Let $\mathcal{T}=(\xi_{m,k})_{1\leq|m|\leq s,1\leq k\leq D_{|m|}}$ and $\mathcal{T}'=(\xi'_{m,k})_{1\leq|m|\leq s,1\leq k\leq D'_{|m|}}$ be two towers. Let $\mathcal{T}\uplus\mathcal{T}':=(\xi''_{m,k})_{1\leq|m|\leq s,1\leq k\leq D_{|m|}+D'_{|m|}}$ be the tower defined by $\xi''_{m,k}=\xi_{m,k}$ for $1\leq k\leq D_{|m|}$ and $\xi''_{m,k'+D_{|m|}}=\xi'_{m,k'}$ for $1\leq k'\leq D'_{|m|}$. We have:

Lemma 5.12 (Sunflower lemma for nilcharacters)**.** Let $[s,r_*]\in\text{DR}$, $d,D\in\mathbb{N}_+$, with $d\geq N(s)$, $C,\delta,\epsilon,L>0$, $p\gg_{C,\delta,d,\epsilon,L}1$ be a prime, and $M\colon\mathbb{F}_p^d\to\mathbb{F}_p$ be a non-degenerate quadratic form. Let H be a subset of $\mathbb{F}_p^d$ with $|H|>\epsilon p^d$ and $(\chi_h)_{h\in H}$ be a family of nilcharacters in $\Xi_p^{[s,r_*];C,D}(V(M))$. If $d\geq N(s)$, then there exist

- a subset $H'\subseteq H$ with $|H'|\gg_{C,\delta,d,\epsilon,L}p^d$;
- a dimension vector $\vec{D}=(D_j)_{1\leq j\leq s}$ with $|\vec{D}|=O_{C,\delta,d,\epsilon,L}(1)$ which can be further decomposed as $\vec{D}=\vec{D^{\text{cor}}}+\vec{D^{\text{ped}}}$ and $\vec{D^{\text{ped}}}=\vec{D^{\text{lin}}}+\vec{D^{\text{ind}}}$ for some dimension vectors $\vec{D^{\text{cor}}}=(D^{\text{cor}}_j)_{1\leq j\leq s}$, $\vec{D^{\text{lin}}}=(D^{\text{lin}}_j)_{1\leq j\leq s}$ and $\vec{D^{\text{ind}}}=(D^{\text{ind}}_j)_{1\leq j\leq s}$;
- a core frequency tower $(\vec{D^{\text{cor}}},\mathcal{T}^{\text{cor}})$;

- for each $h \in H'$ petal frequency towers $(\vec{D}^{\text{lin}}, \mathcal{T}_h^{\text{lin}} = (\vec{\xi}^{\text{lin}}_{h,j,k})_{1\leq j\leq s, 1\leq k\leq D_j^{\text{lin}}})$, $(\vec{D}^{\text{ind}}, \mathcal{T}_h^{\text{ind}} = (\vec{\xi}^{\text{lin}}_{h,j,k})_{1\leq j\leq s, 1\leq k\leq D_j^{\text{ind}}})$ and $(\vec{D}^{\text{ped}}, \mathcal{T}_h^{\text{ped}} = (\vec{\xi}^{\text{lin}}_{h,j,k})_{1\leq j\leq s, 1\leq k\leq D_j^{\text{ped}}})$ such that $\mathcal{T}_h^{\text{ped}} = \mathcal{T}_h^{\text{ind}} \uplus \mathcal{T}_h^{\text{lin}}$;
- a filtered nilmanifold G/Γ of degree-rank $\leq [s, r_*]$ of complexity $O_{C,\delta,d,\epsilon,L}(1)$ and a filtered nilmanifold G_0/Γ_0 of degree-rank $\leq [s-1, r_*-1]$ of complexity $O_{C,\delta,d,\epsilon,L}(1)$;
- for each $h \in H'$ a function $F_h \in \text{Lip}((G/\Gamma \times G_0/\Gamma_0) \to \mathbb{S}^D)$ of Lipschitz norm $O_{C,\delta,d,\epsilon}(1)$;
- a vertical frequency $\eta\colon G^{\vec{D},d}_{[s,r_*]} \to \mathbb{R}$ of complexity $O_{C,d}(1)$ with dimension vector $\vec{D}$ on the universal nilmanifold $G^{\vec{D},d}/\Gamma^{\vec{D},d}$ of degree-rank $[s, r_*]$;
- a $O_{C,\delta,d,\epsilon,L}(1)$-rational filtered homomorphism $\phi\colon G^{\vec{D},d}/\Gamma^{\vec{D},d} \to G/\Gamma$

such that the followings hold:

(i) For all $h \in H'$, χ_h has an $(V(M), V(M)^h, O_{C,\delta,d,\epsilon,L}(1))$-universal representation of the form
$$(\vec{D}, \mathcal{T}^{\text{cor}} \uplus \mathcal{T}_h^{\text{ped}}, \eta, F_h, \phi, G/\Gamma \times G_0/\Gamma_0).$$
(ii) For all but δp^3 additive quadruples $h_1 + h_2 = h_3 + h_4$ with $h_1, \dots, h_4 \in H'$, the tower
$$\mathcal{T}^{\text{cor}} \uplus \uplus_{i=1}^{3} \mathcal{T}_{h_i}^{\text{lin}} \uplus \uplus_{i=1}^{4} \mathcal{T}_{h_i}^{\text{ind}}$$
is $(\tau(V(M)^{h_1,h_2,h_3}), L, p)$-independent.

(iii) For all $1 \leq j \leq s$ and $1 \leq k \leq D_j^{\text{lin}}$, the map $h \mapsto \vec{\xi}^{\text{lin}}_{h,j,k}, h \in H'$ is a p-almost linear Freiman homomorphism of complexity $O_{C,\delta,d,\epsilon,L}(1)$.

Proof. First we may raise the dimension of χ_h to D for all $h \in H$ by adding constant zero functions to the new coordinates. By Lemma 5.10, each χ_h has a $(V(M), \mathbb{F}_p^d, O_{C,d}(1))$-universal representation $(\vec{D}_h = (D_{h,j})_{1\leq j\leq s}, \mathcal{T}_h = (\vec{\xi}_{h,j,k})_{1\leq j\leq s, 1\leq k\leq D_{h,j}}, \eta_h)$. By the Pigeonhole Principle, there exists a subset $H'' \subseteq H$ with $|H''| \gg_{C,d,\epsilon} p^d$ such that $\vec{D}_h = \vec{D} = (D_1, \dots, D_s)$ for all $h \in H''$. Since $|\vec{D}| = O_{C,d}(1)$, applying Lemma 5.6 to the j-strings $\vec{\xi}_{h,j,1}, \dots, \vec{\xi}_{h,j,D_j}, h \in H''$ for each $1 \leq j \leq s$, we have that if $p \gg_{C,\delta,d,\epsilon,L} 1$, then there exists a subset $H' \subseteq H''$ with $|H'| \gg_{C,\delta,d,\epsilon,L} p^d$ such that for all $h \in H'$, the tower $(\vec{D}, \mathcal{T}_h)$ is $(\tau(V(M)^h), O_{C,\delta,d,\epsilon,L}(1), p)$-represented by a tower of the form $(\vec{D}^{\text{cor}} + \vec{D}^{\text{ped}}, \mathcal{T}^{\text{cor}} \uplus \mathcal{T}_h^{\text{ped}})$ satisfying Property (i) of Lemma 5.6, and Properties (ii) and (iii) in Lemma 5.12. By Lemma 5.11, we see that for all $h \in H'$, χ_h has a $(V(M), V(M)^h, O_{C,\delta,d,\epsilon,L}(1))$-universal representation of the form
$$(\vec{D}^{\text{cor}} + \vec{D}^{\text{ped}}, \mathcal{T}^{\text{cor}} \uplus \mathcal{T}_h^{\text{ped}}, \eta'_h, F_h, \phi_h, G_h/\Gamma_h \times G_{0,h}/\Gamma_{0,h}).$$
By the Pigeonhole Principle and passing to another subset, we may require $\eta'_h = \eta$, $\phi_h = \phi$, $G_h/\Gamma_h = G/\Gamma$ and $G_{0,h}/\Gamma_{0,h} = G_0/\Gamma_0$ to be independent of h. We are done. □

6. Information on the frequencies of χ_h

In this section, we use the sunflower lemma (Lemma 5.12) obtained in Section 5 to analysis the inequality (4.2). Our goal is to understand the relation between the frequencies of the characters $\chi_{h_1}, \chi_{h_2}, \chi_{h_3}$ and $\chi_{h_1+h_2-h_3}$ when χ_{h_1,h_2,h_3} correlates with an $(s-1)$-step nilsequence. Throughout this section, we assume that the conditions in Proposition 4.1 are satisfies, i.e. there exist a subset $H \subseteq \mathbb{F}_p^d$ with $|H| \geq \epsilon p^d$, some $\chi_0 \in \Xi_p^{(1,s);C,D}((\mathbb{F}_p^d)^2)$, some $\chi_h \in \Xi_p^{[s,r_*];C,D}(\mathbb{F}_p^d)$, and some $\psi_h \in \mathrm{Nil}_p^{s-1;C,1}(\mathbb{F}_p^d)$ for all $h \in H$ such that (4.1) holds for all $h \in H$. We may apply Lemma 5.12 to $(\chi_h)_{h\in H}$ for some $\delta, L > 0$ depending only on C, d, D, ϵ to be chosen later, and for $d \geq N(s)$, $p \gg_{C,d,D,\epsilon} 1$. For convenience we keep the same notions as in Lemma 5.12. Then we have a subset H' of H with $|H'| \gg_{C,d,D,\epsilon} p^d$ such that for all $h \in H'$, χ_h has a $(V(M), V(M)^h, O_{C,d,D,\epsilon}(1))$-universal representation

$$(\vec{D^{\mathrm{cor}}} + \vec{D^{\mathrm{ped}}}, \mathcal{T}^{\mathrm{cor}} \uplus \mathcal{T}_h^{\mathrm{ped}}, \eta).$$

The main effort in this section is to understand the property of the common vertical frequency η, with the main result being Theorem 6.5 to be stated later. By Proposition 4.1, there exists a subset $U \subseteq (H')^3$ with $|U| \gg_{C,d,D,\epsilon} p^{3d}$ such that for all $(h_1, h_2, h_3) \in U$, h_1, h_2, h_3 are linearly independent and M-non-isotropic, and

$$\left|\mathbb{E}_{n\in V(M)^{h_1,h_3,h_3-h_2}} \chi_{h_1,h_2,h_3}(n)\psi_{h_1,h_2,h_3}(n)\right| \gg_{C,d,D,\epsilon} 1 \tag{6.1}$$

for some $\psi_{h_1,h_2,h_3} \in \mathrm{Nil}_p^{s-1;O_{C,d,D,\epsilon}(1),1}(\mathbb{F}_p^d)$, where

$$\chi_{h_1,h_2,h_3}(n) := \chi_{h_1}(n) \otimes \chi_{h_2}(n + h_3 - h_2) \otimes \overline{\chi}_{h_3}(n) \otimes \overline{\chi}_{h_1+h_2-h_3}(n + h_3 - h_2).$$

Let U' be the set of $(h_1, h_2, h_3) \in (H')^3$ such that $(h_1, h_2, h_3, h_1 + h_2 - h_3)$ does not fall in the exceptional set in which Condition (ii) in Lemma 5.12 fails. Since $|U| \gg_{C,d,D,\epsilon} p^{3d}$, we may choose δ sufficiently small depending on C, d, D, ϵ such that $|U'| \geq |U|/2$. In particular, there exists at least one tuple $(h_1, h_2, h_3) \in (H')^3$ which is linearly independent and M-non-isotropic such that (6.1) holds and that

$$\mathcal{T}^{\mathrm{cor}} \uplus \uplus_{i=1}^3 \mathcal{T}_{h_i}^{\mathrm{lin}} \uplus \uplus_{i=1}^4 \mathcal{T}_{h_i}^{\mathrm{ind}}$$

is $(\tau(V(M)^{h_1,h_3,h_3-h_2}), O_{C,d,D,\epsilon}(1), p)$-independent, where $h_4 = h_1 + h_2 - h_3$.

We now fix such a choice of h_1, h_2, h_3. Then $\chi_{h_1,h_2,h_3}(n)\psi_{h_1,h_2,h_3}(n)$ can be written as a degree-rank $\leq [s, r_*]$ nilsequence $n \mapsto F(g \circ \tau(n)\Gamma)$. Here

$$G/\Gamma := \Big(\prod_{i=1}^4 G_{(i)}/\Gamma_{(i)}\Big) \times G_{(0)}/\Gamma_{(0)}$$

for some filtered nilmanifold $G_{(0)}/\Gamma_{(0)}$ of degree-rank $< [s, r_*]$ and complexity $O_{C,d,D,\epsilon}(1)$, some filtered nilmanifolds $G_{(i)}/\Gamma_{(i)}$ of degree-rank $\leq [s, r_*]$ and complexity $O_{C,d,D,\epsilon}(1)$ for $1 \leq i \leq 4$. The polynomial sequence g can be written as

$$g(n) = (g_1(n), g_2(n), g_3(n), g_4(n), g_0(n)) \in \mathrm{poly}_p(\mathbb{Z}^d \to G_{\mathrm{DR}}|\Gamma)$$

such that for all $1 \leq i \leq 4$, we may write $g_i = g_i' g_i''$ for some $g_i \in \mathrm{poly}_p(\mathbb{Z}^d \to (G_{(i)})_{\mathrm{DR}} | \Gamma_{(i)})$ and $g_i', g_i'' \in \mathrm{poly}(\mathbb{Z}^d \to (G_{(i)})_{\mathrm{DR}})$, where

$$\mathrm{Taylor}_m(g_i') = \Big(\pi_{\mathrm{Horiz}_{|m|}(G_{(i)}/\Gamma_{(i)})} \circ \phi_i(\prod_{k=1}^{D_{|m|}} e_{m,k}^{\xi_{h_i,m,k}})\Big), \tag{6.2}$$ [13]

and the j-string

$$(\mathrm{Taylor}_m(g_i''))_{|m|=j}$$

is $(\tau(V(M)^{h_i}), p)$-reducible for all $1 \leq j \leq s$, where $\vec{D} = (D_1, \ldots, D_s), \phi_i \colon G^{\vec{D}}/\Gamma^{\vec{D}} \to G_{(i)}/\Gamma_{(i)}$ is a filtered homomorphism and $\pi_{\mathrm{Horiz}_j(G_{(i)}/\Gamma_{(i)})} \colon (G_{(i)})_j \to \mathrm{Horiz}_j(G_{(i)}/\Gamma_{(i)})$ is the projection to the j-th type-II horizontal torus. Finally, $F \in \mathrm{Lip}(G/\Gamma \to \mathbb{S}^{D^4})$ has Lipschitz norm $O_{C,d,D,\epsilon}(1)$, and

$$F(\phi_1(t_1)x_1, \phi_2(t_2)x_2, \phi_3(t_3)x_3, \phi_4(t_4)x_4, y) = \exp(\eta(t_1 t_2 t_3^{-1} t_4^{-1}))F(x_1, x_2, x_3, x_4, y) \tag{6.3}$$

for all $(x_1, x_2, x_3, x_4, y) \in G/\Gamma$ and $t_1, \ldots, t_4 \in G_{[s,r_*]}^{\vec{D},d}$ (note that shifting χ_{h_i} does not affect the type-II Taylor coefficients of g_i, see the remark following Definition 9.6 of [12]).

By (6.1), we have that,

$$|\mathbb{E}_{n \in V(M)^{h_1,h_3,h_3-h_2}} F(g \circ \tau(n)\Gamma)| \gg_{C,D,d,\epsilon} 1.$$

Since h_1, h_2, h_3 are linearly independent, $V(M)^{h_1,h_3,h_3-h_2}$ can be written as $V(M) \cap (V+c)$ for some affine subspace $V+c$ of $\mathbb{F}_p^d$ of co-dimension 3. Since h_1, h_2, h_3 are M-non-isotropic, by Proposition A.4, $\mathrm{rank}(M|_{V+c}) = d-3$. By Theorem 2.28 (applied to the $\mathbb{N}$-filtration induced by the degree-rank filtration of G), since $p \gg_{C,d,D,\epsilon} 1$ and $d \geq s+16$, we have that

$$\Big| \int_{G_P/\Gamma_P} F(\epsilon_0 x)\, dm_{G_P/\Gamma_P}(x) \Big| \gg_{C,D,d,\epsilon} 1 \tag{6.4}$$

for some $\epsilon_0 \in G$ of complexity $O_{C,D,d,\epsilon}(1)$, some rational subgroup G_P of G which is $O_{C,d,D,\epsilon}(1)$-rational relative to the $O_{C,d,D,\epsilon}(1)$-rational Mal'cev basis $\mathcal{X}$ of G/Γ, and some $\Gamma_P = G_P \cap \Gamma$ with m_{G_P/Γ_P} being the Haar measure of G_P/Γ_P such that for all $1 \leq j \leq s$,

$$\pi_{\mathrm{Horiz}_j(G)}(G_P \cap G_j) \geq \Xi^{\perp}_{V(M)^{h_1,h_3,h_3-h_2},j}(g), \tag{6.5}$$

where $\Xi_{V(M)^{h_1,h_3,h_3-h_2},j}(g)$ is the group of all continuous homomorphisms $\xi_j \colon \mathrm{Horiz}_j(G/\Gamma) \to \mathbb{R}$ such that such that

$$\xi_j(\mathrm{Taylor}_j(g)(m_1, \ldots, m_j)) = \xi_j(\Delta_{m_j} \ldots \Delta_{m_1} g(n) \mod G_{[j,2]}) \in \mathbb{Z}$$

for all $(n, m_1, \ldots, m_j) \in \square_{p,j}(\tau(V(M)^{h_1,h_3,h_3-h_2}))$, and

$$\Xi^{\perp}_{V(M)^{h_1,h_3,h_3-h_2},j}(g) := \{x \in \mathrm{Horiz}_j(G) \colon \xi_j(x) \in \mathbb{Z} \text{ for all } \xi_j \in \Xi_{V(M)^{h_1,h_3,h_3-h_2},j}(g)\}.$$

Lemma 6.1. We have that $\Xi_{V(M)^{h_1,h_3,h_3-h_2},j}(g) = \Xi_{V(M)^{h_1,h_2,h_3},j}(g)$.

[13] For a core frequency, i.e. for $k \leq D_{|m|}^{\mathrm{cor}}$, we write $\xi_{h_i,m,k} = \xi_{m,k}$ for convenience.

Proof. Fix any $\xi_j \in \Xi_{V(M)^{h_1,h_3,h_3-h_2},j}(g)$. Denote
$$f(n) := \sum_{m\in\mathbb{N}^d,|m|=j} \xi_j(\mathrm{Taylor}_m(g))(m!)^* n^m.$$
By Lemma 2.25, $\mathrm{Taylor}_m(g) \in \mathbb{Z}/p$ for all $m \in \mathbb{N}^d$ with $|m| = j$. So f is a homogeneous polynomial of degree j with $\mathbb{Z}/p$ coefficients. By Lemma 2.24, we have
$$\begin{aligned}0 &\equiv \xi_j(\mathrm{Taylor}_j(g)(t_1,\dots,t_j)) \equiv \Delta_{t_j}\dots\Delta_{t_1}\Big(\sum_{m\in\mathbb{N}^d,|m|=j} \xi_j(\mathrm{Taylor}_m(g))\binom{n}{m}\Big)\\ &\equiv \Delta_{t_j}\dots\Delta_{t_1} f(n) \mod \mathbb{Z}\end{aligned}$$
for all $(n, m_1, \dots, m_j) \in \Box_{p,j}(\tau(V(M)^{h_1,h_3,h_3-h_2}))$. This means that f is $(\tau(V(M)^{h_1,h_3,h_3-h_2}), p)$-reducible. By Lemma 5.3, the map induced by f belongs to $J^M_{h_1,h_3,h_3-h_2} = J^M_{h_1,h_2,h_3}$. Again by Lemma 5.3, we have that f is $(\tau(V(M)^{h_1,h_2,h_3}), p)$-reducible. In other words, $\xi_j \in \Xi_{V(M)^{h_1,h_2,h_3},j}(g)$ and thus $\Xi_{V(M)^{h_1,h_3,h_3-h_2},j}(g) \subseteq \Xi_{V(M)^{h_1,h_2,h_3},j}(g)$. For similar reasons, we have that $\Xi_{V(M)^{h_1,h_3,h_3-h_2},j}(g) \supseteq \Xi_{V(M)^{h_1,h_2,h_3},j}(g)$ and so $\Xi_{V(M)^{h_1,h_3,h_3-h_2},j}(g) = \Xi_{V(M)^{h_1,h_2,h_3},j}(g)$. □

Lemma 6.2. For all $h_1 \in (G_{(1)})_{[s,r_*]}$ with $(h_1, id, id, id, id) \in G_P$, we have that $\eta(h_1) \in \mathbb{Z}$.

Proof. Denote $h := (h_1, id, id, id, id) \in G_P$. Then h is in the center of G and so
$$\begin{aligned}\int_{G_P/\Gamma_P} F(\epsilon_0 x)\, dm_{G_P/\Gamma_P}(x) &= \int_{G_P/\Gamma_P} F(h\epsilon_0 x)\, dm_{G_P/\Gamma_P}(x)\\ &= \exp(\eta(h_1)) \int_{G_P/\Gamma_P} F(\epsilon_0 x)\, dm_{G_P/\Gamma_P}(x),\end{aligned}$$
where we used (6.3) in the last equality. We may then deduce from (6.4) that $\eta(h_1) \in \mathbb{Z}$. □

For $1 \le j \le s$, let $V_{123,j}$ denote the subgroup of $\mathrm{Horiz}_j(G_{(1)})\times\mathrm{Horiz}_j(G_{(2)})\times\mathrm{Horiz}_j(G_{(3)})$ generated by
$$(\phi_1(e_{m,k}), \phi_2(e_{m,k}), \phi_3(e_{m,k})), |m| = j, 1 \le k \le D^{\mathrm{cor}}_j$$
and
$$(\phi_1(e_{m,k}), id, id), (id, \phi_2(e_{m,k}), id), (id, id, \phi_3(e_{m,k})), |m| = j, D^{\mathrm{cor}}_j + 1 \le k \le D_j.$$
Define $V_{124,j}$, $V_{134,j}$ and $V_{234,j}$ in a similar way.

Lemma 6.3. There exists $L_0 = L_0(c, d, D, \epsilon) \in \mathbb{N}$ such that if $L \ge L_0$, then for $1 \le j \le s$, the projection of $G_P \cap G_j$ to $\mathrm{Horiz}_j(G_{(1)}) \times \mathrm{Horiz}_j(G_{(2)}) \times \mathrm{Horiz}_j(G_{(3)})$ contains $V_{123,j}$, and similar results hold with $V_{123,j}$ replaced by $V_{124,j}$, $V_{134,j}$ and $V_{234,j}$.

Proof. We only prove for the case $V_{123,j}$ since the proof of other cases are similar. Suppose that the statement fails for some $1 \le j \le s$. Let
$$\pi_j\colon G_P \cap G_j \to \mathrm{Horiz}_j(G_{(1)}) \times \mathrm{Horiz}_j(G_{(2)}) \times \mathrm{Horiz}_j(G_{(3)})$$
and
$$\pi'_j\colon \mathrm{Horiz}_j(G) \to \mathrm{Horiz}_j(G_{(1)}) \times \mathrm{Horiz}_j(G_{(2)}) \times \mathrm{Horiz}_j(G_{(3)})$$

denote the projection maps, and let Z denote the set of $\tilde{\xi}_j \in \mathfrak{N}_j(G/\Gamma)$ which annihilates $\pi'^{-1}_j \circ \pi_j(G_P \cap G_j)$ (recall that $\mathfrak{N}_j(G/\Gamma)$ is the group of all j-th type-II horizontal characters). By duality, there exists $\tilde{\xi}_j \in Z$ which is nontrivial on $\pi'^{-1}_j(V_{123,j})$. Since G_P is $O_{C,d,D,\epsilon}(1)$-rational relative to the Mal'cev basis $\mathcal{X}$ of G/Γ, there exist $t = O_{C,d,D,\epsilon}(1) \in \mathbb{N}$ and $\xi_{j,1}, \dots, \xi_{j,t} \in Z$ of complexity at most $O_{C,d,D,\epsilon}(1)$ such that every $\tilde{\xi}_j \in Z$ is a linear combination of $\xi_{j,1}, \dots, \xi_{j,t}$. So there exists $\xi_j \in \{\xi_{j,1}, \dots, \xi_{j,t}\}$ which is nontrivial on $\pi'^{-1}_j(V_{123,j})$. By (6.5), Lemma 6.1 and duality, $\xi_j \in \Xi_{V(M)^{h_1,h_2,h_3},j}(g)$.

Since ξ_j annihilates $\pi'^{-1}_j \circ \pi_j(G_P \cap G_j)$, we have that

$$\xi_j(x_1, x_2, x_3, x_4, x_0) = \xi_{(1),j}(x_1) + \xi_{(2),j}(x_2) + \xi_{(3),j}(x_3) \tag{6.6}$$

for all $x_i \in \text{Horiz}_j(G_{(i)}), 0 \leq i \leq 4$ for some $\xi_{(i),j} \in \mathfrak{N}_j(G_{(i)}/\Gamma_{(i)})$. Fix any $(n, t_1, \dots, t_j) \in \square_{p,j}(\tau(V(M)^{h_1,h_2,h_3}))$. Since the j-string $(\text{Taylor}_m(g''_i))_{|m|=j}$ is $(\tau(V(M)^{h_i}), p)$-reducible, we have that

$$\sum_{|m|=j} \xi_{(i),j}(\text{Taylor}_m(g''_i))(m!)^* n^m$$

is $(\tau(V(M)^{h_i}), p)$-reducible for all $1 \leq i \leq 3$. Since $(n, t_1, \dots, t_j) \in \square_{p,j}(\tau(V(M)^{h_i}))$ for $1 \leq i \leq 3$, it follows from Lemma 2.24 that

$$\begin{aligned}
&\sum_{i=1}^{3} \xi_{(i),j}(\text{Taylor}_j(g''_i)(t_1, \dots, t_j)) \equiv \sum_{i=1}^{3} \Delta_{t_j} \dots \Delta_{t_1}\Big(\sum_{m \in \mathbb{N}^d, |m|=j} \xi_{(i),j}(\text{Taylor}_m(g''_i))\binom{n}{m}\Big) \\
\equiv &\sum_{i=1}^{3} \Delta_{t_j} \dots \Delta_{t_1}\Big(\sum_{m \in \mathbb{N}^d, |m|=j} \xi_{(i),j}(\text{Taylor}_m(g''_i))(m!)^* n^m\Big) \equiv 0 \mod \mathbb{Z}.
\end{aligned}$$

On the other hand, By the definition of $\Xi_{V(M)^{h_1,h_2,h_3},j}(g)$ and Lemma 2.24, we have

$$\sum_{i=1}^{3} \xi_{(i),j}(\text{Taylor}_j(g_i)(t_1, \dots, t_j)) = \sum_{i=1}^{3} \Delta_{t_j} \dots \Delta_{t_1}\Big(\sum_{m \in \mathbb{N}^d, |m|=j} \xi_{(i),j}(\text{Taylor}_m(g_i))\binom{n}{m}\Big) \in \mathbb{Z}.$$

So by (6.2),

$$\begin{aligned}
&\Delta_{t_j} \dots \Delta_{t_1}\Big(\sum_{m \in \mathbb{N}^d, |m|=j} \Big(\sum_{i=1}^{3} \sum_{k=1}^{D_j} \xi_{(i),j} \circ \phi_i(e_{m,k})\Big)\xi_{h_i,m,k}(m!)^* n^m\Big) \\
\equiv &\sum_{i=1}^{3} \sum_{m \in \mathbb{N}^d, |m|=j} \sum_{k=1}^{D_j} \xi_{(i),j} \circ \phi_i(e_{m,k})\xi_{h_i,m,k}\Big(\Delta_{t_j} \dots \Delta_{t_1}\binom{n}{m}\Big) \\
\equiv &\, \xi_j(\text{Taylor}_j(g')(t_1, \dots, t_j)) \equiv \sum_{i=1}^{3} \xi_{(i),j}(\text{Taylor}_j(g'_i)(t_1, \dots, t_j)) \\
\equiv &\sum_{i=1}^{3} \xi_{(i),j}(\text{Taylor}_j(g_i)(t_1, \dots, t_j)) - \sum_{i=1}^{3} \xi_{(i),j}(\text{Taylor}_j(g''_i)(t_1, \dots, t_j)) \equiv 0 \mod \mathbb{Z}.
\end{aligned}$$

Therefore

$$\sum_{m\in\mathbb{N}^d,|m|=j}\Big(\sum_{i=1}^{3}\sum_{k=1}^{D_j}\xi_{(i),j}\circ\phi_i(e_{m,k})\Big)\xi_{h_i,m,k}(m!)^*n^m \text{ is } (\tau(V(M)^{h_1,h_2,h_3}),p)\text{-reducible.} \tag{6.7}$$

Since ξ_j is of complexity at most $O_{C,d,D,\epsilon}(1)$, so are $\xi_{(i),j}$, $1\leq i\leq 3$. Since ϕ_i is $O_{C,d,D,\epsilon}(1)$-rational, we have that for all $1\leq i\leq 3$ and $D_j^{\text{cor}}+1\leq k\leq D_j$, $\xi_{(i),j}\circ\phi_i(e_{m,k})$ is of complexity $O_{C,d,D,\epsilon}(1)$. So if the tower

$$\mathcal{T}^{\text{cor}}\uplus\uplus_{i=1}^{3}\mathcal{T}_{h_i}^{\text{lin}}\uplus\uplus_{i=1}^{4}\mathcal{T}_{h_i}^{\text{ind}}$$

is $(\tau(V(M)^{h_1,h_2,h_3}),L,p)$-independent for some $L=O_{C,d,D,\epsilon}(1)$, then (6.7) implies that $\xi_{(i),j}\circ\phi_i(e_{m,k})$ vanishes for all $1\leq i\leq 3$, $|m|=j$ and $D_j^{\text{cor}}+1\leq k\leq D_j$, and that $\sum_{i=1}^{3}\xi_{(i),j}\circ\phi_i(e_{m,k})$ vanishes for all $|m|=j$ and $1\leq k\leq D_j^{\text{cor}}$. By (6.6), this implies that ξ_j vanishes on $\pi'^{-1}_j(V_{123,j})$, a contradiction. □

For $1\leq j\leq s$, let $V_{\text{ind},j}$ be the subspace of $\text{Horiz}_j(G_{(1)})\times\text{Horiz}_j(G_{(2)})\times\text{Horiz}_j(G_{(3)})\times\text{Horiz}_j(G_{(4)})$ generated by

$$(\phi_1(e_{m,k}),\phi_2(e_{m,k}),\phi_3(e_{m,k}),\phi_4(e_{m,k})),|m|=j,1\leq k\leq D_j^{\text{cor}}$$

and

$$(\phi_1(e_{m,k}),id,id,id),(id,\phi_2(e_{m,k}),id,id),(id,id,\phi_3(e_{m,k}),id),(id,id,id,\phi_4(e_{m,k}))$$

for $|m|=j$ and $D_j^{\text{cor}}+1\leq k\leq D_j^{\text{cor}}+D_j^{\text{ind}}$.

Lemma 6.4. There exists $L_0=L_0(c,d,D,\epsilon)\in\mathbb{N}$ such that if $L\geq L_0$, then for $1\leq j\leq s$, the projection of $G_P\cap G_j$ to $\text{Horiz}_j(G_{(1)})\times\text{Horiz}_j(G_{(2)})\times\text{Horiz}_j(G_{(3)})\times\text{Horiz}_j(G_{(4)})$ contains $V_{\text{ind},j}$.

Proof. Suppose that the statement fails for some $1\leq j\leq s$. Let

$$\pi_j\colon G_P\cap G_j\to\text{Horiz}_j(G_{(1)})\times\text{Horiz}_j(G_{(2)})\times\text{Horiz}_j(G_{(3)})\times\text{Horiz}_j(G_{(4)})$$

and

$$\pi'_j\colon \text{Horiz}_j(G)\to\text{Horiz}_j(G_{(1)})\times\text{Horiz}_j(G_{(2)})\times\text{Horiz}_j(G_{(3)})\times\text{Horiz}_j(G_{(4)})$$

denote the projection maps. By (6.5) and duality, similar to the argument in Lemma 6.3, we may find some $\xi_j\in\Xi_{V(M)^{h_1,h_2,h_3},j}(g)$ of complexity at most $O_{C,d,D,\epsilon}(1)$ which annihilates $\pi'^{-1}_j\circ\pi_j(G_P\cap G_j)$ and is nontrivial on $\pi'^{-1}_j(V_{\text{ind},j})$.

Since ξ_j annihilates $\pi'^{-1}_j\circ\pi_j(G_P\cap G_j)$, We may write

$$\xi_j(x_1,x_2,x_3,x_4,x_0)=\xi_{(1),j}(x_1)+\xi_{(2),j}(x_2)+\xi_{(3),j}(x_3)+\xi_{(4),j}(x_4) \tag{6.8}$$

for all $x_i\in\text{Horiz}_j(G_{(i)}),0\leq i\leq 4$ for some characters $\xi_{(i),j}\in\mathfrak{N}_j(G_{(i)}/\Gamma_{(i)})$. Fix any $(n,t_1,\dots,t_j)\in\square_{p,j}(\tau(V(M)^{h_1,h_2,h_3}))$. Since the j-string $(\text{Taylor}_m(g''_i))_{|m|=j}$ is $(\tau(V(M)^{h_i}),p)$-reducible, we have that

$$\sum_{|m|=j}\xi_{(i),j}(\text{Taylor}_m(g''_i))(m!)^*n^m$$

is $(\tau(V(M)^{h_i}), p)$-reducible for all $1 \leq i \leq 4$. By Lemma 2.24,

$$\sum_{i=1}^{4} \xi_{(i),j}(\text{Taylor}_j(g_i'')(t_1, \dots, t_j)) \equiv \sum_{i=1}^{4} \Delta_{t_j} \dots \Delta_{t_1}\Big(\sum_{m \in \mathbb{N}^d, |m|=j} \xi_{(i),j}(\text{Taylor}_m(g_i''))\binom{n}{m}\Big)$$
$$\equiv \sum_{i=1}^{4} \Delta_{t_j} \dots \Delta_{t_1}\Big(\sum_{m \in \mathbb{N}^d, |m|=j} \xi_{(i),j}(\text{Taylor}_m(g_i''))(m!)^* n^m\Big) \equiv 0 \mod \mathbb{Z}.$$

On the other hand, By the definition of $\Xi_{V(M)^{h_1,h_2,h_3},j}(g)$ and Lemma 2.24, we have

$$\sum_{i=1}^{4} \xi_{(i),j}(\text{Taylor}_j(g_i)(t_1, \dots, t_j)) = \sum_{i=1}^{4} \Delta_{t_j} \dots \Delta_{t_1}\Big(\sum_{m \in \mathbb{N}^d, |m|=j} \xi_{(i),j}(\text{Taylor}_m(g_i))\binom{n}{m}\Big) \in \mathbb{Z}.$$

So by (6.8),

$$\Delta_{t_j} \dots \Delta_{t_1}\Big(\sum_{m \in \mathbb{N}^d, |m|=j} \Big(\sum_{i=1}^{4} \sum_{k=1}^{D_j} \xi_{(i),j} \circ \phi_i(e_{m,k})\Big)\xi_{h_i,m,k}(m!)^* n^m\Big)$$
$$\equiv \sum_{i=1}^{4} \sum_{m \in \mathbb{N}^d, |m|=j} \sum_{k=1}^{D_j} \xi_{(i),j} \circ \phi_i(e_{m,k})\xi_{h_i,m,k}\Big(\Delta_{t_j} \dots \Delta_{t_1}\binom{n}{m}\Big)$$
$$\equiv \xi_j(\text{Taylor}_j(g')(t_1, \dots, t_j)) \equiv \sum_{i=1}^{4} \xi_{(i),j}(\text{Taylor}_j(g_i')(t_1, \dots, t_j))$$
$$\equiv \sum_{i=1}^{4} \xi_{(i),j}(\text{Taylor}_j(g_i)(t_1, \dots, t_j)) - \sum_{i=1}^{4} \xi_{(i),j}(\text{Taylor}_j(g_i'')(t_1, \dots, t_j)) \equiv 0 \mod \mathbb{Z}.$$

Therefore,

$$\sum_{m \in \mathbb{N}^d, |m|=j} \Big(\sum_{i=1}^{4} \sum_{k=1}^{D_j} \xi_{(i),j} \circ \phi_i(e_{m,k})\Big)\xi_{h_i,m,k}(m!)^* n^m \text{ is } (\tau(V(M)^{h_1,h_2,h_3}), p)\text{-reducible.} \tag{6.9}$$

Since the map $h \mapsto \xi_{h,m,k}, h \in H'$ is a p-almost linear Freiman homomorphism for all $D_j^{\text{cor}} + D_j^{\text{ind}} + 1 \leq k \leq D_j$ and $|m| = j$, we have that

$$\xi_{h_4,m,k} \equiv \xi_{h_1,m,k} + \xi_{h_2,m,k} - \xi_{h_3,m,k} \mod \mathbb{Z}. \tag{6.10}$$

Since ξ_j is of complexity at most $O_{C,d,D,\epsilon}(1)$, so are $\xi_{(i),j}, 1 \leq i \leq 4$. Since ϕ_i is $O_{C,d,D,\epsilon}(1)$-rational, we have that for all $1 \leq i \leq 4$ and $D_j^{\text{cor}} + 1 \leq k \leq D_j$, $\xi_{(i),j} \circ \phi_i(e_{m,k})$ is of complexity $O_{C,d,D,\epsilon}(1)$. So if the tower

$$\mathcal{T}^{\text{cor}} \uplus \uplus_{i=1}^{3} \mathcal{T}_{h_i}^{\text{lin}} \uplus \uplus_{i=1}^{4} \mathcal{T}_{h_i}^{\text{ind}}$$

is $(\tau(V(M)^{h_1,h_2,h_3}), L, p)$-independent for some $L = O_{C,d,D,\epsilon}(1)$, then we may deduce from (6.9) and (6.10) that $\xi_{(i),j} \circ \phi_i(e_{m,k})$ vanishes for all $1 \leq i \leq 4$, $|m| = j$ and $D_j^{\text{cor}} + 1 \leq k \leq D_j^{\text{cor}} + D_j^{\text{ind}}$, and that $\sum_{i=1}^{4} \xi_{(i),j} \circ \phi_i(e_{m,k})$ vanishes for $|m| = j$ and $1 \leq k \leq D_j^{\text{cor}}$. By (6.8), this implies that ξ_j vanishes on $\pi'^{-1}_j(V_{\text{ind},j})$, a contradiction. □

We may now summarize the discussion in this section by the following theorem:

Theorem 6.5. Let the notations be the same as in Section 6 and let $\delta = \delta(C, d, D, \epsilon)$, $L = L_0(C, d, D, \epsilon)$. Let $w \in G^{\vec{D},d}$ be any iterated commutator of $e_{m_1,j_1}, \dots, e_{m_{r_*},j_{r_*}}$ for some $1 \leq |m_1|, \dots, |m_{r_*}| \leq s$ with $|m_1| + \cdots + |m_{r_*}| = s$ and for some $1 \leq j_\ell \leq D_{|m_\ell|}$ for all $1 \leq \ell \leq r_*$.

(i) If $j_\ell > D^{\mathrm{cor}}_{|m_\ell|}$ for at least two values of ℓ, then $\eta(w) \in \mathbb{Z}$;
(ii) If $D^{\mathrm{cor}}_{|m_\ell|} < j_\ell \leq D^{\mathrm{cor}}_{|m_\ell|} + D^{\mathrm{ind}}_{|m_\ell|}$ for at least one value of ℓ, then $\eta(w) \in \mathbb{Z}$.

Proof. Let $h_1, h_2, h_3 \in H'$ be chosen as in this section. Suppose that $j_\ell > D^{\mathrm{cor}}_{|m_\ell|}$ for at least two values of ℓ. By Corollary 11.4 of [12], we have that Lemma 6.3 implies that $(\phi_1(w), id, id, id, id) \in G_P$. So $\eta(w) \in \mathbb{Z}$ by Lemma 6.2.

Similarly, suppose that $D^{\mathrm{cor}}_{|m_\ell|} < j_\ell \leq D^{\mathrm{cor}}_{|m_\ell|} + D^{\mathrm{ind}}_{|m_\ell|}$ for at least one value of ℓ. By Corollary 11.6 of [12], we have that Lemma 6.4 implies that $(\phi_1(w), id, id, id, id) \in G_P$. So $\eta(w) \in \mathbb{Z}$ by Lemma 6.2. □

7. The construction of a degree-$(1, s)$ nilcharacter

In this section, we complete Step 3 described in Section 1.2 by showing that the nilcharacters $\chi_h(n)$ described in Proposition 4.1 can be expressed in the form $\chi(h, n)$ for some nilcharacters χ of multi-degree $(1, s)$. To be more precise, we show:

Proposition 7.1. Let $d, D \in \mathbb{N}_+$, $C, \epsilon > 0$, $[s, r_*] \in \mathrm{DR}$, $p \gg_{C,d,D,\epsilon} 1$ be a prime, $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a non-degenerate quadratic form, and $f\colon \mathbb{F}_p^d \to \mathbb{C}$ be a function bounded by 1. Suppose that there exist a subset H of $\mathbb{F}_p^d$ with $|H| \geq \epsilon p^d$, some $\chi_0 \in \Xi_p^{(1,s);C,D}((\mathbb{F}_p^d)^2)$, some $\chi_h \in \Xi_p^{[s,r_*];C,D}(\mathbb{F}_p^d)$, and some $\psi_h \in \mathrm{Nil}_p^{s-1;C,1}(\mathbb{F}_p^d)$ for all $h \in H$ such that

$$\left|\mathbb{E}_{n \in V(M)^h} f(n+h)\overline{f}(n)\chi_0(h, n) \otimes \chi_h(n)\psi_h(n)\right| > \epsilon \tag{7.1}$$

for all $h \in H$. If $d \geq N(s)$, then there exist a subset $H' \subseteq H$ with $|H'| \gg_{C,d,D,\epsilon} p^d$, some $\chi \in \Xi_p^{(1,s);O_{C,d,D,\epsilon}(1),O_{C,d,D,\epsilon}(1)}((\mathbb{F}_p^d)^2)$, some $\chi'_h \in \Xi_p^{<[s,r_*];O_{C,d,D,\epsilon}(1),O_{C,d,D,\epsilon}(1)}(\mathbb{F}_p^d)$ for all $h \in H'$ such that

$$\left|\mathbb{E}_{n \in V(M)^h} f(n+h)\overline{f}(n)\chi(h, n) \otimes \chi'_h(n)\psi_h(n)\right| \gg_{C,d,D,\epsilon} 1$$

for all $h \in H'$.

The rest of the section is devoted to the proof of Proposition 7.1.

Step 1. We first use the sunflower lemma to represent all of χ_h in a uniform way. Passing to a subset if necessary, by Lemma A.7, we may assume without loss of generality that h is M-non-isotropic for all $h \in H$. Let $\delta = \delta(C, d, D, \epsilon), L = L(C, d, D, \epsilon) > 0$ to be chosen later. Since $p \gg_{C,d,D,\epsilon} 1$ and $d \geq N(s)$, there exists a subset H' of H with $|H'| \gg_{C,d,D,\epsilon} p^d$ such that all the requirements of Lemma 5.12 are satisfied for $(\chi_h)_{h \in H'}$. For convenience we use the same notations as in Lemma 5.12.

Note that if we replace the function F_h which defines χ_h in Lemma 5.12 by a function F' such that $\|F_h - F'\|_{L^\infty} < \epsilon/2$. Then (7.1) still holds with the right hand side replaced by $\epsilon/2$. On the other hand, there exists $K = O_{C,d,D,\epsilon}(1)$ such that all of $F_h, h \in H'$ belong to the set

S of all Lipschitz functions of Lipschitz norm at most K. Note that S is equicontinuous and pointwise bounded, and thus is relatively compact by Arzelà-Ascoli Theorem. So there exist open balls $B_1, \dots, B_N$ of radius $\epsilon/2$ centered at $F_1, \dots, F_N$ respectively for some $N = O_{C,d,D,\epsilon}(1)$ which covers S. By the Pigeonhole Principle, there exists a subset H'' of H' of cardinality $\gg_{C,d,D,\epsilon} p^d$ such that all of $F_h, h \in H''$ lie in the same ball B centered at F. Then $\|F - F_h\|_{L^\infty} < \epsilon/2$. In conclusion, replacing the right hand side of (7.1) by $\epsilon/2$ if necessary and denoting H'' as H' for convenience, we may assume without loss of generality that (7.1) holds and that $F_h = F$ for all $h \in H'$. We now choose $\delta = \delta(C, d, D, \epsilon)$ to be sufficiently small and $L = L(C, d, D, \epsilon)$ to be sufficiently large such that Theorem 6.5 applies to the family $(\chi_h)_{h\in H'}$.

For all $h \in H'$, we may write $\chi_h = \chi'_h \circ \tau$, where

$$\chi'_h(n) = F(g_h(n)\Gamma, g_{0,h}(n)\Gamma_0), \text{ for all } n \in \mathbb{Z}^d, h \in H',$$

$g_h \in \text{poly}_p(\mathbb{Z}^d \to G_{\mathbb{N}}|\Gamma)$, $g_{0,h} \in \text{poly}_p(\mathbb{Z}^d \to (G_0)_{\mathbb{N}}|\Gamma_0)$, F is a Lipschitz function of complexity $O_{C,d,D,\epsilon}(1)$ with a vertical frequency $\eta\colon G^{\vec{D},d}_{[s,r_*]} \to \mathbb{R}$ of complexity $O_{C,d,D,\epsilon}(1)$ in the sense that

$$F(\phi(t)x, x_0) = \exp(\eta(t))F(x, x_0) \tag{7.2}$$

for all $t \in G^{\vec{D},d}_{[s,r_*]}$, $x \in G/\Gamma$ and $x_0 \in G_0/\Gamma_0$, and that the j-string

$$\Big(\text{Taylor}_m(g_h) \cdot \Big(\pi_{\text{Horiz}_j(G/\Gamma)} \circ \phi\big(\prod_{k=1}^{D_j^{\text{cor}}} e_{m,k}^{\xi_{m,k}} \prod_{k=D_j^{\text{cor}}+1}^{D_j} e_{m,k}^{\xi_{h,m,k}}\big)\Big)^{-1}\Big)_{|m|=j} \tag{7.3}$$

is $(\tau(V(M)^h), p)$-reducible for all $1 \le j \le s$.

Step 2. We next show that by modifying the polynomial sequence g_h appropriately, we can make (7.3) take values in Γ.

Let $\mathcal{X}$ be a Mal'cev basis adapted to the degree filtration $(G_{[j,1]})_{j\in\mathbb{N}}$. Assume that $\mathcal{X}$ can be partitioned into $\cup_{1\le x\le s}\mathcal{X}_x$ such that $G_{[j,1]}$ is generated by $\cup_{x\ge j}\mathcal{X}_x$. Since $G_{[j,1]}/G_{[j+1,1]}$ is abelian, We may further partition $\mathcal{X} = \cup_{[x,y]\in\text{DR},[x,y]\le[s,r_*]}\mathcal{X}_{x,y}$, where $\mathcal{X}_{x,y} = \{X_{x,y,1}, \dots, X_{x,y,\ell_{x,y}}\}$ is such that $G_{[x,y]}$ is generated by $\cup_{[x',y']\in\text{DR},[x',y']\ge[x,y]}\mathcal{X}_{x',y'}$. By Lemma 2.15, we may write

$$g_h(n) = \prod_{1\le x\le s} \prod_{1\le y\le x} \prod_{z=1}^{\ell_{x,y}} \exp(P^h_{x,y,z}(n)X_{x,y,z})$$

for some polynomial $P^h_{x,y,z}\colon \mathbb{Z}^d \to \mathbb{R}$ of degree at most x. Assume that

$$P^h_{x,1,z}(n) = \sum_{m\in\mathbb{N}^d,|m|\le x} a^h_{x,z,m}\binom{n}{m} \tag{7.4}$$

for some $a^h_{x,z,m} \in \mathbb{R}$ for all $1 \le x \le s$, $1 \le z \le \ell_{x,1}$. By Lemma 2.22, for all $m \in \mathbb{N}^d$, we have that

$$\text{Taylor}_m(g_h) = \prod_{z=1}^{\ell_{|m|,1}} \exp(a^h_{|m|,z,m}X_{|m|,1,z}) \mod G_{[|m|,2]}. \tag{7.5}$$

Assume that

$$\phi\Big(\prod_{k=1}^{D_j^{\mathrm{cor}}} e_{m,k}^{\xi_{m,k}} \prod_{k=D_j^{\mathrm{cor}}+1}^{D_j} e_{m,k}^{\xi_{h,m,k}}\Big) = \prod_{z=1}^{\ell_{|m|,1}} \exp(b^h_{z,m} X_{|m|,1,z}) \mod G_{[|m|,2]} \tag{7.6}$$

for some $b^h_{z,m} \in \mathbb{R}$ for all $m \in \mathbb{N}^d$. Since $g_h \in \mathrm{poly}_p(\mathbb{Z}^d \to G_{\mathbb{N}}|\Gamma)$, by Lemma 2.25, we have that $a^h_{|m|,z,m} \in \mathbb{Z}/p$. Since $\xi_{m,k}, \xi_{h,m,k} \in \mathbb{Z}/p$, we have that $b^h_{z,m} \in \mathbb{Z}/p$. Since (7.3) is $(\tau(V(M)^h), p)$-reducible for all $1 \leq x \leq s$, it follows from (7.5) and (7.6) that the polynomial

$$\sum_{m\in\mathbb{N}^d,|m|=x} (a^h_{x,z,m} - b^h_{z,m})(m!)^* n^m$$

is $(\tau(V(M)^h), p)$-reducible for all $1 \leq x \leq s$ and $1 \leq z \leq \ell_{x,1}$. Since $d \geq N(s)$ and h is M-non-isotropic, by Proposition A.13, we may write

$$\sum_{m\in\mathbb{N}^d,|m|=x} (a^h_{x,z,m} - b^h_{z,m})(m!)^* n^m = f^h_{x,z}(n) + g^h_{x,z}(n) \tag{7.7}$$

for some $\mathbb{Z}/p$-valued polynomial $f^h_{x,z}$ of degree at most x with $f^h_{x,z}(n) \in \mathbb{Z}$ for all $n \in \tau(V(M)^h) + p\mathbb{Z}^d$, and some $\mathbb{Z}/p$-valued polynomial $g^h_{x,z}$ of degree at most $x-1$. By Lemma A.1, modifying $f^h_{x,z}$ and $g^h_{x,z}$ if necessary, we may further require that $g^h_{x,z}$ has $\mathbb{Z}/p$-coefficients.

Denote

$$g''_h(n) := \prod_{1\leq x\leq s} \prod_{z=1}^{\ell_{x,1}} \exp(Q^h_{x,1,z}(n) X_{x,1,z}),$$

where

$$Q^h_{x,1,z}(n) := P^h_{x,1,z}(n) + \sum_{m\in\mathbb{N}^d,|m|=x} (b^h_{z,m} - a^h_{x,z,m})(m!)^* n^m + g^h_{x,z}(n) \tag{7.8}$$

is a polynomial with $\mathbb{Z}/p$-coefficients. Since $Q^h_{x,y,z}$ is a polynomial of degree at most x, we have that the map $n \mapsto \exp(Q^h_{x,y,z}(n) X_{x,y,z})$ belongs to $\mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$, and thus $g''_h \in \mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$ by Corollary B.4 of [12] (however, we caution the readers that g''_h needs not necessarily belong to $\mathrm{poly}_p(\mathbb{Z}^d \to G_{\mathbb{N}}|\Gamma)$). By Lemma 2.22, (7.4) and (7.6),

$$\begin{aligned} \mathrm{Taylor}_m(g''_h) &\equiv \prod_{z=1}^{\ell_{|m|,1}} \mathrm{Taylor}_m(\exp(Q^h_{|m|,1,z}(\cdot) X_{|m|,1,z})) \\ &\equiv \prod_{z=1}^{\ell_{|m|,1}} \exp\Big(((1-m!(m!)^*)a^h_{|m|,z,m} + m!(m!)^* b^h_{z,m}) X_{|m|,1,z}\Big) \qquad (7.9) \\ &\equiv \prod_{z=1}^{\ell_{|m|,1}} \exp(b^h_{z,m} X_{|m|,1,z}) \equiv \phi\Big(\prod_{k=1}^{D_j^{\mathrm{cor}}} e_{m,k}^{\xi_{m,k}} \prod_{k=D_j^{\mathrm{cor}}+1}^{D_j} e_{m,k}^{\xi_{h,m,k}}\Big) \mod G_{[j,2]} \end{aligned}$$

for all $1 \leq j \leq s$ and $m \in \mathbb{N}^d$ with $|m| = j$.

On the other hand, by (7.7) and (7.8), $P^h_{x,1,z}(n) - Q^h_{x,1,z}(n) = f^h_{x,z}(n)$. By the Baker-Campbell-Hausdorff formula, there exists $g'''_h \in \text{poly}(\mathbb{Z}^d \to [G,G]_{\mathbb{N}})$ (with the filtration induced by $G_{\mathbb{N}}$) such that

$$g_h(n) = g''_h(n)g'''_h(n) \prod_{1\leq x\leq s} \prod_{z=1}^{\ell_{x,1}} \exp(f^h_{x,z}(n)X_{x,1,z})$$

for all $n \in \mathbb{Z}^d$. Write $g'_h := g''_h g'''_h \in \text{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$. Since g'''_h takes values in $G_{[0,2]}$, it follows from Lemma 2.22 and (7.9) that

$$\text{Taylor}_m(g'_h) = \text{Taylor}_m(g''_h) = \phi\Big(\prod_{k=1}^{D^{\text{cor}}_j} e^{\xi_{m,k}}_{m,k} \prod_{k=D^{\text{cor}}_j+1}^{D_j} e^{\xi_{h,m,k}}_{m,k}\Big) \mod G_{[j,2]} \tag{7.10}$$

for all $1 \leq j \leq s$ and $m \in \mathbb{N}^d$ with $|m| = j$. Finally, for all $n \in \tau(V(M)^h) + p\mathbb{Z}^d$, since $f^h_{x,z}(n) \in \mathbb{Z}$, we have that $g_h(n)\Gamma = g'_h(n)\Gamma$. So

$$\chi'_h(n) = F(g'_h(n)\Gamma, g_{0,h}(n)\Gamma_0), \text{ for all } n \in \tau(V(M)^h) + p\mathbb{Z}^d, h \in H'.$$

Step 3. Our next step is to show that we can remove the core frequencies $\xi_{m,k}$ in the expression (7.10) by studying a nilcharacter which differs from χ'_h by a nilsequence of lower degree. We use a variation of the method in Section 12 of [12]. Let $\tilde{G}$ be the free Lie group generated by the generators $\tilde{e}_{m,k}, m \in \mathbb{N}^d, 1 \leq |m| \leq s, k \in [1, D^{\text{cor}}_{|m|}] \cup [D^{\text{cor}}_{|m|} + D^{\text{ind}}_{|m|} + 1, D_{|m|}]$ subject to the following relations:

- any iterated commutators $\tilde{e}_{m_1,k_1}, \dots, \tilde{e}_{m_r,k_r}$ vanishes if $|m_1| + \dots + |m_r| > s$;
- any iterated commutators $\tilde{e}_{m_1,k_1}, \dots, \tilde{e}_{m_r,k_r}$ vanishes if $|m_1|+\dots+|m_r| = s$ and $r > r_*$;
- any iterated commutators $\tilde{e}_{m_1,k_1}, \dots, \tilde{e}_{m_r,k_r}$ vanishes if $k_\ell > D^{\text{cor}}_{|m_\ell|}$ for at least two values of ℓ.

We may endow a DR-filtration $\tilde{G}_{\text{DR}}$ on $\tilde{G}$ by setting $\tilde{G}_{[t,r]}$ to be the group generated by iterated commutators $\tilde{e}_{m_1,k_1}, \dots, \tilde{e}_{m_{r'},k_{r'}}$ with $|m_1| + \dots + |m_r| \geq t$ and $r' \geq r$. Let $\tilde{\Gamma}$ be the discrete group generated by all of $\tilde{e}_{m,k}$. Then $\tilde{G}/\tilde{\Gamma}$ is a nilmanifold of degree-rank at most $[s, r_*]$.

Let G^* be the subgroup of $G^{\vec{D},d}$ generated by iterated commutators of $e_{m_1,k_1}, \dots, e_{m_r,k_r}$ for some $r \in \mathbb{N}_+$ such that either $k_\ell > D^{\text{cor}}_{|m_\ell|}$ for at least two values of ℓ, or $D^{\text{cor}}_{|m_\ell|} < k_\ell \leq D^{\text{cor}}_{|m_\ell|} + D^{\text{ind}}_{|m_\ell|}$ for at least one value of ℓ. Then $\tilde{G}$ is isomorphic to the quotient of $G^{\vec{D},d}$ by G^*. Let $\tilde{\phi}\colon G^{\vec{D},d} \to \tilde{G}$ denote the quotient map. It is clear that $\tilde{G}_{[s,r_*]}$ is isomorphic to the quotient of $G^{\vec{D},d}_{[s,r_*]}$ by $G^* \cap G^{\vec{D},d}_{[s,r_*]}$ under the map ϕ.

Let G^{**} be the subgroup of $G^{\vec{D},d}$ generated by iterated commutators of $e_{m_1,j_1}, \dots, e_{m_{r_*},j_{r_*}}$ for some $1 \leq |m_1|, \dots, |m_{r_*}| \leq s$ with $|m_1|+\dots+|m_{r_*}| = s$ and for some $1 \leq j_\ell \leq D_{|m_\ell|}$ for all $1 \leq \ell \leq r_*$ such that either $j_\ell > D^{\text{cor}}_{|m_\ell|}$ for at least two values of ℓ or $D^{\text{cor}}_{|m_\ell|} < j_\ell \leq D^{\text{cor}}_{|m_\ell|} + D^{\text{ind}}_{|m_\ell|}$ for at least one value of ℓ. It is clear that $G^* \cap G^{\vec{D},d}_{[s,r_*]}$ is a subgroup of the central group $G^{\vec{D},d}_{[s,r_*]}$ containing G^{**}.

By Theorem 6.5, $\eta\colon G^{\vec{D},d}_{[s,r_*]} \to \mathbb{R}$ annihilates G^{**} and thus descends to a group homomorphism $\eta'\colon G^{\vec{D},d}_{[s,r_*]}/G^{**} \to \mathbb{R}$. Since $G^{*} \cap G^{\vec{D},d}_{[s,r_*]}$ contains G^{**}, η' induces a group homomorphism $\eta''\colon G^{\vec{D},d}_{[s,r_*]}/(G \cap G^{\vec{D},d}_{[s,r_*]}) \to \mathbb{R}$. Since $\tilde{G}_{[s,r_*]}$ is isomorphic to $G^{\vec{D},d}_{[s,r_*]}/(G \cap G^{\vec{D},d}_{[s,r_*]})$ under the quotient map $\tilde{\phi}$, η'' induces a group homomorphism $\tilde{\eta}\colon \tilde{G}_{[s,r_*]} \to \mathbb{R}$. Finally, since η is a vertical frequency of $G^{\vec{D},d}/\Gamma^{\vec{D},d}$ of complexity $O_{C,d,D,\epsilon}(1)$, we have that $\tilde{\eta}$ is a vertical frequency of $\tilde{G}/\tilde{\Gamma}$ of complexity $O_{C,d,D,\epsilon}(1)$. By (7.2) and similar to the construction in (6.3) of [12], it is not hard to construct a function $\tilde{F} \in \mathrm{Lip}(\tilde{G}/\tilde{\Gamma} \to \mathbb{S}^{O_{C,d,D,\epsilon}(1)})$ of complexity $O_{C,d,D,\epsilon}(1)$ with vertical frequency $\tilde{\eta}$.

For all $D^{\mathrm{cor}}_{|m|} + D^{\mathrm{ind}}_{|m|} + 1 \leq k \leq D_{|m|}$, we may write $\xi_{h,m,k} = \xi_{m,k}(h)$ for some almost p-linear Freiman homomorphism $\xi_{m,k}$ defined on H' given by

$$\xi_{h,m,k} := \xi_{m,k}(h) = \sum_{i=1}^{K} \{\alpha_{m,k,i} \cdot \tau(h)\}\beta_{m,k,i}$$

for some $K = O_{C,d,D,\epsilon}(1)$, $\alpha_{m,k,i} \in (\mathbb{Z}/p)^d, \beta_{m,k,i} \in \mathbb{Z}/p$. Passing to another subset of H' if necessary, we may assume without loss of generality that for all m, k and i, all of $\{\alpha_{m,k,i} \cdot \tau(h)\}, h \in H'$ lie in an interval $I_{m,k,i}$ of length at most $1/10$.

Denote

$$g_0(n) := \prod_{1\leq |m| \leq s} \prod_{k=1}^{D^{\mathrm{cor}}_{|m|}} \tilde{\phi}(e_{m,k})^{\xi_{m,k}\binom{n}{m}}, \tag{7.11}$$

and

$$\tilde{g}_h(n) := \prod_{1\leq |m| \leq s} \prod_{k=D^{\mathrm{cor}}_{|m|}+D^{\mathrm{ind}}_{|m|}+1}^{D_{|m|}} \tilde{\phi}(e_{m,k})^{\xi_{m,k}(h)\binom{n}{m}} = \prod_{1\leq |m| \leq s} \prod_{k=D^{\mathrm{cor}}_{|m|}+D^{\mathrm{ind}}_{|m|}+1}^{D_{|m|}} \prod_{i=1}^{K} \tilde{\phi}(e_{m,k})^{\{\alpha_{m,k,i}\cdot\tau(h)\}\beta_{m,k,i}\binom{n}{m}}. \tag{7.12}$$

Consider the nilcharacter

$$\tilde{\chi}_h(n) := \tilde{F}(g_0(n)\tilde{g}_h(n)\tilde{\Gamma}). \tag{7.13}$$

Clearly, $\tilde{\chi}_h \in \Xi^{[s,r_*];O_{C,d,D,\epsilon}(1),O_{C,d,D,\epsilon}(1)}(\mathbb{Z}^d)$. Again we remark that $g_0\tilde{g}_h$ is not necessarily p-periodic. However, similar to Lemma 12.1 of [12], we have

Lemma 7.2. For all $h \in H'$, we have that the map

$$n \mapsto \chi'_h(n) \otimes \overline{\tilde{\chi}}_h(n)$$

belongs to $\mathrm{Nil}^{<[s;r_*];O_{C,d,D,\epsilon}(1)}(\tau(V(M)^h)+p\mathbb{Z}^d)$. Moreover, χ'_h is an $O_{C,d,D,\epsilon}(1)$-complexity linear combination of the components $\tilde{\chi}_h \otimes \psi'_h$ for some $\psi'_h \in \mathrm{Nil}^{<[s,r_*];O_{C,d,D,\epsilon},O_{C,d,D,\epsilon}}(\tau(V(M)^h) + p\mathbb{Z}^d)$.

Proof. We may rewrite the sequence $n \mapsto \chi'_h(n) \otimes \overline{\tilde{\chi}}_h(n)$ as

$$n \mapsto F'_h(f_h(n)\Gamma') \text{ for all } n \in \tau(V(M)^h) + p\mathbb{Z}^d, \tag{7.14}$$

where $G' := G \times G_0 \times \tilde{G}$, $\Gamma' := \Gamma \times \Gamma_0 \times \tilde{\Gamma}$,

$$f_h(n) := (g'_h(n), g_{0,h}(n), g_0(n)\tilde{g}_h(n))$$

and F'_h is the Lipschitz function on G'/Γ' given by

$$F'_h(x, x_0, y) := F_h(x, x_0) \otimes \overline{F}(y).$$

We define a DR-filtration G'_{DR} on G' as follows. For $[i, r] \in \mathrm{DR}$ with $r \geq 1$, let $G'_{[i,r]}$ be the Lie group generated by $G_{[i,r+1]} \times (G_0)_{[i,r]} \times \tilde{G}_{[i,r+1]}$ and $\{(\phi(g), id, \tilde{\phi}(g)) \colon g \in G^{\vec{D},d}_{[s,r]}\}$, where we take the convention that $[i, i+1] = [i+1, 0]$. Set also $G'_{[i,0]} := G'_{[i,1]}$. It is not hard to see that this defines a degree-rank filtration.

We first show that f_h is a polynomial with respect to the filtration G'_{DR}. Indeed, the sequence $n \mapsto (id, g_{0,h}(n), id)$ is already a polynomial with respect to the filtration G'_{DR}. By Corollary B.4 of [12], it suffices to verify that the sequence

$$n \mapsto (g'_h(n), id, g_0(n)\tilde{g}_h(n)) \tag{7.15}$$

is a polynomial sequence with respect to the filtration G'_{DR}. By Lemma 2.20, we may write $g'_h(n) = \prod_{0 \leq |m| \leq s} g'_{h,m}{}^{\binom{n}{m}}$ for some $g'_{h,m} \in G_{[|m|,0]}$. By (7.10) and Lemma 2.22, we have

$$g'_{h,m} \equiv \mathrm{Taylor}_m(g'_h) \equiv \phi\Big(\prod_{k=1}^{D^{\mathrm{cor}}_{|m|}} e_{m,k}^{\xi_{m,k}} \prod_{k=D^{\mathrm{cor}}_{|m|}+1}^{D_{|m|}} e_{m,k}^{\xi_{h,m,k}}\Big) \mod G_{[|m|,2]}.$$

By the construction of the filtration G'_{DR}, this implies that

$$\Big(g'_{h,m}, id, \prod_{k=1}^{D^{\mathrm{cor}}_{|m|}} e_{m,k}^{\xi_{m,k}} \prod_{k=D^{\mathrm{cor}}_{|m|}+1}^{D_{|m|}} e_{m,k}^{\xi_{h,m,k}} \mod G^*\Big) \in G'_{[|m|,1]}.$$

Applying Corollary B.4 of [12], we have that the sequence

$$n \mapsto \Big(g'_h(n), id, \prod_{0 \leq |m| \leq s} \Big(\prod_{k=1}^{D^{\mathrm{cor}}_{|m|}} e_{m,k}^{\xi_{m,k}} \prod_{k=D^{\mathrm{cor}}_{|m|}+1}^{D_{|m|}} e_{m,k}^{\xi_{h,m,k}}\Big)^{\binom{n}{m}} \mod G^*\Big)$$

is a polynomial with respect to the filtration G'_{DR}. By the Baker-Campbell-Hausdorff formula, (7.11) and (7.12), we have that

$$n \mapsto \prod_{0 \leq |m| \leq s} \Big(\prod_{k=1}^{D^{\mathrm{cor}}_{|m|}} e_{m,k}^{\xi_{m,k}} \prod_{k=D^{\mathrm{cor}}_{|m|}+1}^{D_{|m|}} e_{m,k}^{\xi_{h,m,k}}\Big)^{\binom{n}{m}} \mod G^*$$

differs from the sequence $n \mapsto g_0(n)\tilde{g}_h(n)$ by a sequence that is polynomial in the shifted filtration $(\tilde{G}_{[d,r+1]})_{[d,r] \in \mathrm{DR}}$ (note that $\tilde{\phi}(e_{m,k}) = id_{\tilde{G}}$ if $D^{\mathrm{cor}}_{|m|} + 1 \leq k \leq D^{\mathrm{cor}}_{|m|} + D^{\mathrm{ind}}_{|m|}$). Therefore, we have that (7.15) is a polynomial sequence with respect to the filtration G'_{DR}.

On the other hand, it is obvious that F'_h is invariant with respect to the action of the central group

$$G'_{[s,r_*]} = \{(\phi(g), id, \tilde{\phi}(g)) \colon g \in G^{\vec{D},d}_{[s,r_*]}\}.$$

We may then quotient G' by $G'_{[s,r_*]}$ to get a representation of (7.14) as a nilsequence of degree-rank $< [s, r_*]$, whose complexity and dimension are clearly $O_{C,d,D,\epsilon}(1)$. In other words, $\chi'_h \otimes \overline{\tilde{\chi}}_h \in \mathrm{Nil}^{<[s,r_*];O_{C,d,D,\epsilon}(1),O_{C,d,D,\epsilon}(1)}(\tau(V(M)^h) + p\mathbb{Z}^d)$.

Finally, since

$$\chi'_h \otimes (\overline{\tilde{\chi}}_h \otimes \tilde{\chi}_h) = (\chi'_h \otimes \overline{\tilde{\chi}}_h) \otimes \tilde{\chi}_h,$$

using the fact that 1 is an $O_{C,d,D,\epsilon}(1)$-complexity linear combination of the components of $\overline{\tilde{\chi}}_h \otimes \tilde{\chi}_h$, we deduce that χ'_h is an $O_{C,d,D,\epsilon}(1)$-complexity linear combination of the components $\tilde{\chi}_h \otimes \psi'_h$ for some $\psi'_h \in \mathrm{Nil}^{<[s,r_*];O_{C,d,D,\epsilon}(1),O_{C,d,D,\epsilon}(1)}(\tau(V(M)^h) + p\mathbb{Z}^d)$. ☐

Step 4. We now realize $(h, n) \mapsto \tilde{\chi}_h(n)$ as a nilcharacter of multi-degree $(1, s)$. The construction is similar to Section 12 of [12]. Since it is convenience for us to treat each term $\{\alpha_{m,k,i} \cdot \tau(h)\}\beta_{m,k,i}$ in the summation of $\xi_{h,m,k}$ separately, a difference between of our construction and the one in [12] is that instead of taking the skew product of G with $\tilde{G}_{\mathrm{ped}}$ (the "pedal" component of $\tilde{G}$), we take the skew product of $\tilde{G}$ with $\tilde{G}^L_{\mathrm{ped}}$ for some large power L. To be more precise, let $\tilde{G}_{\mathrm{ped}}$ be the subgroup of $\tilde{G}$ generated by any iterated commutators of $\tilde{e}_{m_1,k_1}, \dots, \tilde{e}_{m_\ell,k_\ell}$ for all $\ell \geq 1$, $1 \leq |m_i| \leq s$, $1 \leq i \leq \ell$ such that there exists $1 \leq i_0 \leq \ell$ with $k_{i_0} \geq D^{\mathrm{cor}}_{|m_{i_0}|} + D^{\mathrm{ind}}_{|m_{i_0}|} + 1$ and $k_i \leq D^{\mathrm{cor}}_{|m_i|}$ for all $i \neq i_0$. Let I denote the collection of all $(m, k, i) \in \mathbb{N}^d \times \mathbb{N} \times \mathbb{N}$ with $1 \leq |m| \leq s$, $D^{\mathrm{cor}}_{|m|} + D^{\mathrm{ind}}_{|m|} + 1 \leq k \leq D_{|m|}$ and $1 \leq i \leq K$. Using the identities

$$z^{-1}[x, y]z = [z^{-1}xz, z^{-1}yz] \text{ and } z^{-1}xz = x[x, z],$$

it is not hard to see that $\tilde{G}_{\mathrm{ped}}$ is a rational abelian normal subgroup of $\tilde{G}$. Therefore, $\tilde{G}$ acts on $\tilde{G}_{\mathrm{ped}}$ by conjugation, leading to the semidirect product on $\tilde{G} \ltimes \tilde{G}^{|I|}_{\mathrm{ped}}$ given by

$$(g, (g_{m,k,i})_{(m,k,i)\in I}) * (g', (g'_{m,k,i})_{(m,k,i)\in I}) := (gg', (g^{g'}_{m,k,i}g'_{m,k,i})_{(m,k,i)\in I}),$$

where $a^b := b^{-1}ab$.

Let R be the commutative ring of tuples $t = (t_{m,k,i})_{(m,k,i)\in I}$ with $t_{m,k,i} \in \mathbb{R}$, which we endow with the pointwise product and addition. Define an action ρ of R (viewed now as an additive group) on $\tilde{G} \ltimes \tilde{G}^{|I|}_{\mathrm{ped}}$ by

$$\rho(t)(g, (g_{m,k,i})_{(m,k,i)\in I}) := \Big(g \prod_{(m,k,i)\in I} g^{t_{m,k,i}}_{m,k,i}, (g_{m,k,i})_{(m,k,i)\in I}\Big).$$

It is not hard to see that this is a well-defined action. We can then define the semi-direct product $G' := R \ltimes_\rho (\tilde{G} \ltimes \tilde{G}^{|I|}_{\mathrm{ped}})$ by setting

$$\begin{aligned}&(t, (g, (g_{m,k,i})_{(m,k,i)\in I})) *' (t', (g', (g'_{m,k,i})_{(m,k,i)\in I}))\\ =& (t + t', (\rho(t')(g, (g_{m,k,i})_{(m,k,i)\in I})) * (g', (g'_{m,k,i})_{(m,k,i)\in I})).\end{aligned}$$

This is a Lie group. We can endow G' with an $\mathbb{N}^2$-filtration $(G'_{(s_1,s_2)})_{(s_1,s_2)\in\mathbb{N}^2}$ as follows:

- If $s_1 > 1$, then $G'_{(s_1,s_2)} = \{id\}$.
- If $s_1 = 1$ and $s_2 > 0$, then $G'_{(1,s_2)}$ consists of the elements $(0, (g, id^{|I|}))$ with $g \in \tilde{G}_{s_2} \cap \tilde{G}_{\mathrm{ped}}$.

- If $s_1 = 1$ and $s_2 = 0$, then $G'_{(1,0)}$ consists of the elements $(t, (g, id^{|I|}))$ with $t \in R$ and $g \in \tilde{G}_{\text{ped}}$.
- If $s_1 = 0$ and $s_2 > 0$, then $G'_{(0,s_2)}$ consists of the elements $(0, (g, (g_{m,k,i})_{(m,k,i)\in I}))$ with $g \in \tilde{G}_{s_2}$ and $g_{m,k,i} \in \tilde{G}_{s_2} \cap \tilde{G}_{\text{ped}}$.
- If $s_1 = s_2 = 0$, then $G'_{(0,0)} = G'$.

On can verify that this is indeed an $\mathbb{N}^2$-filtration of degree $\leq (1, s)$. Let Γ' be the subgroup of $\tilde{G}$ consisting of tuples $(t, (g, (g_{m,k,i})_{(m,k,i)\in I}))$ with $g \in \tilde{\Gamma}, g_{m,k} \in \tilde{G}_{\text{ped}} \cap \tilde{\Gamma}$ and with all the coefficients of t being integers. It is not hard to see that Γ' is a cocompact subgroup of G', and that the above $\mathbb{N}^2$-filtration of G' is rational with respect to Γ'. Therefore, G'/Γ' is a nilmanifold, and is clearly of complexity $O_{C,d,D,\epsilon}(1)$.

For $(m, k, i) \in I$, let $g_{m,k,i} \in \text{poly}(\mathbb{Z}^d \to \tilde{G}_{\mathbb{N}})$ be given by

$$g_{m,k,i}(n) := \phi(e_{m,k})^{\beta_{m,k,i}\binom{n}{m}}.$$

Then

$$\tilde{g}_h(n) = \prod_{(m,k,i)\in I} g_{m,k,i}(n)^{\{\alpha_{m,k,i}\cdot\tau(h)\}}. \tag{7.16}$$

Consider the map $f : (\mathbb{Z}^d)^2 \to G'$ given by

$$f(h, n) := (\mathbf{0}, (g_0(n), (g_{m,k,i}(n))_{(m,k,i)\in I})) *' (\alpha h, (id, id^{|I|})),$$

where

$$\alpha h := (\alpha_{m,k,i} \cdot h)_{(m,k,i)\in I}.$$

It is not hard to check that $f \in \text{poly}((\mathbb{Z}^d)^2 \to (G')_{\mathbb{N}^2})$ and that

$$f'(h, n)\Gamma' = \Big(\{\alpha h\}, \Big(g_0(n) \prod_{(m,k,i)\in I} g_{m,k,i}(n)^{\{\alpha_{m,k,i}h\}}, (g_{m,k,i}(n))_{(m,k,i)\in I}\Big)\Big)\Gamma'. \tag{7.17}$$

Recall that for all $h \in H'$, each component $\{\alpha_{m,k,i}h\}$ of $\{\alpha h\}$ lies in an interval $I_{m,k,i}$ of length at most $1/10$. Let $2I_{m,k,i}$ be the interval of twice the length and with the same center as $I_{m,k,i}$, and let $\psi_{m,k,i} : \mathbb{R} \to \mathbb{R}$ be a smooth cutoff function supported on $2I_{m,k,i}$ and taking value 1 on $I_{m,k,i}$. We then define a function $F' : G'/\Gamma' \to \mathbb{C}^{O_{C,d,D,\epsilon}(1)}$ by setting

$$F'((t, (g, (g_{m,k,i})_{(m,k,i)\in I}))\Gamma') := \Big(\prod_{(m,k,i)\in I} \psi_{m,k,i}(t_{m,k,i})\Big)\tilde{F}(g\Gamma) \tag{7.18}$$

for all $(g, (g_{m,k,i})_{(m,k,i)\in I}) \in G \ltimes \tilde{G}^{|I|}_{\text{ped}}$ and $t = (t_{m,k,i})_{(m,k,i)\in I}$ with $t_{m,k,i} \in 2I_{m,k,i}$, and set F' to be zero whenever such an representation of $(t, (g, (g_{m,k,i})_{(m,k,i)\in I}))$ does not exist. One can easily verify that F' is a well defined Lipschitz function with Lipschitz norm bounded by $O_{C,d,D,\epsilon}(1)$. Since $\tilde{F}$ has vertical frequency $\tilde{\eta}$, F' has vertical frequency $\eta' : G'_{(1,s)} \to \mathbb{R}$ defined by

$$\eta'(0, (g, id^{|I|})) := \tilde{\eta}(g)$$

for all $g \in \tilde{G}_s$, which is of complexity $O_{C,d,D,\epsilon}(1)$. Combining (7.13), (7.16), (7.17) and (7.18), we have that

$$\tilde{\chi}_h(n) = F'(f'(\tau(h), n)\Gamma') \tag{7.19}$$

for all $h \in H'$ and $n \in \mathbb{Z}^d$.

Step 5. Unfortunately f' is not necessarily p-periodic and F' does not take values in $\mathbb{S}^{O_{C,d,D,\epsilon}(1)}$. To complete the proof of Proposition 7.1, we need to substitute expression (7.19) with a p-periodic nilcharacter. We use the approximate results obtained in Appendix B. By (7.1), (7.19), Lemma 7.2, and the Pigeonhole Principle, for all $h \in H'$, there exists a scalar-valued nilsequence $\alpha_h \in \text{Nil}^{<[s,r_*];O_{C,d,D,\epsilon}(1),1}(\mathbb{F}_p^d)$ such that

$$\mathbb{E}_{h\in H'}\Big|\mathbb{E}_{n\in V(M)^h} f(n+h)\overline{f}(n)\chi_0(h,n)\otimes F'(f'(\tau(h),\tau(n))\Gamma')\alpha_h(n)\psi_h(n)\Big| \gg_{C,d,D,\epsilon} 1.$$

Since $F'(f'\circ\tau(\cdot)\Gamma') \in \text{Nil}^{(1,s);O_{C,d,D,\epsilon}(1),O_{C,d,D,\epsilon}(1)}((\mathbb{F}_p^d)^2)$, by Lemma B.5 and the Pigeonhole Principle, there exist $\chi' \in \Xi_p^{(1,s);O_{C,d,D,\epsilon}(1),O_{C,d,D,\epsilon}(1)}((\mathbb{F}_p^d)^2)$ and $\alpha'_h \in \text{Nil}^{<[s,r_*];O_{C,d,D,\epsilon}(1),1}(\mathbb{F}_p^d)$ for each $h \in H'$ such that

$$\mathbb{E}_{h\in H'}\Big|\mathbb{E}_{n\in V(M)^h} f(n+h)\overline{f}(n)\chi_0(h,n)\otimes\chi(h,n)\alpha'_h(n)\psi_h(n)\Big| \gg_{C,d,D,\epsilon} 1.$$

By the Pigeonhole Principle, there exists a subset H'' of H' with $|H''| \gg_{C,d,D,\epsilon} p^d$ such that

$$\Big|\mathbb{E}_{n\in V(M)^h} f(n+h)\overline{f}(n)\chi_0(h,n)\otimes\chi(h,n)\alpha'_h(n)\psi_h(n)\Big| \gg_{C,d,D,\epsilon} 1$$

for all $h \in H''$. By Corollary B.6 and the Pigeonhole Principle, there exists χ'_h in the set $\Xi_p^{<[s,r_*];O_{C,d,D,\epsilon}(1),O_{C,d,D,\epsilon}(1)}((\mathbb{F}_p^d)^2)$ for all $h \in H''$ such that

$$\Big|\mathbb{E}_{n\in V(M)^h} f(n+h)\overline{f}(n)\chi_0(h,n)\otimes\chi'(h,n)\otimes\chi'_h(n)\psi_h(n)\Big| \gg_{C,d,D,\epsilon} 1$$

for all $h \in H''$. This completes the proof of Proposition 7.1 by setting $\chi := \chi_0 \otimes \chi'$.

8. The symmetric argument for $s = 1$

In Sections 8 and 9, we complete the proof of SGI($s+1$). Recall that we assume that SGI(s) holds for some $s \geq 1$ and we wish to prove SGI($s+1$). By Lemma A.8 and a change of variables, it is not hard to check that we only need to prove SGI($s+1$) for the case when M is a pure quadratic form. So we assume that M is pure throughout Sections 8 and 9. By Example 6.11 of [12], every connected, simply-connected nilpotent Lie group G of degree s induces a filtration of degree-rank $[s,s]$. So SGI(s) and (3.2) implies the initial hypothesis of Proposition 7.1 for the case $r_* = s$ holds. Since $\Xi_{\text{DR}}^{<[s,1]}(\mathbb{F}_p^d) \subseteq \text{Nil}^{s-1}(\mathbb{F}_p^d)$ and $d \geq N(s)$, we may then use Proposition 7.1 inductively combined with Corollary B.3 and the Pigeonhole Principle to conclude that there exist $H \subseteq \mathbb{F}_p^d$ with $|H| \gg_{d,\epsilon} p^d$, some $\chi \in \Xi_p^{(1,s);O_{d,\epsilon}(1),O_{d,\epsilon}(1)}((\mathbb{F}_p^d)^2)$, and some $\psi_h \in \text{Nil}_p^{s-1;O_{d,\epsilon}(1),1}(\mathbb{F}_p^d)$ for all $h \in \mathbb{F}_p^d$ such that

$$|\mathbb{E}_{n\in V(M)^h}\Delta_h f(n)\chi(h,n)\psi_h(n)| \gg_{d,\epsilon} 1 \tag{8.1}$$

for all $h \in H$. Therefore,

$$\mathbb{E}_{h\in\mathbb{F}_p^d}|\mathbb{E}_{n\in V(M)^h}\Delta_h f(n)\chi(h,n)\psi_h(n)| \gg_{d,\epsilon} 1. \tag{8.2}$$

By Lemma A.5, we may assume without loss of generality that all the elements in H are M-non-isotropic.

By Lemma C.4 (vii) and Theorem C.9, there exists $\tilde{\chi} \in \Xi_p^{(1,\dots,1);O_{d,\epsilon}(1),O_{d,\epsilon}(1)}((\mathbb{F}_p^d)^{s+1})$ (with 1 repeated $s+1$ times) which is symmetric in the last s variables such that

$$\chi(h,n) \sim_{O_{d,\epsilon}(1)} \tilde{\chi}(h,n,\dots,n)^{\otimes(s+1)} \mod \Xi_p^{s+1}((\mathbb{F}_p^d)^2). \tag{8.3}$$

Our goal is to first use (8.2) to deduce some symmetric property for $\tilde{\chi}$. Then we use it to show that in (8.2), up to a nilsequence of lower degree, the term $\chi(h,n)$ can be "replaced" by $\tilde{\chi}(n+h,\dots,n+h)\otimes\overline{\tilde{\chi}}(n,\dots,n)$, which can be then absorbed by the term $\Delta_h f$. Finally, we use the Cauchy-Schwartz inequality to complete the proof of SGI$(s+1)$.

The proof for the case $s=1$ and the case $s\geq 2$ are slightly different. We prove SGI(2) in Section 8 and SGI$(s+1)$, $s\geq 2$ in Section 9.

8.1. **A symmetric property for the case** $s=1$**.** We start with the case $s=1$. We first use (8.1) to deduce some symmetric property on $\tilde{\chi}$. Since $s=1$, we may set $\psi_h(n)\equiv 1$ in (8.2). Denote

$$\Lambda_1 := \{(h,n,m)\in(\mathbb{F}_p^d)^3\colon n,m,n+h,m+h\in V(M)\}.$$

Then Λ_1 is a consistent M-set of total co-dimension 4. Since $d\geq 9$, by the (vector-valued) Cauchy-Schwartz inequality and Theorem A.12, we deduce from (8.2) that

$$\begin{aligned}
&1 \ll_{d,\epsilon} \mathbb{E}_{h\in\mathbb{F}_p^d}|\mathbb{E}_{n\in V(M)^h}\Delta_h f(n)\chi(h,n)\psi_h(n)|^2\\
&= |\mathbb{E}_{h\in\mathbb{F}_p^d}\mathbb{E}_{n,m\in V(M)^h} f(n+h)\overline{f}(m+h)\overline{f}(n)f(m)\chi(h,n)\otimes\overline{\chi}(h,m)| \qquad (8.4)\\
&= |\mathbb{E}_{(h,m,n)\in\Lambda_1} f(n+h)\overline{f}(m+h)\overline{f}(n)f(m)\chi(h,n)\otimes\overline{\chi}(h,m)| + O(p^{-1/2}).
\end{aligned}$$

Denote

$$\Lambda_2 := \{(z,n,m)\in(\mathbb{F}_p^d)^3\colon n,m,z-n,z-m\in V(M)\},$$

which is a consistent M-set of total co-dimension 4. Let Λ_2' denote the set of $(z,m)\in(\mathbb{F}_p^d)^2$ with $m,z-m\in V(M)$. For $z\in\mathbb{F}_p^d$, let $\Lambda_2(z)$ denote the set of $n\in\mathbb{F}_p^d$ with $n,z-n\in V(M)$. Replacing h with $z=n+m+h$ and using Theorem A.12, we have that the right hand side of (8.4) is bounded by

(8.5)

$$\begin{aligned}
&|\mathbb{E}_{(z,n,m)\in\Lambda_2} f(z-m)\overline{f}(z-n)\overline{f}(n)f(m)\chi(z-n-m,n)\otimes\overline{\chi}(z-n-m,m)| + O(p^{-1/2})\\
&= |\mathbb{E}_{(z,m)\in\Lambda_2'} f(z-m)f(m)\mathbb{E}_{n\in\Lambda_2(z)}\overline{f}(z-n)\overline{f}(n)\chi(z-n-m,n)\otimes\overline{\chi}(z-n-m,m)| + O(p^{-1/2})\\
&\leq \mathbb{E}_{(z,m)\in\Lambda_2'}|\mathbb{E}_{n\in\Lambda_2(z)}\overline{f}(z-n)\overline{f}(n)\chi(z-n-m,n)\otimes\overline{\chi}(z-n-m,m)| + O(p^{-1/2}).
\end{aligned}$$

Denote

$$\Lambda_3 := \{(z,m,n_1,n_2)\in(\mathbb{F}_p^d)^4\colon m,n_1,n_2,z-m,z-n_1,z-n_2\in V(M)\},$$

which is a consistent M-set of total co-dimension 6. Using a similar method, the square of the right hand side of (8.5) is bounded by $O_{d,\epsilon}(1)$ times

(8.6)
$$\begin{aligned}
&\mathbb{E}_{(z,m)\in\Lambda'_2}|\mathbb{E}_{n\in\Lambda_2(z)}\overline{f}(z-n)\overline{f}(n)\chi(z-n-m,n)\otimes\overline{\chi}(z-n-m,m)|^2+O(p^{-1/2})\\
=&\Big|\mathbb{E}_{(z,m)\in\Lambda'_2}\mathbb{E}_{n_1,n_2\in\Lambda_2(z)}\\
&\prod_{i=1}^{2}C^{i+1}(f(z-n_i)f(n_i))\bigotimes_{i=1}^{2}C^{i}(\chi(z-n_i-m,n_i)\otimes\overline{\chi}(z-n_i-m,m))\Big|+O(p^{-1/2})\\
=&\Big|\mathbb{E}_{(z,m,n_1,n_2)\in\Lambda_3}\\
&\prod_{i=1}^{2}C^{i+1}(f(z-n_i)f(n_i))\bigotimes_{i=1}^{2}C^{i}(\chi(z-n_i-m,n_i)\otimes\overline{\chi}(z-n_i-m,m))\Big|+O(p^{-1/2})\\
\leq&\,\mathbb{E}_{(z,n_1,n_2)\in\Lambda_2}\Big|\mathbb{E}_{m\in\Lambda_2(z)}\bigotimes_{i=1}^{2}C^{i}(\chi(z-n_i-m,n_i)\otimes\overline{\chi}(z-n_i-m,m))\Big|+O(p^{-1/2}).
\end{aligned}$$

Denote

$$\Lambda_4:=\{(z,n_1,n_2,m_1,m_2)\in(\mathbb{F}_p^d)^5\colon n_1,n_2,m_1,m_2,z-n_1,z-n_2,z-m_1,z-m_2\in V(M)\}.$$

which is a consistent M-set of total co-dimension 8. Similarly, since $d\geq 17$, the square of the right hand side of (8.6) is bounded by $O_{d,\epsilon}(1)$ times

$$\begin{aligned}
&\mathbb{E}_{(z,n_1,n_2)\in\Lambda_2}\Big|\mathbb{E}_{m\in\Lambda_2(z)}\bigotimes_{i=1}^{2}C^{i}(\chi(z-n_i-m,n_i)\otimes\overline{\chi}(z-n_i-m,m))\Big|^2+O(p^{-1/2})\\
=&\Big|\mathbb{E}_{(z,n_1,n_2)\in\Lambda_2}\mathbb{E}_{m_1,m_2\in\Lambda_2(z)}\bigotimes_{i,j\in\{1,2\}}C^{i+j}(\chi(z-n_i-m_j,n_i)\otimes\overline{\chi}(z-n_i-m_j,m_j))\Big|+O(p^{-1/2})\\
=&\Big|\mathbb{E}_{(z,n_1,n_2,m_1,m_2)\in\Lambda_4}\bigotimes_{i,j\in\{1,2\}}C^{i+j}(\chi(z-n_i-m_j,n_i)\otimes\overline{\chi}(z-n_i-m_j,m_j))\Big|+O(p^{-1/2}).
\end{aligned}$$

In conclusion, we have that

$$\Big|\mathbb{E}_{(z,n_1,n_2,m_1,m_2)\in\Lambda_4}\bigotimes_{i,j\in\{1,2\}}C^{i+j}(\chi(z-n_i-m_j,n_i)\otimes\overline{\chi}(z-n_i-m_j,m_j))\Big|\gg_{d,\epsilon}1 \tag{8.7}$$

if $p\gg_{d,\epsilon}1$. By Lemma B.1, the map

$$(z,n_1,n_2,m_1,m_2)\mapsto\bigotimes_{i,j\in\{1,2\}}C^{i+j}(\chi(z-n_i-m_j,n_i)\otimes\overline{\chi}(z-n_i-m_j,m_j))$$

belongs to $\Xi_p^{2;O_{d,\epsilon}(1),O_{d,\epsilon}(1)}((\mathbb{F}_p^d)^5)$. On the other hand, Note that Λ_4 is the set Ω_1 in Example B.4 of [19], which is a nice and consistent M-set of total co-dimension 8. Since $d\geq 33$, it follows from Lemma C.7 and (8.7) that

$$\bigotimes_{i,j\in\{1,2\}}C^{i+j}(\chi(z-n_i-m_j,n_i)\otimes\overline{\chi}(z-n_i-m_j,m_j))\sim_{O_{d,\epsilon}(1)}1\mod \Xi_p^2(\Lambda_4). \tag{8.8}$$

By Lemmas C.2, C.4 (iv) and C.5, it is not hard to simplify (8.8) as

$$\chi(n_1-n_2,m_1-m_2)\sim_{O_{d,\epsilon}(1)}\chi(m_1-m_2,n_1-n_2) \mod \Xi_p^2(\Lambda_4). \tag{8.9}$$

In order to better apply (8.9) to study (8.4), we need to do a change of variable with $h_1=m_1-m_2$ and $h_2=n_1-n_2$, and add an extra variable x. let $L\colon (\mathbb{F}_p^d)^6\to(\mathbb{F}_p^d)^5$ be the linear transformation given by

$$L(x,y_1,y_2,y_3,h_1,h_2):=(y_1,y_2+h_1,y_2,y_3+h_2,y_3).$$

Note that $L^{-1}(\Lambda_4)$ consists of $(x,y_1,y_2,y_3,h_1,h_2)\in(\mathbb{F}_p^d)^6$ with

$$y_2+h_1,y_2,y_3+h_2,y_3,y_1-y_2-h_1,y_1-y_2,y_1-y_3-h_2,y_1-y_3\in V(M).$$

By (8.9) and Lemma C.4 (vi), we have that

$$\chi(h_1,h_2)\sim_{O_{d,\epsilon}(1)}\chi(h_2,h_1) \mod \Xi_p^2(L^{-1}(\Lambda_4)) \tag{8.10}$$

(where the sequences are viewed as in the variables (x,y_1,y_2,y_3,h_1,h_2) in $L^{-1}(\Lambda_4)$). By (8.3), (8.10), Lemmas C.2 and C.6, we have that

$$\tilde{\chi}(h_1,h_2)\sim_{O_{d,\epsilon}(1)}\tilde{\chi}(h_2,h_1) \mod \Xi_p^2(L^{-1}(\Lambda_4)). \tag{8.11}$$

Let Λ_5 be the set of $(x,y_1,y_2,y_3,h_1,h_2)\in L^{-1}(\Lambda_4)$ such that $(x,h_1,h_2)\in\Box_2(V(M))$. By (8.11) and Lemma C.4 (iv), we have that

$$\tilde{\chi}(h_1,h_2)\sim_{O_{d,\epsilon}(1)}\tilde{\chi}(h_2,h_1) \mod \Xi_p^2(\Lambda_5). \tag{8.12}$$

By the description of $L^{-1}(\Lambda_4)$, it is not hard to compute that Λ_5 is a consistent M-set of total co-dimension 12.

Remark 8.1. In Section 13 of [12], the authors only kept track of the domain of the parameters h_1,h_2, and treated all other parameters as constants. In our setting, if we treated x,y_1,y_2,y_3 as constants, then there will be a significant loss of information in expressions such as (8.12). This is why in our method we must keep track of the domain of all the parameters x,y_1,y_2,y_3,h_1,h_2 (which makes the computation more complicated). The same remark applies to the proof for the case $s\geq 2$ in Section 9.1.

8.2. **Completion of the proof of SGI(2).** In Section 13 of [12], the authors deduced a conclusion similar to (8.12), and then used it to conclude that $\chi(h,n)$ can be written as $\tilde{\chi}(n+h,\dots,n+h)\otimes\overline{\tilde{\chi}}(n,\dots,n)$ for some nilcharacter Θ of degree $\leq s+1$. However, this method does not apply to our case. This is because that Λ_5 is a sparse subset of $(\mathbb{F}_p^d)^6$, and that equation (8.12) does not provide us any information on whether $\tilde{\chi}(h_1,h_2)$ is equivalent to $\tilde{\chi}(h_2,h_1)$ outside Λ_5. So we need to use a more sophisticated method.

Convention 8.2. Throughout Sections 8 and 9 only, $n\in\mathbb{F}_p^d$ is treated as a special variable. If $F(n,m_1,\dots,m_k)$ is a map with one of the variables being n, then we write

$$\Delta_h F(n,m_1,\dots,m_k):=F(n+h,m_1,\dots,m_k)\otimes\overline{F}(n,m_1,\dots,m_k)$$

for all $h\in\mathbb{F}_p^d$ (i.e. the difference is taken with respect to the special variable n). For example, the expression $\Delta_h\chi(n+m,m+y,n+y)$ is understood as $\chi(n+h+m,m+y,n+h+y)\otimes\overline{\chi}(n+m,m+y,n+y)$.

Step 1: reformulate (8.1). Under a change of variable, we may rewrite the right hand side of (8.4) as

$$|\mathbb{E}_{(n,h_1,h_2)\in\Box_2(V(M))}\Delta_{h_2}\Delta_{h_1}\tilde{f}(n)\otimes\beta(h_1,h_2,n)|+O(p^{-1/2}), \tag{8.13}$$

where $\tilde{f}(n) := f(n)\tilde{\chi}(n,n)$ and

$$\beta(h_1,h_2,n) := \Delta_{h_2}(\chi(h_1,n)\otimes\overline{\tilde{\chi}}(n+h_1,n+h_1)\otimes\tilde{\chi}(n,n)).$$

We may write $\Lambda_5 = V(\mathcal{J})$ for some independent $(M,6)$-family $\mathcal{J}\subseteq\mathbb{F}_p[n,y_1,y_2,y_3,h_1,h_2]$ of total dimension 12. Let $(\mathcal{J}',\mathcal{J}'')$ be an $\{n,h_1,h_2\}$-decomposition of $\mathcal{J}$. For convenience denote $\mathbf{h} := (h_1,h_2)$ and $\mathbf{y} := (y_1,y_2,y_3)$. It is not hard to compute that $(n,\mathbf{y},\mathbf{h})\in V(\mathcal{J}')$ if and only if $(n,\mathbf{h})\in\Box_2(V(M))$. For $(n,\mathbf{h})\in(\mathbb{F}_p^d)^3$, let $\Lambda_5(n,\mathbf{h})$ denote the set of $\mathbf{y}\in(\mathbb{F}_p^d)^3$ such that $(n,\mathbf{y},\mathbf{h})\in V(\mathcal{J}'')$. Since Λ_5 is a consistent M-set of total co-dimension 12, by Theorem A.12, (8.4) and (8.13),

$$\begin{aligned}1 &\ll_{d,\epsilon} |\mathbb{E}_{(n,h_1,h_2)\in\Box_2(V(M))}\Delta_{h_2}\Delta_{h_1}\tilde{f}(n)\otimes\beta(h_1,h_2,n)|+O(p^{-1/2})\\ &=|\mathbb{E}_{(n,\mathbf{h})\in\Box_2(V(M))}\mathbb{E}_{\mathbf{y}\in\Lambda_5(n,\mathbf{h})}\Delta_{h_2}\Delta_{h_1}\tilde{f}(n)\otimes\beta(\mathbf{h},n)|+O(p^{-1/2})\\ &=|\mathbb{E}_{(n,\mathbf{y},\mathbf{h})\in\Lambda_5}\Delta_{h_2}\Delta_{h_1}\tilde{f}(n)\otimes\beta(\mathbf{h},n)|+O(p^{-1/2}),\end{aligned} \tag{8.14}$$

Step 2: use (8.3) to convert β into a lower degree nilsequence. Using (8.3), (8.12) and Lemma C.5, it is not hard to compute that

$$\beta(h_1,h_2,n)\sim_{O_{d,\epsilon}(1)}\tilde{\chi}(h_1,h_2)\otimes\overline{\tilde{\chi}}(h_2,h_1)\sim_{O_{d,\epsilon}(1)} 1 \mod \Xi_p^2(\Lambda_5).$$

Therefore, $\beta(h_1,h_2,n)\in\text{Nil}^{1;O_{d,\epsilon}(1)}(\Lambda_5)$. Then $\beta(h_1,h_2,n)\in\text{Nil}^{J;O_{d,\epsilon}(1)}(\Lambda_5)$, where

$$J := \{(a,b_1,b_2,b_3,c_1,c_2)\in\mathbb{N}^5\colon a+b_1+b_2+b_3+c_1+c_2\leq 1\}.$$

For $1\leq i\leq 2$, let

$$J_i := \{(a,b_1,b_2,b_3,c_1,c_2)\in J\colon c_i=0\}.$$

Then $J=J_1\cup J_2$. Note that any sequence in $\text{Nil}^{J_i}(\Lambda_5)$ is independent of h_i. By (8.14), Lemma B.4 and the Pigeonhole Principle, $for 1\leq i\leq 2$, there exists a scalar valued sequence $\psi_i(n,\mathbf{y},\mathbf{h})$ bounded by 1 which is independent of h_i such that

$$\Big|\mathbb{E}_{(n,\mathbf{y},\mathbf{h})\in\Lambda_5}\Delta_{h_2}\Delta_{h_1}\tilde{f}(n)\prod_{i=1}^{2}\psi_i(n,\mathbf{y},\mathbf{h})\Big|\gg_{d,\epsilon}1$$

if $p\gg_{d,\epsilon}1$. By the definition of $\tilde{f}$ and the Pigeonhole Principle, there exists $\phi'\in\text{Nil}_p^{2;O_{d,\epsilon}(1),1}(\mathbb{F}_p^d)$ such that writing $f'(n) := f(n)\phi(n)$, we have that

$$\Big|\mathbb{E}_{(n,\mathbf{y},\mathbf{h})\in\Lambda_5}\Delta_{h_2}\Delta_{h_1}f'(n)\prod_{i=1}^{2}\psi_i(n,\mathbf{y},\mathbf{h})\Big|\gg_{d,\epsilon}1. \tag{8.15}$$

Step 3: use the Cauchy-Schwartz inequality to remove ψ_i. For $r\leq 3$, we say that *Property-r* holds if there exist a scalar valued function $\psi_i(n,\mathbf{y},\mathbf{h},h'_1,\dots,h'_{r-1})$ bounded by

1 and independent of h_i for all $1 \leq i \leq r$ such that if $p \gg_{d,\epsilon} 1$, then

$$(8.16)\qquad \Big|\mathbb{E}_{(n,\mathbf{y},\mathbf{h},h'_1,\dots,h'_{r-1})\in\Lambda_{5,r}}\Delta_{h_2}\Delta_{h_1}f'(n)\prod_{i=r}^{2}\psi_i(n,\mathbf{y},\mathbf{h},h'_1,\dots,h'_{r-1})\Big| \gg_{d,\epsilon} 1,$$

where $\Lambda_{5,r}$ is the set of all $(n,\mathbf{y},\mathbf{h},h'_1,\dots,h'_{r-1}) \in (\mathbb{F}_p^d)^{r+5}$ such that

$$(n-h'_1-\dots-h'_{r-1},\mathbf{y},h'_1+\epsilon_1h_1,\dots,h'_{r-1}+\epsilon_{r-1}h_{r-1},h_r,\dots,h_2) \in \Lambda_5$$

for all $\epsilon_1,\dots,\epsilon_{r-1} \in \{0,1\}$ (for $r=3$, the term $\prod_{i=r}^{2}\psi_i(n,\mathbf{y},\mathbf{h},h'_1,\dots,h'_{r-1})$ is considered to be constant 1). Since M is pure, Λ_5 is a pure and consistent M-set. By Propositions A.10 and A.11, $\Lambda_{5,r}$ is a pure and consistent M-set of totally co-dimension at most 36.[14]

Clearly, Property-1 holds by (8.15). Now suppose that Property-r holds for some $1 \leq r \leq 2$. For convenience denote $\mathbf{x} := (n,\mathbf{y},h_1,\dots,h_{r-1},h_{r+1},\dots,h_2,h'_1,\dots,h'_{r-1})$, and let $(\mathbf{x};h)$ be the vector obtained by inserting h in the middle of h_{r-1} and h_{r+1}. Let $\tilde{\Lambda}_{5,r}$ denote the set of $(\mathbf{x},h'_r,h''_r) \in (\mathbb{F}_p^d)^{r+6}$ such that $(\mathbf{x};h'_r),(\mathbf{x};h'_r) \in \Lambda_{5,r}$. Since $\Lambda_{5,r}$ is a pure and consistent M-set, by Propositions A.10 and A.11, $\tilde{\Lambda}_{5,r}$ is a pure and consistent M-set of total co-dimension at most 45. We may then write $\Lambda_{5,r} = V(\mathcal{J})$ for some consistent $(M,r+6)$-family $\mathcal{J} \subseteq \mathbb{F}_p[n,y_1,y_2,y_3,h_1,h_2,h'_1,\dots,h'_{r-1}]$ of total dimension at most 45. Let $(\mathcal{J}',\mathcal{J}'')$ be a $\{n,y_1,y_2,y_3,h_1,\dots,h_{r-1},h_{r+1},\dots,h_2,h'_1,\dots,h'_{r-1}\}$-decomposition of $\mathcal{J}$. Let $\Lambda'_{5,r}$ be the set of $\mathbf{x} \in (\mathbb{F}_p^d)^{r+4}$ such that $(\mathbf{x};h_r) \in V(\mathcal{J}')$ for all $h_r \in \mathbb{F}_p^d$. For $\mathbf{x} \in (\mathbb{F}_p^d)^{r+4}$, let $\Lambda_{5,r}(\mathbf{x})$ denote the set of $h_r \in \mathbb{F}_p^d$ such that $(\mathbf{x};h_r) \in V(\mathcal{J}'')$. Then by pulling out the h_r-independent term

$$\Delta_{h_2}\dots\Delta_{h_{r+1}}\Delta_{h_{r-1}}\dots\Delta_{h_1}\overline{f}'(n)\psi_r(n,\mathbf{y},\mathbf{h},h'_1,\dots,h'_{r-1})$$

in (8.16) and then apply the Cauchy-Schwartz inequality and Theorem A.12, we have that

$$1 \ll_{d,\epsilon} \mathbb{E}_{\mathbf{x}\in\Lambda'_{5,r}}\Big|\mathbb{E}_{h_r\in\Lambda_{5,r}(\mathbf{x})}\Delta_{h_2}\dots\Delta_{h_{r+1}}\Delta_{h_{r-1}}\dots\Delta_{h_1}f'(n+h_r)\prod_{i=r+1}^{2}\psi_i(\mathbf{x};h_r)\Big|^2$$

$$= \Big|\mathbb{E}_{(\mathbf{x},h'_r,h''_r)\in\tilde{\Lambda}_{5,r}}\Delta_{h_2}\dots\Delta_{h_{r+1}}\Delta_{h_{r-1}}\dots\Delta_{h_1}(f'(n+h''_r)\overline{f}'(n+h'_r))\prod_{i=r+1}^{2}\psi_i(\mathbf{x};h''_r)\overline{\psi}_i(\mathbf{x};h'_r)\Big| + O(p^{-1/2})$$

$$= \Big|\mathbb{E}_{(\mathbf{x};h'_r),(\mathbf{x};h'_r+h_r)\in\Lambda_{5,r}}\Delta_{h_2}\dots\Delta_{h_{r+1}}\Delta_{h_r}\Delta_{h_{r-1}}\dots\Delta_{h_1}f'(n+h'_r)\prod_{i=r+1}^{2}\psi_i(\mathbf{x};h'_r+h_r)\overline{\psi}_i(\mathbf{x};h'_r)\Big| + O(p^{-1/2}).$$

For all $r+1 \leq i \leq 2$, since ψ_i is independent of h_i and bounded by 1, $\psi_i(\mathbf{x};h'_r+h_r)\overline{\psi}_i(\mathbf{x};h'_r)$ is also independent of h_i and bounded by 1. On the other hand, note that $(\mathbf{x};h'_r),(\mathbf{x};h'_r+h_r) \in \Lambda_{5,r}$ if and only if

$$(n-h'_1-\dots-h'_{r-1},\mathbf{y},h'_1+\epsilon_1h_1,\dots,h'_{r-1}+\epsilon_{r-1}h_{r-1},h'_r,h_{r+1},\dots,h_2) \in \Lambda_5$$

and

$$(n-h'_1-\dots-h'_{r-1},\mathbf{y},h'_1+\epsilon_1h_1,\dots,h'_{r-1}+\epsilon_{r-1}h_{r-1},h'_r+h_r,h_{r+1},\dots,h_2) \in \Lambda_5$$

[14]The total co-dimension of $\Lambda_{5,r}$ can be computed explicitly, but here we do not care about its precise value.

for all $\epsilon_1, \dots, \epsilon_{r-1} \in \{0,1\}$. Writing $n' = n + h'_r$, this is equivalent of saying that

$$(n' - h'_1 - \dots - h'_r, \mathbf{y}, h'_1 + \epsilon_1 h_1, \dots, h'_r + \epsilon_r h_r, h_{r+1}, \dots, h_2) \in \Lambda_5$$

for all $\epsilon_1, \dots, \epsilon_{r-1} \in \{0,1\}$, or equivalently, $(n', h_1, h_2, h'_1, \dots, h'_r) \in \Lambda_{5,r+1}$. So we have that Property-$(r+1)$ holds.

Inductively, we have that Property-3 holds. Since $\Lambda_{5,3}$ is a pure and consistent M-set of total co-dimension at most 45, we may write $\Lambda_{5,3} = V(\tilde{\mathcal{J}})$ for some consistent $(M,8)$-family $\tilde{\mathcal{J}} \subseteq \mathbb{F}_p[n, y_1, y_2, y_3, h_1, h_2, h'_1, h'_2]$ of total dimension at most 45. For convenience denote $\mathbf{h}' := (h'_1, h'_2)$. Let $(\tilde{\mathcal{J}}', \tilde{\mathcal{J}}'')$ be an $\{n, h_1, h_2\}$-decomposition of $\tilde{\mathcal{J}}$. Let $\Lambda''_{5,3}$ denote the set of $(n, \mathbf{h}) \in (\mathbb{F}_p^d)^3$ such that $(n, \mathbf{y}, \mathbf{h}, \mathbf{h}') \in V(\tilde{\mathcal{J}}')$ for all $(\mathbf{y}, \mathbf{h}') \in (\mathbb{F}_p^d)^5$. It is not hard to compute that $V(\tilde{\mathcal{J}}') = \Box_2(V(M))$. For $(n, \mathbf{h}) \in (\mathbb{F}_p^d)^3$, let $\Lambda''_{5,3}(n, \mathbf{h})$ denote the set of $(\mathbf{y}, \mathbf{h}') \in (\mathbb{F}_p^d)^5$ such that $(n, \mathbf{y}, \mathbf{h}, \mathbf{h}') \in V(\tilde{\mathcal{J}}'')$. It follows from Property-3 and Theorem A.12 that

$$1 \ll_{d,\epsilon} |\mathbb{E}_{(n,\mathbf{y},\mathbf{h},\mathbf{h}') \in \Lambda_{5,3}} \Delta_{h_2}\Delta_{h_1} f'(n)| = |\mathbb{E}_{(n,\mathbf{h}) \in \Box_2(V(M))} \Delta_{h_2}\Delta_{h_1} f'(n) \mathbb{E}_{(\mathbf{y},\mathbf{h}') \in \Lambda''_{5,3}(n,\mathbf{h})} 1| + O(p^{-1/2})$$
$$= |\mathbb{E}_{(n,\mathbf{h}) \in \Box_2(V(M))} \Delta_{h_2}\Delta_{h_1} (f\phi')(n)| + O(p^{-1/2}) = \|f\phi'\|^4_{U^2(V(M))} + O(p^{-1/2}).$$

Since SGI(1) holds by assumption, if $p \gg_{d,\epsilon} 1$, then there exists $\phi'' \in \mathrm{Nil}_p^{1;O_{d,\epsilon}(1),1}(\mathbb{F}_p^d)$ such that

$$|\mathbb{E}_{n \in V(M)} f(n)\phi'(n)\phi''(n)| \gg_{d,\epsilon} 1.$$

This completes the proof of SGI(2) since $\phi'\phi'' \in \mathrm{Nil}_p^{2;O_{d,\epsilon}(1),1}(\mathbb{F}_p^d)$.

9. The symmetric argument for $s \geq 2$

9.1. **A symmetric property for the case $s \geq 2$.** In this section, we prove SGI$(s+1)$ for $s \geq 2$. Again we first use (8.1) to deduce some symmetric property on $\tilde{\chi}$. By (8.1), Corollary B.6 and the Pigeonhole Principle, for all $h \in \mathbb{F}_p^d$, there exists $\varphi_h \in \Xi_p^{s-1;O_{d,\epsilon}(1),O_{d,\epsilon}(1)}(\mathbb{F}_p^d)$ such that

$$|\mathbb{E}_{n \in V(M)^h} \Delta_h f(n) \chi(h,n) \otimes \varphi_h(n)| \gg_{d,\epsilon} 1 \tag{9.1}$$

for all $h \in H$. For convenience denote $h_4 := h_1 + h_2 - h_3$. By Proposition 4.2, if $p \gg_{d,\epsilon} 1$ and $d \geq 9$, then the absolute value of the average of the sequence

$$(n, h_1, h_2, h_3) \mapsto \chi(h_1, n) \otimes \chi(h_2, n + h_3 - h_2) \otimes \overline{\chi}(h_3, n) \otimes \overline{\chi}(h_4, n + h_3 - h_2)$$
$$\otimes \varphi_{h_1}(n) \otimes \varphi_{h_2}(n + h_3 - h_2) \otimes \overline{\varphi}_{h_3}(n) \otimes \overline{\varphi}_{h_4}(n + h_3 - h_2)$$

along the set $\{(n, h_1, h_2, h_3) \in (\mathbb{F}_p^d)^4 \colon n \in V(M)^{h_1, h_3, h_3 - h_2}\}$ is $\gg_{d,\epsilon} 1$, where we set $\varphi_h \equiv 0$ for $h \notin H$.[15] By the change of variable $(h_1, h_2, h_3, h_4) = (h_0 + a, h_0 + b, h_0 + a + b, h_0)$, we have that

(9.2)
$$|\mathbb{E}_{h_0,a,b \in \mathbb{F}_p^d, n \in V(M)^{h_0+a, h_0+a+b, a}} \tau(h_0, a, b, n) \otimes \varphi_{h_0+a}(n) \otimes \varphi_{h_0+b}(n+a) \otimes \overline{\varphi}_{h_0+a+b}(n) \otimes \overline{\varphi}_{h_0}(n+a)|$$

[15] Note that φ_h is not a nilcharacter for $h \notin H$ since it does not have absolute value 1. But this does not affect the proof.

is $\gg_{d,\epsilon} 1$, where

$$\tau(h_0,a,b,n) := \chi(h_0+a,n)\otimes\chi(h_0+b,n+a)\otimes\overline{\chi}(h_0+a+b,n)\otimes\overline{\chi}(h_0,n+a). \tag{9.3}$$

Note that the square of (9.2) is bounded by $O_{d,\epsilon}(1)$ times

(9.4)

$$\begin{aligned}
&|\mathbb{E}_{h_0,a\in\mathbb{F}_p^d}\mathbb{E}_{n\in V(M)^{h_0+a,a}}\mathbb{E}_{b\in V(M(h_0+a+n+\cdot))}\tau(h_0,a,b,n)\otimes\varphi_{h_0+a}(n)\otimes\varphi_{h_0+b}(n+a)\\
&\quad\otimes\overline{\varphi}_{h_0+a+b}(n)\otimes\overline{\varphi}_{h_0}(n+a)|^2+O(p^{-1/2})\\
\ll_{d,\epsilon}\;&\mathbb{E}_{h_0,a\in\mathbb{F}_p^d}\mathbb{E}_{n\in V(M)^{h_0+a,a}}|\mathbb{E}_{b\in V(M(h_0+a+n+\cdot))}\tau(h_0,a,b,n)\otimes\varphi_{h_0+b}(n+a)\otimes\overline{\varphi}_{h_0+a+b}(n)|^2+O(p^{-1/2})\\
\ll_{d,\epsilon}\;&|\mathbb{E}_{h_0,a\in\mathbb{F}_p^d}\mathbb{E}_{n\in V(M)^{h_0+a,a}}\mathbb{E}_{b,b'\in V(M(h_0+a+n+\cdot))}\tau(h_0,a,b,n)\otimes\overline{\tau}(h_0,a,b',n)\\
&\quad\otimes\varphi_{h_0+b}(n+a)\otimes\overline{\varphi}_{h_0+b'}(n+a)\otimes\overline{\varphi}_{h_0+a+b}(n)\otimes\varphi_{h_0+a+b'}(n)|+O(p^{-1/2})\\
\ll_{d,\epsilon}\;&|\mathbb{E}_{h_0,a,b,b'\in\mathbb{F}_p^d}\mathbb{E}_{n\in V(M)^{h_0+a,h_0+a+b,h_0+a+b',a}}\tau(h_0,a,b,n)\otimes\overline{\tau}(h_0,a,b',n)\\
&\quad\otimes\varphi_{h_0+b}(n+a)\otimes\overline{\varphi}_{h_0+b'}(n+a)\otimes\overline{\varphi}_{h_0+a+b}(n)\otimes\varphi_{h_0+a+b'}(n)|+O(p^{-1/2}),
\end{aligned}$$

where we used Theorem A.12 and the fact that the sets

$$\{(h_0,a,b,n)\in(\mathbb{F}_p^d)^4\colon n\in V(M)^{h_0+a,h_0+a+b,a}\}$$

and

$$\{(h_0,a,b,b',n)\in(\mathbb{F}_p^d)^4\colon n\in V(M)^{h_0+a,h_0+a+b,h_0+a+b',a}\}$$

are consistent M-sets of total co-dimensions 4 and 5 respectively. Substituting $c:=a+b+b'$ in (9.4), we have that

(9.5)

$$|\mathbb{E}_{h_0,c,b,b'\in\mathbb{F}_p^d}\mathbb{E}_{n\in V(M)^{h_0+c-b-b',h_0+c-b',h_0+c-b,c-b-b'}}\alpha(h_0,c,b,b',n)\otimes\varphi'_{h_0,b,c}(n)\otimes\overline{\varphi'}_{h_0,b',c}(n)|\gg_{d,\epsilon}1,$$

provided that $p\gg_{d,\epsilon}1$, where

$$\alpha(h_0,c,b,b',n):=\tau(h_0,c-b-b',b,n)\otimes\overline{\tau}(h_0,c-b-b',b',n) \tag{9.6}$$

and $\varphi'_{h_0,b,c}(n):=\varphi_{h_0+b}(n+c-b-b')\otimes\varphi_{h_0+c-b}(n)$. Let Λ_1 be the set of $(h_0,c,b_1,b_2,b',n)\in(\mathbb{F}_p^d)^6$ such that

$$n,n+h_0+c-b',n+h_0+c-b_i-b',n+h_0+c-b_i,n+c-b_i-b'\in V(M)\text{ for }i=1,2,$$

and Λ'_1 be the set of $(h_0,c,b_1,b_2,n)\in(\mathbb{F}_p^d)^5$ such that

$$n,n+h_0+c-b_1,n+h_0+c-b_2\in V(M).$$

For $(h_0,c,b_1,b_2,n)\in(\mathbb{F}_p^d)^5$, let $\Lambda_1(h_0,c,b_1,b_2,n)$ denote the set of b' such that

$$n+h_0+c-b',n+h_0+c-b_i-b',n+c-b_i-b'\in V(M)\text{ for }i=1,2.$$

Let A be the matrix associated to M. Note that $(h_0,c,b_1,b_2,n,b')\in\Lambda_1$ if and only if

- $h_0,c,b_1\in\mathbb{F}_p^d$;
- $((b_1-b_2)A)\cdot h_0=0$;
- $M(n)=M(n+c+h_0-b_1)=M(n+c+h_0-b_2)=0$;

- $M(n+h_0+c-b')=M(n+h_0+c-b_1-b')=M(n+h_0+c-b_2-b')=M(n+c-b_1-b')=0.$[16]

This implies that Λ_1 is a consistent M-set of total co-dimension 8, and that Λ_1' is a consistent M-set of total co-dimension 4. By Theorem A.12, the square of left hand side of (9.5) is bounded by $O_{d,\epsilon}(1)$ times

$$\begin{aligned}(9.7)\quad &|\mathbb{E}_{h_0,c,b'\in\mathbb{F}_p^d}\mathbb{E}_{n\in V(M)^{h_0+c-b'}}\mathbb{E}_{b\in\mathbb{F}_p^d\colon n+h_0+c-b-b',n+h_0+c-b,n+c-b-b'\in V(M)}\\ &\quad\alpha(h_0,c,b,b',n)\otimes\varphi'_{h_0,b,c}(n)\otimes\overline{\varphi'}_{h_0,b',c}(n)|^2+O(p^{-1/2})\\ \ll_{d,\epsilon}\ &\mathbb{E}_{h_0,c,b'\in\mathbb{F}_p^d}\mathbb{E}_{n\in V(M)^{h_0+c-b'}}|\mathbb{E}_{b\in\mathbb{F}_p^d\colon n+h_0+c-b-b',n+h_0+c-b,n+c-b-b'\in V(M)}\\ &\quad\alpha(h_0,c,b,b',n)\otimes\varphi'_{h_0,b,c}(n)|^2+O(p^{-1/2})\\ \ll_{d,\epsilon}\ &|\mathbb{E}_{h_0,c,b'\in\mathbb{F}_p^d}\mathbb{E}_{n\in V(M)^{h_0+c-b'}}\mathbb{E}_{b_1,b_2\in\mathbb{F}_p^d\colon n+h_0+c-b_i-b',n+h_0+c-b_i,n+c-b_i-b'\in V(M)\text{ for }i=1,2}\\ &\quad\alpha(h_0,c,b_1,b',n)\otimes\overline{\alpha}(h_0,c,b_2,b',n)\otimes\varphi'_{h_0,b_1,c}(n)\otimes\overline{\varphi'}_{h_0,b_2,c}(n)|+O(p^{-1/2})\\ \ll_{d,\epsilon}\ &|\mathbb{E}_{(h_0,c,b_1,b_2,n)\in\Lambda_1'}\mathbb{E}_{b'\in\Lambda_1(h_0,c,b_1,b_2,n)}\\ &\quad\alpha(h_0,c,b_1,b',n)\otimes\overline{\alpha}(h_0,c,b_2,b',n)\otimes\varphi'_{h_0,b_1,c}(n)\otimes\overline{\varphi'}_{h_0,b_2,c}(n)|+O(p^{-1/2})\\ \ll_{d,\epsilon}\ &\mathbb{E}_{(h_0,c,b_1,b_2,n)\in\Lambda_1'}|\mathbb{E}_{b'\in\Lambda_1(h_0,c,b_1,b_2,n)}\alpha(h_0,c,b_1,b',n)\otimes\overline{\alpha}(h_0,c,b_2,b',n)|+O(p^{-1/2}),\end{aligned}$$

where we used the fact that the set

$$\{(h_0,c,b,b',n)\in(\mathbb{F}_p^d)^5\colon n\in V(M)^{h_0+c-b-b',h_0+c-b',h_0+c-b,c-b-b'}\}$$

is a consistent M-set of total co-dimensions 5.

Let Λ_2 denote the set of $(h_0,c,b_1,b_1',b_2,b_2',n)\in(\mathbb{F}_p^d)^7$ such that (h_0,c,b_1,b_2,b_1',n) and (h_0,c,b_1,b_2,b_2',n) belong to Λ_1. Since M is pure, by Propositions A.10 and A.11, Λ_2 is a pure and consistent M-set and is of total co-dimension at most 28. So by Theorem A.12, the square of the right hand side of (9.7) is bounded by $O_{d,\epsilon}(1)$ times

$$\begin{aligned}&\mathbb{E}_{(h_0,c,b_1,b_2,n)\in\Lambda_1'}|\mathbb{E}_{b'\in\Lambda_1(h_0,c,b_1,b_2,n)}\alpha(h_0,c,b_1,b',n)\otimes\overline{\alpha}(h_0,c,b_2,b',n)|^2+O(p^{-1/2})\\ (9.8)\quad\ll_{d,\epsilon}\ &|\mathbb{E}_{(h_0,c,b_1,b_2,n)\in\Lambda_1'}\mathbb{E}_{b_1',b_2'\in\Lambda_1(h_0,c,b_1,b_2,n)}\alpha'(h_0,c,b_1,b_1',b_2,b_2',n)|+O(p^{-1/2})\\ \ll_{d,\epsilon}\ &|\mathbb{E}_{(h_0,c,b_1,b_1',b_2,b_2',n)\in\Lambda_2}\alpha'(h_0,c,b_1,b_1',b_2,b_2',n)|+O(p^{-1/2}),\end{aligned}$$

where

$$\begin{aligned}(9.9)\quad&\alpha'(h_0,c,b_1,b_1',b_2,b_2',n)\\ &:=\alpha(h_0,c,b_1,b_1',n)\otimes\overline{\alpha}(h_0,c,b_2,b_1',n)\otimes\overline{\alpha}(h_0,c,b_1,b_2',n)\otimes\alpha(h_0,c,b_2,b_2',n).\end{aligned}$$

We have now successfully eliminated the φ_h term in (9.1), and we could use Lemma C.7 to deduce from (9.8) that α' coincides with a nilsequence of degree at most s on Λ_2. However, in order to better use (9.8) to derive the information we need, we need to make some further transformations for α' before applying Lemma C.7.

Firstly, we need to take sufficiently many derivatives of α in the variable n in order to "annihilate" the appearance of other variables. Since Λ_2 is a consistent M-set of

[16] In particular, the condition $M(n+c-b_2-b')=0$ is a consequence of other restrictions by Lemma A.3.

total co-dimension at most 28, we may write $\Lambda_2 = V(\mathcal{J})$ for some consistent $(M, 7)$-family $\mathcal{J} \subseteq \mathbb{F}_p[h_0, c, b_1, b'_1, b_2, b'_2, n]$ of total dimension at most 28. Let $(\mathcal{J}', \mathcal{J}'')$ be an $\{h_0, c, b_1, b'_1, b_2, b'_2\}$-decomposition of $\mathcal{J}$. Let Λ'_2 be the set of $\mathbf{y} := (h_0, c, b_1, b'_1, b_2, b'_2) \in (\mathbb{F}_p^d)^6$ such that $(\mathbf{y}, n) \in V(\mathcal{J}')$ for all $n \in \mathbb{F}_p^d$. For $\mathbf{y} \in (\mathbb{F}_p^d)^6$, let $\Lambda_2(\mathbf{y})$ denote the set of $n \in \mathbb{F}_p^d$ such that $(\mathbf{y}, n) \in V(\mathcal{J}'')$. Since Λ_2 is a pure and consistent M-set of total co-dimension at most 28, so is Λ'_2 by Propositions A.10 and A.11. For convenience write $\mathbf{h}' = (h_1, \dots, h_{s-1}) \in (\mathbb{F}_p^d)^{s-1}$. Let Λ_3 denote the set of $(\mathbf{y}, n, \mathbf{h}') \in (\mathbb{F}_p^d)^{s+6}$ such that $(\mathbf{y}, n + \epsilon \cdot \mathbf{h}') \in \Lambda_2$ for all $\epsilon \in \{0, 1\}^{s-1}$, where

$$\epsilon \cdot \mathbf{h}' := (\epsilon_1, \dots, \epsilon_{s-1}) \cdot (h_1, \dots, h_{s-1}) = \epsilon_1 h_1 + \cdots + \epsilon_{s-1} h_{s-1}.$$

By repeatedly using Propositions A.10 and A.11, we have that Λ_3 is a pure and consistent M-set of total co-dimension at most $\binom{s+7}{2}$. By Theorem A.12, the 2^{s-1}-th power of the right hand side of (9.8) is bounded by $O_{d,\epsilon}(1)$ times

$$\begin{aligned}
&|\mathbb{E}_{\mathbf{y}\in\Lambda'_2}\mathbb{E}_{n\in\Lambda_2(\mathbf{y})}\alpha'(\mathbf{y}, n)|^{2^{s-1}} + O(p^{-1/2}) \ll_{d,\epsilon} \mathbb{E}_{\mathbf{y}\in\Lambda'_2}|\mathbb{E}_{n\in\Lambda_2(\mathbf{y})}\alpha'(\mathbf{y}, n)|^{2^{s-1}} + O(p^{-1/2}) \\
\ll_{d,\epsilon} &\left|\mathbb{E}_{\mathbf{y}\in\Lambda'_2}\mathbb{E}_{(n,\mathbf{h}')\in\Box_{s-1}(\Lambda_2(\mathbf{y}))} \bigotimes_{\epsilon\in\{0,1\}^{s-1}} C^{|\epsilon|}\alpha'(\mathbf{y}, n + \epsilon \cdot \mathbf{h}')\right| + O(p^{-1/2}) \\
\ll_{d,\epsilon} &\left|\mathbb{E}_{(\mathbf{y},n,\mathbf{h}')\in\Lambda_3} \bigotimes_{\epsilon\in\{0,1\}^{s-1}} C^{|\epsilon|}\alpha'(\mathbf{y}, n + \epsilon \cdot \mathbf{h}')\right| + O(p^{-1/2}).
\end{aligned}$$

Combining all the computations above, we have that

$$|\mathbb{E}_{(\mathbf{y},n,\mathbf{h}')\in\Lambda_3}\sigma(\mathbf{y}, n, \mathbf{h}')| \gg_{d,\epsilon} 1, \tag{9.10}$$

where

$$\sigma(\mathbf{y}, n, \mathbf{h}') := \bigotimes_{\epsilon\in\{0,1\}^{s-1}} C^{|\epsilon|}\alpha'(\mathbf{y}, n + \epsilon \cdot \mathbf{h}'). \tag{9.11}$$

We next provide a better description for the set Λ_3 under a change of variables. Let $L\colon (\mathbb{F}_p^d)^{s+6} \to (\mathbb{F}_p^d)^{s+6}$ be the bijective linear transformation which maps $(h_0, c, b_1, b_2, b'_1, b'_2, n, \mathbf{h}')$ to

$$(n + c - b_1 - b'_1, n + c - b_1 - b'_1 + h_0, n + c - b_1 + h_0, n + c - b'_1 + h_0, n, \mathbf{h}', b_1 - b_2, b'_1 - b'_2).$$

One can compute that

$$\begin{aligned}
&L^{-1}(w_1, w_2, w_3, w_4, n, \mathbf{h}', h_s, h_{s+1}) \\
&= (w_2 - w_1, w_1 - 2w_2 + w_3 + w_4 - n, w_4 - w_2, w_4 - w_2 - h_s, w_3 - w_2, w_3 - w_2 - h_{s+1}, n, \mathbf{h}').
\end{aligned} \tag{9.12}$$

Denote $\sigma' := \sigma \circ L^{-1}$, $\Lambda_4 := L(\Lambda_3)$. By Propositions A.10 and A.11, Λ_4 is a pure and consistent M-set of total co-dimension at most $\binom{s+7}{2}$.

Then (9.10) implies that

$$|\mathbb{E}_{(w_1,w_2,w_3,w_4,n,\mathbf{h}',h_s,h_{s+1})\in\Lambda_4}\sigma'(w_1, w_2, w_3.w_4, n, \mathbf{h}', h_s, h_{s+1})| \gg_{d,\epsilon} 1. \tag{9.13}$$

We need a precise description for the set Λ_4. Note that $(h_0, c, b_1, b_2, b'_1, b'_2, n, \mathbf{h}') \in \Lambda_3$ if and only if the following vectors are in $V(M)$:

$$n+\epsilon'\cdot\mathbf{h}', n+\epsilon'\cdot\mathbf{h}'+h_0+c-b'_i, n+\epsilon'\cdot\mathbf{h}'+h_0+c-b_i, n+\epsilon'\cdot\mathbf{h}'+h_0+c-b'_i-b_j, n+\epsilon'\cdot\mathbf{h}'+c-b'_i-b_j$$

for all $i, j \in \{1, 2\}$ and $\epsilon' \in \{0, 1\}^{s-1}$. This is equivalent of saying that

$$n+\epsilon'\cdot\mathbf{h}', w_4+\epsilon'\cdot\mathbf{h}'+\epsilon_{s+1}h_{s+1}, w_3+\epsilon'\cdot\mathbf{h}'+\epsilon_s h_s, w_2+\epsilon'\cdot\mathbf{h}'+\epsilon_s h_s+\epsilon_{s+1}h_{s+1}, w_1+\epsilon'\cdot\mathbf{h}'+\epsilon_s h_s+\epsilon_{s+1}h_{s+1}$$

for all $\epsilon' \in \{0, 1\}^{s-1}$ and $\epsilon_s, \epsilon_{s+1} \in \{0, 1\}$, which is further equivalent of saying that $(w_1, \mathbf{h}', h_s, h_{s+1}), (w_2, \mathbf{h}', h_s, h_{s+1}) \in \Box_{s+1}(V(M))$, $(w_3, \mathbf{h}', h_s) \in \Box_s(V(M))$, $(w_4, \mathbf{h}', h_{s+1}) \in \Box_s(V(M))$ and $(n, \mathbf{h}') \in \Box_{s-1}(V(M))$.

In conclusion, Λ_4 is the set Ω_2 in Example B.4 of [19] (with the variable w_5 replaced by n), which is a nice and consistent M-set of total co-dimension $(s^2 + 11s + 12)/2$. By Lemma B.1, we have $\sigma, \sigma' \in \Xi_p^{s+1;O_{d,\epsilon}(1),O_{d,\epsilon}(1)}((\mathbb{F}_p^d)^{s+6})$. Since $d \geq N(s)$ and $p \gg_{d,\epsilon} 1$, it follows from Lemma C.7 and (9.13) that

$$\sigma'(w_1, w_2, w_3, w_4, n, \mathbf{h}', h_s, h_{s+1}) \sim_{O_{d,\epsilon}(1)} 1 \mod \Xi_p^{s+1}(\Lambda_4).$$

Therefore, by Lemma C.4 (vi),

$$\sigma(h_0, c, b_1, b'_1, b_2, b'_2, n, \mathbf{h}') \sim_{O_{d,\epsilon}(1)} 1 \mod \Xi_p^{s+1}(\Lambda_3). \tag{9.14}$$

9.2. **Calculation of the equivalence classes.** We now simplify (9.14) up to the equivalence relation $\sim$ by using the nilcharacters defined in (9.3), (9.6), (9.9) and (9.11). We first recapture the definitions:

$$\begin{aligned}
&\tau(h_0, a, b, n) := \chi(h_0 + a, n) \otimes \chi(h_0 + b, n + a) \otimes \overline{\chi}(h_0 + a + b, n) \otimes \overline{\chi}(h_0, n + a),\\
&\alpha(h_0, c, b, b', n) := \tau(h_0, c - b - b', b, n) \otimes \overline{\tau}(h_0, c - b - b', b', n),\\
&\alpha'(h_0, c, b_1, b'_1, b_2, b'_2, n)\\
&\ := \alpha(h_0, c, b_1, b'_1, n) \otimes \overline{\alpha}(h_0, c, b_2, b'_1, n) \otimes \overline{\alpha}(h_0, c, b_1, b'_2, n) \otimes \alpha(h_0, c, b_2, b'_2, n)\\
&\sigma(h_0, c, b_1, b'_1, b_2, b'_2, n, \mathbf{h}') := \bigotimes_{\epsilon\in\{0,1\}^{s-1}} C^{|\epsilon|}\alpha'(h_0, c, b_1, b'_1, b_2, b'_2, n + \epsilon \cdot \mathbf{h}').
\end{aligned} \tag{9.15}$$

By Lemma C.4 (iv), (vi), Lemma C.5, (8.3) and (9.15), for all $i, j \in \{1, 2\}$, we have that

$$\begin{aligned}
&\tau(h_0, c - b_i - b'_j, b_i, n)\\
\sim_{O_{d,\epsilon}(1)}\ &\chi(h_0 + c - b_i - b'_j, n) \otimes \chi(h_0 + b_i, n + c - b_i - b'_j)\\
&\quad\otimes \overline{\chi}(h_0 + c - b'_j, n) \otimes \overline{\chi}(h_0, n + c - b_i - b'_j)\\
\sim_{O_{d,\epsilon}(1)}\ &\chi(b_i, n + c - b_i - b'_j) \otimes \overline{\chi}(b_i, n) \mod \Xi_p^{s+1}(\Lambda_3),
\end{aligned}$$

where all the sequences are viewed as nilcharacters in the variables $(h_0, c, b_1, b'_1, b_2, b'_2, n, \mathbf{h}')$. Similarly, for all $i, j \in \{1, 2\}$, we have that

$$\begin{aligned}&\alpha(h_0, c, b_i, b'_j, n)\\ \sim_{O_{d,\epsilon}(1)}\ &\tau(h_0, c - b_i - b'_j, b_i, n) \otimes \overline{\tau}(h_0, c - b_i - b'_j, b'_j, n)\\ \sim_{O_{d,\epsilon}(1)}\ &\chi(b_i, n + c - b_i - b'_j) \otimes \overline{\chi}(b_i, n) \otimes \overline{\chi}(b'_j, n + c - b_i - b'_j) \otimes \chi(b'_j, n)\\ \sim_{O_{d,\epsilon}(1)}\ &\chi(b_i - b'_j, n + c - b_i - b'_j) \otimes \overline{\chi}(b_i - b'_j, n) \mod \Xi_p^{s+1}(\Lambda_3)\end{aligned}$$

and

$$\begin{aligned}&\alpha'(h_0, c, b_1, b_2, b'_1, b'_2, n)\\ \sim_{O_{d,\epsilon}(1)}\ &\bigotimes_{i,j\in\{1,2\}} C^{i+j}(\chi(b_i - b'_j, n + c - b_i - b'_j) \otimes \overline{\chi}(b_i - b'_j, n))\\ \sim_{O_{d,\epsilon}(1)}\ &\bigotimes_{i,j\in\{1,2\}} C^{i+j}\chi(b_i - b'_j, n + c - b_i - b'_j) \mod \Xi_p^{s+1}(\Lambda_3).\end{aligned} \tag{9.16}$$

Lemma 9.1. We have that

$$\begin{aligned}&\sigma(h_0, c, b_1, b_2, b'_1, b'_2, n, \mathbf{h}')\\ \sim_{O_{d,\epsilon}(1)}\ &C^{s-1}(\tilde{\chi}(b'_1 - b'_2, b_1 - b_2, \mathbf{h}') \otimes \overline{\tilde{\chi}}(b_1 - b_2, b'_1 - b'_2, \mathbf{h}'))^{\otimes(s+1)!} \mod \Xi_p^{s+1}(\Lambda_3).\end{aligned} \tag{9.17}$$

Proof. By (9.15) and (9.16), we have that

$$\begin{aligned}\sigma(h_0, c, b_1, b_2, b'_1, b'_2, n, \mathbf{h}') &= \bigotimes_{\epsilon\in\{0,1\}^{s-1}} C^{|\epsilon|}\alpha'(h_0, c, b_1, b_2, b'_1, b'_2, n + \epsilon \cdot \mathbf{h}')\\ &\sim_{O_{d,\epsilon}(1)} \bigotimes_{i,j\in\{1,2\}} C^{i+j} \bigotimes_{\epsilon\in\{0,1\}^{s-1}} C^{|\epsilon|}\chi(b_i - b'_j, n + c - b_i - b'_j + \epsilon \cdot \mathbf{h}') \mod \Xi_p^{s+1}(\Lambda_3).\end{aligned} \tag{9.18}$$

We first claim that for all $i, j \in \{1, 2\}$,

$$\begin{aligned}&\bigotimes_{\epsilon\in\{0,1\}^{s-1}} C^{|\epsilon|}\chi(b_i - b'_j, n + c - b_i - b'_j + \epsilon \cdot \mathbf{h}')^{\otimes 2}\\ \sim_{O_{d,\epsilon}(1)}\ &C^{s-1}\Big(\tilde{\chi}(b_i - b'_j, n + c - b_i - b'_j, \mathbf{h}')^{\otimes 2} \bigotimes_{\ell=1}^{s-1} \tilde{\chi}(b_i - b'_j, h_\ell, \mathbf{h}')\Big)^{\otimes(s+1)!} \mod \Xi_p^{s+1}(\Lambda_3).\end{aligned} \tag{9.19}$$

For convenience denote $h_0 = n + c - b_i - b'_j$. By (8.3), Lemma C.4 (iv) and (vi), Lemma C.5, and the fact that $\tilde{\chi}$ is symmetric in the last s variables,

$$\begin{aligned}&\bigotimes_{\epsilon\in\{0,1\}^{s-1}} C^{|\epsilon|}\chi(b_i - b'_j, n + c - b_i - b'_j + \epsilon \cdot \mathbf{h}')^{\otimes 2}\\ \sim_{O_{d,\epsilon}(1)}\ &\bigotimes_{i_0+\dots+i_{s-1}=s} \Big(\tilde{\chi}(b_i - b'_j, h_0, \dots, h_0, \dots, h_{s-1}, \dots, h_{s-1})^{\otimes 2C_{i_0,\dots,i_{s+1}}}\Big)^{\otimes(s+1)} \mod \Xi_p^{s+1}(\Lambda_3)\end{aligned}$$

for some $C_{i_0,\dots,i_{s-1}} \in \mathbb{Z}$, where in the last s variables of $\tilde{\chi}(b_i - b'_j, h_0, \dots, h_0, \dots, h_{s-1}, \dots, h_{s-1})$, h_j appears i_j times for all $0 \leq j \leq s+1$. Moreover, $C_{i_0,\dots,i_{s-1}}$ are the unique integers such that the polynomial

$$F(x_0, \dots, x_{s-1}) := \sum_{\epsilon=(\epsilon_1,\dots,\epsilon_{s-1})\in\{0,1\}^{s-1}} (-1)^{|\epsilon|}(x_0 + \sum_{i=1}^{s-1} \epsilon_i x_i)^s \in \mathbb{F}_p[x_0, \dots, x_{s-1}]$$

can be written as

$$F(x_0, \dots, x_{s-1}) := \sum_{i_0+\dots+i_{s-1}=s} C_{i_0,\dots,i_{s-1}} x_0^{i_0} \dots x_{s-1}^{i_{s-1}}.$$

In other words,

$$C_{i_0,\dots,i_{s-1}} = \sum_{\epsilon=(\epsilon_1,\dots,\epsilon_{s-1})\in E_{i_0,\dots,i_{s-1}}} (-1)^{|\epsilon|} \frac{s!}{i_0! \dots i_{s-1}!},$$

where $E_{i_0,\dots,i_{s-1}}$ is the set of $\epsilon = (\epsilon_1, \dots, \epsilon_{s-1}) \in \{0,1\}^{s-1}$ such that $\epsilon_j = 1$ whenever $i_j \geq 1$ for all $1 \leq j \leq s-1$. If one of $i_1, \dots, i_{s-1}$ is 0, say $i_\ell = 0$, then by separating those $(\epsilon_1, \dots, \epsilon_{s-1}) \in E_{i_0,\dots,i_{s-1}}$ with $\epsilon_\ell = 0$ and those with $\epsilon_\ell = 1$ in the sum, we have that $C_{i_0,\dots,i_{s-1}} = 0$. So if $C_{i_0,\dots,i_{s-1}} \neq 0$, then all of $i_1, \dots, i_{s-1}$ are at least 1. Since $i_0 + \dots + i_{s-1} = s$, we have that the only nontrivial terms $C_{i_0,\dots,i_{s-1}}$ are $C_{1,\dots,1}$, which equals to $(-1)^{s-1}s!$, and $C_{0,1,\dots,1,2,1,\dots,1}$ (where the coordinate 2 corresponds to the unique i_ℓ which equals to 2), which equals to $(-1)^{s-1}s!/2$. Since $\tilde{\chi}$ is symmetric in the last s variables, this implies (9.19) and proves the claim.

After some simple computations, one can deduce (9.17) from (9.18) and (9.19) by the claim and Lemmas C.4 (iv), C.5 and C.6. We leave the details to the interested readers. □

Since Λ_4 is a nice and consistent M-set of total co-dimension $(s^2 + 11s + 12)/2$, so is Λ_3 by Propositions A.10 (iii) and A.11. Combining Lemmas C.6, 9.1 and equation (9.14), we have that

$$\tilde{\chi}(b'_1 - b'_2, b_1 - b_2, h_1, \dots, h_{s-1}) \sim_{O_{d,\epsilon}(1)} \tilde{\chi}(b_1 - b_2, b'_1 - b'_2, h_1, \dots, h_{s-1}) \mod \Xi_p^{s+1}(\Lambda_3).$$

By Lemma C.4 (vi), we have that

$$\tilde{\chi}(h_s, h_{s+1}, h_1, \dots, h_{s-1}) \sim_{O_{d,\epsilon}(1)} \tilde{\chi}(h_{s+1}, h_s, h_1, \dots, h_{s-1}) \mod \Xi_p^{s+1}(\Lambda_4),$$

where all the sequences are viewed as nilcharacters in the variables $(w_1, \dots, w_4, n, h_1, \dots, h_{s+1})$. Since $\tilde{\chi}$ is symmetric in the last s variables, we have that

$$\tilde{\chi}(h_1, \dots, h_{s+1}) \sim_{O_{d,\epsilon}(1)} \tilde{\chi}(h_{\sigma(1)}, \dots, h_{\sigma(s+1)}) \mod \Xi_p^{s+1}(\Lambda_4) \tag{9.20}$$

for any permutation $\sigma\colon \{1, \dots, s+1\} \to \{1, \dots, s+1\}$.

9.3. **Completion of the proof of SGI($s+1$) for $s \geq 2$.** Our final step is to use (9.20) to complete the proof.

Step 1: reformulate (8.1). Since every $h \in H$ is M-non-isotropic, $V(M)^h$ can be written as $V(M) \cap (V+c)$ for some affine subspace $V+c$ of $\mathbb{F}_p^d$ of co-dimension 1 such that $V \cap V^{\perp_M} = \{\mathbf{0}\}$. By Proposition A.4, $\text{rank}(M|_{V+c}) = d-1$. So by (8.1) and Proposition D.1, $\|\Delta_h f \cdot \chi(h,\cdot)\|_{U^s(V(M)^h)} \gg_{d,\epsilon} 1$ for all $h \in H$ and thus

$$\mathbb{E}_{h\in\mathbb{F}_p^d}\|\Delta_h f \cdot \chi(h,\cdot)\|_{U^s(V(M)^h)} \gg_{d,\epsilon} 1.$$

Denote $\tilde{f}(n) := f(n)\tilde{\chi}(n,\dots,n)$ for all $n \in \mathbb{F}_p^d$. Since $\square_{s+1}(V(M))$ is an M-set of total co-dimension $(s^2+3s+4)/2$ and $d \geq s^2+3s+5$, by Theorem A.12,

$$\begin{aligned} 1 &\ll_{d,\epsilon} \mathbb{E}_{h\in\mathbb{F}_p^d}\|\Delta_h f \cdot \chi(h,\cdot)\|^{2^s}_{U^s(V(M)^h)} \\ &= |\mathbb{E}_{(n,h_1,\dots,h_{s+1})\in\square_{s+1}(V(M))}\Delta_{h_{s+1}}\dots\Delta_{h_1}\tilde{f}(n)\otimes\beta(h_1,\dots,h_{s+1},n)| + O_{d,\epsilon}(p^{-1/2}), \end{aligned} \tag{9.21}$$

where (using Convention 8.2)

$$\beta(h_1,\dots,h_{s+1},n) := \Delta_{h_{s+1}}\dots\Delta_{h_2}(\chi(h_1,n)\otimes\overline{\tilde{\chi}}(n+h_1,\dots,n+h_1)\otimes\tilde{\chi}(n,\dots,n)).$$

Step 2: use (9.20) to convert β into a lower degree nilsequence. For convenience from now on we write the variables in Λ_4 as $(n,w_2,\dots,w_5,h_1,\dots,h_{s+1})$ instead of $(w_1,\dots,w_4,n,h_1,\dots,h_{s+1})$, and denote $\mathbf{h} := (h_1,\dots,h_{s+1})$ and $\mathbf{w} := (w_2,\dots,w_5)$.

Since Λ_4 is a consistent M-set is of total co-dimension $(s^2+11s+12)/2$, we may write $\Lambda_4 = V(\mathcal{J})$ for some consistent $(M,s+6)$-family $\mathcal{J} \subseteq \mathbb{F}_p[n,w_2,\dots,w_5,h_1,\dots,h_{s+1}]$ of total dimension $(s^2+11s+12)/2$. Let $(\mathcal{J}',\mathcal{J}'')$ be an $\{n,h_1,\dots,h_{s+1}\}$-decomposition of $\mathcal{J}$. Let Λ_4' be the set of $(n,\mathbf{h}) \in (\mathbb{F}_p^d)^{s+2}$ such that $(n,\mathbf{w},\mathbf{h}) \in V(\mathcal{J}')$ for all $\mathbf{w} \in (\mathbb{F}_p^d)^4$. For $(n,\mathbf{h}) \in (\mathbb{F}_p^d)^{s+2}$, let $\Lambda_4(n,\mathbf{h})$ denote the set of $\mathbf{w} \in (\mathbb{F}_p^d)^4$ such that $(n,\mathbf{w},\mathbf{h}) \in V(\mathcal{J}'')$. By the description of Λ_4, we have that $\Lambda_4' = \square_{s+1}(V(M))$. So by Theorem A.12 and (9.21), we have

$$\begin{aligned} &|\mathbb{E}_{(n,\mathbf{w},\mathbf{h})\in\Lambda_4}\Delta_{h_{s+1}}\dots\Delta_{h_1}\tilde{f}(n)\otimes\beta(\mathbf{h},n)| \\ \gg_{d,\epsilon}\ &|\mathbb{E}_{(n,\mathbf{h})\in\square_{s+1}(V(M))}\Delta_{h_{s+1}}\dots\Delta_{h_1}\tilde{f}(n)\otimes\beta(\mathbf{h},n)\mathbb{E}_{\mathbf{w}\in\Lambda_4(n,\mathbf{h})}1| + O_{d,\epsilon}(p^{-1/2}) \gg_{d,\epsilon} 1. \end{aligned} \tag{9.22}$$

Using (8.3), (9.20), Lemma C.4 (iii), (iv), (vi) and Lemma C.5, it is not hard to compute that (the details are left to the interested readers)

$$\beta(h_1,\dots,h_{s+1},n) \sim_{O_{d,\epsilon}(1)} 1 \mod \Xi_p^{s+1}(\Lambda_4),$$

where now the variables in Λ_4 are labeled as $(n,w_2,\dots,w_5,h_1,\dots,h_{s+1})$. Therefore, $\beta(h_1,\dots,h_{s+1},n)$ belongs to $\text{Nil}^{s;O_{d,\epsilon}(1),O_{d,\epsilon}(1)}(\Lambda_4)$ and thus belongs to $\text{Nil}^{J;O_{d,\epsilon}(1),O_{d,\epsilon}(1)}(\Lambda_4)$, where

$$J := \{(a_1,\dots,a_4,b,c_1,\dots,c_{s+1}) \in \mathbb{N}^{s+6} \colon a_1+\dots+a_4+b+c_1+\dots+c_{s+1} \leq s\}.$$

For all $1 \leq i \leq s+1$, let

$$J_i := \{(a_1,\dots,a_4,b,c_1,\dots,c_{s+1}) \in J \colon c_i = 0\}.$$

Then $J = \cup_{i=1}^{s+1} J_i$. Note that any sequence in $\text{Nil}^{J_i}(\Lambda_4)$ is independent of h_i. By (9.22), Lemma B.4 and the Pigeonhole Principle, for $1 \leq i \leq s+1$, there exists a scalar valued sequence $\psi_i(\mathbf{w}, n, \mathbf{h})$ bounded by 1 which is independent of h_i such that

$$\left|\mathbb{E}_{(n,\mathbf{w},\mathbf{h})\in\Lambda_4}\Delta_{h_{s+1}}\dots\Delta_{h_1}\tilde{f}(n)\prod_{i=1}^{s+1}\psi_i(n,\mathbf{w},\mathbf{h})\right| \gg_{d,\epsilon} 1$$

since $p \gg_{d,\epsilon} 1$. By the definition of $\tilde{f}$ and the Pigeonhole Principle, there exists a scalar valued nilsequence ϕ' of step at most $s+1$ of complexity $O_{d,\epsilon}(1)$ such that writing $f' := f\phi'$, we have that

$$\left|\mathbb{E}_{(n,\mathbf{w},\mathbf{h})\in\Lambda_4}\Delta_{h_{s+1}}\dots\Delta_{h_1}f'(n)\prod_{i=1}^{s+1}\psi_i(n,\mathbf{w},\mathbf{h})\right| \gg_{d,\epsilon} 1. \tag{9.23}$$

Step 3: use the Cauchy-Schwartz inequality to remove ψ_i. For $1 \leq r \leq s+2$, we say that Property-r holds if there exist a scalar valued function $\psi_i(n, \mathbf{w}, \mathbf{h}, h'_1, \dots, h'_{r-1})$ bounded by 1 and independent of h_i for all $1 \leq i \leq r$ such that

$$\left|\mathbb{E}_{(n,\mathbf{w},\mathbf{h},h'_1,\dots,h'_{r-1})\in\Lambda_{4,r}}\Delta_{h_{s+1}}\dots\Delta_{h_1}f'(n)\prod_{i=r}^{s+1}\psi_i(n,\mathbf{w},\mathbf{h},h'_1,\dots,h'_{r-1})\right| \gg_{d,\epsilon} 1, \tag{9.24}$$

where $\Lambda_{4,r}$ is the set of all $(n, \mathbf{w}, \mathbf{h}, h'_1, \dots, h'_{r-1}) \in (\mathbb{F}_p^d)^{s+r+5}$ such that

$$(n - h'_1 - \dots - h'_{r-1}, \mathbf{w}, h'_1 + \epsilon_1 h_1, \dots, h'_{r-1} + \epsilon_{r-1} h_{r-1}, h_r, \dots, h_{s+1}) \in \Lambda_4$$

for all $\epsilon_1, \dots, \epsilon_{r-1} \in \{0, 1\}$, and the term $\prod_{i=r}^{s+1} \psi_i(n, \mathbf{w}, \mathbf{h}, h'_1, \dots, h'_{r-1})$ is understood as 1 if $r = s+2$.

Clearly, Property-1 holds by (9.23). Now suppose that Property-r holds for some $1 \leq r \leq s+1$. For convenience denote $\mathbf{x} := (n, \mathbf{w}, h_1, \dots, h_{r-1}, h_{r+1}, \dots, h_{s+1}, h'_1, \dots, h'_{r-1})$, and let $(\mathbf{x}; h)$ be the vector obtained by inserting h in the middle of h_{r-1} and h_{r+1}. Let $\tilde{\Lambda}_{4,r}$ denote the set of $(\mathbf{x}, h'_r, h''_r) \in (\mathbb{F}_p^d)^{s+r+6}$ such that $(\mathbf{x}; h'_r), (\mathbf{x}; h''_r) \in \Lambda_{4,r}$. Since Λ_4 is a pure and consistent M-set, by Propositions A.10 and A.11, $\Lambda_{4,r}$ and $\tilde{\Lambda}_{4,r}$ are pure and consistent M-sets of total co-dimension at most $\binom{2s+8}{2}$. We may then write $\Lambda_{4,r} = V(\mathcal{J})$ for some consistent $(M, s+r+5)$-family $\mathcal{J} \subseteq \mathbb{F}_p[n, w_2, \dots, w_5, h_1, \dots, h_{s+1}, h'_1, \dots, h'_{r-1}]$ of total dimension at most $\binom{2s+8}{2}$. Let $(\mathcal{J}', \mathcal{J}'')$ be an $\{n, w_2, \dots, w_5, h_1, \dots, h_{r-1}, h_{r+1}, \dots, h_{s+1}, h'_1, \dots, h'_{r-1}\}$-decomposition of $\mathcal{J}$. Let $\Lambda'_{4,r}$ be the set of $\mathbf{x} \in (\mathbb{F}_p^d)^{s+r+4}$ such that $(\mathbf{x}; h_r) \in V(\mathcal{J}')$ for all $h_r \in \mathbb{F}_p^d$. For $\mathbf{x} \in (\mathbb{F}_p^d)^{s+r+4}$, let $\Lambda_{4,r}(\mathbf{x})$ denote the set of $h_r \in \mathbb{F}_p^d$ such that $(\mathbf{x}; h_r) \in V(\mathcal{J}'')$. Since $d \geq (2s+7)(2s+6)+1$, by pulling out the h_r-independent term

$$\Delta_{h_{s+1}}\dots\Delta_{h_{r+1}}\Delta_{h_{r-1}}\dots\Delta_{h_1}\overline{f'}(n)\psi_r(n,\mathbf{w},h_1,\dots,h_{s+1},h'_1,\dots,h'_{r-1})$$

in (9.24) and apply the Cauchy-Schwartz inequality and Theorem A.12, we have that

$$\begin{aligned}
1 &\ll_{d,\epsilon} \mathbb{E}_{\mathbf{x}\in\Lambda'_{4,r}}\Big|\mathbb{E}_{h_r\in\Lambda_{4,r}(\mathbf{x})}\Delta_{h_{s+1}}\dots\Delta_{h_{r+1}}\Delta_{h_{r-1}}\dots\Delta_{h_1}f'(n+h_r)\prod_{i=r+1}^{s+1}\psi_i(\mathbf{x};h_r)\Big|^2\\
&=\Big|\mathbb{E}_{(\mathbf{x};h'_r,h''_r)\in\tilde{\Lambda}_{4,r}}\Delta_{h_{s+1}}\dots\Delta_{h_{r+1}}\Delta_{h_{r-1}}\dots\Delta_{h_1}(f'(n+h''_r)\overline{f'}(n+h'_r))\cdot\prod_{i=r+1}^{s+1}\psi_i(\mathbf{x};h''_r)\overline{\psi}_i(\mathbf{x};h'_r)\Big|\\
&=\Big|\mathbb{E}_{(\mathbf{x};h'_r),(\mathbf{x};h'_r+h_r)\in\Lambda_{4,r}}\Delta_{h_{s+1}}\dots\Delta_{h_{r+1}}\Delta_{h_r}\Delta_{h_{r-1}}\dots\Delta_{h_1}f'(n+h'_r)\cdot\prod_{i=r+1}^{s+1}\psi_i(\mathbf{x};h'_r+h_r)\overline{\psi}_i(\mathbf{x};h'_r)\Big|.
\end{aligned}$$

For all $r+1\leq i\leq s+1$, since ψ_i is independent of h_i and bounded by 1, $\psi_i(\mathbf{x};h'_r+h_r)\overline{\psi}_i(\mathbf{x};h'_r)$ is independent of h_i and bounded by 1. On the other hand, note that $(\mathbf{x};h'_r)$, $(\mathbf{x};h'_r+h_r)\in\Lambda_{4,r}$ if and only if

$$(n-h'_1-\dots-h'_{r-1},\mathbf{w},h'_1+\epsilon_1h_1,\dots,h'_{r-1}+\epsilon_{r-1}h_{r-1},h'_r,h_{r+1},\dots,h_{s+1})\in\Lambda_4$$

and

$$(n-h'_1-\dots-h'_{r-1},\mathbf{w},h'_1+\epsilon_1h_1,\dots,h'_{r-1}+\epsilon_{r-1}h_{r-1},h'_r+h_r,h_{r+1},\dots,h_{s+1})\in\Lambda_4$$

for all $\epsilon_1,\dots,\epsilon_{r-1}\in\{0,1\}$. Writing $m=n+h'_r$, this is equivalent of saying that

$$(m-h'_1-\dots-h'_r,\mathbf{w},h'_1+\epsilon_1h_1,\dots,h'_r+\epsilon_rh_r,h_{r+1},\dots,h_{s+1})\in\Lambda_4$$

for all $\epsilon_1,\dots,\epsilon_{r-1}\in\{0,1\}$, or equivalently, $(m,\mathbf{w},h_1,\dots,h_{s+1},h'_1,\dots,h'_r)\in\Lambda_{4,r+1}$. So we have that Property-$(r+1)$ holds.

Inductively, we have that Property-$(s+2)$ holds. Since Λ_4 is a pure and consistent M-set, by Propositions A.10 and A.11, $\Lambda_{4,s+2}=V(\tilde{\mathcal{J}})$ for some pure and consistent $(M,2s+7)$-family $\tilde{\mathcal{J}}\subseteq\mathbb{F}_p[n,w_2,\dots,w_5,h_1,\dots,h_{s+1},h'_1,\dots,h'_{s+1}]$ of total dimension at most $\binom{2s+8}{2}$. Let $(\tilde{\mathcal{J}}',\tilde{\mathcal{J}}'')$ be an $\{n,h_1,\dots,h_{s+1}\}$-decomposition of $\tilde{\mathcal{J}}$. For convenience denote $\mathbf{h}'=(h'_1,\dots,h'_{s+1})$. Let $\Lambda''_{4,s+2}$ be the set of $(n,\mathbf{h})\in(\mathbb{F}_p^d)^{s+2}$ such that $(n,\mathbf{w},\mathbf{h},\mathbf{h}')\in V(\tilde{\mathcal{J}}')$ for all $(\mathbf{w},\mathbf{h}')\in(\mathbb{F}_p^d)^{s+5}$. For $(n,\mathbf{h})\in(\mathbb{F}_p^d)^{s+2}$, let $\Lambda_{4,r}(n,\mathbf{h})$ denote the set of $(\mathbf{w},\mathbf{h}')\in(\mathbb{F}_p^d)^{s+5}$ such that $(n,\mathbf{w},\mathbf{h},\mathbf{h}')\in V(\tilde{\mathcal{J}}'')$. It is not hard to see that $\Lambda''_{4,s+2}=\Box_{s+1}(V(M))$. By Property-$(s+2)$ and Theorem A.12,

$$\begin{aligned}
1&\ll_{d,\epsilon}|\mathbb{E}_{(n,\mathbf{w},\mathbf{h},\mathbf{h}')\in\Lambda_{4,s+2}}\Delta_{h_{s+1}}\dots\Delta_{h_1}f'(n)|\\
&=|\mathbb{E}_{(n,\mathbf{h})\in\Box_{s+1}(V(M))}\Delta_{h_{s+1}}\dots\Delta_{h_1}f'(n)\mathbb{E}_{(\mathbf{w},\mathbf{h}')\in\Lambda_{4,s+2}(x,\mathbf{h})}1|+O(p^{-1/2})\\
&=|\mathbb{E}_{(n,\mathbf{h})\in\Box_{s+1}(V(M))}\Delta_{h_{s+1}}\dots\Delta_{h_1}(f\phi')(n)|+O(p^{-1/2})=\|f\phi'\|_{U^{s+1}(V(M))}^{2^{s+1}}+O(p^{-1/2}).
\end{aligned}$$

Since SGI(s) holds by assumption, there exists $\phi''\in\text{Nil}_p^{s;O_{d,\epsilon}(1),1}(\mathbb{F}_p^d)$ such that

$$|\mathbb{E}_{n\in V(M)}f(n)\phi'(n)\phi''(n)|\gg_{d,\epsilon}1.$$

This completes the proof of SGI($s+1$) since $\phi'\phi''\in\text{Nil}_p^{s+1;O_{d,\epsilon}(1),1}(\mathbb{F}_p^d)$.

Appendix A. Results from previous papers

In this appendix, we collect some results proved in [19, 20] on quadratic forms which are used in this paper.

A.1. **Liftings of polynomials.** We refer the readers to Section 1.3 for the definition of liftings of polynomials.

Lemma A.1 (Lemma 2.8 of [19])**.** Let $d \in \mathbb{N}$, p be a prime, and $f \in \text{poly}(\mathbb{Z}^d \to \mathbb{Z})$ be an integer valued polynomial of degree smaller than p. There exist integer valued polynomial $f_1 \in \text{poly}(\mathbb{Z}^d \to \mathbb{Z})$ and integer coefficient polynomial $f_2 \in \text{poly}(\mathbb{Z}^d \to \mathbb{Z})$ both having degrees at most $\deg(f)$ such that $\frac{1}{p}f = f_1 + \frac{1}{p}f_2$.

We refer the readers to Appendix A.4 for the definition of d-integral linear transformations.

Lemma A.2 (Lemma 9.1 of [19])**.** Every d-integral linear transformation $\tilde{L}\colon (\mathbb{Z}^d)^k \to (\frac{1}{p}\mathbb{Z}^d)^{k'}$ induces a d-integral linear transformation $L\colon (\mathbb{F}_p^d)^k \to (\mathbb{F}_p^d)^{k'}$. Conversely, every d-integral linear transformation $L\colon (\mathbb{F}_p^d)^k \to (\mathbb{F}_p^d)^{k'}$ admits a regular lifting $\tilde{L}\colon (\mathbb{Z}^d)^k \to (\frac{1}{p}\mathbb{Z}^d)^{k'}$ which is a d-integral linear transformation.

Moreover, we have that $p\tilde{L} \circ \tau \equiv \tau \circ L \mod p(\mathbb{Z}^d)^{k'}$.

A.2. **The rank of quadratic forms.**

Lemma A.3 (Lemma 4.2 of [19])**.** Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form associated with the matrix A, and $x, y, z \in \mathbb{F}_p^d$. Suppose that $M(x) = M(x+y) = M(x+z) = 0$. Then $M(x+y+z) = 0$ if and only if $(yA) \cdot z = 0$.

Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form associated with the matrix A and V be a subspace of $\mathbb{F}_p^d$. Let $V^{\perp_M}$ denote the set of $\{n \in \mathbb{F}_p^d\colon (mA) \cdot n = 0 \text{ for all } m \in V\}$. A subspace V of $\mathbb{F}_p^d$ is *M-isotropic* if $V \cap V^{\perp_M} \neq \{\mathbf{0}\}$. We say that a tuple $(h_1, \dots, h_k)$ of vectors in $\mathbb{F}_p^d$ is *M-isotropic* if the span of $h_1, \dots, h_k$ is an M-isotropic subspace. We say that a subspace or tuple of vectors is *M-non-isotropic* if it is not M-isotropic.

Proposition A.4 (Proposition 4.8 of [19])**.** Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form and V be a subspace of $\mathbb{F}_p^d$ of co-dimension r, and $c \in \mathbb{F}_p^d$.

(i) We have $\dim(V \cap V^{\perp_M}) \leq \min\{d - \text{rank}(M) + r, d - r\}$.
(ii) The rank of $M|_{V+c}$ equals to $d - r - \dim(V \cap V^{\perp_M})$ (i.e. $\dim(V) - \dim(V \cap V^{\perp_M})$).
(iii) The rank of $M|_{V+c}$ is at most $d - r$ and at least $\text{rank}(M) - 2r$.
(iv) $M|_{V+c}$ is non-degenerate (i.e. $\text{rank}(M|_{V+c}) = d - r$) if and only if V is not an M-isotropic subspace.

Lemma A.5 (Lemma 4.11 of [19])**.** Let $d, k \in \mathbb{N}_+$, p be a prime, and $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a non-degenerate quadratic form.

(i) The number of tuples $(h_1, \dots, h_k) \in (\mathbb{F}_p^d)^k$ such that $h_1, \dots, h_k$ are linearly dependent is at most $kp^{(d+1)(k-1)}$.
(ii) The number of M-isotropic tuples $(h_1, \dots, h_k) \in (\mathbb{F}_p^d)^k$ is at most $O_{d,k}(p^{kd-1})$.

A.3. **Some basic counting properties.** We refer the readers to Section 1.3 for the notations used in this section.

Lemma A.6 (Lemma 4.10 of [19])**.** Let $P \in \text{poly}(\mathbb{F}_p^d \to \mathbb{F}_p)$ be of degree at most r. Then $|V(P)| \leq O_{d,r}(p^{d-1})$ unless $P \equiv 0$.

Lemma A.7 (Lemma 4.14 and Corollary C.7 of [19])**.** Let $d \in \mathbb{N}_+, r, s \in \mathbb{N}$ and p be a prime. Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form and $V + c$ be an affine subspace of $\mathbb{F}_p^d$ of co-dimension r. If either $\text{rank}(M|_{V+c})$ or $\text{rank}(M) - 2r$ is at least $s^2 + s + 3$, then

$$|\Box_s(V(M) \cap (V + c))| = p^{(s+1)(d-r)-(\frac{s(s+1)}{2}+1)}(1 + O_s(p^{-1/2})).$$

In particular, If either $\text{rank}(M|_{V+c})$ or $\text{rank}(M) - 2r$ is at least 3, then

$$|V(M) \cap (V + c)| = p^{d-r-1}(1 + O(p^{-1/2})).$$

Lemma A.8 (Lemma 4.18 of [19])**.** Let $s \in \mathbb{N}$, $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form associated with the matrix A, and $V + c$ be an affine subspace of $\mathbb{F}_p^d$ of dimension r. For $n, h_1, \dots, h_s \in \mathbb{F}_p^d$, we have that $(n, h_1, \dots, h_s) \in \Box_s(V(M) \cap (V + c))$ if and only if

- $n \in V + c, h_1, \dots, h_s \in V$;
- $n \in V(M)^{h_1,\dots,h_s}$;
- $(h_i A) \cdot h_j = 0$ for all $1 \leq i, j \leq s, i \neq j$.

In particular, let $\phi\colon \mathbb{F}_p^r \to V$ be any bijective linear transformation, $M'(m) := M(\phi(m) + c)$. We have that $(n, h_1, \dots, h_s) \in \Box_s(V(M) \cap (V + c))$ if and only if $(n, h_1, \dots, h_s) = (\phi(n') + c, \phi(h_1'), \dots, \phi(h_s'))$ for some $(n', h_1', \dots, h_s') \in \Box_s(V(M'))$.

A.4. ***M*-families and *M*-sets.** We say that a linear transformation $L\colon (\mathbb{F}_p^d)^k \to (\mathbb{F}_p^d)^{k'}$ is *d-integral* if there exist $a_{i,j} \in \mathbb{F}_p$ for $1 \leq i \leq k$ and $1 \leq j \leq k'$ such that

$$L(n_1, \dots, n_k) = \Big(\sum_{i=1}^{k} a_{i,1} n_i, \dots, \sum_{i=1}^{k} a_{i,k'} n_i\Big)$$

for all $n_1, \dots, n_k \in \mathbb{F}_p^d$. Let $L\colon (\mathbb{F}_p^d)^k \to \mathbb{F}_p^d$ be a d-integral linear transformation given by $L(n_1, \dots, n_k) = \sum_{i=1}^k a_i n_i$ for some $a_i \in \mathbb{F}_p, 1 \leq i \leq k$. We say that L is the d-integral linear transformation *induced* by $(a_1, \dots, a_k) \in \mathbb{F}_p^k$.

Similarly, we say that a linear transformation $L\colon (\mathbb{Z}^d)^k \to (\frac{1}{p}\mathbb{Z}^d)^{k'}$ is *d-integral* if there exist $a_{i,j} \in \mathbb{Z}/p$ for $1 \leq i \leq k$ and $1 \leq j \leq k'$ such that

$$L(n_1, \dots, n_k) = \Big(\sum_{i=1}^{k} a_{i,1} n_i, \dots, \sum_{i=1}^{k} a_{i,k'} n_i\Big)$$

for all $n_1, \dots, n_k \in \mathbb{Z}^d$.

Let $d \in \mathbb{N}_+$, p be a prime, $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form with A being the associated matrix. We say that a function $F\colon (\mathbb{F}_p^d)^k \to \mathbb{F}_p$ is an *(M, k)-integral quadratic function* if

$$F(n_1, \dots, n_k) = \sum_{1 \leq i \leq j \leq k} b_{i,j} (n_i A) \cdot n_j + \sum_{1 \leq i \leq k} v_i \cdot n_i + u \quad \text{(A.1)}$$

for some $b_{i,j}, u \in \mathbb{F}_p$ and $v_i \in \mathbb{F}_p^d$. We say that an (M, k)-integral quadratic function F is *pure* if F can be written in the form of (A.1) with $v_1 = \cdots = v_k = \mathbf{0}$. We say that an (M, k)-integral quadratic function $F\colon (\mathbb{F}_p^d)^k \to \mathbb{F}_p$ is *nice* if

$$F(n_1, \dots, n_k) = \sum_{1\leq i\leq k'} b_i(n_{k'}A) \cdot n_i + u$$

for some $0 \leq k' \leq k$, $b_i, u \in \mathbb{F}_p$.

For F given in (A.1), denote

$$v_M(F) := (b_{k,k}, b_{k,k-1}, \dots, b_{k,1}, v_k, b_{k-1,k-1}, \dots, b_{k-1,1}, v_{k-1}, \dots, b_{1,1}, v_1, u) \in \mathbb{F}_p^{\binom{k+1}{2}+kd+1},$$

and

$$v'_M(F) := (b_{k,k}, b_{k,k-1}, \dots, b_{k,1}, v_k, b_{k-1,k-1}, \dots, b_{k-1,1}, v_{k-1}, \dots, b_{1,1}, v_1) \in \mathbb{F}_p^{\binom{k+1}{2}+kd}.$$

Informally, we say that $b_{i,j}$ is the n_in_j-coefficient, v_i is the n_i-coefficient, and u is the constant term coefficient for these vectors.

An (M, k)-*family* is a collections of (M, k)-integral quadratic functions. Let $\mathcal{J} = \{F_1, \dots, F_r\}$ be an (M, k)-family.

- We say that $\mathcal{J}$ is *pure* if all of $F_1, \dots, F_r$ are pure.
- We say that $\mathcal{J}$ is *consistent* if $(0, \dots, 0, 1)$ does not belong to the span of $v_M(F_1)$, $\dots$, $v_M(F_r)$, or equivalently, there is no linear combination of $F_1, \dots, F_r$ which is a constant nonzero function, or equivalently, for all $c_1, \dots, c_r \in \mathbb{F}_p$, we have
$$c_1v'_M(F_1) + \cdots + c_rv'_M(F_r) = \mathbf{0} \Rightarrow c_1v_M(F_1) + \cdots + c_rv_M(F_r) = \mathbf{0}.$$
- We say that $\mathcal{J}$ is *independent* if $v'_M(F_1), \dots, v'_M(F_r)$ are linearly independent, or equivalently, there is no nontrivial linear combination of $F_1, \dots, F_r$ which is a constant function, or equivalently, for all $c_1, \dots, c_r \in \mathbb{F}_p$, we have
$$c_1v'_M(F_1) + \cdots + c_rv'_M(F_r) = \mathbf{0} \Rightarrow c_1 = \cdots = c_r = 0.$$
- We say that $\mathcal{J}$ is *nice* if there exist some bijective d-integral linear transformation $L\colon (\mathbb{F}_p^d)^k \to (\mathbb{F}_p^d)^k$ and some $v \in (\mathbb{F}_p^d)^k$ such that $F_i(L(\cdot) + v)$ is nice for all $1 \leq i \leq r$.

The dimension of the span of $v'_M(F_1), \dots, v'_M(F_r)$ is called the *dimension* of an (M, k)-family $\{F_1, \dots, F_r\}$.

When there is no confusion, we call an (M, k)-family to be an M-*family* for short.

We say that a subset Ω of $(\mathbb{F}_p^d)^k$ is an M-*set* if there exists an (M, k)-family $\{F_i\colon (\mathbb{F}_p^d)^k \to \mathbb{F}_p^d\colon 1 \leq i \leq r\}$ such that $\Omega = \cap_{i=1}^r V(F_i)$. We call either $\{F_1, \dots, F_r\}$ or the ordered set $(F_1, \dots, F_r)$ an M-*representation* of Ω.

Let $\mathbf{P} \in$ {pure, nice, independent, consistent}. We say that Ω is $\mathbf{P}$ if one can choose the M-family $\{F_1, \dots, F_r\}$ to be $\mathbf{P}$. We say that the M-representation $(F_1, \dots, F_r)$ is $\mathbf{P}$ if $\{F_i\colon 1 \leq i \leq r\}$ is $\mathbf{P}$. We say that r is the *dimension* of the M-representation $(F_1, \dots, F_r)$. The *total co-dimension* of a consistent M-set Ω, denoted by $r_M(\Omega)$, is the minimum of the dimension of the independent M-representations of Ω. It was shown in Proposition C.8 of [19] that $r_M(\Omega)$ is independent of the choice of the independent M-representation if d and p are sufficiently large.

We say that an independent M-representation $(F_1,\dots,F_r)$ of Ω is *standard* if the matrix $\begin{bmatrix} v_M(F_1) \\ \dots \\ v_M(F_r) \end{bmatrix}$ is in the reduced row echelon form (or equivalently, the matrix $\begin{bmatrix} v'_M(F_1) \\ \dots \\ v'_M(F_r) \end{bmatrix}$ is in the reduced row echelon form). If $(F_1,\dots,F_r)$ is a standard M-representation of Ω, then we may relabeling $(F_1,\dots,F_r)$ as

$$(F_{k,1},\dots,F_{k,r_k},F_{k-1,1},\dots,F_{k-1,r_{k-1}},\dots,F_{1,1},\dots,F_{1,r_1})$$

for some $r_1,\dots,r_k\in\mathbb{N}$ such that $F_{i,j}$ is non-constant with respect to n_i and is independent of $n_{i+1},\dots,n_k$. We also call $(F_{i,j}\colon 1\leq i\leq k, 1\leq j\leq r_i)$ a *standard M-representation* of Ω. The vector $(r_1,\dots,r_k)$ is called the *dimension vector* of this representation.

Convention A.9. We allow (M,k)-families, M-representations and dimension vectors to be empty. In particular, $(\mathbb{F}_p^d)^k$ is considered as a nice and consistent M-set with total co-dimension zero.

Proposition A.10 (Proposition B.3 of [19])**.** Let $d,k,r\in\mathbb{N}_+$, p be a prime, $M\colon\mathbb{F}_p^d\to\mathbb{F}_p$ be a quadratic form and $\mathcal{J}=\{F_1,\dots,F_r\}$, $F_i\colon(\mathbb{F}_p^d)^k\to\mathbb{F}_p$, $1\leq i\leq r$ be an (M,k)-family.

(i) For any subset $I\subseteq\{1,\dots,r\}$, $\{F_i\colon i\in I\}$ is an (M,k)-family.

(ii) Let $I\subseteq\{1,\dots,r\}$ be a subset such that all of $F_i, i\in I$ are independent of the last d-variables. Then writing $G_i(n_1,\dots,n_{k-1}):=F_i(n_1,\dots,n_{k-1})$, we have that $\{G_i\colon i\in I\}$ is an $(M,k-1)$-family.

(iii) For any bijective d-integral linear transformation $L\colon(\mathbb{F}_p^d)^k\to(\mathbb{F}_p^d)^k$ and any $v\in(\mathbb{F}_p^d)^k$, $\{F_1(L(\cdot)+v),\dots,F_r(L(\cdot)+v)\}$, $F_i\colon(\mathbb{F}_p^d)^k\to\mathbb{F}_p$, $1\leq i\leq r$ is an (M,k)-family.

(iv) Let $1\leq k'\leq k$ and $1\leq r'\leq r$ be such that $F_1,\dots,F_{r'}$ are independent of $n_{k-k'+1},\dots,n_r$ and that every nontrivial linear combination of $F_{r'+1},\dots,F_r$ dependents nontrivially on some of $n_{k-k'+1},\dots,n_r$. Define $G_i,H_i\colon(\mathbb{F}_p^d)^{k+k'}\to\mathbb{F}_p$, $1\leq i\leq r$ by

$$G_i(n_1,\dots,n_{k+k'}):=F_i(n_1,\dots,n_k) \text{ and } H_i(n_1,\dots,n_{k+k'}):=F_i(n_1,\dots,n_{k-k'},n_{k+1},\dots,n_{k+k'}).$$

Then $\{F_1,\dots,F_{r'},G_{r'+1},\dots,G_r,H_{r'+1},\dots,H_r\}$[17] is an $(M,k+k')$-family.

Moreover, if $\mathcal{J}$ is nice, then so are the sets mentioned in Parts (i)-(iii); if $\mathcal{J}$ is consistent/independent, then so are the sets mentioned in Parts (i)-(iv); if $\mathcal{J}$ is pure, then so are the sets mentioned in Parts (i), (ii), (iv), and so is the set mentioned in Part (iii) when $v=\mathbf{0}$.

Proposition A.11 (Proposition C.8 of [19])**.** Let $d,k,r\in\mathbb{N}_+$, p be a prime, $M\colon\mathbb{F}_p^d\to\mathbb{F}_p$ be a quadratic form and $\Omega\subseteq(\mathbb{F}_p^d)^k$ be a consistent M-set. Suppose that $\Omega=\cap_{i=1}^r V(F_i)$ for some consistent (M,k)-family $\{F_1,\dots,F_r\}$. Let $1\leq k'\leq k$. Suppose that $\text{rank}(M)\geq 2r+1$ and $p\gg_{k,r}1$.

(i) The dimension of all independent M-representations of Ω equals to $r_M(\Omega)$.

(ii) We have that $r_M(\Omega)\leq r$, and that $r_M(\Omega)=r$ if the (M,k)-family $\{F_1,\dots,F_r\}$ is independent.

[17] Here we regard $F_1,\dots,F_{r'}$ as functions of $n_1,\dots,n_{k+k'}$, which acutally depends only on $n_1,\dots,n_{k-k'}$.

(iii) For all $I \subseteq \{1, \dots, r\}$, the set $\Omega' := \cap_{i\in I} V(F_i)$ is a consistent M-set such that $r_M(\Omega') \leq r_M(\Omega)$. Moreover, if the (M,k)-family $\{F_1, \dots, F_r\}$ is independent, then $r_M(\Omega') = |I|$.

(iv) For any bijective d-integral linear transformation $L\colon (\mathbb{F}_p^d)^k \to (\mathbb{F}_p^d)^k$ and $v \in (\mathbb{F}_p^d)^k$, we have that $L(\Omega) + v$ is a consistent M-set and that $r_M(L(\Omega) + v) = r_M(\Omega)$.

(v) Assume that Ω admits a standard M-representation with dimension vector $(r_1, \dots, r_k)$, Then for all $1 \leq k' \leq k$, the set
$$\{(n_1, \dots, n_{k+k'}) \in (\mathbb{F}_p^d)^{k+k'} \colon (n_1, \dots, n_k), (n_1, \dots, n_{k-k'}, n_{k+1}, \dots, n_{k+k'}) \in \Omega\}$$
admits a standard M-representation with dimension vector $(r_1, \dots, r_{k'}, r_{k+1}, \dots, r_{k'})$.

(vi) Assume that Ω admits a standard M-representation with dimension vector $(r_1, \dots, r_k)$. For $I = \{1, \dots, k'\}$ for some $1 \leq k' \leq k$, the I-projection Ω_I of Ω admits a standard M-representation with dimension vector $(r_1, \dots, r_{k'})$.

(vii) If Ω is a nice and consistent M-set or a pure and consistent M-set, then $r_M(\Omega) \leq \binom{k+1}{2}$.

A.5. **Fubini's theorem for M-sets.** Let $n_i = (n_{i,1}, \dots, n_{i,d}) \in \mathbb{F}_p^d$ denote a d-dimensional variable for $1 \leq i \leq k$. For convenience we denote $\mathbb{F}_p^d[n_1, \dots, n_k]$ to be the polynomial ring $\mathbb{F}_p[n_{1,1}, \dots, n_{1,d}, \dots, n_{k,1}, \dots, n_{k,d}]$. Let J, J', J'' be finitely generated ideals of the polynomial ring $\mathbb{F}_p^d[n_1, \dots, n_k]$ and $I \subseteq \{n_1, \dots, n_k\}$. Suppose that $J = J' + J''$, $J' \cap J'' = \{0\}$, all the polynomials in J' are independent of $\{n_1, \dots, n_k\}\backslash I$, and all the non-zero polynomials in J'' depend nontrivially on $\{n_1, \dots, n_k\}\backslash I$. Then we say that J' is an *I-projection* of J and that (J', J'') is an *I-decomposition* of J. It was proved in Proposition C.1 of [19] that the I-projection of J exists and is unique.

Let $\mathcal{J}$, $\mathcal{J}'$ and $\mathcal{J}''$ be finite subsets of $\mathbb{F}_p^d[n_1, \dots, n_k]$, and $I \subseteq \{n_1, \dots, n_k\}$. Let J, J', and J'' be the ideals generated by $\mathcal{J}$, $\mathcal{J}'$ and $\mathcal{J}''$ respectively. If J' is an I-projection of J with (J', J'') being an I-decomposition of J, then we say that $\mathcal{J}'$ is an *I-projection* of $\mathcal{J}$ and that $(\mathcal{J}', \mathcal{J}'')$ is an *I-decomposition* of $\mathcal{J}$. Note that the I-projection $\mathcal{J}'$ of $\mathcal{J}$ is not unique. However, by Proposition C.1 of [19], the ideal generated by $\mathcal{J}'$ and the set of zeros $V(\mathcal{J}')$ are unique.

If I is a subset of $\{1, \dots, k\}$, for convenience we say that J' is an *I-projection* of J if J' is an $\{n_i \colon i \in I\}$-projection of J. Similarly, we say that (J', J'') is an *I-decomposition* of J if (J', J'') is an $\{n_i \colon i \in I\}$-decomposition of J. Here J, J', J'' are either ideals or finite subsets of the polynomial ring.

We now define the projections of M-sets. Let $\mathcal{J} \subseteq \mathbb{F}_p[n_1, \dots, n_k]$ be a consistent (M,k)-family and let $\Omega = V(\mathcal{J}) \subseteq (\mathbb{F}_p^d)^k$. Let $I \cup J$ be a partition of $\{1, \dots, k\}$ (where I and J are non-empty), and $(\mathcal{J}_I, \mathcal{J}_I')$ be an I-decomposition of $\mathcal{J}$. Let Ω_I denote the set of $(n_i)_{i\in I} \in (\mathbb{F}_p^d)^{|I|}$ such that $f(n_1, \dots, n_k) = 0$ for all $f \in \mathcal{J}_I$ and $(n_j)_{j\in J} \in (\mathbb{F}_p^d)^{|J|}$. Note that all $f \in \mathcal{J}_I$ are independent of $(n_j)_{j\in J}$, and that Ω_I is independent of the choice of the I-decomposition. We say that Ω_I is an *I-projection* of Ω.

For $(n_i)_{i\in I} \in (\mathbb{F}_p^d)^{|I|}$, let $\Omega_I((n_i)_{i\in I})$ be the set of $(n_j)_{j\in J} \in (\mathbb{F}_p^d)^{|J|}$ such that $f(n_1, \dots, n_k) = 0$ for all $f \in \mathcal{J}_I'$. By construction, for any $(n_i)_{i\in I} \in \Omega_I$, we have that $(n_j)_{j\in J} \in \Omega_I((n_i)_{i\in I})$ if and only if $(n_1, \dots, n_k) \in \Omega$. So for all $(n_i)_{i\in I} \in \Omega_I$, $\Omega_I((n_i)_{i\in I})$ is independent of the choice of the I-decomposition.

Theorem A.12 (Theorem C.3 of [19])**.** Let $d, k \in \mathbb{N}_+$, $r_1, \dots, r_k \in \mathbb{N}$ and p be a prime. Set $r := r_1 + \dots + r_k$. Let $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form with $\mathrm{rank}(M) \geq 2r + 1$, and $\Omega \subseteq (\mathbb{F}_p^d)^k$ be a consistent M-set admitting a standard M-representation of Ω with dimension vector $(r_1, \dots, r_k)$.

(i) We have $|\Omega| = p^{dk-r}(1 + O_{k,r}(p^{-1/2}))$;

(ii) If $I = \{1, \dots, k'\}$ for some $1 \leq k' \leq k - 1$, then Ω_I is a consistent M-set admitting a standard M-representation with dimension vector $(r_1, \dots, r_{k'})$. Moreover, for all but at most $(k - k')r'_{k'} p^{dk'+r'_{k'}-1-\mathrm{rank}(M)}$ many $(n_i)_{i\in I} \in (\mathbb{F}_p^d)^{k'}$, we have that $\Omega_I((n_i)_{i\in I})$ has a standard M-representation with dimension vector $(r_{k'+1}, \dots, r_k)$ and that $|\Omega_I((n_i)_{i\in I})| = p^{d(k-k')-(r_{k'+1}+\dots+r_k)}(1 + O_{k,r}(p^{-1/2}))$, where $r'_{k'} := \max_{k'+1\leq i\leq k} r_i$.

(iii) If $k \geq 2$, and $I \cup J$ is a partition of $\{1, \dots, k\}$ (where I and J are non-empty), then for any function $f\colon \Omega \to \mathbb{C}$ with norm bounded by 1, we have that

$$\mathbb{E}_{(n_1,\dots,n_k)\in\Omega} f(n_1, \dots, n_k) = \mathbb{E}_{(n_i)_{i\in I}\in\Omega_I} \mathbb{E}_{(n_j)_{j\in J}\in\Omega_I((n_i)_{i\in I})} f(n_1, \dots, n_k) + O_{k,r}(p^{-1/2}).$$

A.6. **Intrinsic definitions for polynomials.** We refer the readers to Section 1.3 for the notations used in this section.

Proposition A.13 (Corollary 7.6 of [19])**.** Let $d, k, s \in \mathbb{N}_+$, $p \gg_d 1$ be a prime, and $M\colon \mathbb{Z}^d \to \mathbb{Z}/p$ be a p-non-degenerate quadratic form associated with the matrix A which induces a non-degenerate quadratic form $\tilde{M}\colon \mathbb{F}_p^d \to \mathbb{F}_p$. Let $m_1, \dots, m_k \in \mathbb{Z}^d$ be such that $\iota(m_1), \dots, \iota(m_k) \in \mathbb{F}_p^d$ are linearly independent and $\tilde{M}$-isotropic. If $d \geq k + s + 3$, then for any polynomial $g \in \mathrm{poly}_p(\mathbb{Z}^d \to \mathbb{Z}/p|\mathbb{Z})$ of degree at most s, the following are equivalent:

(i) for all $(n, h_1, \dots, h_s) \in \Box_{p,s}(V_p(M)^{m_1,\dots,m_k})$, we have that $\Delta_{h_s} \dots \Delta_{h_1} g(n) \in \mathbb{Z}$;

(ii) we have

$$g(n) = M(n)g_0(n) + \sum_{i=1}^{k} (M(n + h_i) - M(n))g_i(n) + g'(n) + g''(n)$$

for some homogeneous integer coefficient polynomials $g_0, \dots, g_k \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$ with $\deg(g_0) = s - 2$ and $\deg(g_1) = \dots = \deg(g_k) = s - 1$, some integer valued polynomial $g'' \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$ with $\deg(g'') = s$, and some $g' \in \mathrm{poly}_p(\mathbb{Z}^d \to \mathbb{Z}/p|\mathbb{Z})$ of degree at most $s - 1$;

(iii) we have

$$g(n) = \frac{1}{p}((nA) \cdot n)g_0(n) + \sum_{i=1}^{k} \frac{1}{p}((nA) \cdot h_i)g_i(n) + g'(n) + g''(n)$$

for some homogeneous integer coefficient polynomials $g_0, \dots, g_k \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$ with $\deg(g_0) = s - 2$ and $\deg(g_1) = \dots = \deg(g_k) = s - 1$, some integer valued polynomial $g'' \in \mathrm{poly}(\mathbb{Z}^d \to \mathbb{Z})$ with $\deg(g'') = s$, and some $g' \in \mathrm{poly}_p(\mathbb{Z}^d \to \mathbb{Z}/p|\mathbb{Z})$ of degree at most $s - 1$.

A.7. **Strong factorization property for nice M-sets.** Let Ω be a non-empty finite set and G/Γ be an $\mathbb{N}$-filtered nilmanifold. A sequence $\mathcal{O}\colon \Omega \to G/\Gamma$ is *δ-equidistributed* on G/Γ if for all Lipschitz function $F\colon G/\Gamma \to \mathbb{C}$, we have that

$$\Big|\frac{1}{|\Omega|}\sum_{n\in\Omega} F(\mathcal{O}(n)) - \int_{G/\Gamma} F\, dm_{G/\Gamma}\Big| \leq \delta\|F\|_{\mathrm{Lip}},$$

where $m_{G/\Gamma}$ is the Haar measure of G/Γ.

Theorem A.14 (Theorem 11.2 of [19])**.** Let $d, k \in \mathbb{N}_+, r, s \in \mathbb{N}$, $C > 0$, $\mathcal{F}\colon \mathbb{R}_+ \to \mathbb{R}_+$ be a growth function, $p \gg_{C,d,\mathcal{F},k,r,s} 1$ be a prime, $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form, and $\Omega \subseteq (\mathbb{F}_p^d)^k$ satisfying one of the following assumptions:

(i) $r = 0$ and $\Omega = (\mathbb{F}_p^d)^k$;
(ii) $k = 1, r = 0$ and $\Omega = V(M) \cap (V + c)$ for some affine subspace $V + c$ of $\mathbb{F}_p^d$ with $\mathrm{rank}(M|_{V+c}) \geq s + 13$;
(iii) M is non-degenerate, $d \geq \max\{4r + 1, 4k + 3, 2k + s + 11\}$ and Ω is a nice and consistent M-set of total co-dimension r.

Let G/Γ be an s-step $\mathbb{N}$-filtered nilmanifold of complexity at most C, and let $g \in \mathrm{poly}_p(\Omega \to G_{\mathbb{N}}|\Gamma)$. There exist some $C \leq C' \leq O_{C,d,\mathcal{F},k,r,s}(1)$, a proper subgroup G' of G which is C'-rational relative to $\mathcal{X}$, and a factorization

$$g(n) = \epsilon g'(n)\gamma(n) \text{ for all } n \in \mathbb{F}_p^d$$

such that $\epsilon \in G$ is of complexity $O_{C'}(1)$, $g' \in \mathrm{poly}_p(\Omega \to G'_{\mathbb{N}}|\Gamma')$, $g'(\mathbf{0}) = id_G$ and $(g'(n)\Gamma)_{n\in\Omega}$ is $\mathcal{F}(C')^{-1}$-equidistributed on G'/Γ', where $\Gamma' := G'\cap\Gamma$, and that $\gamma \in \mathrm{poly}(\Omega \to G_{\mathbb{N}}|\Gamma)$, $\gamma(\mathbf{0}) = id_G$.

A.8. **Additive combinatorial properties for M-ideals.** Let $d \in \mathbb{N}_+, p$ be a prime and $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form associated with the matrix A. We say that an ideal I of the polynomial ring $\mathbb{F}_p[x_1, \dots, x_d]$ is an *M-ideal* if $I = \langle (nA)\cdot n, (h_1A)\cdot n, \dots, (h_kA)\cdot n\rangle$ (which are viewed as polynomials in the variable n) for some $k \in \mathbb{N}$ and $h_1, \dots, h_k \in \mathbb{F}_p^d$. Denote

$$J^M := \langle (nA)\cdot n\rangle \text{ and } J^M_{h_1,\dots,h_k} := \langle (nA)\cdot n, (h_1A)\cdot n, \dots, (h_kA)\cdot n\rangle.$$

For a subspace V of $\mathbb{F}_p^d$, denote

$$J^M_V := \langle (nA)\cdot n, (hA)\cdot n\colon h \in V\rangle.$$

Definition A.15 (Almost linear function and Freiman homomorphism)**.** Let $d, K \in \mathbb{N}_+$, p be a prime, H be a subset of $\mathbb{F}_p^d$, and $\xi\colon H \to \mathbb{Z}/p$ be a map. We say that ξ is a *p-almost linear function* of *complexity* at most K if there exist $\alpha_i \in (\mathbb{Z}/p)^d$ and $\beta_i \in \mathbb{Z}/p$ for all $1 \leq i \leq K$ such that for all $h \in H$,

$$\xi(h) := \sum_{i=1}^{K}\{\alpha_i\cdot\tau(h)\}\beta_i.$$ [18]

[18] In particular, $\sum_{i=1}^{K}\{\alpha_i\cdot\tau(h)\}\beta_i$ is required to take values in $\mathbb{Z}/p$.

We say that a p-almost linear function ξ is a *p-almost linear Freiman homomorphism* if for all $h_1, h_2, h_3, h_4 \in H$ with $h_1 + h_2 = h_3 + h_4$, we have that

$$\xi(h_1) + \xi(h_2) \equiv \xi(h_3) + \xi(h_4) \mod \mathbb{Z}.$$

We say that a map $\xi\colon H \to \mathbb{F}_p$ is an *almost linear function/Freiman homomorphism* of complexity at most K if there exists a p-almost linear function/Freiman homomorphism $\xi'\colon H \to \mathbb{Z}/p$ of complexity at most K such that $\xi(h) = \iota(p\xi'(h))$ for all $h \in H$.

Theorem A.16 (Theorem 1.5 of [20])**.** Let $d, s \in \mathbb{N}_+$, with $d \geq N(s)$ (recall (1.1) for the definition), $\delta > 0$ and p be a prime, $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a non-degenerate quadratic form, $H \subseteq \mathbb{F}_p^d$ with $|H| > \delta p^d$, and ξ_i be a map from H to $\mathrm{HP}_d(s)$ for $1 \leq i \leq 4$. Suppose that there exists $U \subseteq \{(h_1, h_2, h_3, h_4) \in H^4\colon h_1 + h_2 = h_3 + h_4\}$ with $|U| > \delta|H|^3$ such that for all $(h_1, h_2, h_3, h_4) \in U$,

$$\xi_1(h_1) + \xi_2(h_2) - \xi_3(h_3) - \xi_4(h_4) \in J^M_{h_1,h_2,h_3}.$$

If $p \gg_{\delta,d} 1$, then there exist $H' \subseteq H$ with $|H'| \gg_{\delta,d} p^d$, some $g \in \mathrm{HP}_d(s)$, and an almost linear Freiman homomorphism $T\colon H' \to \mathbb{F}_p^d$ of complexity $O_{\delta,d}(1)$ such that

$$\xi_1(h) - T(h) - g \in J^M_h \text{ for all } h \in H'.$$

Let $V_1, \dots, V_k$ be subspaces of $\mathbb{F}_p^d$. We say that $V_1, \dots, V_k$ are *linearly independent* if for all $v_i \in V_i, 1 \leq i \leq k$, $\sum_{i=1}^k v_i = \mathbf{0}$ implies that $v_1 = \cdots = v_k = \mathbf{0}$. In this case we also say that $(V_1, \dots, V_k)$ is a *linearly independent tuple* (or a *linearly independent pair* if $k = 2$).

Let $V_1, \dots, V_k$ be subspaces of $\mathbb{F}_p^d$ and $v_1, \dots, v_{k'} \in \mathbb{F}_p^d$. We say that $v_1, \dots, v_{k'}, V_1, \dots, V_k$ are *linearly independent* if for all $c_i \in \mathbb{F}_p, 1 \leq i \leq k$ and $v'_{i'} \in V_{i'}, 1 \leq i' \leq k'$, $\sum_{i=1}^k c_i v_i + \sum_{i'=1}^{k'} v'_{i'} = \mathbf{0}$ implies that $c_1 = \cdots = c_k = 0$ and $v'_1 = \cdots = v'_{k'} = \mathbf{0}$. In this case we also say that $(v_1, \dots, v_{k'}, V_1, \dots, V_k)$ is a *linearly independent tuple* (or a *linearly independent pair* if $k + k' = 2$).

Proposition A.17 (Propositions 3.6 and 3.7 of [20])**.** Let $m, s \in \mathbb{N}$, $d, N, r \in \mathbb{N}_+$, $p \gg_d 1$ be a prime, $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a non-degenerate quadratic form, and $V, V_1, \dots, V_N$ be subspaces of $\mathbb{F}_p^d$ such that $\dim(V) = m$ and $\dim(V_i) \leq r, 1 \leq i \leq N$. Suppose that either

- $V_1, \dots, V_N$ are linearly independent and $N \geq s + m + 1$; or
- $V, V_1, \dots, V_N$ are linearly independent and $N \geq s + 1$,

and that either $\mathrm{rank}(M|_{V^{\perp_M}}) \geq 2N(r - 1) + 7$ or $d \geq 2m + 2N(r - 1) + 7$. Then for all $f \in \mathbb{F}_p[x_1, \dots, x_d]$ of degree at most s, we have that

$$f \in \cap_{i=1}^N J^M_{V+V_i} \Leftrightarrow f \in J^M_V.$$

Proposition A.18 (A special case of Proposition 4.6 of [20])**.** Let $d \in \mathbb{N}_+$, $k, s \in \mathbb{N}$, $\delta > 0$, $p \gg_{\delta,d,k} 1$ be a prime, H be a subset of $\mathbb{F}_p^d$ with $|H| > \delta p^d$ and $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a non-degenerate quadratic form. Let $F\colon \mathbb{F}_p^d \to \mathrm{HP}_d(s)$ be a map such that

$$F(x) \equiv F(y) \mod J^M_{x,y}$$

for all $x, y \in H$. If $d \geq s + 8$, then there exists $G \in \mathrm{HP}_d(s)$ such that

$$F(x) \equiv G \mod J_x^M$$

for all $x \in H$.

Appendix B. Approximation properties for nilsequences

In this section, we provide some useful properties which allow us to approximate a nilsequence or nilcharacter by one with better properties.

The following lemma demonstrates the type of operations we can apply to nilcharacters. This is a generalization of Lemma E.3 of [12]. Compared with Lemma E.3, our statement needs to take the periodicity property of nilcharacters into consideration, and our statement is more quantitative since we do not use the non-standard analysis approach as in [12].

Lemma B.1. Let $d, D, k, k' \in \mathbb{N}_+$, $C > 0$, I be the degree, multi-degree, or degree-rank ordering with $\dim(I)|k$, and J be a finite downset of I. Let $\Omega \subseteq \mathbb{F}_p^k$ and $\chi, \chi' \in \Xi_p^{J;C,D}(\Omega)$. Then for all $h \in \mathbb{F}_p^k$ and $q \in \mathbb{F}_p\backslash\{0\}$, we have that $\chi \otimes \chi' \in \Xi_p^{J;2C,D^2}(\Omega)$, $\chi(\cdot + h) \in \Xi_p^{J;C,D}(\Omega - h)$, $\chi(q\cdot) \in \Xi_p^{J;C,D}(q^{-1}\Omega)$ and $\overline{\chi} \in \Xi_p^{J;C,D}(\Omega)$.

If I is the degree filtration, then for all linear transformation $L \colon \mathbb{F}_p^{k'} \to \mathbb{F}_p^k$, we have that $\chi \circ L \in \Xi_p^{J;C,D}(L^{-1}(\Omega))$.

Proof. Let $((G/\Gamma)_I, g, F, \eta)$ and $((G'/\Gamma')_I, g', F', \eta')$ be $\Xi_p^J(\Omega)$-representations of χ and χ' respectively of complexities at most C and dimensions D', D'' respectively. Then

$$\chi \otimes \chi'(n) = F \otimes F'(g(\tau(n))\Gamma, g'(\tau(n))\Gamma')$$

for all $n \in \Omega$. Note that $(G \times G')/(\Gamma \times \Gamma')$ is an I-filtered nilmanifold of complexity at most $2C$ and degree $\subseteq J$, and $F \otimes F'$ is a function taking values in $\mathbb{S}^{D'D''}$ with Lipschitz norm bounded by $2C$ and has a vertical frequency (η, η') of complexity at most $2C$. Moreover, since $g \in \mathrm{poly}_p(\mathbb{Z}^k \to G_I|\Gamma)$ and $g' \in \mathrm{poly}_p(\mathbb{Z}^k \to G'_I|\Gamma')$, we have that $(g(\cdot), g'(\cdot)) \in \mathrm{poly}_p(\mathbb{Z}^k \to (G \times G')_I|\Gamma \times \Gamma')$. So $\chi \otimes \chi' \in \Xi_p^{J;2C,D^2}(\Omega)$.

Since $g \circ \tau \in \mathrm{poly}_p(\mathbb{F}_p^k \to G_I|\Gamma)$, by Proposition 2.18, there exist $g', g'' \in \mathrm{poly}_p(\mathbb{F}_p^k \to G_I|\Gamma)$ such that $g'(n)\Gamma = g \circ \tau(n + h)\Gamma$ for all $n \in \Omega - h$ and that $g''(n)\Gamma = g \circ \tau(qn)\Gamma$ for all $n \in q^{-1}\Omega$. From this it is not hard to see that $\chi(\cdot + h) \in \Xi_p^{J;C,D}(\Omega - h)$ and $\chi(q\cdot) \in \Xi_p^{J;C,D}(q^{-1}\Omega)$. We also have $\overline{\chi} \in \Xi_p^{J;C,D}(\Omega)$ since $\overline{\chi}(n) = \overline{F}(g(n)\Gamma)$ for all $n \in \Omega$.

If I is the degree filtration, then by Proposition 2.18, there exists $g''' \in \mathrm{poly}_p(\mathbb{F}_p^k \to G_I|\Gamma)$ such that $g'''(n)\Gamma = g \circ \tau \circ L(n)\Gamma$ for all $n \in L^{-1}(\Omega)$. From this it is not hard to see that $\chi \circ L \in \Xi_p^{J;C,D}(L^{-1}(\Omega))$. □

Given any non-periodic nilsequence, the method of Manners [17] allows us to approximate it with a p-periodic one. We summarize this approach in the following theorem, and tailor it for our purposes:

Theorem B.2 (Manners' approximation theorem)**.** Let $D, k \in \mathbb{N}_+$, $C > 0$, p be a prime, I be the degree, multi-degree, or degree-rank ordering with $\dim(I)|k$, G/Γ be a nilmanifold with flirtation G_I of complexity at most C, $g \in \text{poly}(\mathbb{Z}^k \to G_I)$ be a polynomial, and $F \in \text{Lip}(G/\Gamma \to \mathbb{C}^D)$ be of Lipschitz norm at most C. Let $\phi\colon \mathbb{T}^k \to [0,1]$ be a smooth function supported on a box of edge length $1/2$. Then there exist a nilmanifold $\tilde{G}/\tilde{\Gamma}$ with flirtation $\tilde{G}_I$ of complexity at most $O_{C,k}(1)$ and degree the same as G/Γ, a p-periodic polynomial $\tilde{g} \in \text{poly}_p(\mathbb{Z}^k \to \tilde{G}_I|\tilde{\Gamma})$, and a function $\tilde{F} \in \text{Lip}(\tilde{G}/\tilde{\Gamma} \to \mathbb{C}^D)$ of Lipschitz norm at most $O_{C,k}(1)$ such that

$$\tilde{F}(\tilde{g}(n)\tilde{\Gamma}) = \phi(n/p) \otimes F(g(n \mod p\mathbb{Z}^k)\Gamma)$$

for all $n \in \mathbb{Z}^k$, where $n \mod p\mathbb{Z}^k$ is the representative of n in $\{0,\dots,p-1\}^k$.

Proof. Consider the polynomial group $\text{poly}(\mathbb{R}^k \to G_I)$ with the pointwise multiplication. By Lemma 2.20, there is an isomorphism $\text{poly}(\mathbb{Z}^k \to G_I) \simeq \text{poly}(\mathbb{R}^k \to G_I)$ given by restrictions. Similar to Lemma C.1 of [12], $\text{poly}(\mathbb{Z}^k \to \Gamma_I)$ is a discrete and cocompact subgroup of $\text{poly}(\mathbb{R}^k \to G_I)$, where Γ_I is the filtration induced by G_I.

Let $T\colon \mathbb{R}^k \to \text{poly}(\mathbb{R}^k \to G_I)$, $T(t,p) := p(\cdot + t)$ denote the shift action. Then T induces a semi-product structure $\tilde{G} := \mathbb{R}^k \ltimes_T \text{poly}(\mathbb{R}^k \to G_I)$ with the group operation given by

$$(t,g) * (t',g') := (t+t', g' \cdot T(t',g)).$$

Using the type-I Taylor expansion in Lemma 2.20, it is clear that the group $\text{poly}(\mathbb{R}^k \to G_I)$ is connected and simply connected. So $\tilde{G}$ is connected and simply connected. Similar to the argument in Theorem 1.5 of [17], $\tilde{\Gamma} := \mathbb{Z}^k \ltimes_T \text{poly}(\mathbb{R}^k \to \Gamma_I)$ is a discrete and cocompact subgroup of $\tilde{G}$.

We now define a filtration $\tilde{G}_I$ as follows. For $i \in I$, let G_I^{+i} denote the filtration given by $G_j^{+i} := G_{i+j}$ for all $j \in I$. If I is the multi-degree ordering, then set $\tilde{G}_0 = \tilde{G}$, $\tilde{G}_i = \mathbb{R}^k \ltimes_T \text{poly}(\mathbb{R}^k \to G_I^{+i})$ for $|i| = 1$ and $\tilde{G}_i = \{\mathbf{0}\} \ltimes_T \text{poly}(\mathbb{R}^k \to G_I^{+i})$ for $|i| \geq 2$. If I is the degree ordering, then set $\tilde{G}_0 = \tilde{G}$, $\tilde{G}_i = \mathbb{R}^k \ltimes_T \text{poly}(\mathbb{R}^k \to G_I^{+i})$ for $i = 1$ and $\tilde{G}_i = \{\mathbf{0}\} \ltimes_T \text{poly}(\mathbb{R}^k \to G_I^{+i})$ for $i \geq 2$. If I is the degree-rank ordering, then let $G_{\mathbb{N}}$ be the degree filtration generated by $G_{[i,0]}$ and set $\tilde{G}'_0 = \tilde{G}$, $\tilde{G}'_i = \mathbb{R}^k \ltimes_T \text{poly}(\mathbb{R}^k \to G_{\mathbb{N}}^{+i})$ for $i = 1$ and $\tilde{G}'_i = \{\mathbf{0}\} \ltimes_T \text{poly}(\mathbb{R}^k \to G_{\mathbb{N}}^{+i})$ for $i \geq 2$. We then set $\tilde{G}_{\text{DR}}$ to be the degree-rank filtration induced by $\tilde{G}'_{\mathbb{N}}$. In all the cases, $\tilde{G}_I$ is an I-filtration and defines a nilmanifold $\tilde{G}/\tilde{\Gamma}$ of complexity $O_{C,k}(1)$ and having the same degree as G/Γ.

Let $h \in \text{poly}(\mathbb{R}^k \to G_I)$ be the rescaling of g given by $h(n) := g(pn)$ (where we view g as a polynomial in $\text{poly}(\mathbb{R}^k \to G_I)$). Let $\tilde{g}\colon \mathbb{Z}^k \to \tilde{G}$ be the map given by

$$\tilde{g}(n) := (\mathbf{0}, h(n)) * (n/p, id_G).$$

It is not hard to compute that $\tilde{g}$ belongs to $\text{poly}(\mathbb{Z}^k \to \tilde{G}_I)$ in all the cases (recall that $\text{poly}(\mathbb{Z}^k \to G_{\text{DR}}) = \text{poly}(\mathbb{Z}^k \to (G_{[i,0]})_{i\in\mathbb{N}})$ when I is the degree-rank ordering). Moreover, since

$$\tilde{g}(n+pm) = (\mathbf{0},h) * (n/p+m, id_G) = (\mathbf{0},h) * (n/p, id_G) * (m, id_G) \in \tilde{g}(n)\tilde{\Gamma}$$

for all $m, n \in \mathbb{Z}^k$, we have that $\tilde{g}$ is p-periodic.

Finally, we define $\tilde{F}$ on the fundamental domain $[0,1)^k \ltimes_T \text{poly}(\mathbb{R}^k \to G_I)/\text{poly}(\mathbb{Z}^k \to \Gamma_I)$ by

$$\tilde{F}(t, g \cdot \text{poly}(\mathbb{Z}^k \to \Gamma_I)) := \phi(t) \otimes F(g(\mathbf{0})\Gamma)$$

(which is obviously well defined), and extend periodically by $\tilde{\Gamma}$, namely

$$\tilde{F}(t, g\cdot \mathrm{poly}(\mathbb{Z}^k \to \Gamma_I)) := \phi(\{t\}) \otimes F(g(-\lfloor t \rfloor)\Gamma)$$

for $t \in \mathbb{R}^k$. It is clear that the Lipschitz norm of $\tilde{F}$ is $O_{C,k}(1)$. Then

$$\begin{aligned}\tilde{F}(\tilde{g}(n)\tilde{\Gamma}) &= \tilde{F}((n/p, h(\cdot + n/p))\tilde{\Gamma}) \\ &= \phi(\{n/p\}) \otimes F(h(n/p - \lfloor n/p \rfloor)) = \phi(\{n/p\}) \otimes F(g(n \mod p\mathbb{Z}^k)).\end{aligned} \tag{B.1}$$

This finishes the proof. ◻

As a consequence of Theorem B.2, we have the following result which allows us to approximate nilsequences by periodic ones:

Corollary B.3 (Approximating a nilsequence by periodic ones)**.** Let $C > 0$, $D, k \in \mathbb{N}_+$, I be the degree, multi-degree, or degree-rank ordering with $\dim(I)|k$, J be a down set of I, and p be a prime. For any $\Omega \subseteq \mathbb{Z}^k$ and $\phi \in \mathrm{Nil}^{J;C,D}(\Omega)$, there exist $\phi_1, \dots, \phi_{20^k} \in \mathrm{Nil}_p^{J;O_{C,k}(1),D}(\Omega)$ such that for all $n \in \Omega$, we have

$$\phi(n \mod p\mathbb{Z}^k) = \sum_{m=1}^{20^k} \phi_m(n).$$

Similarly, for any prime p, any subset $\Omega' \subseteq \mathbb{F}_p^k$ and any $\phi' \in \mathrm{Nil}^{J;C,D}(\Omega')$, there exist $\phi'_1, \dots, \phi'_{20^k} \in \mathrm{Nil}_p^{J;O_{C,k}(1),D}(\Omega')$ such that for all $n \in \Omega'$, we have

$$\phi'(n) = \sum_{m=1}^{20^k} \phi'_m(n).$$

Proof. Since any nilsequence defined on Ω naturally extends to a nilsequence defined on $\mathbb{Z}^k$, we may assume without loss of generality that $\Omega = \mathbb{Z}^k$. Similarly, we may assume without loss of generality that $\Omega' = \mathbb{F}_p^k$.

We choose 20^k closed boxes $I_1, \dots, I_{20^k}$ in $\mathbb{T}^k$, each of edge length exactly $1/10$, whose interiors cover $\mathbb{T}^k$. Let $\rho_m, 1 \leq m \leq 20^k$ be a smooth partition of unity on $\mathbb{T}^k$ adapted to $I_m, 1 \leq m \leq 20^k$. Then each ρ_m is supported on a box J_m of edge $1/10$.

Denote

$$\phi_m(n) := \rho_m(\{n/p\}) \otimes \phi(n \mod p\mathbb{Z}^k)$$

for all $1 \leq m \leq 20^k$. Then

$$\sum_{m=1}^{20^k} \phi_m(n) = \phi(n \mod p\mathbb{Z}^k).$$

By Theorem B.2, we have that $\phi_m \in \mathrm{Nil}_p^{J;O_{C,k}(1),D}(\mathbb{Z}^k)$.

For any $\phi' \in \mathrm{Nil}^{J;C,D}(\mathbb{F}_p^k)$, we may write $\phi' = \phi \circ \tau$ for some $\phi \in \mathrm{Nil}^{J;C,D}(\mathbb{Z}^k)$. So there exist $\phi_1, \dots, \phi_{20^k} \in \mathrm{Nil}_p^{J;O_{C,k}(1),D}(\mathbb{Z}^k)$ such that

$$\phi(n \mod p\mathbb{Z}^k) = \sum_{m=1}^{20^k} \phi_m(n)$$

for all $n \in \mathbb{Z}^k$. Then $\phi_1 \circ \tau, \dots, \phi_{20^k} \circ \tau \in \mathrm{Nil}_p^{J;O_{C,k}(1),D}(\mathbb{F}_p^k)$ and

$$\phi'(n) = \phi(\tau(n)) = \sum_{m=1}^{20^k} \phi_m(\tau(n)) = \sum_{m=1}^{20^k} \phi'_m(n)$$

for all $n \in \mathbb{F}_p^k$. □

The next lemma allows us to approximate nilsequences of degree $J \cup J'$ by the combinations of (p-periodic) nilsequences of degrees J and J'.

Lemma B.4 (Splitting a nilsequence into nilsequences of smaller degrees)**.** Let $k \in \mathbb{N}_+$, $C, \epsilon > 0$, I be the degree, multi-degree, or degree-rank ordering with $\dim(I)|k$, J, J' be finite downsets of I, p be a prime, and $\Omega \subseteq \mathbb{F}_p^k$. For any $\psi \in \mathrm{Nil}^{J \cup J';C,1}(\Omega)$,

$$\Big\|\psi(n) - \sum_{j=1}^{K} \psi_j(n)\psi'_j(n)\Big\|_{\ell^\infty(\Omega)} < \epsilon$$

for some $K := K(C, \epsilon, J, J', k) \in \mathbb{N}_+$, $\psi_j \in \mathrm{Nil}_p^{J;O_{C,\epsilon,J,J',k}(1),1}(\Omega)$ and $\psi'_j \in \mathrm{Nil}_p^{J';O_{C,\epsilon,J,J',k}(1),1}(\Omega)$ for all $1 \leq j \leq K$.

Proof. The idea of the proof is similar to Lemma E.4 of [12]. However, our case is more intricate in the sense that we need a more quantitative approximation, and that we need to make the nilsequences to be p-periodic.

By Corollary B.3, it suffices to prove this lemma under the milder restriction that $\psi_j \in \mathrm{Nil}^{J;O_{C,\epsilon,J,J',k}(1),1}(\Omega)$ and $\psi'_j \in \mathrm{Nil}^{J';O_{C,\epsilon,J,J',k}(1),1}(\Omega)$, which does not require p-periodicity. Let $((G/\Gamma)_I, g, F)$ be a $\mathrm{Nil}^{J \cup J'}(\Omega)$-representation of ψ of complexity at most C and dimension 1. For each $j \in J \cup J'$, let $e_{j,1}, \dots, e_{j,d_j}$ be a basis of generators for Γ_j. We may then lift G to the *universal nilpotent Lie group* that is formally generated by $e_{j,1}, \dots, e_{j,d_j}$ under the only restriction that any iterated commutator of $e_{j_1,i_1}, \dots, e_{j_r,i_r}$ is trivial if $j_1 + \cdots + j_r \notin J \cup J'$. Similarly, we can lift Γ, F and g (with the help of Corollary B.10 of [12]). Therefore, since the universal nilpotent Lie group is $O(C)$-rational relative to $\mathcal{X}$, we may assume without loss of generality that G is universal.

The degree $\subseteq J \cup J'$ nilmanifold G/Γ projects down to the degree $\subseteq J$ nilmanifold $G/G_{>J}\Gamma$, which is $O_{C,J,J'}(1)$-rational relative to $\mathcal{X}$, where $G_{>J}$ is the group generated by the G_j for all $j \in J' \backslash J$. Similarly, we have a projection from G/Γ to the degree $\subseteq J'$ nilmanifold $G/G_{>J'}\Gamma$ which is $O_{C,J,J'}(1)$-rational relative to $\mathcal{X}$. The algebras $\mathrm{Lip}(G/G_{>J}\Gamma \to \mathbb{C})$ and $\mathrm{Lip}(G/G_{>J'}\Gamma \to \mathbb{C})$ pull back to sub algebras of $\mathrm{Lip}(G/\Gamma \to \mathbb{C})$. By the universality of G, $G_{>J}$ and $G_{>J'}$ are disjoint. So the union of these two algebras separate points in G/Γ.

We next approximate F pointwise by the union of the pullbacks of $\mathrm{Lip}(G/G_{>J}\Gamma \to \mathbb{C})$ and $\mathrm{Lip}(G/G_{>J'}\Gamma \to \mathbb{C})$. We remark that we can not directly use Stone-Weierstrass theorem since we need this approximation to be quantitative.

For convenience denote $X = G/\Gamma$, $Y_1 = G/G_{>J}\Gamma$, $Y_2 = G/G_{>J'}\Gamma$. Let $\pi_1 \colon X \to Y_1$ and $\pi_2 \colon X \to Y_2$ denote the projection maps. Denote $\pi = \pi_1 \times \pi_2$ and $Y = Y_1 \times Y_2$. We first claim that $\pi \colon X \to Y$ is an injection. Indeed, for all $x, x' \in X$, if $\pi_i(x) = \pi_i(x')$ for $i = 1, 2$, then $g(\pi_i(x)) = g(\pi_i(x'))$ for all $g \in \text{Lip}(Y_i \to \mathbb{C})$. Since the union of the pullbacks of $\text{Lip}(Y_1 \to \mathbb{C})$ and $\text{Lip}(Y_2 \to \mathbb{C})$ separate points in G/Γ, we have that $x = x'$. This proves the claim.

Let d_X, d_{Y_1} and d_{Y_2} denote the metrices associated to X, Y_1 and Y_2 induced by their Mal'cev basis respectively. It is clear that there exists $L = O_{C,J,J'}(1) > 1$ such that

$$L^{-1}d_X(x,x') \leq d_{Y_i}(\pi_i(x), \pi_i(x')) \leq Ld_X(x,x')$$

for all $x, x' \in X$ and $i = 1, 2$. Let $S := \pi(X) \subseteq Y$, and let $F' \colon S \to \mathbb{C}$ be given by $F'(y) = F(x)$ for all $x \in X, y \in Y$ with $\pi(x) = y$. Since π is an injection, F' is well defined. Moreover, for all $x, x' \in X, y, y' \in Y, y \neq y'$ with $\pi(x) = y$ and $\pi(x') = y'$, we have that

$$\frac{|F'(y) - F'(y')|}{d_Y(y,y')} = \frac{|F(x) - F(x')|}{d_X(x,x')} \cdot \frac{d_X(x,x')}{d_Y(\pi(x), \pi(x'))} \leq CL^2 = O_{C,J,J'}(1),$$

where d_Y is the product metric of d_{Y_1} and d_{Y_2}. By McShane's theorem [18], writing

$$F''(y) := \sup_{z \in S}(F'(z) - CL^2 d_Y(y,z))$$

for all $y \in Y$, we have that $F''|_S = F'$ and that $\|F''\|_{\text{Lip}(Y)} = O_{C,J,J'}(1)$.

Let $\delta > 0$ to be chosen later. It is not hard to see that there exist $N = N(C, \delta, J, J') \in \mathbb{N}_+, K \leq N$ and $y_{1,1}, \dots, y_{1,K} \in Y_1$ such that writing $U(y_{1,j})$ to be the open ball in Y_1 of radius δ centered at $y_{1,j}$ for all $1 \leq j \leq K$, we have that $U(y_{1,1}), \dots, U(y_{1,K})$ cover Y_1, and that there exist a partition of unity $f_1, \dots, f_K \colon Y_1 \to [0,1]$ subordinate to this cover with $\|f_j\|_{\text{Lip}(Y_1)} \leq N$ for all $1 \leq j \leq K$. Let

$$g(x) := \sum_{j=1}^{K} f_j(\pi_1(x))F''(y_{1,j}, \pi_2(x))$$

for all $x \in X$. Then for all $x \in X$, writing $y_1 = \pi_1(x)$ and $y_2 = \pi_2(x)$, we have that

$$|g(x) - F(x)| = \Big|\sum_{j=1}^{K} f_j(y_1)(F''(y_{1,j}, y_2) - F''(y_1, y_2))\Big|$$

$$\leq \sum_{j=1}^{K} \mathbf{1}_{y_1 \in U(y_{1,j})} f_j(y_1) d_{Y_1}(y_1, y_{1,j}) \|F''\|_{\text{Lip}(Y)} \leq \sum_{j=1}^{K} \mathbf{1}_{y_1 \in U(y_{1,j})} f_j(y_1) O_{C,J,J'}(\delta) = O_{C,J,J'}(\delta) < \epsilon$$

if we set δ to be sufficiently small depending on C, ϵ, J, J'. Then $N = N(C, \delta, J, J') = O_{C,\epsilon,J,J'}(1)$.

Finally, for all $1 \leq j \leq K$, recall that $\|f_j\|_{\text{Lip}(Y_1)} \leq N = N(C, \delta, J, J') = O_{C,\epsilon,J,J'}(1)$ and $\|F''(y_{1,j}, \cdot)\|_{\text{Lip}(Y_2)} = O_{C,J,J'}(1)$. So $\|f_j \circ \pi_1\|_{\text{Lip}(X)}, \|F''(y_{1,j}, \pi_2(\cdot))\|_{\text{Lip}(X)} = O_{C,\epsilon,J,J'}(1)$. We are done by setting $\psi_j := f_j \circ \pi_1 \circ g \circ \tau$ and $\psi'_j := F''(y_{1,j}, \pi_2 \circ g \circ \tau(\cdot))$. □

The next lemma allows us to approximate a nilsequence with nilcharacters. The method we use is similar to the discussion in Section 6 and Proposition 8.3 of [12].

Lemma B.5 (Approximating a nilsequence with nilcharacters)**.** Let $D, k \in \mathbb{N}_+$, $C, \epsilon > 0$, I be the degree, multi-degree, or degree-rank ordering with $\dim(I)|k$, $s \in I$, p be a prime, and $\Omega \subseteq \mathbb{F}_p^k$. For any $\psi \in \mathrm{Nil}^{s;C,D}(\Omega)$, there exist $N = N(C, D, \epsilon, k, s) \in \mathbb{N}$, and for all $1 \leq j \leq N$ a linear transformation $T_j\colon \mathbb{C}^{O_{C,D,\epsilon,k,s}(1)} \to \mathbb{C}^D$ of suitable dimensions of complexity $O_{C,D,\epsilon,k,s}(1)$, a nilcharacter $\chi_j \in \Xi_p^{s;O_{C,D,\epsilon,k,s}(1),O_{C,D,\epsilon,k,s}(1)}(\Omega)$, and a nilsequence $\psi_j \in \mathrm{Nil}_p^{<s;O_{C,D,\epsilon,k,s}(1),1}(\Omega)$ such that

$$\Big\|\psi - \sum_{j=1}^{N} T_j(\psi_j \otimes \chi_j)\Big\|_{\ell^\infty(\Omega)} < \epsilon.$$

Proof. It suffices to show this for scalar valued nilsequences ψ. By Corollary B.3, we may assume without loss of generality that $\psi \in \mathrm{Nil}_p^{s;C,1}(\Omega)$. Let $((G/\Gamma)_I, g, F)$ be a $\mathrm{Nil}_p^s(\Omega)$-representation of ψ of complexity at most C and dimension 1. For convenience denote $X = G/\Gamma$ and $X' = G/G_s\Gamma$. Let $\pi\colon X \to X'$ be the quotient map. Note that filtering G/G_s with the groups G_i/G_s, we may view X' as an I-filtered nilmanifold of degree $\prec s$, whose complexity is obviously $O_{C,s}(1)$. The fibers of the map π are isomorphic to $T := G_s/(G_s \cap \Gamma)$. Since G_s is abelian, T is a torus, and so G/Γ is a torus bundle over X' with the structure group T.

Let $\delta > 0$ to be chosen later, and $\sum_{j=1}^{K} \varphi_j$ be a smooth partition of unity on X', where each $\varphi_j \in \mathrm{Lip}(X' \to \mathbb{C})$ is supported on an open ball B_j of radius δ (with respect to the metric induced by the $O_{C,s}(1)$-rational Mal'cev basis of X'). Then $K = O_{C,\delta,s}(1)$. This induces a partition $\psi = \sum_{j=1}^{K} \psi_j$, where

$$\psi_j(n) := F(g(n)\Gamma)\varphi_j(\pi(g(n)\Gamma)) := F_j(g(n)\Gamma) \text{ for all } n \in \Omega.$$

Then F_j is compactly supported in the cylinder $\pi^{-1}(B_j)$ an has Lipschitz constant $O_{C,\delta,s}(1)$.

We now pick $\delta > 0$ in a way such that for each $1 \leq j \leq K$, there is a smooth section $\iota_j\colon B_j \to G$ which partially inverts the projection from G to X'. Then we can take $\delta = O_{C,s}(1)$. Fix some $1 \leq j \leq K$. We can then parametrize any element x of $\pi^{-1}(B_j)$ uniquely as $\iota_j(x_0)t\Gamma$ for some $x_0 \in B_j$ and $t \in T$ (note that $t\Gamma$ is well defined as an element of G/Γ).

Now fix $1 \leq j \leq K$. We can now view the Lipschitz function F_j as a compactly supported Lipschitz function in $B_j \times T$ (again with Lipschitz constant $O_{C,s}(1)$). For convenience denote $r := \dim(G_s)$ and $R := \dim(G)$ (then $T \simeq \mathbb{T}^r$). Then there exists a smooth function $F'_j\colon X \to \mathbb{C}$ with

$$|F_j(x) - F'_j(x)| \leq \epsilon/2K \text{ for all } x \in X, \text{ and } \|F'_j\|_{C^{2R}(X)} = O_{C,\epsilon,s}(1).$$

Let $\phi\colon G \to \mathbb{R}^R$ be the Mal'cev coordinate map. For $h \in \mathbb{Z}^r$, let $\xi_h\colon G_s \to \mathbb{R}$ denote the vertical character such that $\xi_h(t) = (\mathbf{0}, h) \cdot \phi(t)$ for all $t \in G_s$. We may apply Fourier decomposition in the T direction to write

$$F'_{j,h}(\iota_j(x_0)t\Gamma) := \int_T e(-\xi_h(u))F'_j(\iota_j(x_0)tu\Gamma)\, dm_T(u(G_s \cap \Gamma))$$

for all $(x_0, t) \in B_j \times T$ and $u \in G_s$, where m_T is the Haar measure of T. We write $F'_{j,h}(x) := 0$ if x can not be written in the form $\iota_j(x_0)t\Gamma$. It is not hard to see that $\|F'_{j,h}\|_{\mathrm{Lip}(X)} = O_{C,\epsilon,s}(1)$ and

$$F'_{j,h}(\iota_j(x_0)t\Gamma) = e(\xi_h(t))F'_{j,h}(\iota_j(x_0)\Gamma) \text{ for all } x_0 \in B_j, t \in G_s. \tag{B.2}$$

Moreover, since $\|F'_j\|_{C^{2R}(X)} = O_{C,\epsilon,s}(1)$, we have that $\|F'_{j,h}\|_{L^\infty} = O_{C,\epsilon,s}((1+|h|)^{-2R})$ and that

$$F'_j(\iota_j(x_0)t\Gamma) = \sum_{h\in\mathbb{Z}^r} F'_{j,h}(\iota_j(x_0)t\Gamma) \text{ for all } (x_0,t)\in B_j\times T.$$

So there exists $m = O_{C,\epsilon,s}(1)$ such that

$$\Big|F_j(x) - \sum_{h\in\mathbb{Z}^r,|h|\leq m} F'_{j,h}(x)\Big| < \epsilon/2K \text{ for all } x\in X.$$

Fix also $h\in\mathbb{Z}^r, |h|\leq m$. Let $F''_{j,h}\colon X'\to\mathbb{C}$ be the function given by

$$F''_{j,h}(y) := F'_{j,h}(\iota_j(x_0)\Gamma) = \int_T e(-\xi_h(t))F'_j(\iota_j(x_0)t\Gamma)\, dm_T(t(G_s\cap\Gamma))$$

if $y = \iota_j(x_0)\Gamma$ for some $x_0\in B_j$, and $F''_{j,h}(y) := 0$ if not such expression exists. Then $F''_{j,h}$ is a compactly supported function on X' with Lipschitz norm bounded by $O_{C,\epsilon,s}(1)$. On the other hand, by an argument similar to pages 1253–1255 of [12], there exist $D' = O_{C,k,s}(1)$ and a vector valued function $f_{j,h} = (f_{j,h,1},\dots,f_{j,h,D'})\in \mathrm{Lip}(X\to\mathbb{S}^{D'})$ supported on $\pi^{-1}(B_j)$ such that the map $\iota(x_0)t\Gamma\mapsto e(u\cdot t), (x_0,t)\in B_j\times T$ can be viewed as linear combination of the components of $f_{j,h}$ of complexity $O_{C,\epsilon,j,s}(1)$, and that for all $1\leq j\leq D'$,

$$f_{j,h,j}(\iota(x_0)tu\Gamma) = e(\xi_h(u))f_{j,h,j}(\iota(x_0)t\Gamma) \text{ for all } u\in G_s, (x_0,t)\in B_j\times T.$$

It then follows from (B.2) that

$$F'_{j,h}(x) = T_{j,h}(F''_{j,h}(\pi(x))\otimes f_{j,h}(x))$$

for some linear transformation $T_{j,h}$ of suitable dimension of complexity $O_{C,\epsilon,k,s}(1)$ for all $x\in X$. By the triangle inequality, we have that

$$\Big|F(x) - \sum_{j=1}^{K}\sum_{h\in\mathbb{Z}^r,|h|\leq m} T_{j,h}(F''_{j,h}(\pi(x))\otimes f_{j,h}(x))\Big| < \epsilon \tag{B.3}$$

for all $x\in X$. The conclusion follows by replacing x with $g(\tau(n))\Gamma$ in (B.3). □

As a consequence of Lemma B.5, we have

Corollary B.6. Let $D, D', k\in\mathbb{N}_+$, $C,\epsilon>0$, I be the degree, multi-degree, or degree-rank ordering with $\dim(I)|k$, $s\in I$, p be a prime, and $\Omega\subseteq\mathbb{F}_p^k$. For all $\psi\in \mathrm{Nil}^{s;C,D}(\Omega)$ and functions $f\colon\Omega\to\mathbb{C}^{D'}$ with $|f|\leq 1$, if

$$|\mathbb{E}_{n\in\Omega} f(n)\otimes\psi(n)| > \epsilon,$$

then there exists $\chi\in\Xi_p^{s;O_{C,D,D',\epsilon,k,s}(1),O_{C,D,D',\epsilon,k,s}(1)}(\Omega)$ such that

$$|\mathbb{E}_{n\in\Omega} f(n)\otimes\chi(n)| \gg_{C,D,D',\epsilon,k,s} 1.$$

The proof of Corollary B.6 is almost identical to the proof of Corollary E.6 of [12], except that we replace the applications of Lemmas E.3, E.4 and E.5 in [12] by that of Lemmas B.1, B.4 and B.5 in this paper. We omit the details.

Finally, we prove an approximation property for a special type of nilcharacters. Let $d, D, s \in \mathbb{N}_+$. We say that $\chi\colon (\mathbb{Z}^d)^2 \to \mathbb{C}^D$ is a *linearized p-periodic* $(1, s)$*-function* if there exist $C > 0$ such that

$$\chi(h, n) = c(n)^h \psi(h)$$

for some functions $\psi\colon \mathbb{Z}^d \to \mathbb{C}^{D'}, D' \leq D$ and $c = (c_1, \dots, c_d), c_1, \dots, c_d\colon \mathbb{Z}^d \to \mathbb{S}$ such that $\psi(h + ph') = \psi(h)$, $|\psi| \leq C$, $c(n + pn') = c(n)$, $c(n)^{ph} = 1$ for all $h, h', n, n' \in \mathbb{Z}^d$, and that for all $h, \ell \in \mathbb{Z}^d$, the map $n \mapsto c(n - \ell)\overline{c}(n)^h$ belongs to $\mathrm{Nil}_p^{s-1;C}(\mathbb{Z}^d)$, where $c(n)^h := c_1(n)^{h_1} \dots c_d(n)^{h_d}$. In this case we say that χ is of *complexity* at most C and *dimension D*.

We say that $\chi\colon (\mathbb{F}_p^d)^2 \to \mathbb{C}^D$ is a *linearized* $(1, s)$*-function* of complexity at most C and dimension D if $\chi = \chi' \circ \tau$ for some linearized p-periodic $(1, s)$-function $\chi\colon (\mathbb{Z}^d)^2 \to \mathbb{C}^D$ of complexity at most C.

The following lemma allows us to approximate p-periodic nilcharacters of multi-degree $(1, s)$ by linearized $(1, s)$-functions.

Proposition B.7 (Approximating a $(1, s)$-step nilcharacter by linearized $(1, s)$-functions)**.** Let $d, D, s \in \mathbb{N}_+$, $C, \epsilon > 0$ and p be a prime. Let $\Omega \subseteq (\mathbb{F}_p^d)^2$ and $\chi \in \Xi_p^{(1,s);C,D}(\Omega)$. Then we can approximate χ to within ϵ in the uniform norm by an $O_{C,d,D,\epsilon,s}(1)$-complexity linear combination of linearized $(1, s)$-functions of complexities $O_{C,d,D,\epsilon,s}(1)$ and of dimension $O_{C,d,D,\epsilon,s}(1)$.

Proof. The proof is similar to Proposition 8.3 of [12]. We may write $\chi = \chi' \circ \tau$ and

$$\chi'(h, n) := F(g(h, n)\Gamma) \text{ for all } (n, h) \in (\mathbb{Z}^d)^2$$

for some $\mathbb{N}^2$-filtered nilmanifold G/Γ of multi-degree$\leq (1, s)$ of complexity at most C, some function $F \in \mathrm{Lip}(G/\Gamma \to \mathbb{S}^{D'})$ of Lipschitz norm at most C having a vertical frequency $\eta\colon G_{(1,s)} \to \mathbb{R}$ of complexity at most C for some $D' \leq D$, and some $g \in \mathrm{poly}_p((\mathbb{Z}^d)^2 \to G_{\mathbb{N}^2}|\Gamma)$.

Let $\pi\colon G/\Gamma \to G/G_{(1,0)}\Gamma$ be the quotient map. Note that filtering $G/G_{(1,0)}$ with the groups $G_{(0,i)}G_{(1,0)}/G_{(1,0)}$, we may view $G/G_{(1,0)}\Gamma$ as an $\mathbb{N}$-filtered degree s nilmanifold, whose complexity is obviously $O_{C,s}(1)$. The fibers of the map π are isomorphic to $T := G_{(1,0)}/\Gamma_{(1,0)}$. Since $G_{(1,0)}$ is abelian, T is a torus, and so G/Γ is a torus bundle over $G/G_{(1,0)}\Gamma$ with the structure group T.

Let $\delta > 0$ to be chosen later, and $\sum_{k=1}^K \varphi_k$ be a smooth partition of unity on $G/G_{(1,0)}\Gamma$, where each $\varphi_k \in \mathrm{Lip}(G/G_{(1,0)}\Gamma \to \mathbb{C})$ is supported on an open ball B_k of radius δ (with respect to the metric induced by the $O_{C,s}(1)$-rational Mal'cev basis of $G/G_{(1,0)}\Gamma$). Then $K = O_{C,\delta,s}(1)$. This induces a partition $\chi' = \sum_{k=1}^K \chi_k$, where

$$\chi_k(h, n) = F(g(h, n)\Gamma)\varphi_k(\pi(g(h, n)\Gamma)) := F_k(g(h, n)\Gamma) \text{ for all } (n, h) \in (\mathbb{Z}^d)^2.$$

Then $\tilde{F}_k$ is compactly supported in the cylinder $\pi^{-1}(B_k)$.

We now pick $\delta > 0$ in a way such that for each $1 \leq k \leq K$, there is a smooth section $\iota_k\colon B_k \to G$ which partially inverts the projection from G to $G/(G_{(1,0)}\Gamma)$. Then we can take $\delta = O_{C,s}(1)$. Fix some $1 \leq k \leq K$. We can then parametrize any element x of $\pi^{-1}(B_k)$

uniquely a $\iota_k(x_0)t\Gamma$ for some $x_0 \in B_k$ and $t \in T$ (note that $t\Gamma$ is well defined as an element of G/Γ).

We can now view the Lipschitz function F_k as a compactly supported Lipschitz function in $B_k \times T$. Similar to the approach of Lemma B.5, we can approximate F_k on Ω uniformly up to an error ϵ/K by a sum $\sum_{k'=1}^{K'} F_{k,k'}$, where $K' = O_{C,d,D,\epsilon,s}(1)$ and each $F_{k,k'} \in \mathrm{Lip}(B_k \times T \to \mathbb{C}^{O_{C,d,D,\epsilon,s}(1)})$ is bounded by $O_{C,d,D,\epsilon,s}(1)$ in Lipschitz norm, compactly supported, and has a vertical character $\xi_{k'} : T \to \mathbb{R}$ of complexity $O_{C,d,D,\epsilon,s}(1)$ such that

$$F_{k,k'}(\iota_k(x_0)t\Gamma) = \exp(\xi_{k'}(t))F_{k,k'}(\iota_k(x_0)\Gamma) \tag{B.4}$$

for all $x_0 \in B_k$ and $t \in T$. It then suffices to show that for each k, k', the sequence

$$\chi_{k,k'}(h,n) := F_{k,k'}(g(h,n)\Gamma)$$

is a p-periodic linearized $(1,s)$-function.

Fix k and k'. For convenience denote $h := (h_1, \dots, h_d)$. By Lemma 2.20 and the Baker-Campbell-Hausdorff formula, we may write

$$g(h,n) := g_0(n) \prod_{i=1}^{d} g_i(n)^{h_i} \tag{B.5}$$

for some $g_0 \in \mathrm{poly}(\mathbb{Z}^d \to G_{\mathbb{N}})$ (with respect to the $\mathbb{N}$-filtration $G_{\mathbb{N}} = (G_{(i,0)})_{i\in\mathbb{N}}$ of G) and $g_i \in \mathrm{poly}(\mathbb{Z}^d \to (G_{(1,0)})_{\mathbb{N}})$ for $1 \leq i \leq d$ (with respect to the $\mathbb{N}$-filtration $(G_{(1,0)})_{\mathbb{N}} = (G_{(1,i)})_{i\in\mathbb{N}}$ of $G_{(1,0)}$). Since g belongs to $\mathrm{poly}_p((\mathbb{Z}^d)^2 \to G_{\mathbb{N}^2}|\Gamma)$, setting $h = \mathbf{0}$ in (B.5), we have that g_0 belongs to $\mathrm{poly}_p(\mathbb{Z}^d \to G_{\mathbb{N}}|\Gamma)$. For any $1 \leq i \leq d$, setting $h_i = 1$ and all other $h_{i'} = 0$ in (B.5), we deduce that g_i belongs to $\mathrm{poly}_p(\mathbb{Z}^d \to G_{\mathbb{N}}|\Gamma)$ for all $1 \leq i \leq d$ (since $G_{(1,0)}$ is contained in the center of G).

For any $1 \leq i \leq d$, since $g \in \mathrm{poly}_p((\mathbb{Z}^d)^2 \to G_{\mathbb{N}^2}|\Gamma)$, for all $n \in \mathbb{Z}^d$, we have that $g((0,\dots,0,p,0,\dots,0),n)^{-1}g((0,\dots,0),n) \in \Gamma$, where p is on the i-th coordinate. By (B.5) and the fact that g_i is commuting with g_j, we have that $g_i(n)^p \in \Gamma$ for all $n \in \mathbb{Z}^d$ and $1 \leq i \leq d$.

Note that $\chi_{k,k'}(h,n)$ is only nonvanishing when $\pi(g_0(n)\Gamma) \in B_k$. Furthermore, in this case we deduce from (B.4) that

$$\chi_{k,k'}(h,n) := \exp\Big(\sum_{i=1}^{d} h_i\xi_{k'}(g_i(n))\Big)F_{k,k'}(g_0(n)\Gamma) = c(n)^h\psi(n), \tag{B.6}$$

where

$$c(n) := (c_1(n), \dots, c_d(n)) := (\exp(\xi_{k'}(g_1(n))), \dots, \exp(\xi_{k'}(g_d(n)))$$

and $\psi(n) := F_{k,k'}(g_0(n)\Gamma)$. Since $g_0 \in \mathrm{poly}_p(\mathbb{Z}^d \to G_{\mathbb{N}}|\Gamma)$, ψ is a p-periodic function bounded by $O_{C,d,D,\epsilon,s}(1)$ (in fact ψ is a p-periodic nilsequence of complexity $O_{C,d,D,\epsilon,s}(1)$). Note that $\xi_{k'}(g_i(n))$ is a polynomial for all $1 \leq i \leq d$. Since $g_i(n)^p \in \Gamma$ for all $n \in \mathbb{Z}^d$, we have that $\xi_{k'}(g_i(n))$ takes values in $\mathbb{Z}/p$. So $c_i(n)^{ph} = 1$ for all $n, h \in \mathbb{Z}^d$.

On the other hand, note that for all $\ell, h \in \mathbb{Z}^d$,

$$c(n-\ell)^h\overline{c}(n)^h = \exp\Big(\sum_{i=1}^{d} h_i(\xi_{k'}(g_i(n-\ell)) - \xi_{k'}(g_i(n)))\Big).$$

Since $G_{(1,0)}$ is abelian, the maps $n \to \xi_{k'}(g_i(n))$ is a polynomial map from $\mathbb{Z}^d$ to $\mathbb{R}$ of degree at most s. So the map $n \to c(n-\ell)^h\overline{c}(n)^h$ is a polynomial map of degree at most $s-1$. Clearly, this map is of complexity $O_{C,d,D,\epsilon,s}(1)$. This map is also p-periodic since c is p-periodic. We are done. □

Appendix C. Equivalence of nilcharacters

In this appendix, we prove some properties for the equivalence relation on nilsequences defined in Definition 2.35. We recall here Definition 2.35 for the convenience of the readers:

Definition C.1 (An equivalence relation for nilcharacters)**.** Let $k \in \mathbb{N}_+$, $C > 0$, p be a prime, Ω be a subset of $\mathbb{F}_p^k$, I be the degree, multi-degree, or degree-rank ordering with $\dim(I)|k$ and let $s \in I$. For $\chi, \chi' \in \Xi_p^s(\Omega)$, we write $\chi \sim_C \chi' \mod \Xi_p^s(\Omega)$ if $\chi \otimes \overline{\chi} \in \mathrm{Nil}^{\prec s;C}(\Omega)$. We write $\chi \sim \chi' \mod \Xi_p^s(\Omega)$ if $\chi \sim_C \chi' \mod \Xi_p^s(\Omega)$ for some $C > 0$.

Most (but not all) of the results in this appendix are very similar to the results in Appendix E of [12], and their proofs are routine. However, since we are working in a completely different setting from [12], it is inevitable for us to write down the proofs in full details to insures the correctness of the statements.

We first show that $\sim$ is an equivalent relation.

Lemma C.2. Let $D, k \in \mathbb{N}_+$, $C, C' > 0$, p be a prime, Ω be a subset of $\mathbb{F}_p^k$, I be the degree, multi-degree, or degree-rank ordering with $\dim(I)|k$, and let $s \in I$. Let $\chi_1, \chi_2, \chi_3 \in \Xi_p^{I;C,D}(\Omega)$. Then

(i) $\chi_1 \sim_{O_C(1)} \chi_1 \mod \Xi_p^s(\Omega)$ and $\chi_1 \otimes \overline{\chi}_1 \sim_{O_C(1)} 1 \mod \Xi_p^s(\Omega)$;
(ii) if $\chi_1 \sim_{C'} \chi_2 \mod \Xi_p^s(\Omega)$, then $\chi_2 \sim_{C'} \chi_1 \mod \Xi_p^s(\Omega)$;
(iii) if $\chi_1 \sim_{C'} \chi_2 \mod \Xi_p^s(\Omega)$ and $\chi_2 \sim_{C'} \chi_3 \mod \Xi_p^s(\Omega)$, then $\chi_1 \sim_{C''} \chi_3 \mod \Xi_p^s(\Omega)$ for some $C'' = O_{C,C',D}(1)$.

Proof. The proof is very similar to Lemma E.7 of [12]. We first prove Part (i). Let $((G/\Gamma)_I, g, F, \eta)$ be a $\Xi_p^s(\Omega)$-representation χ_1 of complexity at most C. Then

$$\chi_1 \otimes \overline{\chi}_1(n) = F \otimes \overline{F}(g(\tau(n))\Gamma, g(\tau(n))\Gamma) := F'(g(\tau(n))\Gamma)$$

for all $n \in \Omega$. It is clear that $F \otimes \overline{F}$ and F' are of complexities $O_C(1)$. Since F' is invariant under G_s, we may quotient G by G_s and represent $F'(g(\tau(n))\Gamma)$ using a nilmanifold of degree $\prec s$ with all the relevant complexities being $O_C(1)$. So $\chi_1 \otimes \overline{\chi}_1 \in \mathrm{Nil}^{\prec s;O_C(1)}(\Omega)$. This proves Part (i).

Part (ii) is obvious and so it remains to prove Part (iii). Note that each component of

$$(\chi_1 \otimes \overline{\chi}_2) \otimes (\chi_2 \otimes \overline{\chi}_3) = \chi_1 \otimes (\overline{\chi}_2 \otimes \chi_2) \otimes \overline{\chi}_3$$

belongs to $\mathrm{Nil}^{<s;O_{C,C',D}(1)}(\Omega)$. Since the trace of $\overline{\chi}_2\otimes\chi_2$ is 1, $\chi_1\otimes\overline{\chi}_3$ is an $O_{C,C',D}(1)$-complexity linear combination of the components of $\chi_1\otimes(\overline{\chi}_2\otimes\chi_2)\otimes\overline{\chi}_3$. So $\chi_1\otimes\overline{\chi}_3$ belongs to $\mathrm{Nil}^{<s;O_{C,C',D}(1)}(\Omega)$. This proves Part (iii). □

Remark C.3. Since $\sim$ is an equivalent relation, in [12], the authors defined the equivalence class of a nilcharacter χ under the relation $\sim$ as the *symbol* of χ. However, although $\sim$ is an equivalent relation, $\sim_C$ is not an equivalent relation for any fixed $C>0$. Since we do not use the non-standard analysis approach, instead of working with the relation $\sim$, we use the more quantitative relation $\sim_C$. Therefore, we will not use the notion of symbol introduced in [12].

Here are some basic properties for the relation $\sim$.

Lemma C.4. Let $D,k\in\mathbb{N}_+$, $C>0$, p be a prime, Ω be a subset of $\mathbb{F}_p^k$, I be the degree, multi-degree, or degree-rank ordering with $\dim(I)|k$, and let $s\in I$.

(i) If $\chi,\chi'\in\Xi_p^{s;C,D}(\Omega)$ and $\psi\in\mathrm{Nil}^{<s;C,D}(\Omega)$, and the components of χ' are C-complexity linear combinations of those of $\chi\otimes\psi$, then $\chi\sim_{O_{C,D}(1)}\chi' \mod \Xi_p^s(\Omega)$.

(ii) Conversely, if $\chi\sim_C\chi' \mod \Xi_p^s(\Omega)$ for some $\chi,\chi'\in\Xi_p^{s;C,D}(\Omega)$, then χ is an $O_{C,D}(1)$-complexity linear combinations of $\chi\otimes\psi$ for some $\psi\in\mathrm{Nil}^{<s;C,D^2}(\Omega)$.

(iii) If $\chi,\chi',\chi''\in\Xi_p^{s;C}(\Omega)$ and $\chi\sim_C\chi' \mod \Xi_p^s(\Omega)$, then $\chi\otimes\chi''\sim_{O_C(1)}\chi'\otimes\chi'' \mod \Xi_p^s(\Omega)$.

(iv) If $\chi\in\Xi_p^s(\Omega)$ and $\Omega'\subseteq\Omega$, then $\chi|_{\Omega'}\in\Xi_p^s(\Omega')$. Moreover, if $\chi'\sim_C\chi \mod \Xi_p^s(\Omega)$, then $\chi'|_{\Omega'}\sim_C\chi|_{\Omega'} \mod \Xi_p^s(\Omega')$.

(v) If $\chi\in\Xi_p^{s;C}(\Omega)$, I is either the degree or multi-degree ordering, and $q\in\mathbb{Z}$, $p\nmid q$, then $\chi(\iota(q)\cdot)\sim_{O_{C,q}(1)}\chi^{\otimes q^{|s|}} \mod \Xi_p^s(\iota(q)^{-1}\Omega)$.

(vi) If $\chi\in\Xi_p^s(L(\Omega))$ for some d-integral linear transformation $L\colon(\mathbb{F}_p^d)^k\to(\mathbb{F}_p^d)^{k'}$ and I is the degree filtration, then $\chi\circ L\in\Xi_p^s(\Omega)$. Moreover, if $\chi'\sim_C\chi \mod \Xi_p^s(L(\Omega))$, then $\chi'\circ L\sim_C\chi\circ L \mod \Xi_p^s(\Omega)$.

(vii) If $\chi\in\Xi_p^{s;C,D}(\Omega)$, I is either the degree or multi-degree ordering, and $q\in\mathbb{Z}$, $p>q$, then there exists $\tilde{\chi}\in\Xi_p^{s;O_{C,q,s}(1),O_{D,q,s}(1)}(\Omega)$ such $\chi\sim_{O_{C,q}(1)}\tilde{\chi}^{\otimes q} \mod \Xi_p^s(\Omega)$. Moreover, if $\Omega=\mathbb{F}_p^k$, then we may further require $\tilde{\chi}$ to satisfy to following property: for any linear transformation $L\colon\mathbb{F}_p^k\to\mathbb{F}_k$, if $\chi\circ L(n)=\chi(n)$ for all $n\in\mathbb{F}_p^k$, then $\tilde{\chi}\circ L(n)=\tilde{\chi}(n)$ for all $n\in\mathbb{F}_p^k$.

Proof. Proof of Part (i). Assume that $\chi'=\sum_{i,j}c_{i,j}(\chi\otimes\psi)_{i,j}$ for some vectors $c_{i,j}$ of complexity at most C and dimension at most D^2, where $(\chi\otimes\psi)_{i,j}$ denotes the (i,j)-th entry of $\chi\otimes\psi$. Then

$$\chi\otimes\overline{\chi}'=\sum_{i,j}c_{i,j}\chi\otimes(\overline{\chi}'\otimes\overline{\psi})_{i,j}.$$

Since $\chi\otimes(\overline{\chi}'\otimes\overline{\psi})=(\chi\otimes\overline{\chi}')\otimes\overline{\psi}\in\mathrm{Nil}^{<s;O_{C,D}(1)}(\Omega)$, we have that $\chi\otimes\overline{\chi}'\in\mathrm{Nil}^{<s;O_{C,D}(1)}(\Omega)$. This proves Part (i).

Proof of Part (ii). Note that

$$\chi\otimes(\overline{\chi'}\otimes\chi')=(\chi\otimes\overline{\chi'})\otimes\chi'.$$

By Lemma B.1, $\chi\otimes\overline{\chi'}\in\mathrm{Nil}^{<s;2C,D^2}(\Omega)$. The conclusion follows from the fact that 1 is an $O_D(1)$-complexity linear combination of the components of $\overline{\chi'}\otimes\chi'$.

Proof of Part (iii). Note that $(\chi\otimes\chi'')\otimes\overline{(\chi'\otimes\chi'')}$ is an reordering of the entries of $(\chi\otimes\overline{\chi}')\otimes(\chi''\otimes\overline{\chi}'')$. Since $\chi\otimes\overline{\chi}'\in\mathrm{Nil}^{<s;C}(\Omega)$ and by Lemma C.2 $\chi''\otimes\overline{\chi}''\in\mathrm{Nil}^{<s;O_C(1)}(\Omega)$, we have that $(\chi\otimes\overline{\chi}')\otimes(\chi''\otimes\overline{\chi}'')$ and thus $(\chi\otimes\chi'')\otimes\overline{(\chi'\otimes\chi'')}$ belongs to $\mathrm{Nil}^{<s;O_C(1)}(\Omega)$. This proves Part (iii).

Proof of Part (iv). It follows from the definitions.

Proof of Part (v). Let $((G/\Gamma)_I, g, F, \eta)$ be a $\Xi_p^s(\Omega)$-representation of χ of complexity at most C and dimension D. With a slight abuse of notations, we use q to denote both the integer and its image in $\mathbb{F}_p$ under ι. Since $g\in\mathrm{poly}_p((\mathbb{Z}^k\to G_I|\Gamma)$, we have that $\chi(qn)=F(g(\tau(qn))\Gamma)=F(g(q\tau(n))\Gamma)$ for all $n\in q^{-1}\Omega$. Similar to the proof of Lemma E.8 (v) of [12], one can show that

$$\tilde{\chi}(n):=F(g(qn)\Gamma)\otimes\overline{F}(g(n)\Gamma)^{\otimes q^{|s|}}, n\in\tau(q^{-1}\Omega)+p\mathbb{Z}^k$$

is a nilsequence of step $\prec s$ and complexity at most $O_{C,q}(1)$, namely $\tilde{\chi}\in\mathrm{Nil}^{<s;O_{C,q}(1)}(\tau(\Omega)+p(\mathbb{Z}^d)^k)$. Composing $\tilde{\chi}$ with τ, we have that $\chi(q\cdot)\sim_{O_{C,q}(1)}\chi^{\otimes q^{|s|}} \mod \Xi_p^s(\iota(q)^{-1}\Omega)$.

Proof of Part (vi). The fact that $\chi\circ L$ belongs to $\Xi_p^s(\Omega)$ follows from Lemma B.1 and the inclusion $L^{-1}(L(\Omega))\supseteq\Omega$. If $\chi'\sim_C\chi \mod \Xi_p^s(L(\Omega))$, then writing $\chi=\tilde{\chi}\circ\tau$ and $\chi'=\tilde{\chi}'\circ\tau$ for some $\tilde{\chi},\tilde{\chi}'\in\Xi_p^s(\tau(L(\Omega))+p\mathbb{Z}^{k'})$, we have that

$$\chi(n)\otimes\overline{\chi'}(n)=(\tilde{\chi}\circ\tau(n))\otimes(\overline{\tilde{\chi}'}\circ\tau(n))=\phi\circ\tau(n)$$

for some $\phi\in\mathrm{Nil}^{<s;C}(\tau(L(\Omega))+p\mathbb{Z}^{k'})$ for all $n\in L(\Omega)$. By Lemma A.2, there exists a linear transformation $\tilde{L}\colon\mathbb{Z}^k\to\mathbb{Z}^{k'}$ such that $\tilde{L}\circ\tau\equiv\tau\circ L \mod p\mathbb{Z}^{k'}$. So

$$\begin{aligned}(\chi\circ L(n))\otimes(\overline{\chi'}\circ L(n))&=(\tilde{\chi}\circ\tau\circ L(n))\otimes(\overline{\tilde{\chi}'}\circ\tau\circ L(n))\\ &=(\tilde{\chi}\circ\tilde{L}\circ\tau)\otimes(\overline{\tilde{\chi}'}\circ\tilde{L}\circ\tau)=\phi\circ\tilde{L}\circ\tau(n)\end{aligned}$$

for all $n\in\Omega$. Since $\phi\in\mathrm{Nil}^{<s;C}(\tau(L(\Omega))+p\mathbb{Z}^{k'})$, we have $\phi\circ\tilde{L}\in\mathrm{Nil}^{<s;C}(\tau(\Omega)+p\mathbb{Z}^k)$. So $\chi'\circ L\sim_C\chi\circ L \mod \Xi_p^s(\Omega)$.

Proof of Part (vii). We present a proof which is different from Lemma E.8 (vii) of [12]. Let $((G/\Gamma)_I, g, F, \eta)$ be a $\Xi_p^s(\Omega)$-representation of χ of complexity at most C and dimension at most D. Let q^* be the unique integer in $\{1,\dots,p-1\}$ such that $qq^*\equiv 1 \mod p\mathbb{Z}$ (such q^* exists since $p>q$). With a slight abuse of notations, we use q, q^* to denote both the integer and their images in $\mathbb{F}_p$ under ι. Since g belongs to $\mathrm{poly}_p(\mathbb{Z}^k\to G_I|\Gamma)$, so does $g(q^*\cdot)$. Writing $\tilde{\chi}(n):=F(g(q^*\tau(n))\Gamma)$ for all $n\in\mathbb{F}_p^k$, we have that $\tilde{\chi}\in\Xi_p^{s;C,D}(\Omega)$ and that

$$\tilde{\chi}(qn)=F(g(q^*\tau(qn))\Gamma)=F(g(\tau(qq^*n))\Gamma)=F(g(\tau(n))\Gamma)=\chi(n) \tag{C.1}$$

for all $n\in\Omega$. By Part (v), $\chi\sim_{O_{C,q}(1)}\chi'^{q^{|s|}} \mod \Xi_p^s(\Omega)$. The claim follows by setting $\tilde{\chi}:=(\chi')^{\otimes q^{|s|-1}}$ (whose complexity is $O_{C,q,s}(1)$ and dimension is $O_{D,q,s}(1)$ by Lemma B.1).

For the "moreover" part, note that for all $n\in\mathbb{F}_p^k$, it follows from (C.1) that

$$\tilde{\chi}(n)=\tilde{\chi}(q^*qn)=\chi(qn).$$

So
$$\tilde{\chi}(L(n)) = \chi(qL(n)) = \chi(L(qn)) = \chi(qn) = \tilde{\chi}(n).$$
□

The following is a generalization of Lemma 13.2 of [12]:

Lemma C.5 (Multi-linearity lemma)**.** Let $d, k, s \in \mathbb{N}_+$, $C > 0$, p be a prime, $L_1, \dots, L_s$, $L'_1 \colon (\mathbb{F}_p^d)^k \to \mathbb{F}_p^d$ be d-integral linear transformations, and let $\chi \in \Xi_p^{(1,\dots,1);C}((\mathbb{F}_p^d)^s)$ (with 1 repeated s times). We have that
$$\chi(L_1(n) + L'_1(n), L_2(n), \dots, L_s(n))$$
$$\sim_{O_C(1)} \chi(L_1(n), L_2(n), \dots, L_s(n)) \otimes \chi(L'_1(n), L_2(n), \dots, L_s(n)) \mod \Xi_p^s((\mathbb{F}_p^d)^k),$$
where $n \in (\mathbb{F}_p^d)^k$. A similar result holds for the other $s-1$ variables.

Proof. By Lemma B.1 and Lemma C.4 (vi) (and by viewing χ as a degree s nilsequence), it suffices to show that

(C.2)
$$\chi(h_1 + h'_1, h_2, \dots, h_s) \sim_{O_{C,L}(1)} \chi(h_1, h_2, \dots, h_s) \otimes \chi(h'_1, h_2, \dots, h_s) \mod \Xi_p^s((\mathbb{F}_p^d)^{s+1}),$$
where both sides of (C.2) are viewed as functions of $(h_1, \dots, h_s, h'_1) \in (\mathbb{F}_p^d)^{s+1}$. We may assume that for all $(h_1, \dots, h_s) \in (\mathbb{F}_p^d)^s$,
$$\chi(h_1, \dots, h_s) = F(g \circ \tau(h_1, \dots, h_s)\Gamma)$$
for some $\mathbb{N}^s$-filtered nilmanifold G/Γ of degree $\leq (1, \dots, 1)$ and complexity at most C endowed with a Mal'cev basis $\mathcal{X}$, some $F \in \mathrm{Lip}(G/\Gamma \to \mathbb{S}^D)$ with Lipschitz norm at most C and with a vertical frequency η for some $D \in \mathbb{N}_+$, and some $g \in \mathrm{poly}_p((\mathbb{Z}^d)^s \to G_{\mathbb{N}^s}|\Gamma)$. It suffices to show that the map

(C.3)
$$(h_1, \dots, h_s, h'_1) \mapsto F(g(h_1 + h'_1, h_2, \dots, h_s)\Gamma) \otimes \overline{F}(g(h_1, \dots, h_s)\Gamma) \otimes \overline{F}(g(h'_1, h_2, \dots, h_s)\Gamma)$$
belongs to $\mathrm{Nil}^{s-1;O_C(1)}((\mathbb{Z}^d)^{s+1})$.

We may write (C.3) in the form
$$\tilde{F}(\tilde{g}(h_1, \dots, h_s, h'_1)\Gamma^3),$$
where
$$\tilde{g}(h_1, \dots, h_s, h'_1) := (g(h_1 + h'_1, \dots, h_s), g(h_1, \dots, h_s), g(h'_1, \dots, h_s))$$
and $\tilde{F} \in \mathrm{Lip}(G^3/\Gamma^3 \to \mathbb{C}^{D^3})$ is given by
$$\tilde{F}(x_1, x_2, x_3) := F(x_1) \otimes \overline{F}(x_2) \otimes \overline{F}(x_3).$$
Then $\tilde{F}$ is a function taking values in $\mathbb{S}^{D^3}$ of Lipschitz norm at most $3C$ with a vertical frequency $(\eta, -\eta, -\eta)$. Since $g \in \mathrm{poly}_p((\mathbb{Z}^d)^s \to G_{\mathbb{N}^s}|\Gamma)$, we have that $\tilde{g} \in \mathrm{poly}_p((\mathbb{Z}^d)^{s+1} \to (G^3)_{\mathbb{N}^s}|\Gamma^3)$.

By Lemma 2.20 and the Baker-Campbell-Hausdorff formula, we may write

$$g(h_1,\dots,h_s)=\prod_{i_1,\dots,i_s\in\mathbb{N}^d,|i_j|\in\{0,1\}} g_{i_1,\dots,i_s}^{\binom{h_1}{i_1}\dots\binom{h_s}{i_s}}$$

for some $g_{i_1,\dots,i_s}\in G_{(|i_1|,\dots,|i_s|)}$. We may now give G^3 an $\mathbb{N}$-filtration by setting $(G^3)_i$ to be the group generated by $G^3_{(j_1,\dots,j_s)}$ for all $j_1,\dots,j_s\in\mathbb{N}$ with $j_1+\dots+j_s>i$ together with the elements (g_1g_2,g_1,g_2) for all $g_1,g_2\in G_{(j_1,\dots,j_s)}$ for some $j_1+\dots+j_s=i$. From the Baker-Campbell-Hausdorff formula one can verify that this is a filtration on G^3 of complexity $O_C(1)$. From the type-I Taylor expansion we also see that $\tilde{g}$ is a polynomial with respect to this filtration. Finally, as F has a vertical frequency $(\eta,-\eta,-\eta)$, $\tilde{F}$ is invariant under $(G^3)_s=\{(g_1g_2,g_1,g_2)\colon g_1,g_2\in G_{(1,\dots,1)}\}$. We may then quotient G^3 by $(G^3)_s$ to conclude that (C.3) is a degree $<s$ nilsequence on $(\mathbb{Z}^d)^{s+1}$ (note that the quotient of $\tilde{g}$ belongs $\mathrm{poly}_p((\mathbb{Z}^d)^{s+1}\to(G^3/G^3_s)_{\mathbb{N}^s}|\Gamma^3/(G^3_s\cap\Gamma^3)))$, whose complexity is certainly $O_C(1)$. □

We now turn to nilcharacters defined on a set admitting a Leibman dichotomy, where the factorization theorem (Theorem A.14) applies. The following is a generalization of Lemma E.13 of [12], which says that if $\chi^{\otimes q}\sim 1 \mod \Xi^s_p(\Omega)$ for some small q, then $\chi\sim 1 \mod \Xi^s_p(\Omega)$.

Lemma C.6 (Torsion-free lemma)**.** Let $d,D,k,q\in\mathbb{N}_+,r,s\in\mathbb{N}$, $C>0$, $p\gg_{C,d,D,k,q,r,s}1$ be a prime, $M\colon\mathbb{F}_p^d\to\mathbb{F}_p$ be a quadratic form, and Ω be a subset of $(\mathbb{F}_p^d)^k$ such that one of the three assumptions (i)–(iii) in Theorem A.14 holds (here r is the parameter appearing in the statement of Theorem A.14). For any $\chi\in\Xi^{s;C,D}_p(\Omega)$, if $\chi^{\otimes q}\in\mathrm{Nil}^{s-1;C}(\Omega)$, then $\chi\in\mathrm{Nil}^{s-1;O_{C,d,D,k,q,r,s}(1)}(\Omega)$.

In particular, if $\chi^{\otimes q}\sim_C 1 \mod \Xi^s_p(\Omega)$, then $\chi\sim_{O_{C,d,D,k,q,r,s}(1)} 1 \mod \Xi^s_p(\Omega)$.

Proof. By Lemma B.1, we have that $\chi^{\otimes q}\in\mathrm{Nil}^{s-1;C,D^q}(\Omega)\cap\Xi_p^{s;O_{C,q}(1),D^q}(\Omega)$. So

$$|\mathbb{E}_{n\in\Omega}\chi(n)^{\otimes q}\otimes\tilde{\chi}(n)|\geq 1$$

for some $\tilde{\chi}\in\mathrm{Nil}^{s-1;C,D^q}(\Omega)$ (we may set $\tilde{\chi}:=\overline{\chi}^{\otimes q}$). By Corollary B.3 and the Pigeonhole Principle, there exist $\chi_0\in\mathrm{Nil}_p^{s-1;O_{C,d}(1),D^q}((\mathbb{F}_p^d)^k)$ and $\epsilon=\epsilon(C,d,k)>0$ such that

$$|\mathbb{E}_{n\in\Omega}\chi(n)^{\otimes q}\otimes\chi_0(n)|\geq\epsilon.$$

Let $((G/\Gamma)_{\mathbb{N}},g,F,\eta)$ be a $\Xi^s_p(\Omega)$-representation of χ of complexity at most C and dimension at most D, and $((G'/\Gamma')_{\mathbb{N}},g',F')$ be a $\mathrm{Nil}_p^{s-1}((\mathbb{F}_p^d)^k)$-representation of χ of complexity at most $O_{C,d}(1)$ and dimension at most D^q. Let $\mathcal{F}$ be a growth function to be chosen later depending only on C,d,D,k,q,r,s. By Theorem A.14, there exists $C\leq C'\leq O_{C,d,D,\mathcal{F},k,q,r,s}(1)=O_{C,d,D,k,q,r,s}(1)$ such that if $p\gg_{C,d,D,\mathcal{F},k,q,r,s}1$, then there exist a proper subgroup W of $G\times G'$ which is C'-rational relative to $\mathcal{X}\times\mathcal{X}'$, and a factorization

$$(g(\tau(n)),g'(\tau(n)))=(t,t')\cdot(h(n),h'(n))\cdot(\gamma(n),\gamma'(n)) \text{ for all } n\in(\mathbb{F}_p^d)^k$$

such that $t\in G$ and $t'\in G'$ are of complexities $O_{C'}(1)$, $(h,h')\in\mathrm{poly}_p(\Omega\to W_{\mathbb{N}}|\Gamma_W)$, $\Gamma_W:=W\cap(\Gamma\times\Gamma')$ with $(h(n)\Gamma,h'(n)\Gamma')_{n\in\Omega}$ being $\mathcal{F}(C')^{-1}$-equidistributed on W/Γ_W, and that $\gamma\in\mathrm{poly}(\Omega\to G_{\mathbb{N}}|\Gamma)$, $\gamma'\in\mathrm{poly}(\Omega\to G'_{\mathbb{N}}|\Gamma)$.

Denoting $\tilde{F}(x,y) := F(tx)^{\otimes q} \otimes F'(t'y)$ for $x \in G/\Gamma$ and $y \in G'/\Gamma'$, we have that

$$|\mathbb{E}_{n\in\Omega}\chi(n)^{\otimes q} \otimes \psi(n)| = |\mathbb{E}_{n\in\Omega}\tilde{F}(h(n)\Gamma, h'(n)\Gamma')| > \epsilon.$$

Since F and F' are bounded in Lipschitz norm by C, and t, t' are of complexities $O_{C'}(1)$, we have that $\tilde{F}$ is bounded in Lipschitz norm by some $Q_{C,D,q}(C')$, where $Q_{C,D,q}$ is viewed as a function of C'. Denote $Y := W/\Gamma_W$ and let m_Y be the Haar measure on Y. So

$$\left|\mathbb{E}_{n\in\Omega}\tilde{F}(h(n)\Gamma, h'(n)\Gamma') - \int_Y \tilde{F}\, dm_Y\right| \leq D^{q+1}Q_{C,D,q}(C')\mathcal{F}(C')^{-1} \leq \epsilon/2$$

provided that $\mathcal{F}$ grows sufficiently fast such that $\mathcal{F}(C') \geq 4\epsilon^{-1}D^{q+1}Q_{C,D,q}(C')$. Clearly $\mathcal{F}$ depends only on C, d, D, k, q, r, s. Thus

$$\left|\int_Y \tilde{F}\, dm_Y\right| > \epsilon/2.$$

Since F has a vertical frequency η, $\tilde{F}$ has a vertical frequency (η^q, id). This means that (η^q, id) must annihilate W_s. On the other hand, since W is a subgroup of $G \times G'$, we have that W_s is a subgroup of $G_s \times \{id_{G'}\}$. Assume that the projection of W_j to the first coordinate is $\tilde{G}_j$ for all $j \in \mathbb{N}$. It is not hard to see that $\tilde{G}_{\mathbb{N}}$ is an $\mathbb{N}$-filtration of complexity $O_{C,C',d,D,k,q,r,s}(1)$. Since $W_s = \tilde{G}_s \times \{id_{G'}\}$, η^q must annihilate $\tilde{G}_s$, which is a subgroup of G_s. Since η is a continuous homomorphism on the connected abelian Lie group G, η itself must also annihilates $\tilde{G}_s$. Let $G'' := \tilde{G}/\tilde{G}_s$, $\Gamma'' := (\tilde{G}\cap\Gamma)/(\tilde{G}_s\cap\Gamma)$, and $\pi\colon \tilde{G} \to G''$ be the projection map. Then we may write $F(t\cdot) = F'' \circ \pi$ for some function $F''\colon G''/\Gamma'' \to \mathbb{C}^{D'}$. Writing $h'' := \pi \circ h \in \mathrm{poly}_p(\Omega \to G''_I|\Gamma'') \subseteq \mathrm{poly}((\mathbb{F}_p^d)^k \to G''_I)$, we have that

$$\chi(n) = F(th(n)\Gamma) = F''(\pi \circ h(n)\Gamma) = F''(h''(n)\Gamma'')$$

for all $n \in \Omega$.

Since $\tilde{G}$ is $O_{C,C',d,D,k,q,r,s}(1)$-rational relative to the C-rational Mal'cev basis $\mathcal{X}$, so is G''. Since F is bounded in Lipschitz norm by C and t is of complexity $O_{C'}(1)$, F'' has Lipschitz norm bounded by $O_{C,C'}(1)$. Since $C' \leq O_{C,d,D,k,q,r,s}(1)$, we have that $\chi \in \mathrm{Nil}^{s-1;O_{C,d,D,k,q,r,s}(1)}(\Omega)$. ∎

The following lemma generalizes of Lemma E.12 of [12], which says that if a degree-s nilchracter χ has nontrivial correlation with a degree-$(s-1)$ nilsequence, then χ is also a degree-$(s-1)$ nilsequence.

Lemma C.7. Let $d, D, k \in \mathbb{N}_+$, $r, s \in \mathbb{N}$, $C, \epsilon > 0$, $p \gg_{C,d,D,\epsilon} 1$ be a prime, $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$ be a quadratic form, and Ω be a subset of $(\mathbb{F}_p^d)^k$ such that one of the three assumptions (i)–(iii) in Theorem A.14 holds (here r is the parameter appearing in the statement of Theorem A.14). Let $\chi \in \Xi_p^{s;C,D}(\Omega)$ and $\psi \in \mathrm{Nil}^{s-1;C,D}(\Omega)$. If

$$|\mathbb{E}_{n\in\Omega}\chi(n) \otimes \psi(n)| > \epsilon,$$

then $\chi \in \mathrm{Nil}^{s-1;O_{C,d,D,\epsilon,k,r,s}(1),D}(\Omega)$. In particular, $\chi \sim_{O_{C,d,D,\epsilon,k,r,s}(1)} 1 \mod \Xi_p^s(\Omega)$.

Proof. By Corollary B.6, there exists $\psi' \in \text{Nil}_p^{s-1;O_{C,d,D,\epsilon,k,r,s}(1),O_{C,d,D,\epsilon,k,r,s}(1)}(\Omega)$ and $\epsilon' = \epsilon'(C, d, D, \epsilon, k, r, s) > 0$ such that

$$|\mathbb{E}_{n\in\Omega}\chi(n) \otimes \psi'(n)| > \epsilon'.$$

Let $((G/\Gamma)_{\mathbb{N}}, g, F, \eta)$ be a $\Xi_p^s(\Omega)$-representation of χ of complexity at most C and dimension at most D, and $((G'/\Gamma')_{\mathbb{N}}, g', F')$ be a $\text{Nil}_p^{s-1}(\Omega)$-representation of ψ' of complexity and dimension at most $O_{C,d,D,\epsilon}(1)$. Let $\mathcal{F}$ be a growth function to be chosen later depending only on $C, d, D, \epsilon, k, r, s$. By Theorem A.14, there exists $C \leq C' \leq O_{C,d,D,\epsilon,\mathcal{F},k,r,s}(1) = O_{C,d,D,\epsilon,k,r,s}(1)$ such that if $p \gg_{C,d,D,\epsilon,\mathcal{F},k,r,s} 1$, then there exist a proper subgroup W of $G\times G'$ which is C'-rational relative to $\mathcal{X} \times \mathcal{X}'$, and a factorization

$$(g \circ \tau(n), g' \circ \tau(n)) = (t, t') \cdot (h(n), h'(n)) \cdot (\gamma(n), \gamma'(n)) \text{ for all } n \in (\mathbb{F}_p^d)^k$$

such that $t \in G$ and $t' \in G'$ are of complexities at most $O_{C'}(1)$, $(h, h') \in \text{poly}_p(\Omega \to W_I|\Gamma_W)$, $\Gamma_W := W \cap (\Gamma \times \Gamma')$ with $(h(n)\Gamma, h'(n)\Gamma')_{n\in\Omega}$ being $\mathcal{F}(C')^{-1}$-equidistributed on $W/\Gamma_W)$, and that $\gamma \in \text{poly}(\Omega \to G_I|\Gamma)$, $\gamma' \in \text{poly}(\Omega \to G'_I|\Gamma)$.

Denoting $\tilde{F}(x, y) := F(tx) \otimes F'(t'y)$ for $x \in G/\Gamma$ and $y \in G'/\Gamma'$, we have that

$$|\mathbb{E}_{n\in\Omega}\chi(n) \otimes \psi'(n)| = |\mathbb{E}_{n\in\Omega}\tilde{F}(h(n)\Gamma, h'(n)\Gamma')| > \epsilon'.$$

Since F and F' are bounded in Lipschitz norm by C and t, t' are of complexities $O_{C'}(1)$, we have that $\tilde{F}$ is bounded in Lipschitz norm by some $Q_{C,D}(C')$, where $Q_{C,D}$ is viewed as a function of C'. Denote $Y := W/\Gamma_W$ and let m_Y be the Haar measure on Y. So

$$\left|\mathbb{E}_{n\in\Omega}\tilde{F}(h(n)\Gamma, h'(n)\Gamma') - \int_Y \tilde{F}\, dm_Y\right| \leq D^2 Q_{C,D}(C')\mathcal{F}(C')^{-1} \leq \epsilon'/2$$

provided that $\mathcal{F}'$ grows sufficiently fast such that $\mathcal{F}(C') \geq 2D^2 Q_{C,D}(C')/\epsilon'$. Clearly, $\mathcal{F}'$ depends only on $C, d, D, \epsilon, k, r, s$. Thus

$$\left|\int_Y \tilde{F}\, dm_Y\right| > \delta/2.$$

Since F has a vertical frequency η, $\tilde{F}$ has a vertical frequency (η, id). This means that (η, id) must annihilate W_s. On the other hand, since W is a subgroup of $G \times G'$, we have that W_s is a subgroup of $G_s \times \{id_{G'}\}$. Assume that the projection of W_j to the first coordinate is $\tilde{G}_j$ for all $j \in \mathbb{N}$. It is not hard to see that $\tilde{G}_{\mathbb{N}}$ is a filtration of complexity $O_{C,C',d,D,\epsilon,k,r,s}(1)$. Since $W_s = \tilde{G}_s \times \{id_{G'}\}$, η must annihilate $\tilde{G}_s$, which is a subgroup of G_s. Let $G'' := \tilde{G}/\tilde{G}_s$, $\Gamma'' := (\tilde{G} \cap \Gamma)/(\tilde{G}_s \cap \Gamma)$, and $\pi \colon \tilde{G} \to G''$ be the projection map. Then we may write $F(t\cdot) = F'' \circ \pi$ for some function $F'' \colon G''/\Gamma'' \to \mathbb{C}^{D'}$. Writing $h'' := \pi \circ h \in \text{poly}_p(\Omega \to G''_{\mathbb{N}}|\Gamma'')$, we have that

$$\chi(n) = F(g \circ \tau(n)\Gamma) = F(th(n)\Gamma) = F''(\pi \circ h(n)\Gamma) = F''(h''(n)\Gamma'')$$

for all $n \in \Omega$. Since $\tilde{G}$ is $O_{C,C',d,D,\epsilon,k,r,s}(1)$-rational relative to the C-rational Mal'cev basis $\mathcal{X}$, so is G''. Since F has Lipschitz norm bounded by C and t is of complexity $O_{C'}(1)$, F'' has Lipschitz norm bounded by $O_{C,C'}(1)$. Since $C' \leq O_{C,d,D,\epsilon,k,r,s}(1)$, we have that $\chi \in \text{Nil}^{s-1;O_{C,d,D,\epsilon,k,r,s}(1),D}(\Omega)$. $\square$

Remark C.8. It is natural to ask whether the assumptions in Lemma C.7 imply that χ belongs to $\text{Nil}_p^{s-1}(\Omega)$ (i.e. we further require χ is p-periodic as a degree-$(s-1)$ nilsequence). Unfortunately we do not know the answer to this question. This is because the factorization result (Theorem A.14) we use in the proof of Lemma C.7 does not ensure that h, h' are p-periodic.

We conclude this section with an analog of Theorem E.10 of [12], which asserts that every nilcharacter can be written as the composition of a linear transformation and a nilcharacter of multi-degree $(1, \dots, 1)$, up to a nilsequence of lower degree.

Theorem C.9. Let $d, D, k \in \mathbb{N}_+$, $s_1, \dots, s_k \in \mathbb{N}_+$, $C > 0$, p be a prime and $\Omega \subseteq (\mathbb{F}_p^d)^k$. Denote $s := (s_1, \dots, s_k)$ and $|s| := s_1 + \dots + s_k$. For all $\chi \in \Xi_p^{s;C,D}(\Omega)$, if $p \gg_{C,d,s} 1$, then there exists $\tilde{\chi} \in \Xi_p^{(1,\dots,1);O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{F}_p^d)^{|s|})$ (with 1 repeated $|s|$ times) such that writing

$$\chi'(n_1, \dots, n_k) := \tilde{\chi}(n_1, \dots, n_1, \dots, n_k, \dots, n_k), n_1, \dots, n_k \in \mathbb{F}_p^d,$$

(with each n_i repeated s_i times), we have that $\chi' \in \Xi_p^{s;O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{F}_p^d)^k)$ and $\chi \sim_{O_{C,d,k,s}(1)} \chi'$ mod $\Xi_p^{|s|}(\Omega)$. Furthermore, one can select

$$\tilde{\chi}(n_{1,1}, \dots, n_{1,s_1}, \dots, n_{k,1}, \dots, n_{k,s_k})$$

to be symmetric with respect to the permutations of $n_{i,1}, \dots, n_{i,s_i}$ for all $1 \leq i \leq k$.

In order to prove Theorem C.9, we need the following lemma:

Lemma C.10. Let $d, D, k \in \mathbb{N}_+$, $s_1, \dots, s_k \in \mathbb{N}_+$, $C > 0$, and p be a prime. Denote $s := (s_1, \dots, s_k)$ and $|s| := s_1 + \dots + s_k$. Let $\chi \in \Xi_p^{(1,\dots,1);C,D}((\mathbb{F}_p^d)^{|s|})$ (with 1 repeated $|s|$ times).

(i) Let $\sigma\colon \{1, \dots, |s|\} \to \{1, \dots, |s|\}$ be a permutation and denote

$$\chi'(n_1, \dots, n_{|s|}) := \chi(n_{\sigma(1)}, \dots, n_{\sigma(|s|)})$$

for all $(n_1, \dots, n_{|s|}) \in (\mathbb{F}_p^d)^{|s|}$. Then $\chi' \in \Xi_p^{(1,\dots,1);C,D}((\mathbb{F}_p^d)^{|s|})$.

(ii) Let

$$\chi'(n_1, \dots, n_k) := \chi(n_1, \dots, n_1, \dots, n_k, \dots, n_k)$$

for all $(n_1, \dots, n_k) \in (\mathbb{F}_p^d)^k$, where each n_i appears s_i times in the above expression. Then $\chi' \in \Xi_p^{s;C,D}((\mathbb{F}_p^d)^k)$.

Lemma C.10 can be proved directly by using the type-I Taylor expansion (Lemma 2.20) and the Baker-Campbell-Hausdorff formula. We leave the details to the interested readers.

Proof of Theorem C.9. The outline of the proof is similar to Theorem E.10 of [12]. In our setting, the proof is more difficult in the following two senses. The first is that we are dealing with multi-dimensional nilcharacters, which make the constructions more complicated. The second is that we need some additional efforts to ensure that the nilcharacters $\tilde{\chi}$ and χ' we construct are p-periodic.

By Lemma C.4 (iv), extending the domain of ξ from Ω to $(\mathbb{F}_p^d)^k$ if necessary, we may assume without loss of generality that $\Omega = (\mathbb{F}_p^d)^k$. By Lemmas C.2, C.4 (vii), C.6 and

C.10, it suffices to construct $\tilde{\chi} \in \Xi_p^{(1,\dots,1);O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{F}_p^d)^{|s|})$ which is symmetric with respect to the permutations of $n_{i,1},\dots,n_{i,s_i}$ for all $1 \leq i \leq k$ such that

$$\chi' \sim_{O_{C,d,s}(1)} \chi^{\otimes s!} \mod \Xi_p^{|s|}((\mathbb{F}_p^d)^{|s|}),$$

where $s! := s_1!\dots s_k!$. Assume that $\chi(n) = \xi \circ \tau(n), n \in (\mathbb{F}_p^d)^k$ for some $\xi \in \Xi_p^{s;C,D}((\mathbb{Z}^d)^k)$. We may write $\xi(n) = F(g(n)\Gamma)$ for all $n \in (\mathbb{Z}^d)^k$, where G/Γ is a nilmanifold of multi-degree s and complexity at most C, $g \in \mathrm{poly}_p((\mathbb{Z}^d)^k \to G_{\mathbb{N}^k}|\Gamma)$, and $F \in \mathrm{Lip}(G/\Gamma \to \mathbb{S}^{D'})$ is of Lipschitz norm at most C for some $D' \leq D$, and has a vertical frequency η of complexity at most C.

We first build a multi-degree $(1,\dots,1)$ nilpotent group $\tilde{G}$. Let $\log \tilde{G}$ be the Lie algebra given by the direct sum (as a real vector space)

$$\log \tilde{G} := \oplus_{J \subseteq \{1,\dots,|s|\}} \log G_{\|J\|},$$

where $\|J\| \in \mathbb{N}^k$ is the vector

$$\|J\| := (|J \cap \{s_1 + \dots + s_{i-1} + 1, \dots, s_1 + \dots + s_{i-1} + s_i\}|)_{1 \leq i \leq k}.$$

For each $J \subseteq \{1,\dots,|s|\}$, there is a natural vector space embedding $\iota_J\colon \log G_{\|J\|} \to \log \tilde{G}$. We then endow $\log \tilde{G}$ with a Lie bracket structure such that for all $J, J' \subseteq \{1,\dots,|s|\}$ and $x_J \in \log G_{\|J\|}$, $x_{J'} \in \log G_{\|J'\|}$,

$$[\iota_J(x_J), \iota_{J'}(x_{J'})] = 0$$

if $J \cap J' \neq \emptyset$ and

$$[\iota_J(x_J), \iota_{J'}(x_{J'})] = \iota_{J \cup J'}([x_J, x_{J'}])$$

otherwise. As is shown in Theorem E.10 of [12], such a definition complies with the Lie bracket axioms.

We then define the $\mathbb{N}^{|s|}$-filtration of $\log \tilde{G}$. For any $(a_1,\dots,a_{|s|}) \in \mathbb{N}^{|s|}$, let $\log \tilde{G}_{(a_1,\dots,a_{|s|})}$ denote the sub Lie algebra of $\log \tilde{G}$ generated by $\iota_J(x_J)$ for all J with $\mathbf{1}_J(j) \geq a_j$ for all $1 \leq j \leq |s|$, and for all $x_J \in G_{\|J\|}$. As is shown in Theorem E.10 of [12], this is an $\mathbb{N}^{|s|}$-filtration of multi-degree $(1,\dots,1)$ and it induces an $\mathbb{N}^{|s|}$-filtration of multi-degree $(1,\dots,1)$ of the Lie group $\tilde{G}$ by taking the exponential function.

Let $\tilde{\Gamma}$ be the group generated by $\exp(W!\iota_J(\log \gamma_j))$ for some $W \in \mathbb{N}$ to be chosen latter for all $J \subseteq \{1,\dots,|s|\}$ and $\gamma_j \in \Gamma_{\|J\|}$. Again by Theorem E.10 of [12], $\tilde{G}/\tilde{\Gamma}$ is a nilmanifold and that $\tilde{\Gamma}_{(1,\dots,1)}$ is contained in $\iota_{(1,\dots,1)}(\log \Gamma_s)$ if $W \gg_{d,s} 1$. Fix such a W. It is not hard to see that $\tilde{G}/\tilde{\Gamma}$ is of complexity $O_{C,d,s}(1)$. Let $\tilde{\eta}$ be the vertical frequency on $\tilde{G}_{(1,\dots,1)}$ given by

$$\tilde{\eta}(\exp(\iota_{(1,\dots,1)}(\log g_s))) := \eta(g_s).$$

Since $\tilde{\Gamma}_{(1,\dots,1)} \subseteq \iota_{(1,\dots,1)}(\log \Gamma_s)$ and $G_{(1,\dots,1)}$ is a central subgroup of G, $\tilde{\eta}$ is indeed a vertical frequency. Moreover, $\tilde{\eta}$ is of complexity $O_{C,d,s}(1)$.

We then construct a function $\tilde{F} \in \mathrm{Lip}(\tilde{G}/\tilde{\Gamma} \to \mathbb{S}^{O_{C,d,s}(1)})$ with vertical frequency $\tilde{\eta}$. Such a function can be constructed using partitions of unity as in (6.3) of [12], and so we omit the details. Moreover, we can make the complexity and dimension of $\tilde{F}$ to be $O_{C,d,s}(1)$.

Next we define a polynomial sequence $\tilde{g}$ as follows. For all $j = (j_1, \dots, j_k)$, $j_i \in \mathbb{N}^d$, denote $\tilde{j} := (|j_1|, \dots, |j_k|) \in \mathbb{N}^k$. For convenience denote $n := (n_1, \dots, n_k)$ for $n \in (\mathbb{Z}^d)^k$. By Lemma 2.20 and the Baker-Campbell-Hausdorff formula, we may write

$$g(n_1, \dots, n_k) = \prod_{j=(j_1,\dots,j_k)\in(\mathbb{N}^d)^k} g_j^{n_1^{j_1}\dots n_k^{j_k}}$$

for some $g_j \in G_j$, where $j = (j_1, \dots, j_k)$, $j_i \in \mathbb{N}^d$ ranges over all elements in $(\mathbb{N}^d)^k$ such that $\tilde{j} \preceq s$, and these elements are arranged in some arbitrary order.

For each $J \subseteq \{1, \dots, |s|\}$ and $j = (j_1, \dots, j_k) = (j_{i,t})_{1\leq i\leq k, 1\leq t\leq d}$, let $P(J, j)$ be the set of tuples $(I_{i,t})_{1\leq i\leq k,1\leq t\leq d}$, $I_{i,t} \subseteq J \cap \{s_1 + \dots + s_{i-1} + 1, \dots, s_1 + \dots + s_i\}$ with $|I_{i,t}| = j_{i,t}$ such that $J \cap \{s_1 + \dots + s_{i-1} + 1, \dots, s_1 + \dots + s_i\} = \cup_{t=1}^d I_{i,t}$. Clearly, if $P(J, j)$ is non-empty, then $\|J\| = \tilde{j}$. For $j = (j_1, \dots, j_k) = (j_{i,t})_{1\leq i\leq k,1\leq t\leq d}$, $j_{i,t} \in \mathbb{N}$, denote $j! := \prod_{1\leq i\leq k,1\leq t\leq d} j_{i,t}!$. We then write

$$\begin{aligned}&\tilde{g}(n_1, \dots, n_{|s|})\\ :=& \prod_{j=(j_1,\dots,j_k)\in(\mathbb{N}^d)^k, \tilde{j}\preceq s} \exp\Big(j! \sum_{J\subseteq\{1,\dots,|s|\}\colon \|J\|=\tilde{j}} \ \sum_{(I_{i,t})_{1\leq i\leq k,1\leq t\leq d}\in P(J,j)} \big(\prod_{i=1}^k \prod_{t=1}^d \prod_{\ell\in I_{i,t}} n_{\ell,t}\big)\iota_J(\log g_j)\Big)\end{aligned}$$

for all $n_1, \dots, n_{|s|} \in \mathbb{F}_p^d$, where $n_{\ell,t}$ is the t-th entry of n_ℓ. Since each monomial

$$(n_1, \dots, n_{|s|}) \mapsto \exp\Big(j! \sum_{J\subseteq\{1,\dots,|s|\}\colon \|J\|=\tilde{j}} \ \sum_{(I_{i,t})_{1\leq i\leq k,1\leq t\leq d}\in P(J,j)} \big(\prod_{i=1}^k \prod_{t=1}^d \prod_{\ell\in I_{i,t}} n_{\ell,t}\big)\iota_J(\log g_j)\Big)$$

belongs to $\mathrm{poly}((\mathbb{Z}^d)^{|s|} \to \tilde{G})$, so does $\tilde{g}$ by Corollary B.4 of [12].

Set

$$\tilde{\xi}(n_1, \dots, n_{|s|}) := \tilde{F}(\tilde{g}(\tau(n_1, \dots, n_{|s|}))\tilde{\Gamma}) \text{ for all } (n_1, \dots, n_{|s|}) \in (\mathbb{Z}^d)^{|s|}$$

and

$$\xi'(n_1, \dots, n_k) := \tilde{\xi}(n_1, \dots, n_1, \dots, n_k, \dots, n_k) \text{ for all } (n_1, \dots, n_k) \in (\mathbb{Z}^d)^k,$$

(with each n_i repeated s_i times). From the construction, we see that $\tilde{\xi}$ is a nilcharacter on $(\mathbb{Z}^d)^{|s|}$ of multi-degree $(1, \dots, 1)$ (with 1 repeated $|s|$ times) and ξ' is a nilcharacter on $(\mathbb{Z}^d)^k$ of multi-degree s, and that both of them are of complexity and dimension at most $O_{C,d,s}(1)$. It is also clear that $\tilde{\xi}$ is symmetric with respect to the permutations of $n_{s_1+\dots+s_{i-1}+1}, \dots, n_{s_1+\dots+s_i}$ for all $1 \leq i \leq k$. However, we caution the readers that $\tilde{\xi}$ and ξ' need not to be p-periodic.

Claim. The map

$$n \mapsto \xi(n)^{\otimes s!} \otimes \overline{\xi'}(n) \text{ for all } n \in (\mathbb{Z}^d)^k \tag{C.4}$$

belongs to $\mathrm{Nil}^{<s;O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{Z}^d)^k)$.

It is not hard to see that $\tilde{j}! = j!|P(J, j)|$ whenever $\|J\| = \tilde{j}$. So we may expand (C.4) as

$$(F^{\otimes s!} \otimes \overline{\tilde{F}})\Big(\prod_{j=(j_1,\dots,j_k)\in(\mathbb{N}^d)^k, \tilde{j}\preceq s} \Big(g_j, \exp\Big(\tilde{j}! \sum_{J\subseteq\{1,\dots,|s|\}\colon \|J\|=\tilde{j}} \iota_J(\log g_j)\Big)\Big)^{n_1^{j_1}\dots n_k^{j_k}} (\Gamma \times \tilde{\Gamma})\Big).$$

The function $(F^{\otimes s!}\otimes\overline{\tilde{F}})$ is a Lipschitz function of complexity $O_{C,d,s}(1)$ on the nilmanifold $(G\times\tilde{G})/(\Gamma\times\tilde{\Gamma})$ of complexity $O_{C,d,s}(1)$. Let G^* be the subgroup of $G\times\tilde{G}$ defined by

$$G^*:=\{(g,\exp(s!\iota_{(1,\dots,1)}(\log g)))\colon g\in G_s\}\leq G_s\times\tilde{G}_{(1,\dots,1)}.$$

This is a rational central subgroup of $G\times\tilde{G}$. As F and $\tilde{F}$ have vertical frequencies η and $\tilde{\eta}$ respectively, $F^{\otimes s!}\otimes\overline{\tilde{F}}$ is invariant under the group action G^* and thus descends to a Lipschitz function F' on the nilmanifold G'/Γ', where $G':=(G\times\tilde{G})/G^*$ and Γ' is the projection of $\Gamma\times\tilde{\Gamma}$ to G. We thus have

$$\xi(n)^{\otimes s!}\otimes\overline{\xi'}(n)=F'\Big(\prod_{j=(j_1,\dots,j_k)\in(\mathbb{N}^d)^k,\tilde{j}\leq s}(g'_j)^{n^j}\Gamma'\Big),\tag{C.5}$$

where g'_j is the projection of $\Big(g_j,\exp\Big(\tilde{j}!\sum_{J\subseteq\{1,\dots,|s|\}\colon \|J\|=\tilde{j}}\iota_J(\log g_j)\Big)\Big)$ to G'.

We now give G' an $\mathbb{N}^k$-filtration as follows. For all $u\in\mathbb{N}^k, u\leq s$, let G'_u be the group generated by

$$\Big(h_j,\exp\Big(\tilde{j}!\sum_{J\subseteq\{1,\dots,|s|\}\colon \|J\|=\tilde{j}}\iota_J(\log h_j)\Big)\Big)\mod G^*$$

for all $j\in(\mathbb{N}^d)^k$ with $\tilde{j}=u$ and $h_j\in G_{\|J\|}$, and

$$(h_j,id),(id,\exp(\iota_J(\log h_j)))\mod G^*$$

for all $J\subseteq\{1,\dots,s\}$ with $\|J\|=\tilde{j}\succ u$ and $h_j\in G_{\|J\|}$. By the Baker-Campbell-Hausdorff formula, one can show that this is a filtration of degree $\prec s$. (Here we used the fact that for any $K\subseteq\{1,\dots,|s|\}$ with $\|K\|=\tilde{j}+\tilde{j}'$, then number of partitions $K=J\cup J'$ with $\|J\|=\tilde{j}$ and $\|J'\|=\tilde{j}'$ is $\frac{(\tilde{j}+\tilde{j})!}{\tilde{j}!\tilde{j}'!}$, which cancels the $u!$ term appearing in the definition of G'_u.) Moreover, $g'_j\in G'_j$ be construction. So the right hand side of (C.5) is a nilsequence of degree $\prec s$, whose complexity and dimension are certainly bounded by $O_{C,d,s}(1)$. This proves the claim.

Due to the fact that $\tilde{\xi}$ is not p-periodic, we can not directly apply the claim and set $\tilde{\chi}:=\tilde{\xi}\circ\tau$ to complete the proof. Our strategy is to approximate $\tilde{\xi}$ with p-periodic nilcharacters. For convenience denote $n:=(n_1,\dots,n_k)\in(\mathbb{Z}^d)^k$. Since $\xi^{\otimes s!}\otimes\overline{\xi}'$ takes values in $\mathbb{S}^{O_{C,d,s}(1)}$, by the claim, there exists $\psi\in\mathrm{Nil}^{\prec s;O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{Z}^d)^k)\subseteq\mathrm{Nil}^{|s|-1;O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{Z}^d)^k)$ (one can take $\psi:=\overline{\xi}^{\otimes s!}\otimes\xi'$) such that

$$|\mathbb{E}_{n\in[p]^{dk}}\xi(n)^{\otimes s!}\otimes\overline{\tilde{\xi}}(n_1,\dots,n_1,\dots,n_k,\dots,n_k)\otimes\psi(n)|=1,$$

where we denote $[p]:=\{0,\dots,p-1\}$. Note that $\tilde{\xi}$ belongs to $\mathrm{Nil}^{(1,\dots,1);O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{Z}^d)^{|s|})$. By Corollary B.6 (applied to the set of $(n_{1,1},\dots,n_{1,s_1},\dots,n_{k,1},\dots,n_{k,s_k})\in(\mathbb{F}_p^d)^{|s|}$ such that $n_{i,j}=n_{i,j'}$ for all $1\leq i\leq k, 1\leq j,j'\leq s_i$), there exists $\tilde{\xi}'\in\Xi_p^{(1,\dots,1);O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{Z}^d)^{|s|})$ such that

$$|\mathbb{E}_{n\in[p]^{dk}}\xi(n)^{\otimes s!}\otimes\overline{\tilde{\xi}'}(n_1,\dots,n_1,\dots,n_k,\dots,n_k)\otimes\psi(n)|\gg_{C,d,s}1.\tag{C.6}$$

Denote

$$\tilde{\alpha}(n_{1,1},\dots,n_{1,s_1},\dots,n_{k,1},\dots,n_{k,s_k}) := \bigotimes_{\sigma_1,\dots,\sigma_k} \tilde{\xi}'(n_{1,\sigma_1(1)},\dots,n_{1,\sigma_1(s_1)},\dots,n_{k,\sigma_k(1)},\dots,n_{k,\sigma_k(s_k)})$$

for all $n_{i,j} \in \mathbb{Z}^d$, where $\sigma_1,\dots,\sigma_k$ ranges over all the permutations $\sigma_i\colon \{1,\dots,s_i\} \to \{1,\dots,s_i\}$, $1 \leq i \leq k$. By Lemmas B.1 and C.10, we have $\tilde{\alpha} \in \Xi_p^{(1,\dots,1);O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{Z}^d)^{|s|})$. By Lemma C.4 (vii), there exists $\tilde{\beta} \in \Xi_p^{(1,\dots,1);O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{Z}^d)^{|s|})$ such that

$$\tilde{\beta}^{\otimes s!} \sim_{O_{C,d,s}(1)} \tilde{\alpha} \mod \Xi_p^{(1,\dots,1)}((\mathbb{Z}^d)^{|s|}).$$

Since $\tilde{\alpha}$ is symmetric with respect to the permutations of $n_{s_1+\dots+s_{i-1}+1},\dots,n_{s_1+\dots+s_i}$ for all $1 \leq i \leq k$ by construction, it follows from Lemma C.4 (vii) that we may also require $\tilde{\beta}$ to be symmetric with respect to these permutations.

Set

$$\alpha'(n_1,\dots,n_k) := \tilde{\alpha}(n_1,\dots,n_1,\dots,n_k,\dots,n_k)$$

and

$$\beta'(n_1,\dots,n_k) := \tilde{\beta}(n_1,\dots,n_1,\dots,n_k,\dots,n_k)$$

for all $n_1,\dots,n_k \in \mathbb{Z}^d$. Since $\tilde{\alpha},\tilde{\beta} \in \Xi_p^{(1,\dots,1);O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{Z}^d)^{|s|})$, by Lemma C.10, we have that $\alpha',\beta' \in \Xi_p^{s;O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{Z}^d)^k)$. By Lemma C.4 (vi), we have that

$$\beta'^{\otimes s!} \sim_{O_{C,d,s}(1)} \alpha' \mod \Xi_p^{|s|}((\mathbb{Z}^d)^k).$$

So

$$\begin{aligned}&\tilde{\xi}'(n_1,\dots,n_1,\dots,n_k,\dots,n_k)^{\otimes s!}\\ =\ &\alpha'(n_1,\dots,n_k) \sim_{O_{C,d,s}(1)} \beta'(n_1,\dots,n_k)^{\otimes s!} \mod \Xi_p^{|s|}((\mathbb{Z}^d)^k).\end{aligned}$$

By Lemmas C.2, C.4 (ii), C.6, (C.6) and the Pigeonhole Principle, there exists $\psi' \in \mathrm{Nil}^{|s|-1;O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{Z}^d)^k)$ such that

$$|\mathbb{E}_{n\in[p]^{dk}}\xi(n)^{\otimes s!} \otimes \overline{\beta'}(n) \otimes \psi'(n)| \gg_{C,d,s} 1.$$

Denoting $\tilde{\chi} := \tilde{\beta} \circ \tau \in \Xi_p^{(1,\dots,1);O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{F}_p^d)^{|s|})$ and $\chi' := \beta' \circ \tau \in \Xi_p^{s;O_{C,d,s}(1),O_{C,d,s}(1)}((\mathbb{F}_p^d)^k)$, we have that

$$\chi'(n_1,\dots,n_k) = \tilde{\chi}(n_1,\dots,n_1,\dots,n_k,\dots,n_k) \text{ for all } n_1,\dots,n_k \in \mathbb{F}_p^d$$

and that

$$|\mathbb{E}_{n\in(\mathbb{F}_p^d)^k}\xi(\tau(n))^{\otimes s!} \otimes \overline{\chi'}(n) \otimes \psi'(\tau(n))| \gg_{C,d,s} 1. \tag{C.7}$$

Since $\chi = \xi \circ \tau$, By Lemma C.7 and (C.7), we have that $\chi^{\otimes s!} \sim_{O_{C,d,s}(1)} \chi' \mod \Xi_p^{|s|}((\mathbb{F}_p^d)^{|s|})$. Finally, since $\tilde{\beta}$ is symmetric with respect to the permutations of $n_{s_1+\dots+s_{i-1}+1},\dots,n_{s_1+\dots+s_i}$ for all $1 \leq i \leq k$, so is $\tilde{\chi}$. This completes the proof. □

Appendix D. The converse of the spherical Gowers inverse theorem

In this appendix, we provide the proof of the converse direction of SGI(s).

Proposition D.1 (Converse of SGI(s))**.** Let $r, s \in \mathbb{N}$, $d, D \in \mathbb{N}_+$, and $C, \epsilon > 0$. There exist $\delta := \delta(C, d, D, \epsilon) > 0$ and $p_0 := p_0(C, d, D, \epsilon) \in \mathbb{N}$ such that for every prime $p \geq p_0$, every quadratic form $M\colon \mathbb{F}_p^d \to \mathbb{F}_p$, every affine subspace $V + c$ of $\mathbb{F}_p^d$ of co-dimension r, every function $f\colon \mathbb{F}_p^d \to \mathbb{C}^D$ bounded by 1, and every $\psi \in \mathrm{Nil}^{s;C,D}(V(M) \cap (V + c))$, if $\mathrm{rank}(M|_{V+c}) \geq s^2 + 3s + 5$ and

$$\left|\mathbb{E}_{n\in V(M)\cap(V+c)} f(n) \otimes \psi(n)\right| > \epsilon, \tag{D.1}$$

then $\|f\|_{U^{s+1}(V(M)\cap(V+c))} > \delta$.

We need the following lemma for the proof of Proposition D.1.

Lemma D.2. Let $d, D, s \in \mathbb{N}_+$, $C > 0$ and p be a prime. If $\chi \in \Xi_p^{s;C,D}(\mathbb{F}_p^d)$, then for all $h \in \mathbb{F}_p^d$, the sequence $n \mapsto \chi(n + h) \otimes \overline{\chi}(n), n \in \mathbb{F}_p^d$ belongs to $\mathrm{Nil}_p^{s-1;O_{C,D,s}(1)}(\mathbb{F}_p^d)$.

Proof. Let $((G/\Gamma)_I, g, F, \eta)$ be a $\Xi_p^s(\mathbb{F}_p^d)$-representation of χ of complexity at most C and dimension D. Write

$$\chi'(n) := F(g(n + \tau(h))\Gamma) \otimes \overline{F}(g(n)\Gamma) \text{ for all } n \in \mathbb{Z}^d.$$

Since $g \in \mathrm{poly}_p(\mathbb{Z}^d \to G_{\mathbb{N}}|\Gamma)$, it is clear that

$$\chi'(\tau(n)) = F(g(\tau(n)+\tau(h))\Gamma)\otimes\overline{F}(g(\tau(n))\Gamma) = F(g(\tau(n+h))\Gamma)\otimes\overline{F}(g(\tau(n))\Gamma) = \chi(n+h)\otimes\overline{\chi}(n)$$

for all $n \in \mathbb{F}_p^d$. It suffices to show that $\chi' \in \mathrm{Nil}_p^{s-1;O_{C,D,s}(1)}(\mathbb{Z}^d)$.

Our method is similar to the discussion in Section 8.3 of [19]. Let $(G^\square/\Gamma^\square)_{\mathbb{N}}$, $(G^\square)_i$, G_s^Δ and $(\overline{G^\square}/\overline{\Gamma^\square})_{\mathbb{N}}$ be defined as in Section 8.3 of [19]. Let $g_h^\square(n) := (g(n + \tau(h)), g(n)) \in \mathrm{poly}_p(\mathbb{Z}^d \to (G^\square)_{\mathbb{N}}|\Gamma^\square)$ and $F_h^\square(n) := F \otimes \overline{F}$. Since F has vertical frequency η, $F_h^\square$ vanishes on G_s^Δ. So $F_h^\square$ descents into a Lipschitz function on $\overline{G^\square}/\overline{\Gamma^\square}$ of complexity $O_{C,d,D,\epsilon}(1)$. So letting $\overline{g^\square}$ denote the projection of $g_h^\square$ on $\overline{G^\square}$, we have that $\overline{g^\square} \in \mathrm{poly}_p(\mathbb{Z}^d \to (\overline{G^\square})_{\mathbb{N}}|\overline{\Gamma^\square})$ and thus

$$\chi'(n) = F_h^\square(g_h^\square(n)\Gamma_h^\square) = \overline{F^\square}(\overline{g_h^\square}(n)\overline{\Gamma^\square})$$

belongs to $\mathrm{Nil}_p^{s-1;O_{C,D,s}(1)}(\mathbb{Z}^d)$. □

Proof of Proposition D.1. The scheme of the proof is similar to Proposition 1.4 of [11]. There is nothing to prove when $s = 0$ since degree 0 nilsequences are constants. Suppose now that the conclusion holds for some $s - 1 \in \mathbb{N}$ and we prove it for s. By Corollary B.6, the induction hypothesis, (D.1) and the Pigeonhole Principle, there exist $\psi' \in \Xi_p^{s;O_{C,d,D,\epsilon}(1),O_{C,d,D,\epsilon}(1)}(V(M) \cap (V + c))$ such that

$$\left|\mathbb{E}_{n\in V(M)\cap(V+c)} f(n) \otimes \psi'(n)\right| \gg_{C,d,D,\epsilon} 1. \tag{D.2}$$

Let $((G/\Gamma)_{\mathbb{N}}, g, F, \eta)$ be a $\Xi_p^s(V(M) \cap (V + c))$-representation of ψ' of complexity and dimension at most $O_{C,d,D,\epsilon}(1)$. Write $\tilde{g} = g \circ \tau$. Let $\phi\colon \mathbb{F}_p^{d-r} \to V$ be any bijective linear transformation and set $M'(m) := M(\phi(m) + c)$. Then $(n, h) \in \Box_1(V(M) \cap (V + c))$ if and only if $n = \phi(m) + c, h = \phi(h')$ for some $(m, h') \in V(M')$. By taking the square of (D.2), we have that

(D.3)

$$\begin{aligned} &1 \ll_{C,d,D,\epsilon} |\mathbb{E}_{(n,h)\in\Box_1(V(M)\cap(V+c))}\Delta_h f(n) \otimes F(\tilde{g}(n+h)\Gamma) \otimes \overline{F}(\tilde{g}(n)\Gamma)| \\ &= |\mathbb{E}_{(m,h')\in\Box_1(V(M'))}\Delta_{\phi(h')} f(\phi(m)+c) \otimes F(\tilde{g}(\phi(m+h')+c)\Gamma) \otimes \overline{F}(\tilde{g}(\phi(m)+c)\Gamma)| \\ &= \mathbb{E}_{h'\in\mathbb{F}_p^{d-r}}|\mathbb{E}_{m\in V(M')^{h'}}\Delta_{\phi(h')} f(\phi(m)+c) \otimes F(\tilde{g}(\phi(m+h')+c)\Gamma) \otimes \overline{F}(\tilde{g}(\phi(m)+c)\Gamma)| + O(p^{-1/2}), \end{aligned}$$

where we used Theorem A.12 in the last equality since $\text{rank}(M') = \text{rank}(M|_{V+c}) \geq 5$ and $\Box_1(V(M'))$ is an M'-set of total co-dimension 2. So if $p \gg_{C,d,D,\epsilon} 1$, then there exists $H' \subseteq \mathbb{F}_p^{d-r}$ of cardinality $\gg_{C,d,D,\epsilon} p^{d-r}$ such that for all $h' \in H'$,

$$|\mathbb{E}_{m\in V(M')^{h'}}\Delta_{\phi(h')} f(\phi(m)+c) \otimes F(\tilde{g}(\phi(m+h')+c)\Gamma) \otimes \overline{F}(\tilde{g}(\phi(m)+c)\Gamma)| \gg_{C,d,D,\epsilon} 1.$$

Let A' be the $(d-r) \times (d-r)$ matrix associated with M' and H'' be the set of $h' \in H'$ such that $h'A' \neq \mathbf{0}$. Then $|H' \backslash H''| \leq p^{d-r-\text{rank}(A')}$. Since $\text{rank}(A') = \text{rank}(M|_{V+c}) \geq 1$, we have that $|H''| \gg_{C,d,D,\epsilon} p^{d-r}$. For $h' \in H''$, note that $V(M')^{h'}$ is the intersection of $V(M)$ with $U_{h'} := \{m \in \mathbb{F}_p^{d-r} : M'(m+h') = M'(m)\}$. Since $h'A' \neq \mathbf{0}$, $U_{h'}$ is an affine subspace of $\mathbb{F}_p^{d-r}$ of co-dimension 1. By Proposition A.4,

$$\text{rank}(M'|_{U_{h'}}) \geq \text{rank}(M') - 2 = \text{rank}(M|_{V+c}) - 2 \geq s^2 + 3s + 3 \geq (s-1)^2 + 3(s-1) + 5.$$

On the other hand, for all $h' \in H''$, by Proposition 2.18, it is not hard to see that the map

$$m \mapsto F(\tilde{g}(\phi(m)+c)\Gamma)$$

belongs to $\Xi_p^{s;O_{C,d,D,\epsilon}(1),O_{C,d,D,\epsilon}(1)}(\mathbb{F}_p^{d-r})$. By Lemma D.2, we have that the map

$$m \mapsto F(\tilde{g}(\phi(m+h')+c)\Gamma) \otimes \overline{F}(\tilde{g}(\phi(m)+c)\Gamma)$$

belongs to $\text{Nil}_p^{s-1;O_{C,d,D,\epsilon}(1),O_{C,d,D,\epsilon}(1)}(\mathbb{F}_p^{d-r})$. By induction hypothesis,

$$\|\Delta_{h'} f(\phi(\cdot)+c)\|_{U^s(V(M')^{h'})} \gg_{C,d,D,\epsilon} 1$$

for all $h' \in H''$. Note that $\square_{s+1}(V(M'))$ is an M'-set of total co-dimension $\frac{(s+1)(s+2)}{2}+1$. If $p \gg_{C,d,D,\epsilon} 1$, then

$$\begin{aligned}
&\|f\|^{2^{s+1}}_{U^{s+1}(V(M)\cap(V+c))}\\
&=\mathbb{E}_{(n,h_1,\dots,h_{s+1})\in\square_{s+1}(V(M)\cap(V+c))}\prod_{\epsilon=(\epsilon_1,\dots,\epsilon_{s+1})\in\{0,1\}^{s+1}}C^{|\epsilon|}f\Big(n+\sum_{i=1}^{s+1}\epsilon_i h_i\Big)\\
&=\mathbb{E}_{(m,h'_1,\dots,h'_{s+1})\in\square_{s+1}(V(M'))}\prod_{\epsilon=(\epsilon_1,\dots,\epsilon_{s+1})\in\{0,1\}^{s+1}}C^{|\epsilon|}f\Big(\phi\Big(m+\sum_{i=1}^{s+1}\epsilon_i h'_i\Big)+c\Big)\\
&=\mathbb{E}_{h'_{s+1}\in\mathbb{F}_p^{d-r}}\mathbb{E}_{(m,h'_1,\dots,h'_s)\in\square_s(V(M')^{h'_{s+1}})}\prod_{\epsilon=(\epsilon_1,\dots,\epsilon_s)\in\{0,1\}^s}C^{|\epsilon|}\overline{\Delta_{h'_{s+1}}f}\Big(\phi\Big(m+\sum_{i=1}^{s+1}\epsilon_i h'_i\Big)+c\Big)+O_s(p^{-\frac{1}{2}})\\
&=\mathbb{E}_{h'_{s+1}\in\mathbb{F}_p^{d-r}}\|\overline{\Delta_{h'_{s+1}}f}(\phi(\cdot)+c)\|^{2^s}_{U^s(V(M')^{h'_{s+1}})}+O_s(p^{-\frac{1}{2}})\gg_{C,d,D,\epsilon}1,
\end{aligned}$$

where we used Lemma A.8 in the second equality, and Theorem A.12 in the third equality (since $\text{rank}(M')=\text{rank}(M|_{V+c})\geq s^2+3s+5$). This finishes the proof. □

References

[1] P. Candela, D. González-Sánchez, and B. Szegedy. On higher-order fourier analysis in characteristic p. *Ergodic Theory Dynam. Systems*, pages 1–70, 2021.

[2] P. Candela and O. Sisask. Convergence results for systems of linear forms on cyclic groups, and periodic nilsequences. *SIAM J. Discrete Math.*, **28**(2):786–810, 2014.

[3] N. Frantzikinakis and B. Host. Higher order Fourier analysis of multiplicative functions and applications. *J. Amer. Math. Soc.*, **30**:67–157, 2017.

[4] T. Gowers. A new proof of Szemerédi's theorem for arithmetic progressions of length four. *Geom. Funct. Anal.*, **8**:529–551, 1998.

[5] T. Gowers. A new proof of Szemerédi theorem. *Geom. Funct. Anal.*, **11**:465–588, 2001.

[6] B. Green and T. Tao. An inverse theorem for the Gowers U^3-norm, with applications. *Proc. Edinburgh Math. Soc.*, **51**:71–153, 2008.

[7] B. Green and T. Tao. The primes contain arbitrarily long arithmetic progressions. *Ann. of Math.*, **167**:481–547, 2008.

[8] B. Green and T. Tao. An arithmetic regularity lemma, an associated counting lemma, and applications, in "an irregular mind: Szemerédi is 70". *Bolyai Society Math. Studies*, **21**, 2010.

[9] B. Green and T. Tao. Linear equations in the primes. *Ann. of Math.*, **171**:1753–1850, 2010.

[10] B. Green and T. Tao. The quantitative behaviour of polynomial orbits on nilmanifolds. *Ann. of Math.*, **175**(2):465–540, 2012.

[11] B. Green, T. Tao, and T. Ziegler. An inverse theorem for the Gowers U^4-norm. *Glasg. Math. J.*, **53**:1–50, 2011.

[12] B. Green, T. Tao, and T. Ziegler. An inverse theorem for the Gowers U^{s+1}-norm. *Ann. of Math.*, **176**(2):1231–1372, 2012.

[13] B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math.*, **161**(1):397–488, 2005.

[14] A. Jamneshan, O. Shalom, and T. Tao. The structure of totally disconnected Host-Kra-Ziegler factors, and the inverse theorem for the U^k gowers uniformity norms on finite abelian groups of bounded torsion. arXiv:2303.04860, 2023.

[15] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifolds. *Ergodic Theory Dynam. Systems*, **25**(1):201–213, 2005.
[16] N. Lyall, A. Magyar, and H. Parshall. Spherical configurations over finite fields. *Amer. J. Math.*, **142**(2):373–404, 2020.
[17] F. Manners. Periodic nilsequences and inverse theorems on cyclic groups. arXiv: 1404.7742.
[18] E. J. McShane. Extension of range of functions. *Bulletin of the American Mathematical Society*, **40**(12):837–842, 1934.
[19] W. Sun. Spherical higher order Fourier analysis over finite fields I: equidistribution for nilsequences. arXiv: 2312.06651.
[20] W. Sun. Spherical higher order Fourier analysis over finite fields II: additive combinatorics for shifted ideals. arXiv: 2312.06650.
[21] W. Sun. Spherical higher order Fourier analysis over finite fields IV: an application to the geometric Ramsey conjecture. arXiv: 2312.06649.
[22] B. Szegedy. On higher order Fourier analysis. arXiv: 1203.2260.
[23] T. Tao and T. Ziegler. Conjecture for the Gowers norm over finite fields in low characteristic. *Ann. Comb.*, **16**:121–188, 2012.
[24] T. Tao and T. Ziegler. The inverse conjecture for the Gowers norm over finite fields via the correspondence principle. *Annals of Combinatorics*, **16**(1):121–188, 2012.

(Wenbo Sun) Department of Mathematics, Virginia Tech, 225 Stanger Street, Blacksburg, VA, 24061, USA
Email address: swenbo@vt.edu