CrossMark

Multiple correlation sequences and nilsequences

Nikos Frantzikinakis

Abstract We study the structure of multiple correlation sequences defined by measure preserving actions of commuting transformations. When the iterates of the transformations are integer polynomials we prove that any such correlation sequence is the sum of a nilsequence and an error term that is small in uniform density; this was previously known only for measure preserving actions of a single transformation. We then use this decomposition result to give convergence criteria for multiple ergodic averages and deduce some rather surprising results, for instance we infer convergence for actions of commuting transformations from the special case of actions of a single transformation. Our proof of the decomposition result differs from previous works of Bergelson, Host, Kra, and Leibman, as it does not rely on the theory of characteristic factors. It consists of a simple orthogonality argument and the main tool is an inverse theorem of Host and Kra for general bounded sequences.

Mathematics Subject Classification   Primary 37A30 ; Secondary 05D10 · 11B30 · 37A05

N. Frantzikinakis (✉)
Department of Mathematics, University of Crete, Voutes University Campus,
71003 Heraklion, Greece
e-mail: frantzikinakis@gmail.com

1 Introduction and results

1.1 Main result

Throughout this article, a *system* is a probability space $(X, \mathcal{X}, \mu)$ together with invertible, measure preserving transformations $T_1, \ldots, T_\ell \colon X \to X$ that commute. A multiple correlation sequence is a sequence of the form

$$\int T_1^{n_1} f_1 \cdot \ldots \cdot T_\ell^{n_\ell} f_\ell \, d\mu$$

where $(X, \mathcal{X}, \mu, T_1, \ldots, T_\ell)$ is a system, $f_1, \ldots, f_\ell \in L^\infty(\mu)$, and $n_1, \ldots, n_\ell \in \mathbb{Z}$. The study of the limiting behavior of averages of such sequences, where the iterates are restricted to certain subsets of $\mathbb{Z}^\ell$, has been an indispensable tool in ergodic Ramsey theory and in particular in proving various far reaching extensions of Szemerédi's theorem on arithmetic progressions. Although the precise structure of the multiple correlation sequences is unknown even when $n_1 = \cdots = n_\ell = n$, there is a widespread belief that modulo negligible terms the building blocks are sequences with algebraic structure (see [7, Problem 1] for a related conjecture).

Definition ([5]) For $\ell \in \mathbb{N}$, an *ℓ-step nilsequence* is a complex valued sequence of the form $(F(g^n \Gamma))$, where $F \in C(X)$, $X = G/\Gamma$, G is an ℓ-step nilpotent Lie group, Γ is a discrete cocompact subgroup, and $g \in G$. A *0-step nilsequence* is a constant sequence.

When $T_i = T^i$, $i = 1, \ldots, \ell$, following the discovery of characteristic factors with algebraic structure for some closely related multiple ergodic averages, Bergelson et al. proved the following beautiful result (see also [17] for related work for $\ell = 3$):

Theorem ([5, Theorem 1.9]) *For $\ell \in \mathbb{N}$, let $(X, \mathcal{X}, \mu, T)$ be an ergodic system and $f_1, \ldots, f_\ell \in L^\infty(\mu)$ be functions with $\|f_i\|_\infty \leq 1$. Then we have the decomposition*

$$\int T^n f_1 \cdot \ldots \cdot T^{\ell n} f_\ell \, d\mu = a_{st}(n) + a_{er}(n), \quad n \in \mathbb{N},$$

where

(i) *$(a_{st}(n))$ is a uniform limit of $(\ell - 1)$-step nilsequences with $\|a_{st}\|_\infty \leq 1$;*
(ii) $\lim_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N-1} |a_{er}(n)|^2 = 0.$

This result was extended by Leibman to cover polynomial iterates in [14] and not necessarily ergodic transformations in [15]. The proofs of these results

depend in an essential way on the fact that characteristic factors for some suitable multiple ergodic averages are inverse limits of nilsystems. This is no longer true for correlation sequences involving actions of commuting transformations, which is why efforts to prove decomposition results for such sequences did not bring any results so far. In fact, characteristic factors for commuting actions are known to be extremely complex (for related work see [2,3]) which has raised suspicions that decomposition results in this more general setup may involve sequences very different from nilsequences. Our main result settles this rather elusive problem; we show that modulo error terms that are small in uniform density, correlation sequences of actions of commuting transformations are nilsequences.

Theorem 1.1 *For $\ell \in \mathbb{N}$ let $(X, \mathcal{X}, \mu, T_1, \ldots, T_\ell)$ be a system and $f_1, \ldots, f_\ell \in L^\infty(\mu)$ be functions with $\|f_i\|_\infty \leq 1$. Then for every $\varepsilon > 0$ we have the decomposition*

$$\int T_1^n f_1 \cdot \ldots \cdot T_\ell^n f_\ell \, d\mu = a_{st}(n) + a_{er}(n), \quad n \in \mathbb{N}, \qquad (1)$$

where

(i) $(a_{st}(n))$ is an $(\ell - 1)$-step nilsequence with $\|a_{st}\|_\infty \leq 1$;
(ii) $\lim_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N-1} |a_{er}(n)|^2 \leq \varepsilon$.

Remark We do not know if a strengthening similar to the one in [5, Theorem 1.9] holds where one uses uniform limits of nilsequences in (i) and takes $\varepsilon = 0$ in (ii).

Our argument is rather versatile and does not rely on the theory of characteristic factors; we rather focus on some distinctive properties correlation sequences as in (1) satisfy (see Theorem 1.3). The idea that starts the proof comes from answering the following natural question: "Can a multiple correlation sequence as in (1) be asymptotically orthogonal to all $(\ell - 1)$-step nilsequences?"

On the one hand, using an inverse theorem of Host and Kra (see Theorem 2.1), one gets that any such sequence has to be U_ℓ-uniform. On the other hand, by successively applying van der Corput's lemma one sees that a sequence of the form (1) is asymptotically orthogonal to all U_ℓ-uniform sequences. Hence, any sequence that provides a positive answer to our question has to be asymptotically orthogonal to itself, that is, has to converge to 0 in density.

With this idea in mind, we prove our main result as follows: Given a sequence $(a(n))$ as in (1), we consider the $(\ell - 1)$-step nilsequence, call it a_{st}, that lies "closest" to $(a(n))$ with respect to the semi-norm $\|\cdot\|_2$ defined in (3). Then

$a_{er} := a - a_{st}$ is asymptotically orthogonal to all $(\ell - 1)$-step nilsequences, and arguing as before, we get that a_{st} and a_{er} have the asserted properties. A slight complication appears because for $\ell \geq 2$ the space of $(\ell - 1)$-step nilsequences (or uniform limits of such sequences) is not $\|\cdot\|_2$-complete; this is the reason why we are led to an error term a_{er} that is small, but not zero, in uniform density. For our argument to work we also have to make sure that various limits of uniform Cesàro averages exist; to guarantee this, we use a result of Austin [1].

Using a variant of the previous argument and a result of Walsh [19] we get:

Theorem 1.2 *Let $\ell, m \in \mathbb{N}$ and $p_{i,j} \in \mathbb{Z}[t]$, $i = 1, \ldots, \ell$, $j = 1, \ldots, m$, be polynomials. Then there exists $k \in \mathbb{N}$, $k = k(\ell, m, \max \deg(p_{i,j}))$, such that for every system $(X, \mathcal{X}, \mu, T_1, \ldots, T_\ell)$, functions $f_1, \ldots, f_m \in L^\infty(\mu)$ with $\|f_i\|_\infty \leq 1$, and $\varepsilon > 0$, we have the decomposition*

$$\int \left(\prod_{i=1}^\ell T_i^{p_{i,1}(n)} \right) f_1 \cdot \ldots \cdot \left(\prod_{i=1}^\ell T_i^{p_{i,m}(n)} \right) f_m \, d\mu = a_{st}(n) + a_{er}(n), \quad n \in \mathbb{N},$$

(2)

where

(i) *$(a_{st}(n))$ is a k-step nilsequence with $\|a_{st}\|_\infty \leq 1$;*
(ii) *$\lim_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N-1} |a_{er}(n)|^2 \leq \varepsilon$.*

1.2 A more general framework

It turns out that Theorem 1.1 is a manifestation of a more general principle which asserts that if a sequence is asymptotically orthogonal to all U_ℓ-uniform sequences and satisfies some necessary regularity conditions, then it admits a decomposition like the one in Theorem 1.1. To make this more precise we introduce some notation (see Sect. 2.1 for the definition of the uniformity seminorms).

Definition Let $\ell \in \mathbb{N}$. We say that the bounded sequence $a \colon \mathbb{N} \to \mathbb{C}$ is

(i) *ℓ-anti-uniform* if there exists $C := C(\ell, a)$ such that

$$\limsup_{N-M \to \infty} \left| \frac{1}{N-M} \sum_{n=M}^{N-1} a(n)b(n) \right| \leq C \|b\|_{U_\ell(\mathbb{N})}$$

for every $b \in \ell^\infty$.

(ii) ℓ-*regular* if the limit

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} a(n)\psi(n)$$

exists for every $(\ell-1)$-step nilsequence $(\psi(n))$.

Theorem 1.3 *For $\ell \in \mathbb{N}$ let $a\colon \mathbb{N} \to \mathbb{C}$ be a sequence with $\|a\|_\infty \leq 1$ that is ℓ-anti-uniform and ℓ-regular. Then for every $\varepsilon > 0$ we have the decomposition*

$$a(n) = a_{st}(n) + a_{er}(n), \quad n \in \mathbb{N},$$

where

(i) $(a_{st}(n))$ *is an $(\ell-1)$-step nilsequence with $\|a_{st}\|_\infty \leq 1$;*
(ii) $\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} |a_{er}(n)|^2 \leq \varepsilon$.

Remark For general ℓ-regular sequences a similar result is proved in [12, Theorem 2.19] with an error term that is small with respect to the seminorm $\|\cdot\|_{U_\ell(\mathbb{N})}$.

A sequence $(a(n))$ that satisfies the asserted decomposition has to be ℓ-regular. It also has to satisfy the estimate defining the ℓ-anti-uniformity property if one introduces an arbitrarily small error term ε on the right hand side and allows C to depend on ε (this follows from [12, Theorem 2.14]).

Theorem 1.3 fails if we use standard Cesàro averages to define the notions of anti-uniformity and regularity (and leave the definition of $\|\cdot\|_{U_\ell(\mathbb{N})}$ as is); the sequence $(e^{i\sqrt{n}})$, illustrates this. The same sequence shows that anti-uniformity does not imply regularity $((e^{i\sqrt{n}})$ is 2-anti-uniform but not 1-regular).

1.3 Applications

On $\ell^\infty(\mathbb{N})$ we define the seminorm $\|\cdot\|_2$ by

$$\|a\|_2^2 := \limsup_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} |a(n)|^2. \tag{3}$$

For $\ell \in \mathbb{N}$ we consider the following subspaces of $\ell^\infty(\mathbb{N})$:

$$\mathcal{A}_\ell := \Big\{(\psi(n))\colon \psi \text{ is an } (\ell-1)\text{-step nilsequence}\Big\};$$

$$\mathcal{B}_\ell := \left\{ \left(\int T^{k_1 n} f_1 \cdot \ldots \cdot T^{k_\ell n} f_\ell \, d\mu \right) : (X, \mathcal{X}, \mu, T) \text{ is a system,} \right.$$
$$\left. f_i \in L^\infty(\mu), k_i = \ell! / i \right\};$$

$$\mathcal{C}_\ell := \left\{ \left(\int T_1^n f_1 \cdot \ldots \cdot T_\ell^n f_\ell \, d\mu \right) : (X, \mathcal{X}, \mu, T_1, \ldots, T_\ell) \text{ is a system and} \right.$$
$$\left. f_i \in L^\infty(\mu) \right\}.$$

After Proposition 2.4 we explain why in the definition of $\mathcal{B}_\ell$ we use the exponents $k_1, \ldots, k_\ell$ instead of $1, \ldots, \ell$. The space $\mathcal{A}_\ell$ is linear since if for $i = 1, 2$, $(F_i(g_i^n \Gamma_i))$ are $(\ell - 1)$-step nilsequences on G_i / Γ_i, then their sum is the $(\ell - 1)$-step nilsequence $(F(g^n \Gamma))$ on G / Γ, where $G = G_1 \times G_2$, $\Gamma := \Gamma_1 \times \Gamma_2$, $g := (g_1, g_2)$, $F(g\Gamma) := F_1(g_1 \Gamma_1) + F_2(g_2 \Gamma_2)$. To see that the space $\mathcal{C}_\ell$ is linear (similarly for $\mathcal{B}_\ell$), let $a, b \in \mathcal{C}_\ell$ be defined by the systems $(X_i, \mathcal{X}_i, \mu_i, T_i)$ and the functions $f_1^i, \ldots, f_\ell^i, i = 1, 2$. Then $c := (a + b)/2$ is also a multiple correlation sequence defined by the system $(X, \mathcal{X}, \mu, T)$, where $X = X_1 \cup X_2$ (considered as disjoint subsets) with the corresponding σ-algebra $\mathcal{X}$, $\mu := (\mu_1 + \mu_2)/2$, T equals T_1 on X_1 and T_2 on X_2, and $f_i := f_i^1 \mathbf{1}_{X_1} + f_i^2 \mathbf{1}_{X_2}, i = 1, \ldots, \ell$.

It is a rather striking fact that, modulo sequences that are small in uniform density, the three subspaces $\mathcal{A}_\ell, \mathcal{B}_\ell, \mathcal{C}_\ell$ coincide.

Theorem 1.4 *For every $\ell \in \mathbb{N}$ we have*

$$\overline{\mathcal{A}_\ell} = \overline{\mathcal{B}_\ell} = \overline{\mathcal{C}_\ell}$$

where the closure is taken with respect to the seminorm $\|\cdot\|_2$ defined in (3).

It is not hard to see that the first equality fails if we consider closures with respect to the $\|\cdot\|_\infty$ norm. The second equality may still hold under such circumstances but this is not something we can prove with the methods developed so far.

The next two results illustrate some rather surprising principles: (i) convergence results for actions of a single transformation automatically imply stronger convergence results for actions of commuting transformations; and (ii) convergence results involving linear iterates automatically imply stronger convergence results involving polynomial iterates.

Theorem 1.5 *Let (r_n) be a strictly increasing sequence of integers such that $r_n = O(n)$. Then for every $\ell \in \mathbb{N}$ the following statements are equivalent:*
 (i) *For every $(\ell{-}1)$-step nilsequence $(\psi(n))$ the limit $\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \psi(r_n)$ exists.*

(ii) *For every system $(X, \mathcal{X}, \mu, T)$, functions $f_1, \ldots, f_\ell \in L^\infty(\mu)$, and for $k_i = \ell!/i$, $i = 1, \ldots, \ell$, the following limit exists*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \int T^{k_1 r_n} f_1 \cdot \ldots \cdot T^{k_\ell r_n} f_\ell \, d\mu.$$

(iii) *For every system $(X, \mathcal{X}, \mu, T_1, \ldots, T_\ell)$ and functions $f_1, \ldots, f_\ell \in L^\infty(\mu)$ the following limit exists*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \int T_1^{r_n} f_1 \cdot \ldots \cdot T_\ell^{r_n} f_\ell \, d\mu.$$

Remark Equivalently, the growth condition $r_n = O(n)$ holds if the set $R := \{r_1, r_2, \ldots\}$ has positive lower natural density.

In the previous result we have established an equivalence for every fixed $\ell \in \mathbb{N}$, in the next result we have to assume that a certain property is known for every $\ell \in \mathbb{N}$ in order to establish an equivalence [this is needed for the equivalence of (ii) and (iii)].

Theorem 1.6 *Let (r_n) be a strictly increasing sequence of integers such that $r_n = O(n)$. Then the following statements are equivalent:*

(i) *For every $\ell \in \mathbb{N}$ and ℓ-step nilsequence $(\psi(n))$ the limit $\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \psi(r_n)$ exists.*

(ii) *For every $\ell \in \mathbb{N}$, system $(X, \mathcal{X}, \mu, T)$, and functions $f_1, \ldots, f_\ell \in L^\infty(\mu)$, the following limit exists*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \int T^{r_n} f_1 \cdot \ldots \cdot T^{\ell r_n} f_\ell \, d\mu.$$

(iii) *For every $\ell \in \mathbb{N}$, polynomials $p_1, \ldots, p_\ell \in \mathbb{Z}[t]$, system $(X, \mathcal{X}, \mu, T_1, \ldots, T_\ell)$, and functions $f_1, \ldots, f_\ell \in L^\infty(\mu)$, the following limit exists*

$$\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N} \int T_1^{p_1(r_n)} f_1 \cdot \ldots \cdot T_\ell^{p_\ell(r_n)} f_\ell \, d\mu.$$

Similar results hold if in (i)–(iii) of Theorems 1.5 and 1.6 one replaces the limit $\lim_{N \to \infty} \frac{1}{N} \sum_{n=1}^{N}$ with the limit $\lim_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N-1}$ and the growth assumption on (r_n) with the assumption that the range of this sequence has positive lower Banach density. Furthermore, the same method can be used to prove convergence criteria for weighted averages where for a given

bounded sequence of complex numbers (w_n) one replaces in (i)–(iii) of Theorems 1.5 and 1.6 the averaging operation $\frac{1}{N} \sum_{n=1}^{N}$ with the averaging operation $\frac{1}{N} \sum_{n=1}^{N} w_n$.

1.4 Conjectures

The growth assumption on (r_n) in Theorems 1.5 and 1.6 is crucial for our argument to work as the proofs use Theorem 1.1 which is not helpful for sequences that grow faster than linearly. Nevertheless, we believe that the following is true:

Conjecture 1 *In Theorems 1.5 and 1.6 the growth assumption on (r_n) is superfluous.*

We also believe in the following strengthening of the second identity in Theorem 1.4:

Conjecture 2 *For every $\ell \in \mathbb{N}$ we have $\overline{\mathcal{B}_\ell} = \overline{\mathcal{C}_\ell}$ where the closure is taken with respect to the norm $\|\cdot\|_\infty$.*

1.5 Notation

We denote by $\mathbb{N}$ the set of positive integers.

If $(a(n))$ is a bounded sequence we denote by $\limsup_{N-M\to\infty} |\frac{1}{N-M} \sum_{n=M}^{N-1} a(n)|$ the limit (it exists by subadditivity) $\lim_{N\to\infty} \sup_{M\in\mathbb{N}} \left|\frac{1}{N} \sum_{n=M}^{M+N-1} a(n)\right|$.

2 Proofs of results

2.1 Uniformity seminorms and the Host-Kra inverse theorem

We give a slight variant of the uniformity seminorms defined by Host and Kra [12].

Definition Let $\ell \in \mathbb{N}$ and $a \colon \mathbb{N} \to \mathbb{C}$ be a bounded sequence.

(i) Given a sequence of intervals $\mathbf{I} = (I_N)$ with lengths tending to infinity, we say that the sequence $(a(n))$ is *distributed regularly along* $\mathbf{I}$ if the limit

$$\lim_{N\to\infty} \frac{1}{|I_N|} \sum_{n\in I_N} a_1(n + h_1) \cdot \ldots \cdot a_r(n + h_r)$$

exists for every $r \in \mathbb{N}$ and $h_1, \ldots, h_r \in \mathbb{N}$, where a_i is either a or $\bar{a}$.

(ii) If **I** is as in (i) and $(a(n))$ is distributed regularly along **I**, we define inductively

$$\|a\|_{\mathbf{I},1} := \lim_{N\to\infty} \left| \frac{1}{|I_N|} \sum_{n\in I_N} a(n) \right|;$$

and for $\ell \geq 2$ (one can show as in [12, Proposition 4.3] that the next limit exists)

$$\|a\|_{\mathbf{I},\ell}^{2^\ell} := \lim_{H\to\infty} \frac{1}{H} \sum_{h=1}^{H} \|\sigma_h a \cdot \bar{a}\|_{\mathbf{I},\ell-1}^{2^{\ell-1}}$$

where σ_h is the shift transformation defined by $(\sigma_h a)(n) := a(n+h)$.

(iii) If $(a(n))$ is a bounded sequence we let

$$\|a\|_{U_\ell(\mathbb{N})} := \sup_{\mathbf{I}} \|a\|_{\mathbf{I},\ell}$$

where the sup is taken over all sequences of intervals **I** with lengths tending to infinity along which the sequence $(a(n))$ is distributed regularly.

An application of Lemma 2.2 shows that $\|a\|_{\mathbf{I},1}$, as defined here, is smaller than the corresponding quantity defined in [12] (they can be different though). Furthermore, the inductive formula is identical in both cases (see [12, Proposition 4.4]), hence $\|\cdot\|_{U_\ell(\mathbb{N})}$, as defined here, is a seminorm that is smaller than the corresponding seminorm defined in [12]. In fact, it can be shown that the two seminorms coincide but we will not need this.

Using the main structural result in [11], Host and Kra proved an inverse theorem that will be a key ingredient in the proof of Theorem 1.3. We state a slight variant of it next ([12, Theorem 2.16] gives a stronger lower bound but it does not allow to assume that $\|b\|_\infty \leq 1$). Its proof amounts to a simple modification of the argument given in [12, Theorem 2.16]; we give the details for completeness.

Theorem 2.1 ([12, Theorem 2.16]) *Let $a\colon \mathbb{N} \to \mathbb{C}$ be a sequence of complex numbers with $\|a\|_\infty \leq 1$ and $\ell \in \mathbb{N}$. Then for every $\varepsilon > 0$ there exists an $(\ell - 1)$-step nilsequence $(b(n))$ with $\|b\|_\infty \leq 1$ such that*

$$\limsup_{N-M\to\infty} \left| \frac{1}{N-M} \sum_{n=M}^{N-1} a(n)b(n) \right| \geq \|a\|_{U_\ell(\mathbb{N})}^{2^\ell} - \varepsilon.$$

Remark It is crucial that the seminorms were defined using uniform and not standard Cesàro averages as in the latter case it is shown in [12, Paragraph 2.4.3]

that the corresponding inverse theorem fails. For standard Cesàro averages a finitary inverse theorem was proved in [10] but it is not clear whether it has an infinitary variant that is useful for our purposes.

Proof We refer the reader to [12] for notation used in this argument. In what follows we assume that the seminorms $\|a\|_{\mathbf{I},\ell}$ are defined as in [12].

Let $0 < \varepsilon < 1$. By [12, Proposition 6.2] there exists a sequence of intervals $\mathbf{I} = (I_N)$ with lengths tending to infinity and an $(\ell-1)$-step nilsequence $(c(n))$ of the form $c(n) = F(g^n\Gamma)$, where F is a continuous function on an $(\ell-1)$-step nilmanifold $X = G/\Gamma$ and $g \in G$ is an element that acts ergodically on X, such that the sequences $a - c$ and a satisfy property $\mathcal{P}(\ell)$ on $\mathbf{I}$ and moreover we have the estimates

$$\|a - c\|_{\mathbf{I},\ell} \leq \varepsilon, \quad \|a\|_{\mathbf{I},\ell} \geq \|a\|_{U_\ell(\mathbb{N})} - \varepsilon. \tag{4}$$

Furthermore, we have $\|F\|_\infty \leq 1$, this is because in the proof of [12, Proposition 6.2] the function F is defined as a conditional expectation of a function bounded by 1. We let $b(n) := H(g^n\Gamma)$, where $H := \mathcal{D}_\ell F$, and check that the asserted properties are satisfied.

First note that $(b(n))$ is an $(\ell-1)$-step nilsequence and since $\|F\|_\infty \leq 1$ we have $\|H\|_\infty \leq 1$, hence $\|b\|_\infty \leq 1$. Furthermore, by [12, Corollary 5.3] we have $H \in C(X)$, hence $F \cdot H \in C(X)$, and since g acts ergodically on X we have

$$\lim_{N\to\infty} \frac{1}{|I_N|} \sum_{n\in I_N} c(n)b(n) = \int F \cdot H \, dm_X = \|F\|_\ell^{2^\ell} = \|c\|_{\mathbf{I},\ell}^{2^\ell}$$

where we used the identity $\int F \cdot \mathcal{D}_\ell F \, dm_X = \|F\|_\ell^{2^\ell}$ and [12, Corollary 3.11] to justify the last two identities. By (4) and the triangle inequality this is greater or equal than

$$(\|a\|_{\mathbf{I},\ell} - \varepsilon)^\ell \geq (\|a\|_{U_\ell(\mathbb{N})} - 2\varepsilon)^\ell \geq \|a\|_{U_\ell(\mathbb{N})}^\ell - k_\ell\varepsilon$$

for some positive integer k_ℓ. On the other hand, by [12, Theorem 2.13] we have

$$\limsup_{N\to\infty} \left| \frac{1}{|I_N|} \sum_{n\in I_N} (a(n) - c(n))b(n) \right| \leq \|a - c\|_{\mathbf{I},\ell} \|b\|_\ell^* \leq \varepsilon$$

where we used (4) and that $\|b\|_\ell^* = \|\mathcal{D}_\ell F\|_\ell^* = \|F\|_\ell^{2^\ell-1} \leq 1$ (the second identity follows from [12, Equation (14)]). Combining the previous bounds we get the asserted result.   □

2.2 Proof of Theorem 1.3

Let $\ell \in \mathbb{N}$ and $(a(n))$ be an ℓ-regular and ℓ-anti-uniform sequence with $\|a\|_\infty \leq 1$. We first remark that the limit

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} |a(n)|^2 \quad \text{exists.} \tag{5}$$

This follows from our anti-uniformity assumption and [12, Theorem 2.19] (it applies since $(a(n))$ is ℓ-regular) which states that for every $\epsilon > 0$ we have a decomposition $a = a_1 + a_2$ where a_1 is an $(\ell - 1)$-step nilsequence and $\|a_2\|_{U_\ell(\mathbb{N})} \leq \epsilon$. Writing $|a(n)|^2 = a\bar{a}_1 + a\bar{a}_2$ one checks the asserted convergence at once.

We let

$$Y := \left\{ (\psi(n)) \colon \psi \text{ is an } (\ell - 1)\text{-step nilsequence} \right\}$$

and

$$X := \text{span}\{Y, a\}.$$

On $X \times X$ we define the bilinear form

$$\langle f, g \rangle := \lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} f(n)\bar{g}(n).$$

Note that the limit exists for $f, g \in X$. This is the case if f or g is equal to a because of our regularity assumption and (5), and when both f and g are in Y because limits of uniform Cesàro averages of nilsequences exist [13, 16]. This bilinear form induces the seminorm

$$\|f\|_2 := \sqrt{\langle f, f \rangle}.$$

This is the restriction on X of the seminorm (3) defined on $\ell^\infty(\mathbb{N})$.

Let $\varepsilon > 0$. There exists $y_0 \in Y$ such that

$$\|a - y_0\|_2^2 \leq d^2 + \delta^2 \tag{6}$$

where

$$d := \inf\{\|a - y\|_2 : y \in Y\}, \quad \delta := (\varepsilon/(4C))^{2^\ell}, \tag{7}$$

and $C := C(\ell, a)$ is the constant determined by our ℓ-anti-uniformity assumption on a. We can assume that $C \geq 1$. Furthermore, we can assume without loss of generality that

$$\|y_0\|_\infty \leq 1. \tag{8}$$

Indeed, let $y_0 := (F(g^n \Gamma))$ where $X = G/\Gamma$ is a nilmanifold, $g \in G$, and $F \in C(X)$. Then the sequence $\tilde{y}_0 := (\tilde{F}(g^n \Gamma))$, where $\tilde{F} := F \cdot \mathbf{1}_{|F| \leq 1} + e^{2\pi i \arg(F)} \cdot \mathbf{1}_{|F| \geq 1} \in C(X)$, is a nilsequence, $\|\tilde{y}_0\|_\infty \leq 1$, and as $\|a\|_\infty \leq 1$ we get that $|a(n) - \tilde{y}_0(n)| \leq |a(n) - y_0(n)|$ for every $n \in \mathbb{N}$, hence $\|a - \tilde{y}_0\|_2 \leq \|a - y_0\|_2$.

It follows from (6) that for every $y \in Y$ we have

$$-\delta^2 \leq \|a - (y_0 + \delta y)\|_2^2 - \|a - y_0\|_2^2 = -2\delta \mathrm{Re}(\langle a - y_0, y \rangle) + \delta^2 \|y\|_2^2.$$

Hence,

$$\mathrm{Re}(\langle a - y_0, y \rangle) \leq \delta \text{ for every } y \in Y \text{ with } \|y\|_2 \leq 1.$$

Inserting $-y$ and $\pm iy$ in place of y we deduce that

$$\sup_{y \in Y: \, \|y\|_2 \leq 1} |\langle a - y_0, y \rangle| \leq 2\delta. \tag{9}$$

Since the set $\{y \in Y: \|y\|_2 \leq 1\}$ contains all $(\ell - 1)$-step nilsequences that are bounded by 1, we deduce from Theorem 2.1 that

$$\|a - y_0\|_{U_\ell(\mathbb{N})} \leq (2\delta)^{2^{-\ell}}. \tag{10}$$

We let

$$a_{st} := y_0, \quad a_{er} := a - y_0.$$

Then

$$a = a_{st} + a_{er}$$

and $(a_{st}(n))$ is an $(\ell - 1)$-step nilsequence with $\|a_{st}\|_\infty \leq 1$ by (8). Since a is ℓ-anti-uniform we get using (10) and the definition of δ in (7) that

$$|\langle a, a_{er} \rangle| \leq C \|a_{er}\|_{U_\ell(\mathbb{N})} \leq \varepsilon/2.$$

Furthermore, (9) gives

$$|\langle a_{st}, a_{er} \rangle| \leq \varepsilon/2.$$

Combining the last two estimates we deduce that

$$\|a_{er}\|_2^2 = \langle a_{er}, a_{er} \rangle \leq |\langle a, a_{er} \rangle| + |\langle a_{st}, a_{er} \rangle| \leq \varepsilon.$$

This completes the proof of Theorem 1.3.

2.3 Proof of Theorem 1.1

In view of Theorem 1.3, it suffices to prove that for every $\ell \in \mathbb{N}$ the sequence $a \colon \mathbb{N} \to \mathbb{C}$ defined by

$$a(n) := \int T_1^n f_1 \cdot \ldots \cdot T_\ell^n f_\ell \, d\mu, \quad n \in \mathbb{N}, \tag{11}$$

is ℓ-anti-uniform and ℓ-regular.

2.3.1 Anti-uniformity

Throughout, we can and will assume that $\|f_i\|_\infty \leq 1$ for $i = 1, \ldots, \ell$. The ℓ-anti-uniformity follows by successive applications of the following Hilbert space variant of van der Corput's estimate (for a proof see [4]).

Lemma 2.2 *Let (v_n) be a bounded sequence of vectors in an inner product space and (I_N) be a sequence of intervals with lengths tending to infinity. Then*

$$\limsup_{N \to \infty} \left\| \frac{1}{|I_N|} \sum_{n \in I_N} v_n \right\|^2 \leq 4 \limsup_{H \to \infty} \frac{1}{H} \sum_{h=1}^{H} \limsup_{N \to \infty} \left| \frac{1}{|I_N|} \sum_{n \in I_N} \langle v_{n+h}, v_n \rangle \right|.$$

It suffices to show that for every $\ell \in \mathbb{N}$ and every sequence of intervals $\mathbf{I} := (I_N)$ with lengths tending to infinity, any sequence $(a(n))$ given by (11) satisfies the estimate

$$\limsup_{N \to \infty} \left| \frac{1}{|I_N|} \sum_{n \in I_N} a(n)b(n) \right| \leq 4 \|b\|_{U_\ell(\mathbb{N})}$$

for every $b \in \ell^\infty(\mathbb{N})$. Using a diagonal argument and passing to a subsequence of (I_N) (if necessary) we can and will assume that the sequence $(b(n))$ is distributed regularly along the sequence $\mathbf{I}$. It suffices to establish that for any sequence $(a(n))$ as in (11) which is bounded by 1 and any $b \in \ell^\infty(\mathbb{N})$ which

is distributed regularly along a sequence of intervals **I**, we have

$$\limsup_{N\to\infty} \left| \frac{1}{|I_N|} \sum_{n\in I_N} a(n)b(n) \right| \le 4 \, \|b\|_{\mathbf{I},\ell} . \tag{12}$$

We prove this by induction on ℓ. For $\ell = 1$ the result holds trivially. Suppose that $\ell \ge 2$ and the statement holds for $\ell - 1$.

We compose with T_ℓ^{-n}, use the Cauchy-Schwarz inequality, and then Lemma 2.2 (on the space $L^2(\mu)$) for the sequence

$$v_n := b(n) \cdot \tilde{T}_1^n f_1 \cdot \tilde{T}_2^n f_2 \cdot \ldots \cdot \tilde{T}_{\ell-1}^n f_{\ell-1}, \quad n \in \mathbb{N},$$

where $\tilde{T}_i := T_i T_\ell^{-1}$ for $i = 1, \ldots, \ell - 1$. We deduce that the square of the left hand side in (12) is bounded by

$$\limsup_{N\to\infty} \left\| \frac{1}{|I_N|} \sum_{n\in I_N} v_n \right\|_{L^2(\mu)}^2 \le 4 \limsup_{H\to\infty} \frac{1}{H} \sum_{h=1}^{H} \limsup_{N\to\infty} \left| \frac{1}{|I_N|} \sum_{n\in I_N} \langle v_{n+h}, v_n \rangle \right|. \tag{13}$$

A simple computation gives that

$$\frac{1}{|I_N|} \sum_{n\in I_N} \langle v_{n+h}, v_n \rangle = \frac{1}{|I_N|} \sum_{n\in I_N} b(n+h) \cdot \bar{b}(n) \int \tilde{T}_1^n \tilde{f}_{1,h} \cdot \ldots \cdot \tilde{T}_{\ell-1}^n \tilde{f}_{\ell-1,h} \, d\mu$$

where $\tilde{f}_{j,h} = \tilde{T}_j^h f_j \cdot \bar{f}_j$ for $j = 1, \ldots, \ell - 1$. Note that the maps $\tilde{T}_1, \ldots, \tilde{T}_{\ell-1}$ commute, for $h \in \mathbb{N}$ the sequence $(b(n+h)\bar{b}(n))$ is distributed regularly along **I**, and $\left\| \tilde{f}_{j,h} \right\|_\infty \le 1$ for $j = 1, \ldots, \ell - 1$. Using the induction hypothesis and the defining property of the seminorms we can bound the right hand side in (13) by 16 times

$$\lim_{H\to\infty} \frac{1}{H} \sum_{h=1}^{H} \|\sigma_h b \cdot b\|_{\mathbf{I},\ell-1} \le \lim_{H\to\infty} \left(\frac{1}{H} \sum_{h=1}^{H} \|\sigma_h b \cdot b\|_{\mathbf{I},\ell-1}^{2^{\ell-1}} \right)^{1/2^{\ell-1}} = \|b\|_{\mathbf{I},\ell}^2$$

where $(\sigma_h b)(n) := b(n+h)$. Taking square roots we get the asserted estimate.

2.3.2 Regularity

Let $\ell \in \mathbb{N}$. To prove that $(a(n))$ is ℓ-regular we will use a known mean convergence result for multiple ergodic averages and Proposition 2.4 below. We start with the following result of Green and Tao:

Lemma 2.3 ([9, Lemma 14.2]) *For $\ell \in \mathbb{N}$ let $X = G/\Gamma$ be an $(\ell - 1)$-step nilmanifold. Then there exists a continuous map $P \colon X^\ell \to X$ such that*

$$P(hg\Gamma, h^2 g\Gamma, \ldots, h^\ell g\Gamma) = g\Gamma, \quad \text{for every } g, h \in G. \quad (14)$$

The result in [9, Lemma 14.2] gives $P(g\Gamma, hg\Gamma, h^2 g\Gamma, \ldots, h^{\ell-1} g\Gamma) = h^\ell g\Gamma$. Inserting $h^{-\ell} g$ in place of g, then h^{-1} in place of h, and rearranging coordinates, we get (14).

Proposition 2.4 *For $\ell \in \mathbb{N}$ let $(\psi(n))$ be an $(\ell - 1)$-step nilsequence. Then for every $\varepsilon > 0$ there exists a system $(X, \mathcal{X}, \mu, T)$ and functions $f_1, \ldots, f_\ell \in L^\infty(\mu)$, such that the sequence $(b(n))$, defined by*

$$b(n) := \int T^{k_1 n} f_1 \cdot \ldots \cdot T^{k_\ell n} f_\ell \, d\mu, \quad n \in \mathbb{N}, \quad (15)$$

where $k_i := \ell!/i$ for $i = 1, \ldots, \ell$, satisfies

$$\|\psi - b\|_\infty \leq \varepsilon.$$

Remark To prove a variant of this result that uses the integers $1, \ldots, \ell$ in place of $k_1, \ldots, k_\ell$, one would have to prove a non-trivial variant of Lemma 2.3 that establishes in place of (14) the identity $P(h^{k_1} g\Gamma, h^{k_2} g\Gamma, \ldots, h^{k_\ell} g\Gamma) = g\Gamma$ for every $g, h \in G$.

Combining [6, Theorem A (ii)] with Proposition 2.4 one deduces that for every bounded generalized polynomial $p \colon \mathbb{N} \to \mathbb{R}$ (see definition in [6]) the sequences $(p(n))$ and $(e^{ip(n)})$ can be approximated arbitrarily well in $\|\cdot\|_2$ by a sequence of the form (15).

Proof Let $\varepsilon > 0$ and

$$\psi(n) := F(g^n \Gamma)$$

where $F \in C(X)$, $X = G/\Gamma$ is an $(\ell - 1)$-step nilmanifold, and $g \in G$.

By [13, Paragraph 1.11] we have that X is isomorphic to a subnilmanifold of a nilmanifold $\tilde{X} = \tilde{G}/\tilde{\Gamma}$, where $\tilde{G}$ is a connected and simply connected $(\ell - 1)$-step nilpotent Lie group, $\tilde{\Gamma}$ is a discrete cocompact subgroup of $\tilde{G}$, and all elements of G are represented in $\tilde{G}$. Then $\psi(n) = \tilde{F}(\tilde{b}^n \tilde{\Gamma})$ for some $\tilde{b} \in \tilde{G}$ and $\tilde{F} \in C(\tilde{X})$. Hence, in what follows we can and will assume that the group G is connected.

Using Lemma 2.3 with g^n in place of g and $h := g^m$, $m, n \in \mathbb{N}$, we get that there exists a continuous map $P \colon X^\ell \to X$ such that

$$g^n \Gamma = P(g^{m+n} \Gamma, g^{2m+n} \Gamma, \ldots, g^{\ell m + n} \Gamma) \quad \text{for every } m, n \in \mathbb{N}. \quad (16)$$

Let $g_0 \in G$ be such that $g_0^{\ell!} = g$ (such a g_0 exists since G is connected, hence divisible) and for $i = 1, \ldots, \ell$ let $g_i := g_0^i$. Applying (16) with g_0 in place of g and $\ell!n$ (a multiple of n is needed that is divisible by all the coefficients of m that appear in (16)) in place of n we get

$$\psi(n) = F(g_0^{\ell!n}\Gamma) = \tilde{F}(g_1^{m+k_1 n}\Gamma, g_2^{m+k_2 n}\Gamma, \ldots, g_\ell^{m+k_\ell n}\Gamma) \quad \text{for every } m, n \in \mathbb{N},$$

where $\tilde{F} := F \circ P \in C(X^\ell)$. Averaging over $m \in \mathbb{N}$ we get

$$\psi(n) = \lim_{M \to \infty} \frac{1}{M} \sum_{m=1}^{M} \tilde{F}(g_1^{m+k_1 n}\Gamma, g_2^{m+k_2 n}\Gamma, \ldots, g_\ell^{m+k_\ell n}\Gamma) \quad \text{for every } n \in \mathbb{N}.$$

Since $\tilde{F}$ can be approximated uniformly by linear combinations of functions of the form $\tilde{f}_1 \otimes \cdots \otimes \tilde{f}_\ell$, where for $i = 1, \ldots, \ell$ the function $\tilde{f}_i \in C(X^\ell)$ depends on the coordinate x_i only, we get that $(\psi(n))$ can be approximated in the $\|\cdot\|_\infty$ norm within ε by a finite linear combination of sequences $(a(n))$ of the form

$$a(n) := \lim_{M \to \infty} \frac{1}{M} \sum_{m=1}^{M} \tilde{f}_1(\tilde{g}^{m+k_1 n}\tilde{\Gamma}) \cdot \tilde{f}_2(\tilde{g}^{m+k_2 n}\tilde{\Gamma}) \cdot \ldots \cdot \tilde{f}_\ell(\tilde{g}^{m+k_\ell n}\tilde{\Gamma}), \quad n \in \mathbb{N},$$

$$(17)$$

where $\tilde{X} := X^\ell$, $\tilde{\Gamma} := \Gamma \times \cdots \times \Gamma$, $\tilde{f}_i \in C(\tilde{X})$, and $\tilde{g} := (g_1, \ldots, g_\ell)$. It is known (see [13] for example) that the limit in (17) is equal to

$$\int_{\tilde{Y}} \tilde{f}_1(\tilde{g}^{k_1 n}\tilde{y}) \cdot \tilde{f}_2(\tilde{g}^{k_2 n}\tilde{y}) \cdot \ldots \cdot \tilde{f}_\ell(\tilde{g}^{k_\ell n}\tilde{y}) \, dm_{\tilde{Y}}, \quad n \in \mathbb{N},$$

where $\tilde{Y}$ is the subnilmanifold of $\tilde{X}$ defined by the closure of the set $\{\tilde{g}^m \tilde{\Gamma} : m \in \mathbb{N}\}$. This proves that the sequence $(a(n))$ has the form (15). Since finite linear combinations of sequences of the form (15) still have the form (15) (see Sect. 1.3) the proof is complete. $\qquad \square$

We are now ready to verify that if $(a(n))$ is as in (11), then it is ℓ-regular for every $\ell \in \mathbb{N}$.

By Proposition 2.4, in order to check that the limit $\lim_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N-1} a(n)\psi(n)$ exists for every $(\ell - 1)$-step nilsequence $(\psi(n))$, it suffices to check that the limit

$$\lim_{N-M \to \infty} \frac{1}{N-M} \sum_{n=M}^{N-1} a(n)b(n) \qquad (18)$$

exists for every sequence $(b(n))$ of the form $\int S^{k_1 n} g_1 \cdot \ldots \cdot S^{k_\ell n} g_\ell \, d\nu$, where $k_1, \ldots, k_\ell \in \mathbb{N}$, $(Y, \mathcal{Y}, \nu, S)$ is a system, and $g_1, \ldots, g_\ell \in L^\infty(\nu)$. This fol-

lows from the mean convergence result of Austin [1] (which strengthens the convergence result of Tao [18] to uniform averages) applied to the transformations $\tilde{T}_i := T_i \times S^{k_i}$ acting on $X \times Y$ with the measure $\tilde{\mu} := \mu \times \nu$ and the functions $\tilde{f}_i := f_i \otimes g_i \in L^\infty(\tilde{\mu})$, $i = 1, \ldots, \ell$.

2.4 Proof of Theorem 1.2

Modulo a known convergence result of Walsh [19] the argument is similar to the one used to prove Theorem 1.1, we explain the minor modifications needed next.

To verify k-anti-uniformity for some $k \in \mathbb{N}$ that depends only on ℓ, m and the maximum degree of the polynomials $p_{i,j}$, one has to make successive uses of Lemma 2.2 and apply an inductive argument, often called PET induction, introduced by V. Bergelson in [4]. The details are very similar to those in the proof of [8, Lemma 3.5] and so we omit them.

To verify regularity, we can argue as in the case of linear iterates, using the convergence result of Walsh [19] for averages of expressions of the form (2). At the very last step one needs to verify that if $(a(n))$ is as in (2), then the limit (18) exists for every sequence $(b(n))$ of the form $\int S^{k_1 n} g_1 \cdot \ldots \cdot S^{k_r n} g_r \, d\nu$, where $r \in \mathbb{N}$ is arbitrary, $k_1, \ldots, k_r \in \mathbb{N}$, $(Y, \mathcal{Y}, \nu, S)$ is a system, and $g_1, \ldots, g_r \in L^\infty(\nu)$. The only change needed is to use Walsh's convergence result for the $\ell + r$ commuting measure preserving transformations $T_i \times \mathrm{id}$, $i = 1, \ldots, \ell$, and $\mathrm{id} \times S^{k_j}$, $j = 1, \ldots, r$, acting on $X \times Y$ with the measure $\tilde{\mu} := \mu \times \nu$, and the functions $f_i \otimes 1$, $i = 1, \ldots, \ell$ and $1 \otimes g_j$, $j = 1 \ldots, r$. If the polynomial iterates are chosen appropriately, one verifies that $a(n)b(n)$ is also a multiple correlation sequence with polynomial iterates, hence, by Walsh's convergence result [19], the limit (18) exists.

2.5 Extension to nilpotent groups

Essentially the same argument can be used when the transformations $T_1, \ldots, T_\ell$ generate a nilpotent group; the only extra difficulty occurs in proving k-anti-uniformity for some $k \in \mathbb{N}$ that depends also on the degree of nilpotency of the group generated by $T_1, \ldots, T_\ell$. In this case, the PET induction is somewhat more complicated, but can be handled by modifying the PET induction used in [8, Lemma 3.5] along the lines of the argument used to prove [19, Theorem 4.2].

2.6 Proof of Theorem 1.4

The inclusion $\overline{\mathcal{A}_\ell} \subset \overline{\mathcal{B}_\ell}$ follows from Proposition 2.4. The inclusion $\overline{\mathcal{B}_\ell} \subset \overline{\mathcal{C}_\ell}$ is obvious. The inclusion $\overline{\mathcal{C}_\ell} \subset \overline{\mathcal{A}_\ell}$ follows from Theorem 1.1.

2.7 Proof of Theorems 1.5 and 1.6

The implication $(ii) \Rightarrow (i)$ follows from Proposition 2.4. (for Theorem 1.6 in order to get property (i) for some fixed $\ell \in \mathbb{N}$ we use property (ii) for $\ell!$). The implication $(i) \Rightarrow (iii)$ follows from Theorem 1.1 and the remarks following this theorem. The implication $(iii) \Rightarrow (ii)$ is obvious.

The same argument applies for the extensions mentioned after Theorem 1.6 related to uniform and weighted Cesàro averages.

References

1. Austin, T.: On the norm convergence of nonconventional ergodic averages. Ergod. Theory Dyn. Syst. **30**, 321–338 (2010)
2. Austin, T.: Pleasant extensions retaining algebraic structure, I. J. Anal. Math. (To appear)
3. Austin, T.: Pleasant extensions retaining algebraic structure, II. J. Anal. Math. (To appear)
4. Bergelson, V.: Weakly mixing PET. Ergod. Theory Dyn. Syst. **7**(3), 337–349 (1987)
5. Bergelson, V., Host, B., Kra, B.: With an appendix by I. Ruzsa. Multiple recurrence and nilsequences. Invent. Math. **160**(2), 261–303 (2005)
6. Bergelson, V., Leibman, A.: Distribution of values of bounded generalized polynomials. Acta Math. **198**, 155–230 (2007)
7. Frantzikinakis, N.: Some open problems on multiple ergodic averages. arXiv:1103.3808
8. Frantzikinakis, N., Host, B., Kra, B.: The polynomial multidimensional Szemerédi theorem along shifted primes. Isr. J. Math. **194**(1), 331–348 (2013)
9. Green, B., Tao, T.: An inverse theorem for the Gowers $U^3(G)$-norm. Proc. Edinb. Math. Soc. **51**(1), 73–153 (2008)
10. Green, B., Tao, T., Ziegler, T.: An inverse theorem for the Gowers $U^{s+1}[N]$-norm. Ann. Math. **176**(2), 1231–1372 (2012)
11. Host, B., Kra, B.: Nonconventional ergodic averages and nilmanifolds. Ann. Math. **161**, 397–488 (2005)
12. Host, B., Kra, B.: Uniformity seminorms on l^∞ and applications. J. Anal. Math. **108**, 219–276 (2009)
13. Leibman, A.: Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold. Ergod. Theory Dyn. Syst. **25**(1), 201–213 (2005)
14. Leibman, A.: Multiple polynomial sequences and nilsequences. Ergod. Theory Dyn. Syst. **30**(3), 841–854 (2010)
15. Leibman, A.: Nilsequences, null-sequences, and multiple correlation sequences. Ergod. Theory Dyn. Syst. **35**(1), 176–191 (2015)
16. Lesigne, E.: Sur une nil-variété, les parties minimales associées à une translation sont uniquement ergodiques. Ergod. Theory Dyn. Syst. **11**(2), 379–391 (1991)
17. Meiri, D.: Generalized correlation sequences. Master's thesis, Tel Aviv University (1990). Available at http://taalul.com/David/Math/ma
18. Tao, T.: Norm convergence of multiple ergodic averages for commuting transformations. Ergod. Theory Dyn. Syst. **28**(2), 657–688 (2008)
19. Walsh, M.: Norm convergence of nilpotent ergodic averages. Ann. Math. **175**(3), 1667–1688 (2012)