# Sets of large values of correlation functions for polynomial cubic configurations

V. Bergelson and A. Leibman

*April 8, 2016*


## Abstract

We prove that for any set $E \subseteq \mathbb{Z}$ with upper Banach density $d^*(E) > 0$, the set "of cubic configurations" in $E$ is large in the following sense: for any $k \in \mathbb{N}$ and any $\varepsilon > 0$, the set $\{(n_1, \ldots, n_k) \in \mathbb{Z}^k : d^*(\bigcap_{e_1,\ldots,e_k \in \{0,1\}}(E - (e_1 n_1 + \cdots + e_k n_k))) > d^*(E)^{2^k} - \varepsilon\}$ is an AVIP$_0^*$-set. We then generalize this result to the case "of polynomial cubic configurations" $e_1 p_1(n) + \cdots + e_k p_k(n)$ where the polynomials $p_i : \mathbb{Z}^d \longrightarrow \mathbb{Z}$ are assumed to be sufficiently algebraically independent.


## 0. Introduction

Let $(X, \mathcal{B}, \mu, T)$ be an invertible probability measure preserving system. By a result of Khintchine ([Kh]), for any $A \in \mathcal{B}$ one has

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \mu(A \cap T^n A) \geq \mu(A)^2. \tag{0.1}$$

It follows from (0.1) that for any $\varepsilon > 0$ the set

$$R_\varepsilon(A) = \left\{ n \in \mathbb{Z} : \mu(A \cap T^n A) \geq \mu(A)^2 - \varepsilon \right\}$$

is syndetic.(1) This fact forms a refinement of the classical recurrence theorem of Poincaré and is referred to as *Khintchine's recurrence theorem*.

The limiting relation (0.1) admits a multiparameter generalization (see [B2], [HoK1], [HoK2], and, for a short proof of a rather general result, [BL2], Theorem 0.8). In particular, one has the following theorem:

Partially supported by NSF grants DMS-1162073 and DMS-1500575.

(1) A set $S$ in a discrete abelian group $G$ is called *syndetic* if there exists a finite set $F \subseteq G$ such that $G = F + S$.

**Theorem 0.1.** *For any $k \in \mathbb{N}$ and any $A \in \mathcal{B}$,*

$$\lim_{N-M\to\infty} \frac{1}{(N-M)^k} \sum_{M\le n_1,\dots,n_k\le N-1} \mu\Big(\bigcap_{e_1,\dots,e_k\in\{0,1\}} T^{e_1n_1+\cdots+e_kn_k}A\Big) \ge \mu(A)^{2^k}.$$

**Corollary 0.2.** *For any $k \in \mathbb{N}$, any $A \in \mathcal{B}$ with $\mu(A) > 0$, and any $\varepsilon > 0$ the set*

$$R_\varepsilon^{(k)}(A) = \Big\{(n_1,\dots,n_k) \in \mathbb{Z}^k : \mu\Big(\bigcap_{e_1,\dots,e_k\in\{0,1\}} T^{e_1n_1+\cdots+e_kn_k}A\Big) > \mu(A)^{2^k} - \varepsilon\Big\} \tag{0.2}$$

*is syndetic.*

Via the Furstenberg correspondence principle(2), Corollary 0.2 implies that any set of positive upper Banach density(3) in $\mathbb{Z}$ contains many "$k$-dimensional cubic configurations":

**Corollary 0.3.** *Let $E \subseteq \mathbb{Z}$ be a set with $d^*(E) > 0$. Then for any $k \in \mathbb{N}$ and any $\varepsilon > 0$ the set*

$$R_\varepsilon^{(k)}(E) = \Big\{(n_1,\dots,n_k) \in \mathbb{Z}^k : d^*\Big(\bigcap_{e_1,\dots,e_k\in\{0,1\}} \big(E - (e_1n_1+\cdots+e_kn_k)\big)\Big) > d^*(E)^{2^k} - \varepsilon\Big\} \tag{0.3}$$

*is syndetic.*

The goal of this paper is to refine Theorem 0.1 and Corollary 0.2 in two natural directions. The first direction has to do with multiple recurrence along polynomials. By utilizing results obtained in [BFM], one can show that for any $A \in \mathcal{B}$ with $\mu(A) > 0$, any $\varepsilon > 0$, and any intersective(4) polynomial $p\colon \mathbb{Z}^d \longrightarrow \mathbb{Z}$, the set

$$R_\varepsilon(A;p) = \big\{n \in \mathbb{Z}^d : \mu(A \cap T^{p(n)}A) > \mu(A)^2 - \varepsilon\big\}$$

is syndetic. This leads to a question whether it is true that for any jointly intersective polynomials $p_1,\dots,p_k$ the set

$$R_\varepsilon(A;p_1,\dots,p_k) = \Big\{n \in \mathbb{Z}^d : \mu\Big(\bigcap_{e_1,\dots,e_k\in\{0,1\}} T^{e_1p_1(n)+\cdots+e_kp_k(n)}A\Big) > \mu(A)^{2^k} - \varepsilon\Big\} \tag{0.4}$$

---

(2) The *Furstenberg correspondence principle* says that for any set $E \subseteq \mathbb{Z}$ there exists (an ergodic) invertible probability measure preserving system $(X,\mathcal{B},\mu,T)$ and a set $A \in \mathcal{B}$ such that $\mu(A) = d^*(E)$ and for any $k$ and any $n_1,\dots,n_k \in \mathbb{Z}$, $d^*((E-n_1)\cap\cdots\cap(E-n_k)) \ge \mu(T^{n_1}A\cap\cdots\cap T^{n_k}A)$, where $d^*$ is the upper Banach density (which is defined in the next footnote). See [B4] and [BHoK].

(3) For a subset $E \subseteq \mathbb{Z}^d$, *the upper Banach density $d^*(E)$ of $E$* is the supremum of $\limsup_{N\to\infty}|E\cap\Phi_N|/|\Phi_N|$ over all Følner sequences $(\Phi_N)$ in $\mathbb{Z}^d$.

(4) A polynomial $p\colon \mathbb{Z}^d \longrightarrow \mathbb{Z}$ is *intersective* if for any $m \in \mathbb{N}$ there exists $n \in \mathbb{Z}^d$ such that $m|p(n)$. Several polynomials $p_1,\dots,p_k\colon \mathbb{Z}^d \longrightarrow \mathbb{Z}$ are *jointly intersective* if for any $m \in \mathbb{N}$ there exists $n \in \mathbb{Z}^d$ such that $m|p_i(n)$ for all $i = 1,\dots,k$. The class of intersective polynomials appears naturally in Ergodic Ramsey Theory – these are the ultimate class of polynomials for which the polynomial Szemerédi theorem holds; see [BLLe2].

is syndetic, which, in turn, would imply that for any $E \subseteq \mathbb{Z}$ the set

$$R_\varepsilon(E;p_1,\dots,p_k) = \Big\{n \in \mathbb{Z}^d : d^*\Big(\bigcap_{e_i\in\{0,1\}} \big(E-(e_1p_1(n)+\dots+e_kp_k(n))\big)\Big) > d^*(E)^{2^k} - \varepsilon\Big\} \tag{0.5}$$

is syndetic. (Note that in the case $d = k$ and $p_i(n) = n_i$, $i = 1,\dots,k$, (0.4) and (0.5) give (0.2) and (0.3).) We answer this question positively in the case where the polynomials $p_i$ are "sufficiently algebraically independent". (See Theorem 0.8 and Corollary 0.9 below.)

Another question of interest – both from the point of view of dynamics and combinatorics – is whether "the sets of returns with large inersections", such as $R_\varepsilon^{(k)}(A)$ and $R_\varepsilon(A;p_1,\dots,p_k)$, have a property stronger than that of syndeticity. As we will see, in certain situations the answer to this question is positive. In order to formulate the results obtained in this paper we first have to introduce some relevant notions of largeness. (The reader will find a more detailed discussion of the hierarchy of notions of largeness in the next section.)

**Definition 0.4.** Let $(n_i)$ be a sequence in $\mathbb{Z}^d$. The *IP-set*(5) generated by $(n_i)$ is "the set of finite sums"

$$\mathrm{FS}(n_1,n_2,\dots) = \big\{n_{i_1}+\dots+n_{i_j} : i_1<\dots<i_j,\ j\in\mathbb{N}\big\}.$$

A set $S \subseteq \mathbb{Z}^d$ is called an *IP$^*$-set* if it has a nontrivial intersection with any IP-set. For $r \in \mathbb{N}$ and any $n_1,\dots,n_r \in \mathbb{Z}^d$ the set

$$\mathrm{FS}(n_1,\dots,n_r) = \big\{n_{i_1}+\dots+n_{i_j} : i_1<\dots<i_j,\ j\le r\big\}$$

is called *an IP$_r$-set.* A set $S \subseteq \mathbb{Z}^d$ is said to be an *IP$_r^*$-set* if it has a nontrivial intersection with any IP$_r$-set, and we say that a set is an *IP$_0^*$-set* if it is an IP$_r^*$-set for some $r \in \mathbb{N}$. (Equivalently, one can define *an IP$_0$-set* as a set that contains an IP$_r$-set for all $r$, and an IP$_0^*$-set as a set having a nonempty intersection with every IP$_0$-set.)

Clearly, IP$_0^*$-sets are IP$^*$; it will be shown in the next section that the property of being an IP$_0^*$-set is strictly stronger than that of being an IP$^*$-set. It is also not hard to see that any IP$^*$-set is syndetic, and that the property of being an IP$^*$-set is stronger than that of being syndetic. For example, one can show that the family of IP$^*$-sets in $\mathbb{Z}^d$ has the filter property, meaning that for any finite collection $S_1,\dots,S_k \subseteq \mathbb{Z}^d$ of IP$^*$-sets the intersection $\bigcap_{i=1}^k S_i$ is also an IP$^*$-set. (This follows from Hindman's theorem ([H]), stating that for any finite partition of $\mathbb{Z}^d$ one of the cells of the partition contains an IP-set.)

One can show in an elementary way that the set $R_\varepsilon(A)$ introduced above is IP$_0^*$ (see [B1], Section 5 and Theorem 1.2 in Section 1) and it is natural to inquire whether the sets $R_\varepsilon^{(k)}$ or $R_\varepsilon(A;p_1,\dots,p_k)$ are IP$_0^*$-sets, or, at least, IP$^*$-sets. While it turns out that even for an ergodic $T$ already the set $R_\varepsilon^{(2)}(A)$ may fail to be IP$^*$(6), we will show that the sets $R_\varepsilon^{(k)}(A)$, as well as the sets $R_\varepsilon(A;p_1,\dots,p_k)$ when the polynomials $p_i$ are "sufficiently algebraically independent", are *Almost* IP$_0^*$-sets:

(5) The abbreviation "IP" was introduced in [FuW] and stands for the "Infinite-dimensional Parallelepiped", as well as for "IdemPotent".

(6) For example, $R_\varepsilon^{(2)}(A)$ may have trivial intersection with the IP-set $\{(n,2n),\ n\in\mathbb{N}\}$.

**Definition 0.5.** A set $S \subseteq \mathbb{Z}^d$ is *an AIP*$^*$*-set* if $S = R \setminus N$ where $R$ is an IP$^*$-set and $N$ is a set of zero Banach density, and is *an AIP*$_0^*$*-set* if $S = R \setminus N$ where $R$ is an IP$_0^*$-set and $N$ is a set of zero Banach density.

The property of being AIP$^*$ is still quite a bit stronger than that of being syndetic. In particular, the family of AIP$^*$-sets has the finite intersection property. We will discuss in Section 1 connections between the AIP$^*$-sets and some other important families of large sets.

Here is one of the ergodic-theoretical results obtained in this paper, together with its combinatorial counterpart:

**Theorem 0.6.** *Let $(X, \mathcal{B}, \mu, T)$ be an invertible ergodic probability measure preserving system and let $A \in \mathcal{B}$ with $\mu(A) > 0$. Then for any $k \in \mathbb{N}$ and any $\varepsilon > 0$, the set $R_\varepsilon^{(k)}(A)$ is an AIP$_0^*$-set.*

**Corollary 0.7.** *For any $E \subseteq \mathbb{Z}$ with $d^*(E) > 0$, any $k \in \mathbb{N}$, and any $\varepsilon > 0$, the set $R_\varepsilon^{(k)}(E)$ is an AIP$_0^*$-set.*

Theorem 0.6 generalizes to families of "algebraically independent" polynomial powers of a transformation. We say that polynomials $p_1, \ldots, p_k \colon \mathbb{Z}^d \longrightarrow \mathbb{Z}$ are *algebraically independent up to degree* $c$ if for any nonzero polynomial $P \colon \mathbb{Z}^k \longrightarrow \mathbb{Z}$ of degree $\leq c$, the polynomial $P(p_1(n), \ldots, p_k(n))$ is not equal to zero.

**Examples.** 1. If $p_i$ are polynomials in pairwise disjoint sets of variables (say, $p_1 = p_1(n_1)$ and $p_2 = p_2(n_2)$), they are algebraically independent up to any degree.
2. The polynomials $n, n^{c+1}, n^{(c+1)^2}, \ldots, n^{(c+1)^k}$ on $\mathbb{Z}$ are algebraically independent up to degree $c$.
3. The polynomials $n^2, n^2 + n$ are algebraically independent up to degree 2.

**Theorem 0.8.** *Let $(X, \mathcal{B}, \mu, T)$ be an invertible ergodic probability measure preserving system and let polynomials $p_1, \ldots, p_k \colon \mathbb{Z}^d \longrightarrow \mathbb{Z}$ with zero constant term be algebraically independent up to degree $k$. Then for any $A \in \mathcal{B}$ with $\mu(A) > 0$ and any $\varepsilon > 0$ the set $R_\varepsilon(A; p_1, \ldots, p_k)$ is AIP$_0^*$.*

**Corollary 0.9.** *Let polynomials $p_1, \ldots, p_k \colon \mathbb{Z}^d \longrightarrow \mathbb{Z}$ with zero constant term be algebraically independent up to degree $k$. Then for any set $E \subseteq \mathbb{Z}$ with $d^*(E) > 0$ and any $\varepsilon > 0$ the set $R_\varepsilon(E; p_1, \ldots, p_k)$ is an AIP$_0^*$-set.*

Polynomials with no constant term form a special case of intersective polynomials (defined in footnote (4) above). When dealing with jointly intersective polynomials one encounters "shifted" AIP$^*$-sets:

**Definition 0.10.** *An IP*$_{0,+}^*$*-set* (respectively, *an AIP*$_{0,+}^*$*-set*) in $\mathbb{Z}^d$ is a set of the form $n_0 + S$ where $S$ is an IP$_0^*$-set (respectively, an AIP$_0^*$-set) and $n_0 \in \mathbb{Z}^d$.

**Theorem 0.11.** *Let $(X, \mathcal{B}, \mu, T)$ be an invertible ergodic probability measure preserving system and let polynomials $p_1, \ldots, p_k \colon \mathbb{Z} \longrightarrow \mathbb{Z}$ be jointly intersective and algebraically independent up to degree $k$. Then for any set $A \in \mathcal{B}$ with $\mu(A) > 0$ and any $\varepsilon > 0$ the set $R_\varepsilon(A; p_1, \ldots, p_k)$ is an AIP$_{0,+}^*$-set.*

**Corollary 0.12.** *Let polynomials $p_1, \ldots, p_k : \mathbb{Z} \longrightarrow \mathbb{Z}$ be jointly intersective and algebraically independent up to degree $k$. Then for any set $E \subseteq \mathbb{Z}$ and any $\varepsilon > 0$ the set $R_\varepsilon(E; p_1, \ldots, p_n)$ is an $AIP^*_{0,+}$-set.*

As a matter of fact, the sets $R^{(k)}_\varepsilon(A)$, $R^{(k)}_\varepsilon(E)$, $R_\varepsilon(E; p_1, \ldots, p_k)$, and $R_\varepsilon(A; p_1, \ldots, p_k)$, appearing in Theorems 0.6–0.11 and their corollaries, possess a property stronger than just being $\mathrm{AIP}^*_0$ or $\mathrm{AIP}^*_{0,+}$-sets – they are, in fact, "polynomial $\mathrm{AIP}^*_0$-sets", *$AVIP^*_0$-sets* or, respectively, *$AVIP^*_{0,+}$-sets*. We postpone the definition and the relevant discussion until the next section.

## 1. Hierarchy of large sets, $\mathbf{AVIP}^*_0$-sets, and translations on nilmanifolds

Our goal in this section is to introduce and discuss various notions of largeness for sets in $\mathbb{Z}$ and $\mathbb{Z}^d$, which will allow us to put in a better perspective the results obtained in this paper.

Let $(X, \mathcal{B}, \mu, T)$ be an invertible probability measure system and let $A \in \mathcal{B}$, $\mu(A) > 0$. One pertinent notion of largeness, namely syndeticity, has already appeared in the introduction, and it was mentioned that the sets

$$R_\varepsilon(A) = \big\{n \in \mathbb{Z} : \mu(A \cap T^n A) \geq \mu(A)^2 - \varepsilon\big\},$$

$$R^{(k)}_\varepsilon(A) = \Big\{(n_1, \ldots, n_k) \in \mathbb{Z}^k : \mu\Big(\bigcap_{e_1,\ldots,e_k \in \{0,1\}} T^{e_1 n_1 + \cdots + e_k n_k} A\Big) > \mu(A)^{2^k} - \varepsilon\Big\} \text{ for } k \geq 2,$$

$$R_\varepsilon(A; p) = \big\{n \in \mathbb{Z} : \mu(A \cap T^{p(n)} A) > \mu(A)^2 - \varepsilon\big\}, \text{ where } p \text{ is an intersective polynomial,}$$

are known to be syndetic. In the case of the sets $R_\varepsilon(A)$ one can however show that these sets possess stronger combinatorial properties, which we will describe now.

The set $\{n_j - n_i : 1 \leq i < j\}$ of differences of the elements of an infinite sequence $(n_i)$ of distinct elements of $\mathbb{Z}^d$ is called *a $\Delta$-set*, and the set of differences of an $r$-element sequence of distinct elements is called *a $\Delta_r$-set*. A set $S \subseteq \mathbb{Z}^d$ is called *a $\Delta^*$-set* if it has a nonempty intersection with every $\Delta$-set in $\mathbb{Z}^d$, and, for $r \in \mathbb{N}$, is called *a $\Delta^*_r$-set* if it has a nonempty intersection with every $\Delta_r$-set in $\mathbb{Z}^d$. Also, we say that a set $S$ is *a $\Delta^*_0$-set* if it is a $\Delta^*_r$-set for some $r$. Let us denote by $\mathcal{S}$, $\Delta^*$, $\Delta^*_r$ for $r \in \mathbb{N}$, and $\Delta^*_0$ the families of syndetic, $\Delta^*$, $\Delta^*_r$, and $\Delta^*_0$-sets respectively. Then, for any $r_1, r_2 \in \mathbb{N}$ with $r_1 > r_2 \geq 4$, the following strict inclusions hold:

$$\mathcal{S} \supset \Delta^* \supset \Delta^*_0 \supset \Delta^*_{r_1} \supset \Delta^*_{r_2}.$$

Let us show, for example, that $\Delta^* \neq \Delta^*_0$ in the case $d = 1$. Put $B = \bigcup_{r=1}^\infty \{2^{2^r}, 2 \cdot 2^{2^r}, 3 \cdot 2^{2^r}, \ldots, r2^{2^r}\}$, so that $B$ contains $\Delta_r$-sets for arbitrarily large $r$, but contains no $\Delta$-sets; hence, the complement $S = \mathbb{Z} \setminus B$ of $B$ is a $\Delta^*$-set but not a $\Delta^*_0$-set.

We observe that $\Delta^*_r$-sets naturally appear in the traditional proof of Poincaré's recurrence theorem. Indeed, let $A \in \mathcal{B}$, $\mu(A) > 0$. Given any set $\{n_1, \ldots, n_r\} \subset \mathbb{Z}$ of cardinality $r > \mu(A)^{-1}$, the sets $T^{n_1}A, \ldots, T^{n_r}A$ cannot be all disjoint. Thus for some $1 \leq i < j \leq r$ one has $0 < \mu(T^{n_i}A \cap T^{n_j}A) = \mu(A \cap T^{n_j - n_i}A)$, which immediately implies that the set

$$\big\{n \in \mathbb{Z} : \mu(A \cap T^n A) > 0\big\}$$

is a $\Delta_r^*$-set for all $r > \mu(A)^{-1}$.

A refinement of this short argument allows one to get the following sharpening of Khintchine's recurrence theorem.

**Theorem 1.1.** *Let $(X, \mathcal{B}, \mu, T)$ be an invertible probability measure preserving system. Then for any $A \in \mathcal{B}$ and $\varepsilon > 0$ the set $R_\varepsilon(A) = \{n \in \mathbb{Z} : \mu(A \cap T^n A) > \mu(A)^2 - \varepsilon\}$ is a $\Delta_0^*$-set.*

The assertion of Theorem 1.1 follows from the elementary fact (originally due to Gillis, [G]) that given $\varepsilon > 0$, if $r \in \mathbb{N}$ is large enough, for any sets $A_1, \ldots, A_r \in \mathcal{B}$ of measure $a > 0$ there exist $1 \le i < j \le r$ such that $\mu(A_i \cap A_j) > a^2 - \varepsilon$. (It also follows from this argument that the set $R_\varepsilon(A)$ is an $\Delta_r^*$-set where $r$ depends only on $\mu(A)$.)

Encouraged by Theorem 1.1, one would like to know whether some other natural "sets of large returns", such as

$$R_\varepsilon(A; n^2) = \{n \in \mathbb{Z} : \mu(A \cap T^{n^2} A) > \mu(A)^2 - \varepsilon\}$$

and

$$R_\varepsilon^{(2)}(A) = \{(n, m) \in \mathbb{Z}^2 : \mu(A \cap T^n A \cap T^m A \cap T^{n+m} A) > \mu(A)^4 - \varepsilon\},$$

which are known to be syndetic, also possess stronger properties. The answer to this query is positive, but the issue is more delicate than one might expect. First, it is not true in general that the sets of the form $R_\varepsilon(A; n^2)$ are $\Delta_0^*$ or even $\Delta^*$. To see this, it is enough to consider an irrational rotation on the unit circle and utilize the fact that for any irrational number $\alpha$ and any $\varepsilon > 0$ there exists an infinite sequence $(n_i)$ such that, for any $i \ne j$, $\left|(n_j - n_i)^2 \alpha \bmod 1 - \frac{1}{2}\right| < \varepsilon$.

On the other hand, it was proved in [BFM] that the sets of the form $R_\varepsilon(A; n^2)$ are IP$^*$-sets (see Definition 0.4 above). This fact looks quite satisfactory since the IP$^*$ property is still (much) stronger than that of syndeticity. But, as we have already mentioned in the introduction, the sets $R_\varepsilon^{(2)}(A)$ do not possess this property. It can be shown that these sets are so-called *$C^*$ (central$^*$) sets.*[7] We obtain a stronger result (see Theorems 2.1 below and 3.9 below), which establishes, for the sets $R_\varepsilon^{(k)}(A)$ and $R_\varepsilon(A; p_1, \ldots, p_k)$, with certain algebraic independence restrictions on the polynomials $p_i$, an *almost* IP$^*$property: we show that they are AIP$^*$, and, indeed, AIP$_0^*$-sets. (In fact, we will show that these sets are AVIP$_0^*$-sets, see Definition 1.2 below.)

To visualize the relations between the various classes of sets we have described, let us denote by IP$^*$, IP$_r^*$, IP$_0^*$, AIP$^*$, AIP$_r^*$, AIP$_0^*$, C$^*$, AIP$_+^*$, AIP$_{r,+}^*$, AIP$_{0,+}^*$, and C$_+^*$, the families of sets having the corresponding property. We then have, for any $r_1 > r_2 \ge 2$, the

[7] A subset of $\mathbb{Z}^d$ is called *central* if it is a member of a minimal idempotent in the semigroup $(\beta\mathbb{Z}^d, +)$ (the Stone-Čech compactification of $\mathbb{Z}^d$; for relevant background see [B1] and [B3]). A subset of $\mathbb{Z}^d$ is *$C^*$* (*central$^*$*) if it has a non-trivial intersection with any central set in $\mathbb{Z}^d$. (This notion often appears in the context of multiple recurrence; see, for example, [BD], [BM1], [BM2].)

following diagram of strict inclusions:

$$\begin{array}{ccccccccccccc}
 & & & & \Delta^* & \supset & \Delta^*_0 & \supset & \Delta^*_{r_1} & \supset & \Delta^*_{r_2} \\
 & & & & \cap & & \cap & & \cap & & \cap \\
 & & & & \mathrm{IP}^* & \supset & \mathrm{IP}^*_0 & \supset & \mathrm{IP}^*_{r_1} & \supset & \mathrm{IP}^*_{r_2} \\
 & & & & \cap & & \cap & & \cap & & \cap \\
 & & \mathrm{C}^* & \supset & \mathrm{AIP}^* & \supset & \mathrm{AIP}^*_0 & \supset & \mathrm{AIP}^*_{r_1} & \supset & \mathrm{AIP}^*_{r_2} \\
 & & \cap & & \cap & & \cap & & \cap & & \cap \\
\mathcal{S} & \supset & \mathrm{C}^*_+ & \supset & \mathrm{AIP}^*_+ & \supset & \mathrm{AIP}^*_{0,+} & \supset & \mathrm{AIP}^*_{r_1,+} & \supset & \mathrm{AIP}^*_{r_2,+}.
\end{array} \tag{1.1}$$

The strictness of some of the inclusions in (1.1) is not immediately obvious, and we address this issue in the following comments.

The inclusion $\mathrm{C}^* \supseteq \mathrm{AIP}^*$ follows from the facts that a subset of $\mathbb{Z}^d$ is IP iff it is a member of any idempotent ultrafilter in the semigroup $\beta(\mathbb{Z}^d,+)$, whereas it is central iff it is a member of any minimal idempotent in $\beta(\mathbb{Z}^d,+)$ (see, for example, [B3], Theorem 2.15 and the subsequent remark), and that the family of $\mathrm{C}^*$-sets is stable under the operation of removing subsets of zero Banach density. On the other hand, a construction due to McCutcheon ([M]; see also [MZ]) shows that there are $\mathrm{C}^*$-sets which are not $\mathrm{AIP}^*$, that is, that this inclusion (as well as the inclusion $\mathrm{C}^*_+ \supset \mathrm{AIP}^*_+$) is strict.

The fact that $\mathcal{S} \supsetneq \mathrm{C}^*_+$ is proven in [B3], Theorem 2.20.

To see that, for $r \geq 2$, $\mathrm{IP}^*_r \neq \Delta^*_r$, it is enough to take a $\Delta_r$-set $B$ which contains no $\mathrm{IP}_r$-sets (say, $B = \{1,\ldots,r-1\} \subset \mathbb{Z}$) and put $S = \mathbb{Z}^d \setminus B$; then $S$ is an $\mathrm{IP}^*_r$ but not a $\Delta^*_r$-set.

To see that $\mathrm{IP}^*_{r_1} \neq \mathrm{IP}^*_{r_2}$ for $r_1 > r_2$ (which implies that, for any $r$, $\mathrm{IP}^*_0 \neq \mathrm{IP}^*_r$), take any $\mathrm{IP}_{r_2}$-set $B$ which is not an $\mathrm{IP}_{r_1}$-set and put $S = \mathbb{Z}^d \setminus B$. Similarly, taking $B$ to be a union of $\mathrm{IP}_r$-sets for all $r$ which contains no IP-set (in $\mathbb{Z}$, the set $B = \bigcup_{r=1}^{\infty}\{2^{2^r}, 2\cdot 2^{2^r}, 3\cdot 2^{2^r},\ldots, r2^{2^r}\}$, already used above, works) and putting $S = \mathbb{Z}^d \setminus B$, we get a set which is $\mathrm{IP}^*$ but not $\mathrm{IP}^*_0$.

It is somewhat harder to establish the strictness of the inclusions between the "A"-families of sets, such as $\mathrm{AIP}^*_r$, $\mathrm{AIP}^*_0$, and $\mathrm{AIP}^*$. To prove, in the case $d=1$, that $\mathrm{AIP}^*_{r_1} \neq \mathrm{AIP}^*_{r_2}$ for $r_1 > r_2$, one can check that for any $r \in \mathbb{N}$ the lattice $r\mathbb{Z}$ is an $\mathrm{IP}^*_r$-set but not $\mathrm{AIP}^*_{r-1}$. Indeed, given any $n_1,\ldots,n_r \in \mathbb{Z}$, consider the $r$ elements $m_1 = n_1$, $m_2 = n_1+n_2$, …, $m_r = n_1+\cdots+n_r$ of $\mathrm{FS}(n_1,\ldots,n_r)$; if none of $m_i$ is divisible by $r$, then for some $i<j$ one has $m_i = m_j \bmod r$, so $n_{i+1}+\cdots+n_j \in \mathrm{FS}(n_1,\ldots,n_r) \cap r\mathbb{Z}$. To show that $r\mathbb{Z}$ is not $\mathrm{AIP}^*_{r-1}$, let $N \subset \mathbb{Z}$ be a set of density zero, and assume that $S = r\mathbb{Z} \cup N$ is an $\mathrm{IP}^*_{r-1}$-set. Then for any $n_1,\ldots,n_{r-1} \in \mathbb{Z}$ with $n_i = 1 \bmod r$ for all $i$ we have $\mathrm{FS}(n_1,\ldots,n_{r-1}) \cap N \neq \emptyset$. Let $(n_1,\ldots,n_s)$ be a maximal sequence in $\mathbb{Z}$ for which $\mathrm{FS}(n_1,\ldots,n_s) \cap N = \emptyset$. Then for any $n \in \mathbb{Z}$ with $n = 1 \bmod r$ we must have $n+m \in N$ for some $m \in \mathrm{FS}(n_1,\ldots,n_s) \cup \{0\} \neq \emptyset$, which contradicts the assumption that $d^*(N) = 0$.

The strictness of $\mathrm{AIP}^* \supset \mathrm{AIP}^*_0$ can be proven "dynamically"; let us briefly describe the underlying idea. By Theorem 1.3 below, the set of returns in any distal topological dynamical system is $\mathrm{IP}^*$. Let $(X_0,T_0,x_0)$ be a pointed distal dynamical system with the property that for any pointed nilsystem[8] $(X_1,T_1,x_1)$, either the systems $(X_0,T_0,x_0)$ and

[8] Nilsystems, as well as distal systems, are defined later in this section.

$(X_1, T_1, x_1)$ are disjoint[9], or have a common isometric factor $(X_2, T_2, x_2)$ over which they are relatively disjoint. (An example of such a system $(X_0, T_0)$ is provided by a skew-product transformation $T_0(x, y) = (x + \alpha, y + f(x))$ on a two-dimensional torus $X_0 = \mathbb{T}^2$ for an irrational $\alpha \in \mathbb{T}$ and some $f \in C(\mathbb{T})$; we thank M. Lemanczyk for kindly confirming this to us.) Let $S = R_U(x_0)$ be the set of returns of $x_0$ into its sufficently small neighborhood $U$, $R_U = \{n \in \mathbb{Z}^d : T^n x_0 \in U\}$; then, by Theorem 1.3, $S$ is an IP$^*$-set. Let $P$ be any IP$^*_0$-set; by Theorem 2.6 in [HoK3], $P$ contains a *piecewise* $\mathit{Nil}_{r-1}\mathit{Bohr}_0$*-set*, that is, the intersection $P' \cap Q$ where $P'$ is a set of returns of the point $x_1$ of a pointed nilsystem $(X_1, T_1, x_1)$ into some its neighborhood and $Q$ is a thick[10] set. Then, because of the disjointness or relative disjointness of $(X_0, T_0, x_0)$ and $(X_1, T_1, x_1)$, the set $P' \setminus S$ is syndetic, and so the set $P \setminus S \supseteq (P' \setminus S) \cap Q$ has positive upper Banach density. Hence, we have $(S \cup N) \neq P$ for any set $N$ of zero Banach density, that is, $S$ is not an AIP$^*_0$-set. A similar argument applies to the families of the AIP$^*_{0,+}$ and AIP$^*_+$-sets. Apropos, based on the fact that every $\Delta^*$-set contains a piecewise-Bohr set (see [BFW]), one can show, in a similar way, that IP$^*_0 \supsetneq$ A$\Delta^*$ (where A$\Delta^*$is the family of “almost” $\Delta^*$-sets, i.e. sets of the form $R \setminus N$ where $R$ is a $\Delta^*$-set and $N$ is a set of zero Banach density).

This concludes our discussion of diagram (1.1). However, it turns out that the sets $R^{(k)}_\varepsilon(A)$ and $R_\varepsilon(A; p_1, \ldots, p_k)$ of large values of correlation functions possess an even stronger property than that of AIP$^*_0$, namely, AVIP$^*_0$, “the polynomial” AIP$^*_0$-property, which we will presently introduce. Let $(n_i)_{i\in\Lambda}$ be a collection of elements of $\mathbb{Z}^d$ indexed by a set $\Lambda$. (That is, let $n$ be a mapping $\Lambda \longrightarrow \mathbb{Z}^d$.) The IP-set generated by $(n_i)_{i\in\Lambda}$ can be interpreted as the image of “a linear” mapping $\varphi: \mathcal{F}(\Lambda) \longrightarrow \mathbb{Z}^d$ from the “partial semigroup” $\mathcal{F}(\Lambda)$ of finite subsets of $\mathbb{N}$ under the operation of disjoint unions: $\varphi$ is defined by $\varphi(\alpha) = \sum_{i\in\alpha} n_i$, $\alpha \in \mathcal{F}(\Lambda)$, and has the property $\varphi(\alpha_1 \cup \alpha_2) = \varphi(\alpha_1) + \varphi(\alpha_2)$ whenever $\alpha_1 \cap \alpha_2 = \emptyset$.

For an arbitrary mapping $\varphi: \mathcal{F}(\Lambda) \longrightarrow \mathbb{Z}^d$ and $\gamma \in \mathcal{F}(\Lambda)$ we define *the $\gamma$-derivative* $D_\gamma\varphi: \mathcal{F}(\Lambda \setminus \gamma) \longrightarrow \mathbb{Z}^d$ of $\varphi$ by $D_\gamma\varphi(\alpha) = \varphi(\alpha \cup \gamma) - \varphi(\alpha)$, $\alpha \in \mathcal{F}(\Lambda \setminus \gamma)$. We now have that $\varphi$ is linear, in the above sense, iff $\varphi(\emptyset) = 0$ and $D_\gamma\varphi$ is constant for every $\gamma \in \mathcal{F}(\Lambda)$; and equivalently, iff $\varphi(\emptyset) = 0$ and $D_{\gamma_1}D_{\gamma_2}\varphi = 0$ for any disjoint $\gamma_1, \gamma_2 \in \mathcal{F}(\Lambda)$. We say that a mapping $\varphi: \mathcal{F}(\Lambda) \longrightarrow \mathbb{Z}^d$ is *polynomial of degree* $\leq k$ if $D_{\gamma_1}D_{\gamma_2}\cdots D_{\gamma_{k+1}}\varphi = 0$ for any disjoint $\gamma_1, \ldots, \gamma_{k+1} \in \mathcal{F}(\Lambda)$. It can be shown (see [BL1], Theorem 8.3) that a mapping $\varphi: \mathcal{F}(\Lambda) \longrightarrow \mathbb{Z}^d$ is polynomial of degree $\leq k$ iff it is of the form $\varphi(\alpha) = \sum_{\substack{\beta\subseteq\alpha \\ |\beta|\leq k}} m_\beta$, where $\beta \mapsto m_\beta$ is a mapping from $\mathcal{F}_{\leq k}(\Lambda) = \big\{\beta \in \mathcal{F}(\Lambda) : |\beta| \leq k\big\}$ to $\mathbb{Z}^d$.

(9) *A pointed dynamical system* $(X, T, x)$ is a compact metric space $X$ with a homeomorphism $T: X \longrightarrow X$ and a point $x \in X$ whose orbit under $T$ is dense in $X$. Two pointed dynamical systems $(X_0, T_0, x_0)$ and $(X_1, T_1, x_1)$ are said to be *disjoint* if the orbit of $(x_0, x_1)$ under the action $T_0 \times T_1$ is dense in the product space $X_0 \times X_1$. If the systems $(X_0, T_0, x_0)$ and $(X_1, T_1, x_1)$ have a nontrivial common factor $(X_2, T_2, x_2)$, they cannot be disjoint; we say that these systems are *relatively disjoint* (with respect to $X_2$) if the orbit of $(x_0, x_1)$ is dense in the relative product space $X_0 \times_{X_2} X_1$.

(10) A subset of $\mathbb{Z}$ is said to be *thick* if it contains arbitrarily large intervals in $\mathbb{Z}$.

**Definition 1.2.** *A VIP-set* of degree $\leq k$ in $\mathbb{Z}^d$ is defined as the image $\varphi(\mathcal{F}(\Lambda))$ of a polynomial mapping $\varphi\colon \mathcal{F}(\Lambda) \longrightarrow \mathbb{Z}^d$ of degree $\leq k$ with $\varphi(\emptyset) = 0$. A set $S \subseteq \mathbb{Z}^d$ is said to be *a VIP*-set* if it has a nonzero intersection with any infinite VIP-set in $\mathbb{Z}^d$. A set $S \subseteq \mathbb{Z}^d$ is said to be *a $\mathrm{VIP}_0^*$-set* if for every $k \in \mathbb{N}$ there exists $r$ such that $S$ has a nonzero intersection with any VIP-set in $\mathbb{Z}^d$ of degree $\leq k$ and of cardinality $\geq r$. We define $\mathrm{AVIP}_0^*$-sets as sets of the form $S \setminus N$ where $S$ is a $\mathrm{VIP}_0^*$-set and $N$ is a set of zero Banach density in $\mathbb{Z}^d$. We define $\mathrm{VIP}_{0,+}^*$ and $\mathrm{AVIP}_{0,+}^*$-sets as shifted $\mathrm{VIP}_0^*$ and $\mathrm{AVIP}_0^*$-sets respectively.

Let us denote by $\mathrm{VIP}^*$, $\mathrm{VIP}_0^*$, $\mathrm{AVIP}^*$, and $\mathrm{AVIP}_0^*$ the classes of $\mathrm{VIP}^*$, $\mathrm{VIP}_0^*$, $\mathrm{AVIP}^*$, and $\mathrm{AVIP}_0^*$-sets respectively. We then have the following diagram of strict inclusions

$$\begin{array}{ccc} \mathrm{VIP}^* & \supset & \mathrm{VIP}_0^* \\ \cap & & \cap \\ \mathrm{IP}^* & \supset & \mathrm{IP}_0^* \end{array} \tag{1.2}$$

Again, easy examples prove the strictness of these inclusions. For the inclusion $\mathrm{IP}^* \supsetneq \mathrm{VIP}^*$, for instance, take $S = \mathbb{Z} \setminus B$, where $B = \{n^2 : n \in \mathbb{N}\}$. $B$ is a VIP-set of degree 2 (this is the image of the polynomial mapping $\varphi(\alpha) = |\alpha|^2$ from $\mathcal{F}(\mathbb{N})$ to $\mathbb{Z}$) and does not contain any infinite IP-set (since any infinite IP-set contains infinitely many pairs of elements with the same difference); thus, the set $S$ is an $\mathrm{IP}^*$ but not a $\mathrm{VIP}^*$-set in $\mathbb{Z}$.

As for the "A"-version of the diagram (1.2),

$$\begin{array}{ccc} \mathrm{AVIP}^* & \supseteq & \mathrm{AVIP}_0^* \\ \cap\vert & & \cap\vert \\ \mathrm{AIP}^* & \supset & \mathrm{AIP}_0^* \end{array} \tag{1.3}$$

we believe that all the inclusions in (1.3) are also strict, but this needs a separate investigation. (The dynamical approach described above fails here since both $\mathrm{IP}_0^*$ and $\mathrm{VIP}_0^*$ families of sets come from the same class of dynamical systems, nilsystems.)

We will now describe how $\mathrm{VIP}_0^*$ and $\mathrm{AVIP}_0^*$-sets appear in the context of the problems we deal with. The crucial fact which explains the emergence of $\mathrm{VIP}^*$-sets in our study has to do with the intrinsic "nilpotent" nature of multiple recurrence. *A polynomial multiple correlation sequence* is a sequence (or rather a mapping $\mathbb{Z}^d \longrightarrow \mathbb{R}$) of the form

$$\tau(n) = \mu\big(A_0 \cap T^{p_1(n)} A_1 \cap \cdots \cap T^{p_k(n)} A_k\big),\ n \in \mathbb{Z}^d, \tag{1.4}$$

or, more generally,

$$\tau(n) = \int_X f_0 \cdot T^{p_1(n)} f_1 \cdots T^{p_k(n)} f_k \, d\mu,\ n \in \mathbb{Z}^d, \tag{1.5}$$

where $T$ is a measure preserving transformation of a probability measure space $(X, \mathcal{B}, \mu)$, $p_i$ are polynomials $\mathbb{Z}^d \longrightarrow \mathbb{Z}$, $A_i \in \mathcal{B}$, and $f_i \in L^\infty(X)$. It follows from [L3], Theorem 3 that any polynomial multiple correlation sequence (1.5) induced by an ergodic transformation is *approximable in density* by correlation sequences of the same sort coming from *nilsystems*:

for any $\varepsilon > 0$ there exists a compact *nilmanifold* $X'$ (a compact homogeneous space of a nilpotent Lie group $G$) with normalized Haar measure $\mu'$, a *nilrotation* $T'$ (the translation of $X'$ by an element of $G$), and functions $f'_0, \ldots, f'_k \in C(X')$ such that for the sequence

$$\tau'(n) = \int_{X'} f'_0 \cdot (T')^{p_1(n)} f'_1 \cdots (T')^{p_k(n)} f'_k \, d\mu'$$

one has $d^*\big(\{n \in \mathbb{Z}^d : |\tau(n) - \tau'(n)| > \varepsilon|\}\big) = 0$. (In other words, nilsystems approximate *a characteristic factor* of $(X, T)$(11); see [HoK1] or [Zi].) It follows that when studying the level sets of polynomial multiple correlation sequences, we may confine ourselves to nilsystems as long as we are ready to ignore subsets of $\mathbb{Z}^d$ of zero Banach density and small errors.

Nilrotations are known to be distal transformations(12) (see [AGH], [Ke1], [Ke2]) and therefore, by the following theorem, possess the property of IP$^*$-recurrence:

**Theorem 1.3.** (Cf. [Fu], Theorem 9.11) *Let $X$ be a compact metric space and let $n \mapsto T^n$ be a distal action of $\mathbb{Z}^d$ on $X$ by self-homemomorphism. Then for any $x_0 \in X$ and any neighborhood $U$ of $x_0$ the set $R_U(x_0) = \big\{n \in \mathbb{Z}^d : T^n x_0 \in U\big\}$ is a IP$^*$-set.*

While the IP$^*$ property of the sets of returns $R_U(x_0)$ is universal for all distal systems (and, in fact, characterizes distality – see, for example, Theorem 3.8 in [B3]), nilsystems possess a stronger property:

**Theorem 1.4.** ([BL3], Theorem 0.5) *Let $X$ be a compact nilmanifold, let $n \mapsto T^n$ be an action of $\mathbb{Z}^d$ on $X$ by nilrotations. Then for any $x_0 \in X$ and any neighborhood $U$ of $x_0$, the set $R_U(x_0)$ is a VIP$^*_0$-set.*

Let us remark that the class of VIP$^*_0$-sets is stable under taking polynomial preimages: if $S \subseteq \mathbb{Z}^d$ is a VIP$^*_0$-set and $p\colon \mathbb{Z}^l \longrightarrow \mathbb{Z}^d$ is a polynomial mapping with $p(0) = 0$, then $p^{-1}(S)$ is a VIP$^*_0$-set in $\mathbb{Z}^l$. (This follows from the fact that if $B$ is a VIP-set of degree $\leq k$ in $\mathbb{Z}^l$ then $p(B)$ is a VIP-set of degree $\leq k \deg p$ in $\mathbb{Z}^d$.) It thus follows from Theorem 1.4 that for any polynomial mapping $p\colon \mathbb{Z}^l \longrightarrow \mathbb{Z}^d$ with $p(0) = 0$ the set $R_U(x_0; p) = \big\{n \in \mathbb{Z}^l : T^{p(n)} x_0 \in U\big\}$ is VIP$^*_0$.

Let $X$ be a compact nilmanifold, $n \mapsto T^n$ is an action of $\mathbb{Z}^d$ on $X$ by nilrotations, $h \in C(X)$, and $x_0 \in X$; then the sequence $\varphi(n) = h(T^n x_0)$, $n \in \mathbb{Z}^d$, is called *a basic nilsequence*; a general *nilsequence* is a uniform limit of basic ones. (This terminology was introduced in [BHoK].) For any $c < h(x_0)$, by Theorem 1.4 the set $R = \big\{n \in \mathbb{Z}^d : \varphi(n) > c\big\}$ is an VIP$^*_0$-set. If one has $c < h(x)$ for some $x \neq x_0$ (and the action $T$ is assumed to be ergodic), $R$ is, in general, only an VIP$^*_{0,+}$ (see Definition 0.10).

Because of the "nilpotent nature" of polynomial multiple correlation sequences, it is not surprising that any such sequence is an "almost" nilsequence:

(11) A factor system $(X', \mathcal{B}', \mu', T')$ of a probability measure preserving system $(X, \mathcal{B}, \mu, T)$ is said to be *characteristic* (*with respect to the scheme* $\int_X f_0 \cdot T^{p_1(n)} f_1 \cdots T^{p_k(n)} f_k \, d\mu$) if for any $f_0, f_1, \ldots, f_k \in L^\infty(X)$ one has UC-$\lim_{n \in \mathbb{Z}^d}\big(\int_X f_0 \cdot T^{p_1(n)} f_1 \cdots T^{p_k(n)} f_k \, d\mu - \int_{X'} f_0 \cdot (T')^{p_1(n)} E(f_1|X') \cdots (T')^{p_k(n)} E(f_k|X') \, d\mu'\big) = 0$. (UC-lims are defined below.)

(12) An action $n \mapsto T^n$ of $\mathbb{Z}^d$ on a metric space $(X, \rho)$ by continuous transformation is said to be *distal* if for any distinct $x_1, x_2 \in X$, $\inf_{n \in \mathbb{Z}^d} \rho(T^n x_1, T^n x_2) > 0$.

**Theorem 1.5.** ([L7], Theorem 0.1) *Let $(X,\mathcal{B},\mu,T)$ be an invertible probability measure presering system. For any polynomials $p_1,\dots,p_k\colon\mathbb{Z}^d\longrightarrow\mathbb{Z}$ and any $A_0,\dots,A_k\in\mathcal{B}$, the polynomial multiple correlation sequence $\tau(n)=\mu\big(A_0\cap T^{p_1(n)}A_1\cap\dots\cap T^{p_k(n)}A_k\big)$, $n\in\mathbb{Z}^d$, has the form $\tau=\varphi+\lambda$, where $\varphi$ is a nilsequence and $\lambda$ is a null-sequence*[13].

(See also a recent result of Frantzikinakis in [F], which deals with the case of several commuting transformations.)

**Definition.** For a sequence $(u_n)$, $n\in\mathbb{Z}^d$, in a normed vector space we write $\text{UC-lim}_{n\in\mathbb{Z}^d}u_n$ for $\lim_{N\to\infty}\frac{1}{|\Phi_N|}\sum_{n\in\mathbb{Z}^d}u_n$ if this limit exists for all Følner sequences $(\Phi_N)$ in $\mathbb{Z}^d$ (in which case the limit is the same for all these sequences).

Let $\tau$ be a polynomial multiple correlation sequence with $\text{UC-lim}_{n\in\mathbb{Z}^d}\tau(n)=C$. Represent $\tau$ in the form $\tau=\varphi+\lambda$ where $\varphi$ is a nilsequence, $\varphi(n)=h(T^nx_0)$ for $x_0\in X$, $h\in C(X)$, and $\lambda$ is a null-sequence. Then $\int_X h\,d\mu(X)=\text{UC-lim}_n\varphi(n)=\text{UC-lim}_n\tau(n)=C$. From Theorem 1.4 we now get the following result:

**Theorem 1.6.** *For any polynomial multiple correlation sequence $\tau$ and any $\varepsilon>0$, the set $R=\big\{n\in\mathbb{Z}^d:\tau(n)>C-\varepsilon\big\}$, where $C=\text{UC-lim}_{n\in\mathbb{Z}^d}\tau(n)$, is AVIP$^*_{0,+}$.*

We now see from Theorems 1.6 and 0.1 that the sets $R^{(k)}_\varepsilon(A)$ and $R_\varepsilon(A;p_1,\dots,p_k)$ appearing in Theorems 0.6 and 0.8 are AVIP$^*_{0,+}$-sets; our goal is to show that these sets are in fact "non-shifted" AVIP$^*_0$-sets. Under the notation of Theorem 1.6, to prove that a set $R$ is AVIP$^*_0$ it suffices to represent the corresponding multiple correlation sequence in the form $\varphi+\lambda$, where $\varphi(n)=h(T^nx_0)$ is a nilsequence and $\lambda$ is a null-sequences, and to show that $h(x_0)\geq C$. In order to achieve this goal we will have to take a close look at the orbit closure of the diagonal of a Cartesian power of a nilmanifold.

## 2. The sets $R^{(k)}_\varepsilon(A)$

The following is the main "linear" (i.e. pertaining to polynomials of degree 1) result of the paper:

**Theorem 2.1.** *Let $(X,\mathcal{B},\mu,T)$ be an ergodic probabilty measure preserving system and let $f\in L^\infty(X)$. Then for any $k\in\mathbb{N}$ and $\varepsilon>0$, the set*

$$R^{(k)}_\varepsilon(f)=\Big\{(n_1,\dots,n_k)\in\mathbb{Z}^k:\int_X\prod_{e_1,\dots,e_k\in\{0,1\}}T^{e_1n_1+\dots+e_kn_k}f\,d\mu>\Big(\int_X f\,d\mu\Big)^{2^k}-\varepsilon\Big\}$$

*is AVIP$^*_0$.*

---

[13] We say that a sequence $\lambda(n)$ is *a null-sequence* if $\lambda\longrightarrow 0$ in Banach density, that is, for any $\varepsilon>0$ the set $\{n\in\mathbb{Z}^d:|\lambda(n)|>\varepsilon\}$ has Banach density zero.

Taking $f = 1_A$ in this theorem where $A$ is a measurable subset of $X$ and confining ourselves to VIP-sets of first degree, we get Theorem 0.6; then, applying Furstenberg's correspondence principle, we obtain Corollary 0.7.

The proof of Theorem 2.1 hinges on results from [L1], [L2], [L3], [L4], [L6], [L7], [BL2], and [BL3]. Let us consider a more general situation: let $p_1, \ldots, p_r$ be distinct linear forms $\mathbb{Z}^k \longrightarrow \mathbb{Z}$ and assume that we need to show that for any nonnegative function $f \in L^\infty(X)$ the set

$$\Big\{n \in \mathbb{Z}^k : \int_X \prod_{i=1}^{r} T^{p_i(n)} f \, d\mu > \Big(\int_X f \, d\mu\Big)^r - \varepsilon\Big\}$$

is $\mathrm{AVIP}_0^*$. First, as was described above, we may confine ourselves to nilsystems and assume that $X = G/\Gamma$ is a compact nilmanifold, $\mu$ is the Haar measure, $T \in G$ is a nilrotation of $X$, and $f$ is a continuous function on $X$. Next, we may assume that the nilmanifold $X$ is connected. Indeed, let $X$ have $b > 1$ connected components, $X_1^o, \ldots, X_b^o$; then for any $j$, the translation $T^b$ preserves $X_j^o$ and is ergodic on it. Assume that for any $j = 1, \ldots, b$ the set

$$R_j = \Big\{n \in \mathbb{Z}^k : \int_X \prod_{i=1}^{r} T^{bp_i(n)} f|_{X_j^o} \, d(b\mu) > \Big(\int_X f|_{X_j^o} \, d(b\mu)\Big)^r - \varepsilon/b\Big\},$$

is $\mathrm{AVIP}_0^*$; then the set $R' = \bigcap_{j=1}^{b} R_j$ is also $\mathrm{AVIP}_0^*$, and for any $n \in R'$, by the convexity of the function $t \mapsto t^r$,

$$\int_X \prod_{i=1}^{b} T^{bp_i(n)} f \, d(b\mu) > \big(\int_X f \, d(b\mu)\big)^r - \varepsilon,$$

so $bR' \subseteq R$.

It is easy to see that if $R'$ is an $\mathrm{AVIP}_0^*$-set, then $bR'$ is also $\mathrm{AVIP}_0^*$, so $R$ is $\mathrm{AVIP}_0^*$. So, after replacing $T$ by $T^b$ and $X$ by each of $X_j^o$, $j = 1, \ldots, b$, we may and will assume that $X$ is connected.

Let $D$ be the diagonal $\{(x, \ldots, x),\ x \in X\}$ of $X^r$. For any $m = (m_1, \ldots, m_r) \in \mathbb{Z}^r$ we have

$$\int_X \prod_{i=1}^{r} T^{m_i} f \, d\mu = \int_{D_m} f^{\otimes r} d\mu_{D_m},$$

where $D_m = (T^{m_1}, \ldots, T^{m_k})D$, $f^{\otimes r}(x_1, \ldots, x_r) = f(x_1) \cdots f(x_r)$, and $\mu_{D_m}$ is the normalized Haar measure on $D_m$. Put $p = (p_1, \ldots, p_r)$ and $F = f^{\otimes r}$. We next use results from [L6] and [L7], which describe the behavior of the sequence $\int_{D_{p(n)}} F \, d\mu_{D_{p(n)}}$. Let $Y$ be the orbit closure $\overline{\bigcup_{n\in\mathbb{Z}^k} D_{p(n)}}$ of $D$ in $X^r$ under the action $T^{p(n)} = (T^{p_1(n)}, \ldots, T^{p_1(n)})$; then $Y$ is a subnilmanifold of $X$, and $\text{UC-lim}_{n\in\mathbb{Z}^k} \int_{D_{p(n)}} F \, d\mu_{D_{p(n)}} = \int_Y F \, d\mu_Y$, where $\mu_Y$ is the Haar measure on $Y$. (See [L2], Corollary 1.9.)

Let $G_0$ be the subgroup of $G$ generated by $T$ and the identity component of $G$; then $X$ is a homogeneous space of $G_0$ as well, so, we may and will assume that $G = G_0$. Let $\pi\colon G \longrightarrow X$ be the natural projection, and let $H$ be the minimal closed subgroup of $G^r$ such that $T^{p(n)} \in H$ for all $n$ and $\pi(H) = Y$. Let $K$ be the normal closure in

$H$ of the diagonal $\Delta = \{(g,\dots,g),\ g \in G\}$ of $G^r$. (That is, $K$ is the minimal normal subgroup of $H$ containing $\Delta$.) Let $N = K\backslash Y$, let $\sigma\colon Y \longrightarrow N$ be the natural projection, let $h = E(F|N)$, and let $w = \sigma(K)$. Then the sequence $T^{p(n)}$ acts on $N$, and by [L7], Proposition 0.2, if we put $\tau(n) = \int_{D_{p(n)}} F\,d\mu_{D_{p(n)}}$ and $\tau'(n) = h(T^{p(n)}w)$, $n \in \mathbb{Z}^d$, then for any $\varepsilon > 0$, $d^*\bigl(\{n \in \mathbb{Z}^d : |\tau(n) - \tau'(n)| > \varepsilon|\}\bigr) = 0$. Hence, we only need to prove that the set $\bigl\{n \in \mathbb{Z}^k : h(T^{p(n)}w) > \bigl(\int_X f\,d\mu\bigr)^r - \varepsilon\bigr\}$ is VIP$^*_0$. By Theorem 1.4, for any $\delta > 0$ the set $\bigl\{n \in \mathbb{Z}^k : \mathrm{dist}(T^{p(n)}w, w) < \delta\bigr\}$ is a VIP$^*_0$-set; it therefore remains to show that $h(w) \geq \bigl(\int_X f\,d\mu\bigr)^r$.

Let $W = \pi^{\times r}(K)$, where $\pi^{\times r}$ denotes the Cartesian product of $k$ copies of $\pi$, $\pi^{\times r}(x_1,\dots,x_k) = \bigl(\pi(x_1),\dots,\pi(x_k)\bigr)$. We have $h(w) = \int_W F\,d\mu_W$, where $\mu_W$ is the Haar measure on $W$; we will now determine $K$ and so, $W$. We need to first introduce some additional notation. Let the nilpotency class of $G$ be $c$, and let $G = G_1 > G_2 > \cdots > G_c > G_{c+1} = \{1_G\}$ be the lower central series of $G$. For $g \in G$ and $v = (v_1,\dots,v_r) \in \mathbb{Z}^r$, define $g^v = (g^{v_1},\dots,g^{v_r}) \in G^r$. For a group $A$ and a set $V \subseteq \mathbb{Z}^r$, by $A^V$ we understand the subgroup of $A^r$ generated by the set $\{g^v,\ g \in A,\ v \in V\}$. For $v = (v_1,\dots,v_r)$ and $u = (u_1,\dots,u_r) \in \mathbb{Z}^r$, let $vu = (v_1u_1,\dots,v_ru_r)$. For a subgroup $V$ of $\mathbb{Z}^r$ and $i \in \mathbb{N}$, let $V^{\star i}$ denote the subgroup of $\mathbb{Z}^r$ generated by the products $v_1\cdots v_i$, $v_1,\dots,v_i \in V$.

Now let $p_i(n_1,\dots,n_k) = \sum_{j=1}^k a_{i,j}n_j$, $i = 1,\dots,r$, and let $V$ be the subgroup of $\mathbb{Z}^r$ generated by the vectors $e = (1,\dots,1)$, $(a_{1,1},\dots,a_{r,1})$, $\dots$, $(a_{1,k},\dots,a_{r,k})$. By [L6], Theorem 6.3,

$$H = G^V G_2^{V^{\star 2}} \cdots G_c^{V^{\star c}}.$$

It follows that

$$K = G^e G_2^V G_3^{V^{\star 2}} \cdots G_c^{V^{\star(c-1)}} = \Delta G_2^V G_3^{V^{\star 2}} \cdots G_c^{V^{\star(c-1)}}.$$

Recall that $W = K/(\Gamma \cap K)$. The torus $(G_2^r \cap K)\backslash W$ is naturally isomorphic to $Z = G_2\backslash X$; let $\eta\colon W \longrightarrow Z$ be the natural projection. Then we have

$$\int_W F\,d\mu_W = \int_W f^{\otimes r}d\mu_W = \int_Z \int_{W_z} f^{\otimes r}(w)\,d\mu_{W_z}(w)\,d\mu_Z(z),$$

where for each $z \in Z$, $W_z = (\eta^{\times r})^{-1}(z)$, and $\mu_{W_z}$ is the Haar measure on $W_z$. For any $z \in Z$ we have $W_z = \pi^{\times r}(K_z)$ where

$$K_z = g_z^e G_2^V G_3^{V^{\star 2}} \cdots G_c^{V^{\star(c-1)}}$$

for $g_z \in \pi^{-1}(\eta^{-1}(z))$. Now let $X_2 = \pi(G_2)$, let

$$M = G_2^V G_3^{V^{\star 2}} \cdots G_c^{V^{\star(c-1)}} \subseteq G_2^r, \tag{2.1}$$

and let $L$ be the subnilmanifold $\pi^{\times r}(M)$ of $X_2^r$; let $\mu_{X_2}$ and $\mu_L$ be the Haar measures on $X_2$ and on $L$ respectively. Then for any $z \in Z$, $K_z = g^e M$ and $W_z = g_z^e L$, so

$$\int_{W_z} f^{\otimes r}(w)\,d\mu_{W_z} = \int_L h_z^{\otimes r}\,d\mu_L$$

where $h_z(u)=f(g_z u)$, $u\in L$. It therefore suffices to show that for any nonnegative continuous function $h$ on $X_2$ one has

$$\int_L h^{\otimes r} d\mu_L \geq \Bigl(\int_{X_2} h\, d\mu_{X_2}\Bigr)^r, \tag{2.2}$$

since, if this is the case, we will have

$$\begin{aligned}\int_W f^{\otimes r} d\mu_W &= \int_Z \int_{W_z} f^{\otimes r}(w)\, d\mu_{W_z}(w)\, d\mu_Z(z) = \int_Z \int_L h_z^{\otimes r}\, d\mu_L\, d\mu_Z(z)\\ &\geq \int_Z \Bigl(\int_{X_2} h_z\, d\mu_{X_2}\Bigr)^r d\mu_Z(z) \geq \Bigl(\int_Z \int_{X_2} h_z\, d\mu_{X_2}\, d\mu_Z(z)\Bigr)^r\\ &= \Bigl(\int_Z \int_{W_z} f(w)\, d\mu_{W_z}(w)\, d\mu_Z(z)\Bigr)^r = \Bigl(\int_Z f\, d\mu_Z\Bigr)^r.\end{aligned}$$

The definition of the subgroup $M$ of $G_2^r$ (in formula (2.1)) and of the subnilmanifold $L$ of $X_2^r$ looks similar to the definition of the subgroup $H$ of $G^r$ and of the subnilmanifold $Y$ of $Z^r$, with $G$ replaced by $G_2$; however, this similarity is somewhat deceiving since the series $G_2>G_3>\cdots>G_c$ is not the lower central series of the group $G_2$. We can restore the similarity by considering the group $Q=G_2\Gamma$; the lower central series of $Q$ is $G_2\Gamma_1>G_3\Gamma_2>\cdots>G_c\Gamma_{c-1}>\Gamma_c$, where $\Gamma=\Gamma_1>\Gamma_2>\cdots>\Gamma_c$ is the lower central series of $\Gamma$. The nilmanifold $X_2$ is a homogeneous space of $Q$, $X_2=Q/\Gamma$, and the "canonical" subnilmanifolds $G_{i+1}\Gamma_i/\Gamma_i$, $i=1,\dots,c$, of $X_2$ coincide with the nilmanifolds $G_{i+1}/\Gamma_{i+1}$. Choose an element $P_0$ in the identity component of $G_2$ such that the orbit $\pi(P_0^n)$, $n\in\mathbb{Z}$, is dense in $X_2$. ($P_0$ is any "irrational" element of the identity component of $G_2$; see [L4], Section 1.2.) Let $\Lambda$ be the group generated by $\Gamma$ and $P_0$; then the action of $\Lambda$ on $X_2$ is ergodic.

**Proof of Theorem 2.1.** We will now deal with the special case where $r=2^k$ and $\{p_1,\dots,p_r\}=\bigl\{e_1n_1+\dots+n_ke_k,\ e_i\in\{0,1\}\bigr\}$. It follows from Theorem 6.3 from [L6] that under the action $\bigl(P_1^{e_1}\cdots P_k^{e_k},\ e_1,\dots,e_k\in\{0,1\}\bigr)_{(P_1,\dots,P_k)\in\Lambda^k}$ of $\Lambda^{2^k}$, the orbit closure of the diagonal of $X_2^{2^k}$ is the nilmanifold $\widetilde{L}=\pi^{\times 2^k}(\widetilde{M})$, where

$$\widetilde{M}=G_2^V\Gamma^{U_1}G_3^{V^{\star 2}}\Gamma_2^{U_2}\cdots G_c^{V^{\star(c-1)}}\Gamma_{c-1}^{U_{c-1}}\Gamma_c^{U_c},$$

where $V$ is, as above, the subgroup of $\mathbb{Z}^{2^k}$ generated by the vector-coefficients of the linear form $p=(p_1,\dots,p_{2^k})$, and for each $i$, $U_i$ is a subgroup of $\mathbb{Z}^{2^k}$ that is slightly larger than $V^{\star i}$. ([L6] deals with a single transformation, but the results therein can be easily extended to the case of a group action: $\widetilde{M}$ contains the orbit of the diagonal of $Q^{2^k}$ and is a minimal subgroup of $Q^{2^k}$ whose projection to $X_2^{2^k}$ coincides modulo $Q_2^{2^k}=(G_3\Gamma_2)^{2^k}$ with the closure of this orbit of the diagonal of $X_2^{2^k}$. For each $i$, $V^{\star i}$ has finite index in $U_i$, and $\Gamma_i^{U_i}$ normalizes $G_j^{\star V_j}$ for all $j\geq i$; the exact definition of the groups $U_i$ is cumbersome, and we will not specify them here because, anyway, the groups $\Gamma_i^{U_i}$ vanish after an application of the mapping $\pi^{\times 2^k}$.) Thus, $\widetilde{L}=\pi^{\times 2^k}(M)$ where $M=G_2^VG_3^{V^{\star 2}}\cdots G_c^{V^{\star(c-1)}}$, that is, $\widetilde{L}$

coincides with the nilmanifold $L$ introduced above. By [L1], Theorems 2.17 and 2.19, for any nonnegative continuous function $h$ on $X_2$ one has

$$\operatorname*{UC-lim}_{(P_1,\dots,P_k)\in\Lambda^k}\int_X\prod_{e_1,\dots,e_k\in\{0,1\}}P_1^{e_1}\cdots P_k^{e_k}h\,d\mu_{X_2}=\int_L h^{\otimes 2^k}\,d\mu_L.$$

By [BL2], Theorem 0.5,

$$\operatorname*{UC-lim}_{(P_1,\dots,P_k)\in\Lambda^k}\int_X\prod_{e_1,\dots,e_k\in\{0,1\}}P_1^{e_1}\cdots P_k^{e_k}h\,d\mu_{X_2}\geq\Bigl(\int_{X_2}h\,d\mu_{X_2}\Bigr)^{2^k}.$$

This establishes (2.2) and thus concludes the proof. ■

## 3. The sets $R_\varepsilon(A;p_1,\dots,p_k)$

We will now extend Theorem 2.1 to polynomial sequences of powers of a measure preserving transformation. Let $p_1,\dots,p_k$ be polynomials $\mathbb{Z}^d\longrightarrow\mathbb{Z}$ with zero constant term. Notice that despite the fact that, by Theorem 2.1, the set $R_\varepsilon(f)$ has "polynomial flavor", this theorem gives no information whether there are nonzero $n\in\mathbb{Z}^d$ such that $(p_1(n),\dots,p_k(n))\in R_\varepsilon(f)$ – since the set $R_\varepsilon(f)$ is not a VIP$_0^*$ but only an AVIP$_0^*$-set, and since, in general, the set $\bigl\{(p_1(n),\dots,p_k(n)),\ n\in\mathbb{Z}^d\bigr\}$ has zero Banach density in $\mathbb{Z}^k$.

As in the linear case, when studying polynomial multiple correlation sequences (1.5), after ignoring a subset of zero Banach density in $\mathbb{Z}^d$ and an arbitrarily small error we may assume that $(X,T)$ is a nilsystem. However, in comparison with the "linear" case, the polynomial situation presents additional difficulties. Like in the linear case, the value of the limit $\operatorname{UC-lim}_{n\in\mathbb{Z}^d}\prod_{i=1}^k T^{p_i(n)}f$ is equal to the integral of $f^{\otimes k}$ over the orbit closure $Y$ of the diagonal of $X^k$ under the polynomial action $(T^{p_1(n)},\dots,T^{p_k(n)})$, but now $Y$ is not, generally speaking, a nilmanifold – it is a union of several nilmanifolds, which are visited by the sequence $(T^{p_1(n)},\dots,T^{p_k(n)})$ with potentially different frequencies. (Consider, for example, the polynomial mapping $n^2$ from $\mathbb{Z}$ to $\mathbb{Z}/3\mathbb{Z}$: the orbit closure $Y$ consists here of the points 0 and 1, where 0 is visited by the sequence $(n^2)$ with the frequency 1/3 and 1 with the frequency 2/3.) Also, in the case of polynomial actions, the orbit closure $Y$ does not have such a simple description as in the linear case. (See [L6], Sections 9–13.) The first problem can be got around by passing to a subsequence $(T^{p_1(bn)},\dots,T^{p_k(bn)})$ of $(T^{p_1(n)},\dots,T^{p_k(n)})$ for certain $b\in\mathbb{N}$ so that the orbit closure $Y$ is reduced to a single connected component. The second problem can be avoided by dealing with only "sufficiently algebraically independent" systems of polynomials.

Let $P=\{p_1,\dots,p_k\}$ be a system of polynomials $\mathbb{Z}^d\longrightarrow\mathbb{Z}$; *the complexity* of $P$ is defined as the minimal nonnegative integer $c$ such that, when computing the limit $\operatorname{UC-lim}_{n\in\mathbb{Z}^d}\int_X\prod_{i=1}^k T^{p_i(n)}f_i\,dx$ for a general ergodic probability system $(X,T)$, one may replace $X$ by its factor-nilsystem of nilpotency class $c$ (see [BLLe1], Section B of the introduction, and [L6], Sections 2.7–2.11). If the complexity of $P$ is $c$, then the polynomial multiple correlation sequence (1.5) corresponding to $P$ for an ergodic system can be computed, up to an arbitrarily small error, on a factor-nilsystem of nilpotency class $c+1$.

(See [L7], Section 5.) The complexity of $P$ depends only on the orbit closure of the diagonal $D$ of the power $X^k$ for a general nilsystem $(X,T)$ under the polynomial action $(T^{p_1(n)},\dots,T^{p_k(n)})$ (see [L6], Section 0.7).

We will also need some information about the structure of nilrotations. Let $X$ be a connected nilmanifold, $X=G/\Gamma$, where $G$ is a nilpotent Lie group of nilpotency class $\le c$ and $\Gamma$ is a lattice in $G$, let $G^o$ be the identity component of $G$, and let $Q=[G^o,G^o]\backslash X=G/([G^o,G^o]\Gamma)$. $Q$ is a torus (*the maximal factor-torus* of $X$), and under the action of $G$ has the structure of a $c$-step skew-product system. To see this, let $\widehat{G}=G/[G^o,G^o]$, so that $Q$ is the factor of $\widehat{G}$ by the lattice $\widehat{\Gamma}=\Gamma/([G^o,G^o]\cap\Gamma)$. Let us identify elements of $G$ with their images in $\widehat{G}$. Let $N$ be the identity component of $\widehat{G}$, then $N$ is a connected commutative Lie group; after passing to the universal cover, we may assume that $N\cong\mathbb{R}^d$. Let $N_1=N$, $N_2=[G,N_1]$, $N_3=[G,N_2]$, $\dots$, $N_{c+1}=[G,N_c]=\{1\}$; this is a sequence of nested connected subgroups of $N$ with the property that, for each $i$, the lattice $N_i\cap\widehat{\Gamma}$ is cocompact in $N_i$. For any $\gamma\in\Gamma$ and any $u\in N$ we have $\gamma u=\gamma u\gamma^{-1}$, that is, the action of $\gamma$ on $N$ is a linear transformation $A_\gamma$ of $L$. Moreover, for any $i$ and any $u\in N_i$, $u^{-1}A_\gamma u=[u,\gamma^{-1}]\in N_{i+1}$, or, passing to additive notation, $A_\gamma u\in u+N_{i+1}$, and every element $T=v\gamma\in G$, with $v\in G^o$ and $\gamma\in\Gamma$, acts as an affine transformation, $Tu=v+A_\gamma u$. For each $i$ find a subspace (a closed subgroup) $L_i$ of $N_i$ that is generated by $L_i\cap\widehat{\Gamma}$ and such that $N_i=N_{i+1}\times L_i$. Then $N=L_1\times L_2\times\dots\times L_c$, and every element $T$ of $G$ acts on $u=(u_1,u_2,u_3,\dots,u_c)$ by the formula

$$T(u_1,u_2,u_3,\dots,u_c)=\big(u_1+v_1,u_2+\psi_2(u_1)+v_2,u_3+\psi_3(u_1,u_2)+v_3,\dots,\\ u_c+\psi_c(u_1,\dots,u_{c-1})+v_c\big),$$

where the $\psi_i$ are linear forms and $v_i\in L_i$, $i=1,\dots,c$. Factorizing by $\Gamma$ and defining $Q_i$ to be the image of $L_i$, $i=1,\dots,c$, in $Q$, we obtain that $Q=Q_1\times Q_2\times\dots\times Q_c$, where the $Q_i$ are tori and every element $T$ of $G$ acts on $Q$ by the formula

$$T(z_1,z_2,z_3,\dots,z_c)=\big(z_1+\alpha_1,z_2+\psi_2(z_1)+\alpha_2,z_3+\psi_3(z_1,z_2)+\alpha_3,\dots,\\ z_c+\psi_c(z_1,\dots,z_{c-1})+\alpha_c\big),$$

where the $\psi_i$ are linear forms and $\alpha_i\in Q_i$, $i=1,\dots,c$.

Let us start with the case where the polynomials $p_i$ are linearly independent. When $X$ is a connected nilmanifold $G/\Gamma$ and $T\in G$ is an ergodic nilrotation of $X$, we will use notation from Section 2, with $r$ replaced by $k$: let $G_2=[G,G]$, $Z=G_2\backslash X$, $D$ be the diagonal of $X^k$, $Y$ be the orbit closure of $D$ under the (polynomial) action $T^{p(n)}=(T^{p_1(n)},\dots,T^{p_k(n)})$, $\Delta$ be the diagonal of $G^k$, $H=(\pi^{\times k})^{-1}(Y)\subseteq G^k$, $\pi\colon G\longrightarrow H$ be the natural projection, $K$ be the normal closure of $\Delta$ in $H$, $W=\pi^{\times k}(K)\subseteq X^k$, $F=f^{\otimes k}$, $h=E(F|N)$, and $w=\sigma(K)$; let also $G_2^o=[G^o,G^o]$ and $Q=G_2^o\backslash X$. $Z$ and $Q$ are tori with $T$ acting as an ergodic rotation on $Z$ and as a skew-shift (an affine unipotent transformation) on $Q$. If the polynomials $p_1,\dots,p_k$ are linearly independent, the orbit of every point, and so of the diagonal, of $Z^k$ under the action $(T^{p_1(n)},\dots,T^{p_k(n)})$ is dense, and actually well distributed, in the torus $Z^k$. Since the system $\{p_1,\dots,p_k\}$ has W-complexity (the complexity with respect to the Weyl systems) 1 (see [L6], Theorem 9.7), this implies that the orbit of the diagonal of $Q^k$ is well distributed in the torus $Q^k$ as well. And by [L2], Theorem C and Corollary 1.9, we then have that $Y=X^k$ and the orbit of $D$ is well distributed in $X^k$. (For the case $d=1$ see also [FK], Theorem 1.2.)

**Theorem 3.1.** *Let $(X,\mathcal{B},\mu,T)$ be an invertible probability measure preserving system and let $p_1,\dots,p_k$ be linearly independent polynomials $\mathbb{Z}^d \longrightarrow \mathbb{Z}$ with $p_i(0)=0$, $i=1,\dots,k$. Then for any nonnegative function $f\in L^\infty(X)$ and any $\varepsilon>0$ there exists $b\in\mathbb{N}$ such that*

$$\operatorname*{UC-lim}_{n\in\mathbb{Z}^d}\int_X\prod_{i=1}^k T^{p_i(bn)}f\,d\mu>\Big(\int_X f\,d\mu\Big)^k-\varepsilon.$$

**Proof.** First, let us assume that $T$ is ergodic. Assume, as we may, that $(X,T)$ is a nilsystem, $X=G/\Gamma$, and that the group $G$ is generated by $T$ and the connected component of the identity.

If $X$ is connected, then for any $b\in\mathbb{N}$ the polynomials $p_1(bn),\dots,p_k(bn)$ are linearly independent, so $Y=X^k$ and

$$\operatorname*{UC-lim}_{n\in\mathbb{Z}^d}\prod_{i=1}^k T^{p_i(bn)}f_i=\prod_{i=1}^k\int_X f_i\,d\mu.$$

Assume now that $X$ is disconnected, and let $X_1^o,\dots,X_a^o$ be its connected components. The nilrotation $T$ permutes the components $X_j^o$; thus there exists $b\in\mathbb{N}$ such that $T^b(X_j^o)=X_j^o$ and $T^{p_i(bn)}(X_j^o)=X_j^o$ for all $j=1,\dots,a$, $i=1,\dots,k$, $n\in\mathbb{Z}^d$. (Note that $b$ may differ from $a$ since the polynomials $p_i$ may have non-integer rational coefficients.) For any $j$, since $X_j^o$ is connected and $T$ is ergodic on $X$, the transformation $T^b$ is ergodic on $X_j^o$. The polynomials $b^{-1}p_i(bn)$, $i=1,\dots,k$, are linearly independent, so for any $j$,

$$\begin{aligned}\operatorname*{UC-lim}_{n\in\mathbb{Z}^d}\int_{X_j^o}\prod_{i=1}^k T^{p_i(bn)}f|_{X_j^o}\,d(a\mu)&=\operatorname*{UC-lim}_{n\in\mathbb{Z}^d}\int_{X_j^o}\prod_{i=1}^k (T^b)^{b^{-1}p_i(bn)}f|_{X_j^o}\,d(a\mu)\\&=\Big(\int_{X_j^o}f|_{X_j^o}\,d(a\mu)\Big)^k.\end{aligned}\tag{3.1}$$

Since the function $t\mapsto t^k$ is convex, we have

$$\sum_{j=1}^a\Big(\int_{X_j^o}f|_{X_j^o}\,d(a\mu)\Big)^k\geq\Big(\int_X f\,d\mu\Big)^k,$$

thus, adding the terms in the left part of (3.1) for $j=1,\dots,a$, we get

$$\operatorname*{UC-lim}_{n\in\mathbb{Z}^d}\int_X\prod_{i=1}^k T^{p_i(n)}f\,d\mu\geq\big(\int_X f\,d\mu\big)^k.\tag{3.2}$$

This proves the assertion in the case $T$ is ergodic; note also that if an integer $b$ works for the construction above, then any multiple of $b$ also works, and so, also gives (3.2).

If $T$ is not ergodic, let $X=\int_\Omega X_\omega dw$, $\mu=\int_\Omega \mu_\omega dw$ be the ergodic decomposition of $X$. For each $\omega\in\Omega$ let $b_\omega$ be the minimal integer satisfying the assertion of the theorem for the system $(X_\omega, T|_{X_\omega})$; clearly, $\omega \mapsto b_\omega$ is a measurable function on $\Omega$. For each $l \in \mathbb{N}$ let $\Omega_l = \{\omega \in \Omega : b_\omega \leq l\}$. Then $\Omega_1 \subseteq \Omega_2 \subseteq \cdots$, and $\bigcup_{l\in\mathbb{N}} \Omega_l = \Omega$; choose $l$ such that $w(\Omega \setminus \Omega_l) < \varepsilon$. Put $b = l!$, so that $b_\omega | b$ for all $\omega \in \Omega_l$. Then

$$\operatorname*{UC-lim}_{n\in\mathbb{Z}^d} \int_{X_\omega} \prod_{i=1}^{k} T^{p_i(bn)} f \, d\mu_\omega > \Bigl(\int_{X_\omega} f \, d\mu_\omega\Bigr)^k - \varepsilon$$

for all $\omega \in \Omega_l$, and so, assuming, as we may, that $|f| \leq 1$ and $\varepsilon < 1$, we get

$$\begin{aligned}\operatorname*{UC-lim}_{n\in\mathbb{Z}^d} \int_X \prod_{i=1}^{k} T^{p_i(bn)} f \, d\mu &\geq \int_{\Omega_l} \operatorname*{UC-lim}_{n\in\mathbb{Z}^d} \int_{X_\omega} \prod_{i=1}^{k} T^{p_i(bn)} f \, d\mu_\omega \, dw \\ &> \int_{\Omega_l} \Bigl(\Bigl(\int_{X_\omega} f \, d\mu_\omega\Bigr)^k - \varepsilon\Bigr) dw \geq \Bigl(\int_{\Omega_l} \int_{X_\omega} f \, d\mu_\omega \, dw\Bigr)^k - \varepsilon \\ &> \Bigl(\int_X f \, d\mu - \varepsilon\Bigr)^k - \varepsilon > \Bigl(\int_X f \, d\mu\Bigr)^k - 2^k \varepsilon.\end{aligned}$$

■

Continuing to deal with a system of linearly independent polynomials, we now get the following result:

**Theorem 3.2.** *Let $(X, \mathcal{B}, \mu, T)$ be an invertible ergodic probability measure preserving system and let $p_1, \ldots, p_k$ be linearly independent polynomials $\mathbb{Z}^d \longrightarrow \mathbb{Z}$ with $p_i(0) = 0$, $i = 1, \ldots, k$. Then for any nonnegative function $f \in L^\infty(X)$ and any $\varepsilon > 0$ the set*

$$R_\varepsilon(f) = \Bigl\{n \in \mathbb{Z}^d : \int_X \prod_{i=1}^{k} T^{p_i(n)} f \, d\mu > \Bigl(\int_X f \, d\mu\Bigr)^k - \varepsilon\Bigr\}.$$

*is AVIP$_0^*$.*

**Proof.** We may assume that $(X, T)$ is a nilsystem: $X = G/\Gamma$, $T \in G$, and $f \in C(X)$. Let us start with the case of connected $X$. Since $p_1, \ldots, p_k$ are linearly independent, we have $Y = X^k$ and thus $K = \Delta G_2^k$. By [L7], Proposition 0.2, if we put $\tau(n) = \int_{T^{p(n)}D} F \, d\mu_{T^{p(n)}(D)}$ and $\tau'(n) = h(T^{p(n)}w)$, $n \in \mathbb{Z}^d$, then for any $\varepsilon > 0$, $d^*\bigl(\{n \in \mathbb{Z}^d : |\tau(n) - \tau'(n)| > \varepsilon|\}\bigr) = 0$. Thus, as explained in the discussion after Theorem 2.1, we only have to check that $h(w) \geq \bigl(\int_X f \, d\mu\bigr)^k$. But

$$h(w) = \int_W F \, d\mu_W = \int_{D_Z} E(f|Z)^{\otimes k} \, d\mu_{D_Z} = \int_Z E(f|Z)^k \, d\mu_Z \geq \Bigl(\int_X f \, d\mu\Bigr)^k,$$

where $D_Z$ is the diagonal of $Z^k$.

Now assume that the nilmanifold $X$ has $a$ components $X_1^o, \ldots, X_a^o$. Let $b \in \mathbb{N}$ be such that $T^{p_i(bn)}(X_j^o) = X_j^o$ for all $j = 1, \ldots, a$, $i = 1, \ldots, k$, $n \in \mathbb{Z}^d$. Then for each $j = 1, \ldots, a$ the set

$$R_j = \Bigl\{n \in \mathbb{Z}^k : \int_{X_j^o} \prod_{i=1}^{k} T^{p_i(bn)} f|_{X_j^o} \, d(a\mu) > \Bigl(\int_{X_j^o} f|_{X_j^o} \, d(a\mu)\Bigr)^k - \varepsilon\Bigr\}$$

is AVIP$^*$. Since the function $t \mapsto t^k$ is convex, and since the intersection of finitely many AVIP$_0^*$-sets is AVIP$_0^*$, after summing these integrals for $j = 1, \ldots, a$ we obtain that the set

$$R = \Big\{ n \in \mathbb{Z}^d : \int_X \prod_{i=1}^k T^{p_i(bn)} f \, d\mu > \Big( \int_X f \, d\mu \Big)^k - \varepsilon \Big\}$$

contains $\bigcap_{j=1}^l R_j$ and thus is AVIP$_0^*$ if all $R_j$ are. Next, it is easy to see that if $R \subseteq \mathbb{Z}^d$ is AVIP$_0^*$, then the set $bR$ is also AVIP$_0^*$; since $R_\varepsilon(f) \supseteq bR$, we get that $R_\varepsilon(f)$ is AVIP$_0^*$. ■

**Remark.** In the proof of Theorem 3.2, we only used the linear independence of the polynomials $p_i$ to get $H \supseteq G_2^k$. Thus, in the situation where the maximal nilsystem factor of a measure preserving system is a Kronecker system, in which case $G_2$ is trivial, the assumption in Theorem 3.2 that the $p_i$ are linearly independent can be dropped:

**Theorem 3.3.** (Cf. [HoKM], Theorem 6.1) *Let $(X, \mathcal{B}, \mu, T)$ be an invertible ergodic probability measure preserving system whose maximal nilsystem factor is a Kronecker system and let $p_1, \ldots, p_k$ be polynomials $\mathbb{Z}^d \longrightarrow \mathbb{Z}$ with $p_i(0) = 0$, $i = 1, \ldots, k$. Then for any nonnegative function $f \in L^\infty(X)$ and any $\varepsilon > 0$ the set*

$$R_\varepsilon(f) = \Big\{ n \in \mathbb{Z}^d : \int_X \prod_{i=1}^k T^{p_i(n)} f \, d\mu > \Big( \int_X f \, d\mu \Big)^k - \varepsilon \Big\}.$$

*is AVIP$_0^*$.*

The situation becomes more subtle when the $p_i$ satisfy some linear relations, that is, when we deal with polynomials of the form $\varphi_1(p_1, \ldots, p_k), \ldots, \varphi_r(p_1, \ldots, p_k)$, where $p_i$ are linearly independent polynomials and $\varphi_j$ are linear forms; the problem is that polynomial relations between $p_i$ also start playing a role. (For example, the systems $\{n, 2n, 3n, n^2\}$ and $\{n, 2n, 3n, n^3\}$ produce different orbit closures of the diagonal of the power of nilmanifolds and so, different Cesàro limits of the corresponding expressions! (See [L6], Example 13.16.)) There is, however, a general principle that sometimes allows one to reduce a polynomial problem to the linear case. We say that polynomials $p_1, \ldots, p_k \colon \mathbb{Z}^d \longrightarrow \mathbb{Z}$ are *algebraically independent up to degree $c$* if for any nonzero polynomial $P \colon \mathbb{Z}^k \longrightarrow \mathbb{Z}$ of degree $\leq c$ the polynomial $P(p_1(n), \ldots, p_k(n))$ is not equal to zero. We have the following:

**Proposition 3.4.** *Let $\varphi_1, \ldots, \varphi_r$ be distinct linear forms $\mathbb{Z}^k \longrightarrow \mathbb{Z}$, let $p_1, \ldots, p_k$ be polynomials $\mathbb{Z}^d \longrightarrow \mathbb{Z}$ with zero constant term and algebraically independent up to degree c, let $X$ be a nilmanifold of nilpotency class $\leq c$ let $T_1, \ldots, T_r$ be nilrotations of $X$, and let $x \in X$. Then for any $b \in \mathbb{N}$ that is divisible by the number of connected component of $X$, the orbit closure of $x$ under the polynomial action $T_1^{\varphi_1(p_1(bn), \ldots, p_k(bn))} \cdots T_r^{\varphi_r(p_1(bn), \ldots, p_k(bn))}$, $n \in \mathbb{Z}^d$, coincides with the orbit closure of $x$ under the action $T_1^{b\varphi_1(m_1, \ldots, m_k)} \cdots T_r^{b\varphi_r(m_1, \ldots, m_k)}$ of $\mathbb{Z}^k$.*

**Proof.** The orbit closure $\overline{\{T_1^{\varphi_1(m)}\cdots T_r^{\varphi_r(m)}x,\ m\in\mathbb{Z}^k\}}$ of $x$ is a subnilmanifold $Y$ of $X$. After replacing $\varphi_i(m)$ by $b\varphi_i(m)$, $i=1,\dots,r$, and $p_j(n)$ by $b^{-1}p_j(bn)$, $j=1,\dots,k$, we may assume that $Y$ is connected.

Let $Y=H/\Lambda$, where $H$ is a nilpotent Lie group of nilpotency class $\le c$ and $\Lambda$ is a lattice in $H$, let $H^o$ be the identity component of $H$, and let $Q$ be the maximal factor-torus of $Y$, $Q=[H^o,H^o]\backslash Y=H^o/(\Lambda[H^o,H^o])$, which, under the action of $H$, has the structure of a $c$-step skew-product system: $Q$ is a product of subtori, $Q=Q_1\times Q_2\times\cdots\times Q_c$, and every element $S$ of $H$ acts on $Q$ by the formula

$$S(z_1,z_2,z_3,\dots,z_c)=\big(z_1+\alpha_1,z_2+\psi_2(z_1)+\alpha_2,z_3+\psi_3(z_1,z_2)+\alpha_3,\dots,\\ z_c+\psi_c(z_1,\dots,z_{c-1})+\alpha_c\big),$$

where the $\psi_i$ are linear forms and $\alpha_i\in Q_i$, $i=1,\dots,c$. In these coordinates, the action $T_1^{\varphi_1(m)}\cdots T_r^{\varphi_r(m)}$ on $Q$ takes the form

$$T_1^{\varphi_1(m)}\cdots T_r^{\varphi_r(m)}(z_1,\dots,z_c)=\big(z_1+q_1(m),z_2+q_2(m,z_1),z_3+p_3(m,z_1,z_2),\dots,\\ z_c+q_c(m,z_1,\dots,z_{c-1})\big),$$

where for each $i=1,\dots,c$, $q_i$ is a polynomial in $m$ of degree $\le i$. Let $z=(z_1,\dots,z_c)$ be the projection of $x$ to $Q$, and let $P_i(m)=q_i(m,z_1,\dots,z_{i-1})$, $i=1,\dots,c$. Since the orbit

$$\{T_1^{\varphi_1(m)}\cdots T_r^{\varphi_r(m)}(z),\ m\in\mathbb{Z}^k\}=\{z_1+P_1(m),z_2+P_2(m),\dots,z_c+P_c(m),\ m\in\mathbb{Z}^k\}$$

is dense in $Q$, the polynomials $P_1,\dots,P_c$ are linearly independent. Now, since the polynomials $p_1,\dots,p_k$ are algebraically independent up to degree $c$ and the polynomials $P_1,\dots,P_c$ have degree $\le c$, the polynomials $P_1(p_1(n),\dots,p_k(n)),\dots,P_c(p_1(n),\dots,p_k(n))$ are also linearly independent. Hence, the orbit

$$\{T_1^{\varphi_1(p_1(n),\dots,p_k(n))}\cdots T_r^{\varphi_r(p_1(n),\dots,p_k(n))}z,\ n\in\mathbb{Z}^d\}=\\ \{z_1+P_1(p_1(n),\dots,p_k(n)),z_2+P_2(p_1(n),\dots,p_k(n)),\dots,z_c+P_c(p_1(n),\dots,p_k(n)),\ n\in\mathbb{Z}^d\}$$

of $z$ is also dense in $Q$. By [L2], Theorem C, this implies that the orbit $\{T_1^{\varphi_1(p_1(n),\dots,p_k(n))}\cdots T_r^{\varphi_r(p_1(n),\dots,p_k(n))}x,\ n\in\mathbb{Z}^d\}$ of $x$ is dense in $Y$, which is what we needed. ∎

Let now $N$ be a connected subnilmanifold of $X$. By [L5], Theorem 1.10, $N$ has a dense subset $J$ such that the orbit closures of the points of $J$ are all congruent (that is, are nil-translations of each other) and their union is dense in the orbit closure of $N$. The same $b$ in Proposition 3.4 works for all points of $J$, and if $N$ is disconnected there is an integer $b$ that works for such dense subsets $J_i$ of every component of $N$. This implies that, under the assumptions of Proposition 3.4, the following is also true:

**Proposition 3.5.** *For any subnilmanifold $N$ of $X$ there exists $b\in\mathbb{N}$ such that the orbit closure of $N$ under the action $T_1^{\varphi_1(p_1(bn),\dots,p_k(bn))}\cdots T_r^{\varphi_r(p_1(bn),\dots,p_k(bn))}$, $n\in\mathbb{Z}^d$, coincides with the orbit closure of $N$ under the (linear) action $T_1^{b\varphi_1(m)}\cdots T_r^{b\varphi_r(m)}$, $m\in\mathbb{Z}^k$.*

Let $(X,T)$ be an ergodic invertible probability measure preserving system, which we will (as we may up to an arbitrarily small error) continue to assume to be a nilsystem. Let $f_1,\dots,f_k\in L^\infty(X)$; as in the proof of Theorem 2.1, the limit

$$\operatorname*{UC-lim}_{n\in\mathbb{Z}^d}\int_X\prod_{i=1}^r T^{\varphi_i(p_1(n),\dots,p_k(n))}f_i\,d\mu$$

is determined by the orbit closure of the diagonal $D$ of $X^r$ under the polynomial action $T_1^{\varphi_1(p_1(n),\dots,p_k(n))}\cdots T_d^{\varphi_r(p_1(n),\dots,p_k(n))}$, $n\in\mathbb{Z}^d$. (See, for example, [L6], Section 0.7.) Applying Proposition 3.5 to this action, we see that if $p_1,\dots,p_k$ are algebraically independent up to degree $c+1$, then there exists $b\in\mathbb{N}$ such that the orbit closure of the diagonal $D_{c+1}$ of $X_{c+1}^r$ under the action $\bigl(T^{\varphi_1(p_1(bn),\dots,p_k(bn))},\dots,T^{\varphi_r(p_1(bn),\dots,p_k(bn))}\bigr)$, $n\in\mathbb{Z}^d$, coincides with the orbit closure of $D_{c+1}$ under the action $\bigl(T^{b\varphi_1(m)},\dots,T^{b\varphi_r(m)}\bigr)$, $m\in\mathbb{Z}^k$. This means that the complexity of the system $\bigl\{\varphi_1(p_1(bn),\dots,p_k(bn)),\dots,\varphi_r(p_1(bn),\dots,p_k(bn))\bigr\}$ is $\le c$, and that

$$\operatorname*{UC-lim}_{n\in\mathbb{Z}^d}\int_X\prod_{i=1}^r T^{\varphi_i(p_1(bn),\dots,p_k(bn))}f_i\,d\mu=\operatorname*{UC-lim}_{m\in\mathbb{Z}^k}\int_X\prod_{i=1}^r T^{b\varphi_i(m)}f_i\,d\mu.$$

Using the ergodic decomposition and taking $b$ "divisible enough" so that it works for most components of this decomposition, we may, as in the proof of Theorem 3.1, get rid of the assumption of ergodicity of $T$ to obtain the following result:

**Proposition 3.6.** *Let $(X,\mathcal{B},\mu,T)$ be an invertible probability measure preserving system, let $\{\varphi_1,\dots,\varphi_r\}$ be a system of linear forms $\mathbb{Z}^k\longrightarrow\mathbb{Z}$ of complexity $c$, and let $p_1,\dots,p_k$ be polynomials $\mathbb{Z}^d\longrightarrow\mathbb{Z}$ with zero constant term that are algebraically independent up to degree $c+1$. Then for any $f_1,\dots,f_r\in L^\infty(X)$ and any $\varepsilon>0$ there exists $b\in\mathbb{N}$ such that*

$$\Bigl|\operatorname*{UC-lim}_{n\in\mathbb{Z}^d}\int_X\prod_{i=1}^r T^{\varphi_i(p_1(bn),\dots,p_k(bn))}f_i\,d\mu-\operatorname*{UC-lim}_{m\in\mathbb{Z}^k}\int_X\prod_{i=1}^r T^{b\varphi_i(m)}f_i\,d\mu\Bigr|<\varepsilon.$$

The complexity of the "$k$-dimensional cubic" system of linear forms is $k-1$ (see [BHoK] or [L6], Example 4 in Section 6.7). Therefore, as a special case of Proposition 3.6 we obtain:

**Theorem 3.7.** *Let $(X,\mathcal{B},\mu,T)$ be an invertible probability measure preserving system and let polynomials $p_1,\dots,p_k\colon\mathbb{Z}^d\longrightarrow\mathbb{Z}$ with zero constant term be algebraically independent up to degree $k$. Then for any nonnegative function $f\in L^\infty(X)$ and any $\varepsilon>0$ there exists $b\in\mathbb{N}$ such that*

$$\operatorname*{UC-lim}_{n\in\mathbb{Z}^d}\int_X\prod_{e_1,\dots,e_k\in\{0,1\}}T^{e_1p_1(bn)+\cdots+e_kp_k(bn)}f\,d\mu>\Bigl(\int f\,d\mu\Bigr)^{2^k}-\varepsilon.$$

As a corollary, we obtain that for any $\varepsilon > 0$, the set

$$R_\varepsilon(f; p_1, \dots, p_k) = \Big\{ n \in \mathbb{Z}^d : \int_X \prod_{e_1,\dots,e_k \in \{0,1\}} T^{e_1 p_1(n) + \dots + e_k p_k(n)} f \, d\mu > \Big( \int f \, d\mu \Big)^{2^k} - \varepsilon \Big\}$$

is syndetic in $\mathbb{Z}^d$. If $T$ is ergodic, this set is, actually, an AVIP$_0^*$-set. Indeed, as in the proof of Theorem 2.1, whether the set

$$\Big\{ n \in \mathbb{Z}^d : \int_X \prod_{i=1}^r T^{\varphi_i(p_1(n),\dots,p_k(n))} f \, d\mu > \Big( \int_X f \, d\mu \Big)^r - \varepsilon \Big\}$$

is AVIP$^*$ or not is determined (assuming that $(X, T)$ is a nilsystem) exclusively by the closure $Y$ of the orbit

$$\bigcup_{n \in \mathbb{Z}^d} \big( T^{\varphi_1(p_1(n),\dots,p_k(n))}, \dots, T^{\varphi_d(p_1(n),\dots,p_k(n))} \big) D$$

of the diagonal $D$ of $X^r$. If the complexity of the system $\{\varphi_1, \dots, \varphi_r\}$ is $c$, then the nilpotent class of $X$ may be assumed to be equal to $c+1$. Now, if the polynomials $p_i$ are algebraically independent up to degree $c+1$, then, as discussed above, for some $b \in \mathbb{N}$, the orbit closure of $D$ under the action $\big( T^{\varphi_1(p_1(bn),\dots,p_k(bn))}, \dots, T^{\varphi_r(p_1(bn),\dots,p_k(bn))} \big)$ becomes connected and coincides with the orbit closure of $D$ with respect to the action $\big( T^{b\varphi_1(m_1,\dots,m_k)}, \dots, T^{b\varphi_d(m_1,\dots,m_k)} \big)$. The transformation $T^b$ may not be ergodic on $X$, but in this case $X$ consists of several ergodic components $X_1^o, \dots, X_a^o$, and for each $j$, if the set

$$\Big\{ m \in \mathbb{Z}^k : \int_{X_j^o} \prod_{i=1}^r T^{b\varphi_i(m_1,\dots,m_k)} f|_{X_j^o} \, d(a\mu) > \Big( \int_{X_j^o} f|_{X_j^o} \, d(a\mu) \Big)^r - \varepsilon \Big\}$$

is AVIP$^*$, then the set

$$R_j = \Big\{ n \in \mathbb{Z}^d : \int_{X_j^o} \prod_{i=1}^d T^{b\varphi_i(p_1(n),\dots,p_k(n))} f|_{X_j^o} \, d(a\mu) > \Big( \int_{X_j^o} f|_{X_j^o} \, d(a\mu) \Big)^r - \varepsilon \Big\}$$

is also AVIP$^*$. Since the function $t \mapsto t^r$ is convex, and since the intersection of finitely many AVIP$_0^*$-sets is AVIP$_0^*$, after summing these integrals for $j = 1, \dots, a$ we obtain that the set

$$R = \Big\{ n \in \mathbb{Z}^d : \int_X \prod_{i=1}^r T^{\varphi_i(p_1(bn),\dots,p_k(bn))} f \, d\mu > \Big( \int_X f \, d\mu \Big)^r - \varepsilon \Big\}$$

contains $\bigcap_{j=1}^l R_j$ and thus is AVIP$_0^*$ if $R_j$ are. Next, if $R \subseteq \mathbb{Z}^d$ is AVIP$_0^*$, then the set $bR$ is also AVIP$_0^*$; it follows that the set

$$\Big\{ n \in \mathbb{Z}^d : \int_X \prod_{i=1}^r T^{\varphi_i(p_1(n),\dots,p_k(n))} f \, d\mu > \Big( \int_X f \, d\mu \Big)^r - \varepsilon \Big\}$$

is AVIP$_0^*$. Returning to the case of a general system $(X, T)$, we obtain the following principle:

**Proposition 3.8.** *Let $\{\varphi_1, \dots, \varphi_r\}$ be a system of linear forms $\mathbb{Z}^k \longrightarrow \mathbb{Z}$ of complexity $c$ such that for any ergodic invertible probability measure preserving system $(X, \mathcal{B}, \mu, T)$ the set*

$$\Big\{m \in \mathbb{Z}^k : \int_X \prod_{i=1}^r T^{\varphi_i(m)} f\, d\mu > \Big(\int_X f\, d\mu\Big)^r - \varepsilon\Big\}$$

*is AVIP$_0^*$. Then for any ergodic invertible probability measure preserving system $(X, \mathcal{B}, \mu, T)$, any polynomials $p_1, \dots, p_k \colon \mathbb{Z}^d \longrightarrow \mathbb{Z}$ with zero constant term and algebraically independent up to degree $c+1$, any nonnegative function $f \in L^\infty(X)$, and any $\varepsilon > 0$ the set*

$$\Big\{n \in \mathbb{Z}^d : \int_X \prod_{i=1}^r T^{\varphi_i(p_1(n), \dots, p_k(n))} f\, d\mu > \Big(\int_X f\, d\mu\Big)^r - \varepsilon\Big\}$$

*is AVIP$_0^*$ as well.*

As a special case, we get

**Theorem 3.9.** *Let $(X, \mathcal{B}, \mu, T)$ be an invertible ergodic probability measure preserving system and let polynomials $p_1, \dots, p_k \colon \mathbb{Z}^d \longrightarrow \mathbb{Z}$ with zero constant term be algebraically independent up to degree $k$. Then for any nonnegative function $f \in L^\infty(X)$ and any $\varepsilon > 0$ the set*

$$R_\varepsilon(f; p_1, \dots, p_k) = \Big\{n \in \mathbb{Z}^d : \int_X \prod_{e_1, \dots, e_k \in \{0,1\}} T^{e_1 p_1(n) + \dots + e_k p_k(n)} f\, d\mu > \Big(\int f\, d\mu\Big)^{2^k} - \varepsilon\Big\} \tag{3.3}$$

*is AVIP$_0^*$.*

Taking $f = 1_A$ in (3.3), where $A$ is a measurable subset of $X$, and utilizing Furstenberg's correspondence principle (see footnote 2 in the introduction), we obtain Theorem 0.8 and Corollary 0.9.

We conclude this paper with addressing the situation where the polynomials $p_1, \dots, p_k \colon \mathbb{Z}^d \longrightarrow \mathbb{Z}$ are not assumed to have zero constant term but are jointly intersective (see footnote 4 in the introduction). In this case, for any point $x$ of a nilmanifold $X$, any translations $T_1, \dots, T_k$ of $X$ and any $b \in \mathbb{N}$ there exists $m \in \mathbb{Z}^d$ such that the orbit closure of $x$ under the action $T_1^{p_1(bn+m)} \cdots T_k^{p_k(bn+m)}$, $n \in \mathbb{Z}^d$, is the same as for the action $T_1^{\hat{p}_1(bn)} \cdots T_k^{\hat{p}_k(bn)}$, $n \in \mathbb{Z}^d$, where $\hat{p}_i = p_i - p_i(0)$, $i = 1, \dots, k$. (This fact is a corollary of Proposition 2.4 from [BLLe2].) This allows us to extend Theorem 3.7 in the following way:

**Theorem 3.10.** *Let $(X, \mathcal{B}, \mu, T)$ be an invertible probability measure preserving system and let polynomials $p_1, \dots, p_k \colon \mathbb{Z}^d \longrightarrow \mathbb{Z}$ be jointly intersective and algebraically independent up to degree $k$. Then*
(i) *for any nonnegative function $f \in L^\infty(X)$ and any $\varepsilon > 0$ there exist $b \in \mathbb{N}$ and $m \in \mathbb{Z}^d$ such that*

$$\operatorname*{UC-lim}_{n \in \mathbb{Z}} \int_X \prod_{e_1, \dots, e_k \in \{0,1\}} T^{e_1 p_1(bn+m) + \dots + e_k p_k(bn+m)} f\, d\mu > \Big(\int f\, d\mu\Big)^{2^k} - \varepsilon;$$

(ii) *for any nonnegative function* $f \in L^\infty(X)$ *and any* $\varepsilon > 0$*, the set* $R_\varepsilon(f; p_1, \ldots, p_k)$ *is* $AVIP^*_{0,+}$.

Specializing to $f = 1_A$, where $A$ is a measurable subset of $X$, and utilizing Furstenberg's correspondence principle, we obtain Theorem 0.11 and Corollary 0.12.

**Acknowledgment.** We thank Donald Robertson for helpful remarks on the preliminary draft of this paper.

## Bibliography


[AGH] L. Auslander, L. Green and F. Hahn, *Flows on homogeneous spaces*, Annals of Math. Studies, vol. 53, Princeton Univ. Press, 1963.

[B1] V. Bergelson, Ergodic Ramsey Theory – an update, *Ergodic Theory of* $\mathbb{Z}^d$*-actions,* London Math. Soc. Lecture Note Ser. **228** (1996), 1-61.

[B2] V. Bergelson, The multifarious Poincaré recurrence theorem, *Descriptive set theory and dynamical systems,* London Math. Soc. Lecture Note Ser. **277** (2000), 31-57.

[B3] V. Bergelson, Minimal idempotents and ergodic Ramsey theory, *Topics in Dynamics and Ergodic Theory,* London Math. Soc. Lecture Note Ser. **310** (2003), 8-39.

[B4] V. Bergelson, Ergodic Ramsey theory, *Cont. Math.* **65** (1987), 63-87.

[BD] V. Bergelson and T. Downarowicz, Large sets of integers and hierarchy of mixing properties of measure preserving systems, *Colloq. Math.* **110** (2008), no. 1, 117-150.

[BFM] V. Bergelson, H. Furstenberg, and R. McCutcheon, IP-sets and polynomial recurrence, *Ergodic Theory and Dynam. Systems* **16** (1996), 963-974.

[BFW] V. Bergelson, H. Furstenberg, and B. Weiss, Piecewise-Bohr sets of integers and combinatorial number theory, *Topics in discrete mathematics*, 13-37, Algorithms Combin., 26, Springer, Berlin, 2006.

[BHoK] V. Bergelson, B. Host, and B. Kra, Multiple recurrence and nilsequences. With an appendix by I. Ruzsa, *Inventiones Math.* **160** (2005), no. 2, 261-303.

[BL1] V. Bergelson and A. Leibman, Set-polynomials and polynomial extension of the Hales-Jewett theorem, *Annals of Mathematics* **150** (1999), 33-75.

[BL2] V. Bergelson and A. Leibman, Cubic averages and large intersections, *Contemp. Math.* **631** (2015), 5-20.

[BL3] V. Bergelson and A. Leibman, $\mathrm{IP}^*_r$-recurrence and nilsystems, preprint. Available on arXiv.

[BLLe1] V. Bergelson, A. Leibman, and E. Lesigne, Weyl complexity of a system of polynomials and constructions in combinatorial number theory, *Journal D'Anal. Mathematique* **103** (2007), 47-92.

[BLLe2] V. Bergelson, A. Leibman, and E. Lesigne, Intersective polynomials and the polynomial Szemerédi theorem, *Advances in Math.* **219** (2008), 369-388.

[BM1] V. Bergelson and R. McCutcheon, Central sets and a non-commutative Roth theorem, *Amer. J. Math.* **129** (2007), no. 5, 1251-1275.

[BM2] V. Bergelson and R. McCutcheon, Idempotent ultrafilters, multiple weak mixing and Szemerdi's theorem for generalized polynomials, *J. Anal. Math.* **111** (2010), 77-130.

[F] N. Frantzikinakis, Multiple correlation sequences and nilsequences, preprint. Available at arXiv:1407.0631

[FK] N. Frantzikinakis and B. Kra, Polynomial averages converge to the product of integrals, *Israel J. Math.* **148** (2005), 267-276.

[Fu] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton Univ. Press, 1981.

[FuW] H. Furstenberg and B. Weiss, Topological dynamics and combinatorial number theory, *J. D'Anal. Math.* **34** (1978), 61-85.

[G] J. Gillis, Note on a Property of Measurable Sets, *J. London Math. Soc.* **11** (1936), no. 2, 139-141.

[H] N. Hindman, Finite sums from sequences within cells of a partition of $\mathbb{N}$, *J. Comb. Theory (A)* **17** (1974), 1-11.

[HoK1] B. Host and B. Kra, Non-conventional ergodic averages and nilmanifolds, *Annals of Math.* **161** (2005), no. 1, 397-488.

[HoK2] B. Host and B. Kra, Averaging along cubes, *Modern dynamical systems and applications*, Cambridge University Press, 2004.

[HoK3] B. Host and B. Kra, Nil-Bohr sets of integers, *Ergodic Theory and Dynam. Systems* **31** (2011), 113-142.

[HoKM] B. Host, B. Kra, and A. Maass, Complexity of nilsystems and systems lacking nilfactors, *J. Anal. Math.* **124** (2014), 261-295.

[Ke1] H. B. Keynes, Topological dynamics in coset transformation groups, *Bull. Amer. Math. Soc.* **72** (1966), 1033-1035.

[Ke2] H. B. Keynes, A study of the proximal relation in coset transformation groups, *Trans. Amer. Math. Soc.* **128** (1967), 389-402.

[Kh] A. Y. Khintchine, Eine Verschärfung des Poincaréschen "Wiederkehrsatzes", *Comp. Math.* **1** (1934), 177-179.

[L1] A. Leibman, Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold, *Ergodic Theory and Dynam. Systems* **25** (2005), 201-213.

[L2] A. Leibman, Pointwise convergence of ergodic averages for polynomial actions of $\mathbb{Z}^d$ by translations on a nilmanifold, *Ergodic Theory and Dynam. Systems* **25** (2005), 215-225.

[L3] A. Leibman, Convergence of multiple ergodic averages along polynomials of several variables, *Israel J. of Math.* **146** (2005), 303-315.

[L4] A. Leibman, Rational sub-nilmanifolds of a compact nilmanifold, *Ergodic Theory and Dynamical Systems* **26** (2006), 787-798.

[L5] A. Leibman, Orbits on a nilmanifold under the action of a polynomial sequence of translations, *Ergodic Theory and Dynam. Systems* **27** (2007), 1239-1252.

[L6] A. Leibman, Orbit of the diagonal in the power of a nilmanifold, *Trans. of AMS* **362** (2010), 1619-1658.

[L7] A. Leibman, Nilsequences, nul-sequences, and the multiple correlation sequences, *Ergodic*

*Theory and Dynam. Systems* **31** (2015), 176-191.

[M] R. McCutcheon, private communications.

[MZ] R. McCutcheon and J. Zhou, D sets and IP rich sets in $\mathbb{Z}$, to appear in *Fundamenta Mathematicae*.

[Zi] T. Ziegler, Universal characteristic factors and Furstenberg averages, *J. of AMS* **20** (2007), 53-97.