ERGODIC AVERAGES FOR INDEPENDENT POLYNOMIALS AND APPLICATIONS

NIKOS FRANTZIKINAKIS AND BRYNA KRA

Abstract. Szemerédi's Theorem states that a set of integers with positive upper density contains arbitrarily long arithmetic progressions. Bergelson and Leibman generalized this, showing that sets of integers with positive upper density contain arbitrarily long polynomial configurations; Szemerédi's Theorem corresponds to the linear case of the polynomial theorem. We focus on the case farthest from the linear case, that of rationally independent polynomials. We derive results in ergodic theory and in combinatorics for rationally independent polynomials, showing that their behavior differs sharply from the general situation.

1. Introduction and results in ergodic theory

1.1. **Background.** The celebrated theorem of Szemerédi [12] states that a subset of the integers with positive upper density[1] contains arbitrarily long arithmetic progressions. Furstenberg [4] drew the deep connection between combinatorial questions and ergodic theory, showing that Szemerédi's Theorem follows from an ergodic theorem, now known as the multiple recurrence theorem.

A natural question is to find other configurations that must occur in subsets of the integers with positive upper density. Furstenberg [5] and Sárközy [11] independently proved that if $\Lambda \subset \mathbb{N}$ has positive upper density and $p(n)$ is an *integer polynomial*, meaning it takes integer values on the integers, and if $p(0) = 0$, then there exist $x, y \in \Lambda$ such that $x - y = p(n)$ for some $n \in \mathbb{N}$. Bergelson and Leibman established a far reaching generalization of this result. They showed that if $\Lambda \subset \mathbb{N}$ has positive upper density and $p_1, \ldots, p_k$ are integer polynomials with $p_i(0) = 0$ for $i = 1, \ldots, k$, then there exists $n \in \mathbb{N}$ such that

$$\bar{d}\big(\Lambda \cap (\Lambda + p_1(n)) \cap \cdots \cap (\Lambda + p_k(n))\big) > 0 \ . \tag{1}$$

As with Furstenberg's proof of Szemerédi's Theorem, the Polynomial Szemerédi Theorem follows from an ergodic theorem:

2000 *Mathematics Subject Classification.* Primary: 37A45; Secondary: 37A30, 28D05.

Key words and phrases. Multiple recurrence, multiple ergodic averages, polynomial Szemerédi.

The first author was partially supported by NSF grant DMS-0111298 and the second author by NSF grant DMS-0244994.

[1] If $\Lambda \subset \mathbb{N}$ we define the *upper density* $\bar{d}(\Lambda) = \limsup_{N\to\infty} |\Lambda \cap [1, N]|/N$.

Polynomial Szemerédi Theorem (Bergelson and Leibman [2]). *Let $(X, \mathcal{X}, \mu, T)$ be an invertible measure preserving system and let $p_1, \ldots, p_k$ be integer polynomials with $p_i(0) = 0$ for $i = 1, \ldots, k$. If $A \in \mathcal{X}$ with $\mu(A) > 0$, then*

$$\liminf_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} \mu\big(A \cap T^{p_1(n)}A \cap \ldots \cap T^{p_k(n)}A\big) > 0 \ . \tag{2}$$

Szemerédi's Theorem (and the ergodic theoretic proof by Furstenberg) corresponds to the case that all the polynomials are linear. We focus on the opposite case of *rationally independent* integer polynomials, meaning a set of integer polynomials such that every nontrivial integer combination of the polynomials is not constant. In some sense, this case is typical, since a generic family of integer polynomials is rationally independent. A particular example is any set of polynomials with pairwise distinct degrees. We prove several results, some ergodic and some combinatorial, for families of rationally independent integer polynomials, focusing on the difference between this case and that of a family of linear integer polynomials.

1.2. **Ergodic Results.** Studying the limiting behavior of the multiple ergodic averages associated with (2) has been a central topic in ergodic theory. Very recently, using methods from [7] convergence was established for totally ergodic systems in [8] and for general systems in [10] . The basic approach is to find an appropriate factor system, called a *characteristic factor*, that controls the limiting behavior as $N - M \to \infty$ in $L^2(\mu)$ of the averages

$$\frac{1}{N-M}\sum_{n=M}^{N-1} T^{p_1(n)}f_1 \cdot \ldots \cdot T^{p_k(n)}f_k \ . \tag{3}$$

A characteristic factor is a factor such that the limit of the averages remains unchanged when each function is replaced by its projection on this factor. The next step is to obtain a concrete description for some well chosen characteristic factor in order to prove convergence. For general polynomials, such a characteristic factor can be described as an inverse limit of nilsystems (defined in Section 3.1). We show that characteristic factors for rationally independent integer polynomials have a significantly simpler structure. In particular, in Section 3.2 we show that a characteristic factor for rationally independent polynomials can be chosen to be an inverse limit of rotations on finite abelian groups:

Theorem 1.1. *Let $(X, \mathcal{X}, \mu, T)$ be an ergodic invertible measure preserving system and $p_1, \ldots, p_k$ be rationally independent integer polynomials. Then the rational Kronecker factor $\mathcal{K}_{rat}$ (defined in Section 3.1) is a characteristic factor for the $L^2(\mu)$-convergence of the averages* (3)*, meaning that if $f_1, \ldots, f_k \in L^\infty(\mu)$, the difference*

$$\frac{1}{N-M}\sum_{n=M}^{N-1} T^{p_1(n)}f_1 \cdot \ldots \cdot T^{p_k(n)}f_k - \frac{1}{N-M}\sum_{n=M}^{N-1} T^{p_1(n)}\tilde{f}_1 \cdot \ldots \cdot T^{p_k(n)}\tilde{f}_k \ , \tag{4}$$

where $\tilde{f}_i = \mathbb{E}(f_i \,|\, \mathcal{K}_{rat})$*,* $i = 1, \ldots, k$*, converges to* 0 *in* $L^2(\mu)$ *as* $N - M \to \infty$*.*

For a given measure preserving system $(X, \mathcal{B}, \mu, T)$ and functions $f_0, f_1, \ldots, f_k \in L^\infty(\mu)$, it was shown in [1] that the multicorrelation sequence

$$a_n = \int f_0 \cdot T^n f_1 \cdot \ldots \cdot T^{kn} f_k \; d\mu$$

can be decomposed as a sum of a k-step nilsequence and a sequence that converges to zero in uniform density (all notions defined in Section 4). We note that the original statement in [1] is for $f_0 = f_1 = \ldots = f_k$, but the same proof holds for different functions. Using Theorem 1.1 we prove an analogous result for the multicorrelation sequence of independent polynomial iterates. Moreover, in Section 4 we show that a significantly simpler class of nilsequences suffices for the decomposition:

Theorem 1.2. *Let* $(X, \mathcal{X}, \mu, T)$ *be an invertible ergodic measure preserving system and let* $p_1, \ldots, p_k$ *be rationally independent integer polynomials with highest degree* d*. If* f_0*,* f_1*, …,* $f_k \in L^\infty(\mu)$*,* $n \in \mathbb{N}$ *and*

$$a_n = \int f_0 \cdot T^{p_1(n)} f_1 \cdot \ldots \cdot T^{p_k(n)} f_k \; d\mu \ ,$$

then $\{a_n\}_{n\in\mathbb{N}}$ *is the sum of a* d*-step nilsequence and a sequence that converges to zero in uniform density. Moreover, the* d*-step nilsequence can be chosen to be of the form* $b_n = \phi(S^n e)$*, where* $S\colon G^d \to G^d$ *is a unipotent affine transformation,* G *is a compact abelian group,* $\phi\colon G^d \to \mathbb{C}$ *is continuous, and* e *is the identity element of* G^d*.*

We also use Theorem 1.1 to prove a multiple recurrence result. We show that for a family of rationally independent integer polynomials, the measure of the intersection in (2) is as large as possible “frequently.” More precisely, a set $\Lambda \subset \mathbb{N}$ is *syndetic* if there exists $M \in \mathbb{N}$ such that every interval of length greater than M intersects Λ nontrivially. In Section 3.3 we show:

Theorem 1.3. *Let* $(X, \mathcal{X}, \mu, T)$ *be an invertible measure preserving system,* $p_1, \ldots, p_k$ *be rationally independent integer polynomials with* $p_i(0) = 0$ *for* $i = 1, \ldots, k$*, and* $A \in \mathcal{X}$*. Then for every* $\varepsilon > 0$*, the set*

$$\left\{ n \in \mathbb{N}\colon \quad \mu\left(A \cap T^{p_1(n)} A \cap \cdots \cap T^{p_k(n)} A\right) \geq \mu(A)^{k+1} - \varepsilon \right\}$$

is syndetic.

We stress that we do not assume ergodicity for this result. This sharply contrasts the behavior of a family of linear integer polynomials. For example when $p_i(n) = in$ for $i = 1, \ldots, k$, it was shown in [1] that the analogous result fails for certain ergodic transformations when $k \geq 4$ and also fails for certain nonergodic transformations when $k \geq 2$.

2. Combinatorial Results

Furstenberg [4] established the connection between combinatorial number theory and ergodic theory, showing that regularity properties of subsets of integers with positive density correspond to multiple recurrence properties of measure preserving systems. This is reflected in what has become known as the Correspondence Principle (first introduced in [4] and given in the form below in [2]):

Furstenberg's Correspondence Principle . *Let $\Lambda \subset \mathbb{N}$. There exist a measure preserving system $(X, \mathcal{X}, \mu, T)$ and $A \in \mathcal{X}$ such that $\mu(A) = \bar{d}(\Lambda)$ and*

$$\bar{d}\big(\Lambda \cap (\Lambda + n_1) \cap \cdots \cap (\Lambda + n_m)\big) \geq \mu(A \cap T^{n_1}A \cap \cdots \cap T^{n_r}A)$$

for all $r \in \mathbb{N}$ and all $n_1, \ldots, n_r \in \mathbb{Z}$.

As an immediate corollary of Theorem 1.3 and Furstenberg's Correspondence Principle, for rational independent polynomials we have tight lower bounds for the upper densities in (1) for every $k \in \mathbb{N}$. This result is known to be false for $k \geq 4$ linear polynomials (see [1]):

Theorem 2.1. *Let $\Lambda \subset \mathbb{N}$ and $p_1, \ldots, p_k$ be rationally independent integer polynomials with $p_i(0) = 0$ for $i = 1, \ldots, k$. Then for every $\varepsilon > 0$, the set*

$$\{n \in \mathbb{N}\colon \quad \bar{d}\big(\Lambda \cap (\Lambda + p_1(n)) \cap \cdots \cap (\Lambda + p_k(n))\big) \geq \bar{d}(\Lambda)^{k+1} - \varepsilon\} \tag{5}$$

is syndetic.

We give an example to show that the lower bounds given in (5) are tight. A set $\Lambda \subset \mathbb{N}$ is called normal if its indicator function $\mathbf{1}_\Lambda$ contains every string of zeros and ones of length k with frequency 2^{-k}. For any such set Λ we have that

$$\bar{d}\big(\Lambda \cap (\Lambda + n_1) \cap \cdots \cap (\Lambda + n_k)\big) = \bar{d}(\Lambda)^{k+1} = 1/2^{k+1}$$

for all choices of nonzero distinct integers $n_1, \ldots, n_k$, meaning that (5) cannot be improved.

We remark that Furstenberg's correspondence Principle and, as a consequence, Theorem 2.1 hold if one replaces the upper density $\bar{d}$ with the upper Banach density d^* defined by $d^*(\Lambda) = \lim_{N\to\infty} \sup_{M\in\mathbb{N}} |\Lambda \cap [M, M+N)|/N$ (the limit exists by subadditivity).

Szemerédi's Theorem has the following finite version: given a length k of a progression and density $\delta > 0$, there exists some $N(k, \delta)$ such that for all $N \geq N(k, \delta)$, any subset of $\{1, \ldots, N\}$ having at least δN elements contains an arithmetic progression of length k. In [1], the authors asked if one can strengthen this to showing that for all $k \in \mathbb{N}$, $\delta > 0$ and $\varepsilon > 0$, there exists $N(k, \varepsilon, \delta)$ such that for all $N \geq N(k, \varepsilon, \delta)$, any subset of $\{1, \ldots, N\}$ with at least δN elements contains at least $(1-\varepsilon)\delta^k N$ arithmetic progressions of length k with the same common difference. Their results show that the answer is no for $k \geq 5$ and they show that a weaker condition holds for $k = 3$ and $k = 4$. Green [6] answered the (stronger) question affirmatively for $k = 3$ and $k = 4$ remains open. Given Theorem 2.1, it is natural to ask whether a similar result holds for independent polynomial configurations. We show that this is the case:

Theorem 2.2. *Let $p_1, \ldots, p_k$ be rationally independent integer polynomials with $p_i(0) = 0$ for $i = 1, \ldots, k$. For every $\delta > 0$ and $\varepsilon > 0$ there exists $N(\varepsilon, \delta)$, such that for all $N > N(\varepsilon, \delta)$, any integer subset $\Lambda \subset [1, N]$ with $|\Lambda| \geq \delta N$ contains at least $(1-\varepsilon)\delta^{k+1}N$ configurations of the form $\{x, x + p_1(n), \ldots, x + p_k(n)\}$ for some fixed $n \in \mathbb{N}$.*

Proof. Suppose that the result fails. Then there exist $\delta_0, \varepsilon_0 > 0$, an integer sequence $N_m \to \infty$, and integer subsets $\Lambda_m \subset [1, N_m]$ such that

$$|\Lambda_m| \geq \delta_0 N_m \tag{6}$$

and

$$|\Lambda_m \cap (\Lambda_m + p_1(n)) \cap \cdots \cap (\Lambda_m + p_k(n))| < (1 - \varepsilon_0)\delta_0^{k+1} N_m \tag{7}$$

for every $m, n \in \mathbb{N}$. We construct a measure preserving system that has bad recurrence properties and then obtain a contradiction from Theorem 1.3.

For $m \in \mathbb{N}$ set $\Lambda_m^0 = \Lambda_m^c$ and $\Lambda_m^1 = \Lambda_m$. Using a diagonal argument we can find a subsequence of $\{N_m\}_{m\in\mathbb{N}}$, which for convenience we call again $\{N_m\}_{m\in\mathbb{N}}$, such that the limit

$$\lim_{m\to\infty} \frac{|(\Lambda_m^{i_1} + n_1) \cap (\Lambda_m^{i_2} + n_2) \cap \cdots \cap (\Lambda_m^{i_r} + n_r) \cap [1, N_m]|}{N_m}$$

exists for every $r \in \mathbb{N}$, $n_1, \ldots, n_r \in \mathbb{Z}$, and $i_1, \ldots, i_r \in \{0, 1\}$.

On the sequence space $(X = \{0,1\}^{\mathbb{Z}}, \mathcal{X})$, where $\mathcal{X}$ is the Borel σ-algebra, we define a measure μ on cylinder sets as follows:

$$\mu(\{x_{n_1} = i_1, x_{n_2} = i_2, \ldots, x_{n_r} = i_r\}) = \lim_{m\to\infty} \frac{|(\Lambda_m^{i_1} + n_1) \cap (\Lambda_m^{i_2} + n_2) \cap \cdots \cap (\Lambda_m^{i_r} + n_r) \cap [1, N_m]|}{N_m}$$

where $n_1, n_2, \ldots, n_r \in \mathbb{Z}$, and $i_1, i_2, \ldots, i_r \in \{0, 1\}$. The finite dimensional statistics are consistent and so we can extend this to a probability measure on $\mathcal{X}$ using Kolmogorov's Extension Theorem. Then the shift transformation T defined by

$$T(\{x(j)\}_{j\in\mathbb{Z}}) = \{x(j+1)\}_{j\in\mathbb{Z}}$$

preserves the measure μ and gives rise to a measure preserving system $(X, \mathcal{X}, \mu, T)$. If $A = \{x \colon x(0) = 1\}$, using the definition of μ we see that

$$\begin{aligned}\mu(A \cap T^{p_1(n)}A \cap \cdots \cap T^{p_k(n)}A) &= \mu(\{x_0 = 1, x_{p_1(n)} = 1, \ldots, x_{p_k(n)} = 1\}) \\ &= \lim_{m\to\infty} \frac{|\Lambda_m \cap (\Lambda_m + p_1(n)) \cap \cdots \cap (\Lambda_m + p_k(n))|}{N_m},\end{aligned}$$

for every $n \in \mathbb{N}$. Combining this with (6) and (7) we find that

$$\mu(A \cap T^{p_1(n)}A \cap \cdots \cap T^{p_k(n)}A)) \leq (1-\varepsilon_0)\delta_0^{k+1} \leq (1-\varepsilon_0)\mu(A)^{k+1} \tag{8}$$

for all $n \in \mathbb{N}$. This contradicts Theorem 1.3 and completes the proof. □

3. Characteristic factors and multiple recurrence result

3.1. **Preliminaries.** By a measure preserving system we mean a quadruple $(X, \mathcal{X}, \mu, T)$, where $(X, \mathcal{X}, \mu)$ is a probability space and $T: X \to X$ is a measurable map such that $\mu(T^{-1}A) = \mu(A)$ for all $A \in \mathcal{X}$. Without loss of generality we can assume that the probability space is Lebesgue. A *factor* of the measure preserving system $(X, \mathcal{X}, \mu, T)$ can be defined in any of the following three equivalent ways: it is a T-invariant sub-σ-algebra $\mathcal{D}$ of $\mathcal{X}$, it is a T-invariant sub-algebra $\mathcal{F}$ of $L^\infty(X)$, or it is a system $(Y, \mathcal{Y}, \nu, S)$ and a measurable map $\pi: X' \to Y'$, where X' is a T-invariant set and Y' is an S-invariant set of full measure, such that $\mu \circ \pi^{-1} = \nu$ and $S \circ \pi(x) = \pi \circ T(x)$ for $x \in X'$. By setting $\mathcal{F} = L^\infty(\mathcal{D})$, we see that the first definition implies the second. Conversely, given $\mathcal{F}$ we define $\mathcal{D}$ to be the σ-algebra generated by $\mathcal{F}$-measurable sets. The equivalence between the first and third definition is seen by identifying $\mathcal{D}$ with $\pi^{-1}(\mathcal{Y})$. In a slight abuse of terminology, when any of these conditions holds, we say that Y (or the appropriate σ-algebra of $\mathcal{X}$) is a factor of X and call $\pi: X' \to Y'$ the factor map. If a factor map $\pi: X' \to Y'$ is also injective, then we say that the systems $(X, \mathcal{X}, \mu, T)$ and $(Y, \mathcal{Y}, \nu, S)$ are *isomorphic*.

If $\mathcal{Y}$ is a T-invariant sub-σ-algebra of $\mathcal{X}$ and $f \in L^2(\mu)$, we define the *conditional expectation* $\mathbb{E}(f|\mathcal{Y})$ *of* f *with respect to* $\mathcal{Y}$ to be the orthogonal projection of f onto $L^2(\mathcal{Y})$. We frequently use the identities

$$\int \mathbb{E}(f|\mathcal{Y})\, d\mu = \int f\, d\mu, \quad T\,\mathbb{E}(f|\mathcal{Y}) = \mathbb{E}(Tf|\mathcal{Y})\ .$$

For each $r \in \mathbb{N}$, we define $\mathcal{K}_r$ to be the factor induced by the algebra

$$\{f \in L^\infty(\mu) : T^r f = f\}\ .$$

We define $\mathcal{K}_{rat}$ to be the factor induced by the algebra generated by the functions

$$\{f \in L^\infty(\mu) : T^r f = f \text{ for some } r \in \mathbb{N}\}\ .$$

The Kronecker factor $\mathcal{K}$ is induced by the algebra spanned by the bounded eigenfunctions of T.

The transformation T is *ergodic* if $\mathcal{K}_1$ consists only of constant functions and T is *totally ergodic* if $\mathcal{K}_{rat}$ consists only of constant functions. The von Neumann Ergodic Theorem states that if T is ergodic and $f \in L^2(\mu)$, then

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T^n f = \int f\, d\mu\ , \tag{9}$$

with the convergence taking place in $L^2(\mu)$.

Every measure preserving system $(X,\mathcal{X},\mu,T)$ has an *ergodic decomposition*, meaning that we can write $\mu=\int \mu_t\ d\lambda(t)$, where λ is a probability measure on $[0,1]$ and μ_t are T-invariant probability measures on $(X,\mathcal{X})$ such that the systems $(X,\mathcal{X},\mu_t,T)$ are ergodic for $t\in[0,1]$.

If G is a k-step nilpotent Lie group and Γ is a cocompact subgroup, then $X=G/\Gamma$ is called a *k-step nilmanifold.* There exists a unique probability measure m on X (the *Haar measure*) that is invariant under left translations. If $a\in G$, then the measure preserving system $(X,\mathcal{X},m,T_a)$ defined by the transformation $T_a(g\Gamma)=(ag)\Gamma$ is called a *nilsystem.* Every unipotent affine transformation on a compact abelian Lie group (with the Borel σ-algebra and the Haar measure) induces a system that is isomorphic to a nilsystem, but these are not the only examples of nilsystems.

We say that the system $(X,\mathcal{X},\mu,T)$ is an *inverse limit of a sequence of factors* $(X,\mathcal{X}_j,\mu,T)$ if $\{\mathcal{X}_j\}_{i\in\mathbb{N}}$ is an increasing sequence of T-invariant sub-σ-algebras such that $\bigvee_{j\in\mathbb{N}}\mathcal{X}_j=\mathcal{X}$ up to sets of measure zero. If in addition for every $j\in\mathbb{N}$ the factor system $(X,\mathcal{X}_j,\mu,T)$ is isomorphic to a nilsystem of order k, we say that $(X,\mathcal{X},\mu,T)$ is an *inverse limit of nilsystems of order k.*

3.2. **Characteristic factors.** A key ingredient in the proof of Theorem 1.1 is the following result of the authors:

Theorem 3.1 ([3])**.** *Let $(X,\mathcal{X},\mu,T)$ be an invertible totally ergodic measure preserving system and let $p_1,\ldots,p_k$ be rationally independent integer polynomials. Then for $f_1,\ldots,f_k\in L^\infty(\mu)$ the difference*

$$\frac{1}{N-M}\sum_{n=M}^{N-1}T^{p_1(n)}f_1\cdot\ldots\cdot T^{p_k(n)}f_k-\prod_{i=1}^{k}\int f_i\,d\mu \tag{10}$$

converges to 0 in $L^2(\mu)$ as $N-M\to\infty$.

We note that the result in [3] is only stated for $M=0$, but the same proof gives this uniform version. If $k\geq 2$ and the polynomials $p_1,\ldots,p_k$ are not rationally independent then there exist totally ergodic systems and bounded functions $f_1,\ldots,f_k$, for which the limit of the average in (10) is not constant. This can be easily seen by considering the example of an irrational rotation on the circle.

Before the proof of Theorem 1.1, we prove a Lemma:

Lemma 3.2. *Let $(X,\mathcal{X},\mu,T)$ be a measure preserving system with ergodic decomposition $\mu=\int\mu_t\ d\lambda(t)$. If $f\in L^\infty(\mu)$ satisfies $\mathbb{E}(f\,|\,\mathcal{K}_{rat}(\mu))=0$, then $\mathbb{E}(f\,|\,\mathcal{K}_{rat}(\mu_t))=0$ for λ-a.e. t.*

Proof. Let σ, σ_t be the spectral measures of the function f with respect to the systems $(X,\mathcal{X},\mu,T)$ and $(X,\mathcal{X},\mu_t,T)$, respectively. It is classical that $\mathbb{E}(f\,|\,\mathcal{K}_{rat}(\mu))=0$ if and only if $\sigma(\{r\}) = 0$ for every $r \in \mathbb{Q}$. Since $\sigma = \int \sigma_t \, d\lambda(t)$, we have that

$$0 = \sigma(\{r\}) = \int \sigma_t(\{r\}) \, d\lambda(t)$$

for every $r \in \mathbb{Q}$. Hence, for every $r \in \mathbb{Q}$ we have $\sigma_t(\{r\}) = 0$ for λ-a.e. t. Since $\mathbb{Q}$ is countable it follows that for λ-a.e. t we have $\sigma_t(\{r\}) = 0$ for every $r \in \mathbb{Q}$, and so for λ-a.e. t we have $\mathbb{E}(f \,|\, \mathcal{K}_{rat}(\mu_t)) = 0$. □

Every integer polynomial $p(n)$ of degree at most d admits a representation of the form $p(n) = \sum_{i=0}^{d} c_i \binom{n}{i}$ for some $c_i \in \mathbb{Q}$, $i = 0, \ldots, d$. Since $p(j) \in \mathbb{Z}$, $j = 0, \ldots, d$, it is immediate that $c_i \in \mathbb{Z}$, $i = 0, \ldots, d$. A fact that we frequently use in the sequel is that whenever $p(n)$ is an integer polynomial of degree at most d, then for every $r \in \mathbb{Z}$ the polynomial $q(n) = p(d!n + r)$ has integer coefficients. This follows easily from the aforementioned representation.

Proof of Theorem 1.1. We begin with some easy reductions. Without loss of generality we can assume that the polynomials $p_1, \ldots, p_k$ have integer coefficients. Indeed, suppose that the highest degree of the polynomials $p_1, \ldots, p_k$ is d. Then for every $r \in \mathbb{Z}$ the polynomial family $\{p_i(d!n + r)\}_{i=1,\ldots,k}$ satisfies the assumptions of the theorem and also has integer coefficients. Using the result for $r = 0, \ldots, d! - 1$ and adding, we obtain the result for the family $\{p_i(n)\}_{i=1,\ldots,k}$. Furthermore, since $\mathbb{E}(T^j f \,|\, \mathcal{K}_{rat}) = T^j \mathbb{E}(f \,|\, \mathcal{K}_{rat})$ for $j \in \mathbb{Z}$, we can further assume that $p_i(0) = 0$ for $i = 1, \ldots, k$.

It suffices to show that if $\mathbb{E}(f_1 \,|\, \mathcal{K}_{rat}) = 0$ then the average (3) converges to zero in $L^2(\mu)$ as $N - M \to \infty$. If f is a function with $\mathbb{E}(f \,|\, \mathcal{K}_{rat}) = 0$ for the measure μ, then by Lemma 3.2 the same property holds for almost every measure in the ergodic decomposition of μ. Hence, we can assume that T is ergodic.

From [10] we know that a characteristic factor for $L^2(\mu)$ convergence of the averages (3) is an inverse limit of nilsystems induced by some T-invariant sub-σ-algebras $\{\mathcal{X}_j\}_{j \in \mathbb{N}}$. Since $\mathbb{E}(f_1 \,|\, \mathcal{K}_{rat}(\mathcal{X})) = 0$ implies that $\mathbb{E}(f_1 \,|\, \mathcal{K}_{rat}(\mathcal{X}_j)) = 0$ for $j \in \mathbb{N}$, using a standard approximation argument we can assume that the system is a nilsystem.

The Kronecker factor of an ergodic nilsystem is isomorphic to a rotation on a monothetic compact abelian Lie group G. Every such group has the form $\mathbb{Z}_{d_1} \times \mathbb{T}^{d_2}$ for some positive integer d_1 and nonnegative integer d_2, where $\mathbb{Z}_d$ denotes the cyclic group with d elements. It follows that $\mathcal{K}_{rat} = \mathcal{K}_{r_0}$ for some $r_0 \in \mathbb{N}$. Hence, every ergodic component of the transformation T^{r_0} is totally ergodic. Since $p_i(0) = 0$ and p_i has integer coefficients, we have that $p_i(nr_0) = r_0 q_i(n)$, where $q_i(n)$, for $i = 1, \ldots, k$, is again a polynomial with integer coefficients. From $\mathbb{E}(f_1 \,|\, \mathcal{K}_{rat}) = 0$, it follows that the function f_1 has integral zero on every ergodic component of T^{r_0}. Applying Theorem 3.1 on the (totally) ergodic components of T^{r_0} with the rationally independent polynomials $q_1, \ldots, q_k$, we have that

$$\frac{1}{N-M} \sum_{n=M}^{N-1} T^{p_1(nr_0)} f_1 \cdot \ldots \cdot T^{p_k(nr_0)} f_k \tag{11}$$

converges to 0 in $L^2(\mu)$ as $N-M \to \infty$. Moreover, $\mathbb{E}(f_1 \,|\, \mathcal{K}_{rat}) = 0$ implies that $\mathbb{E}(T^j f_1 \,|\, \mathcal{K}_{rat}) = 0$ for $j \in \mathbb{N}$ and so the limit is zero with $p_i(nr_0+k)$ substituted for $p_i(nr_0)$ in (11) for $k = 0, \ldots, r_0-1$. Adding these, we have that (3) converges to 0 in $L^2(\mu)$ as $N-M \to \infty$. □

3.3. **Multiple recurrence.** We prove Theorem 1.3.

Proof of Theorem 1.3. Suppose that the highest degree of the polynomials $p_1, \ldots, p_k$ is d. Then the polynomial family $\{p_i(d!n)\}_{i=1,\ldots,k}$ satisfies the assumptions of the theorem and has integer coefficients. By applying the result for this family we can assume that the polynomials $p_1, \ldots, p_k$ have integer coefficients.

Let $\varepsilon > 0$. There exists $r \in \mathbb{N}$ such that

$$\|\mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_r) - \mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_{rat})\|_{L^2(\mu)} \leq \frac{\varepsilon}{k+1} \,. \tag{12}$$

By Theorem 1.1,

(13)
$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \mu(A \cap T^{p_1(nr)}A \cap \cdots \cap T^{p_k(nr)}A) =$$
$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \int \mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_{rat}) \cdot T^{-p_1(nr)}\mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_{rat}) \cdot \ldots \cdot T^{-p_k(nr)}\mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_{rat}) \, d\mu \,.$$

For every choice of integers $a_0, \ldots, a_k$, we have

$$\begin{aligned}
&\Big|\int \prod_{i=0}^{k} T^{a_i}\mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_{rat}) \, d\mu - \int \prod_{i=0}^{k} T^{a_i}\mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_r) \, d\mu\Big| \\
&\leq \int \sum_{i=0}^{k} |T^{a_i}\mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_{rat}) - T^{a_i}\mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_r)| \, d\mu \\
&= \sum_{i=0}^{k} \int |\mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_{rat}) - \mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_r)| \, d\mu \\
&\leq \sum_{i=0}^{k} \|\mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_{rat}) - \mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_r)\|_{L^2(\mu)} \leq \varepsilon
\end{aligned}$$

by (12). It follows that the limit in (13) is greater than or equal to

$$\begin{aligned}
&\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \int \mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_r) \cdot T^{-p_1(nr)}\mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_r) \cdot \ldots \cdot T^{-p_k(nr)}\mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_r) \, d\mu - \varepsilon \\
&\qquad = \int \mathbb{E}(\mathbf{1}_A \,|\, \mathcal{K}_r)^{k+1} \, d\mu - \varepsilon \ \geq \mu(A)^{k+1} - \varepsilon,
\end{aligned}$$

where the last equality holds since r divides $p_i(nr)$ for $i=1,\dots,k$, and every $\mathcal{K}_r$-measurable function is T^r invariant. □

4. Correlations of independent polynomial iterates and nilsequences

We now prove the Structure Theorem 1.2 for multicorrelation sequences of independent polynomials. We start with some definitions from [1]:

Definition 4.1. Let $k \geq 1$ be an integer and let $X = G/\Gamma$ be a k-step nilmanifold. Suppose that ϕ is a continuous complex valued function on X, $a \in G$, and $x_0 \in X$. The sequence $\{\phi(a^n x_0)\}_{n\in\mathbb{N}}$ is called a *basic k-step nilsequence*. A *k-step nilsequence* is a uniform limit of basic k-step nilsequences.

Definition 4.2. Let $\{a_n\}_{n\in\mathbb{N}}$ be a bounded sequence of complex numbers. We say that a_n *tends to zero in uniform density*, and write $UD\text{-}\lim a_n = 0$, if

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} |a_n| = 0 \ .$$

Before the proof, we begin with a Lemma:

Lemma 4.3. *Let $(X, \mathcal{X}, \mu, T)$ be a measure preserving system, $p_1, \dots, p_k$ be rationally independent integer polynomials, and $f_0, f_1, \dots, f_k \in L^\infty(\mu)$. Then*

$$UD\text{-}\lim \Big(\int f_0 \cdot T^{p_1(n)} f_1 \cdot \ldots \cdot T^{p_k(n)} f_k \, d\mu - \int \tilde{f}_0 \cdot T^{p_1(n)} \tilde{f}_1 \cdot \ldots \cdot T^{p_k(n)} \tilde{f}_k \, d\mu \Big) = 0 \ , \tag{14}$$

where $\tilde{f}_i = \mathbb{E}(f_i \,|\, \mathcal{K})$, $i = 0, 1, \dots, k$ and $\mathcal{K}$ is the Kronecker factor of the system.

Proof. It suffices to show that if $\mathbb{E}(f_i \,|\, \mathcal{K}) = 0$ for some $i \in \{1, \dots, k\}$, then the UD-limit in (14) is zero. Without loss of generality, we can assume that $i = 1$. We apply Theorem 1.1 to the product system induced by $T\times T$ acting on $X\times X$. From [5] (Lemma 4.18) we know that $f \in \mathcal{K}(X \times X)$ if and only if it has the form

$$f(x, x') = \sum_{n\in\mathbb{N}} c_n \, g_n(x) \cdot h_n(x')$$

where $g_n, h_n \in \mathcal{K}(X)$ and $c_n \in \mathbb{C}$ for $n \in \mathbb{N}$. Since $\mathbb{E}(f_1 \,|\, \mathcal{K}(X)) = 0$, it follows that $\mathbb{E}(f_1 \otimes \bar{f}_1 \,|\, \mathcal{K}(X \times X)) = 0$ which implies that $\mathbb{E}(f_1 \otimes \bar{f}_1 \,|\, \mathcal{K}_{rat}(X \times X)) = 0$. Hence, the average

$$\frac{1}{N-M} \sum_{n=M}^{N-1} (T\times T)^{p_1(n)}(f_1 \otimes \bar{f}_1) \cdot \ldots \cdot (T\times T)^{p_k(n)}(f_k \otimes \bar{f}_k)$$

converges to zero in $L^2(\mu \times \mu)$ as $N - M \to \infty$. It follows that

$$\lim_{N-M\to\infty} \frac{1}{N-M} \sum_{n=M}^{N-1} \Big| \int f_0 \cdot T^{p_1(n)} f_1 \cdot \ldots \cdot T^{p_k(n)} f_k \, d\mu \Big|^2 = 0 \ .$$

and this completes the proof. □

Proof of Theorem 1.2. By Lemma 4.3, we can assume that $\mathcal{X} = \mathcal{K}$. Since the system is ergodic and coincides with its Kronecker factor we can assume that T is a rotation on a compact abelian group G. Every compact abelian group is an inverse limit of compact abelian Lie groups and so using an easy approximation argument, such as the one used in [1] (see page 296), we can further assume that G is Lie.

Suppose now that G is a compact abelian Lie group with Haar measure m and that $T\colon G \to G$ is given by $T(g) = g + a$ for some $a \in G$. For $j = 0, \ldots, d$ we have that $p_j(n) = \sum_{i=0}^{d} c_{i,j} \binom{n}{i}$ for some $c_{i,j} \in \mathbb{Z}$. We construct the advertised transformation $S\colon G^d \to G^d$ and the continuous function $\phi\colon G^d \to \mathbb{C}$ as follows: S is defined by

$$S\big(g_1, g_2, \ldots, g_d\big) = \big(g_1 + a, g_2 + g_1, \ldots, g_d + g_{d-1}\big) \ ,$$

and the continuous function ϕ is defined by

$$\phi(g_1, \ldots, g_d) = \int f_0(g) \cdot \prod_{i=1}^{k} f_i(g + c_{i,0}a + \sum_{j=1}^{d} c_{i,j} g_i) \ dm(g) \ .$$

Note that S is unipotent since all its eigenvalues are 1. It is easy to check that

$$S^n(0, \ldots, 0) = \Big(\binom{n}{1}a, \ldots, \binom{n}{d}a\Big) \ ,$$

and so

$$\begin{aligned}\phi\big(S^n(0, \ldots, 0)\big) =& \phi\Big(\binom{n}{1}a, \ldots, \binom{n}{d}a\Big) \\ =& \int f_0(g) \cdot f_1(g + p_1(n)a) \cdot \ldots \cdot f_k(g + p_k(n)a) \ dm(g) \\ =& \int f_0 \cdot T^{p_1(n)} f_1 \cdot \ldots \cdot T^{p_k(n)} f_k \ dm = a_n \ .\end{aligned}$$

The system (G^d, S) is topologically conjugate to a d-step nilsystem, meaning that there exist a d-step nilmanifold H/Γ, an $a \in H$, and an invertible continuous map $\pi\colon G^d \to H/\Gamma$ such that $S = \pi^{-1} \circ T_a \circ \pi$, where T_a is defined by $T_a(g\Gamma) = (ag)\Gamma$. It follows that

$$\phi\big(S^n(0, \ldots, 0)\big) = \phi'(T_a^n x_0)$$

where $\phi' = \phi \circ \pi^{-1}$ is a continuous function on H/Γ and $x_0 = \pi(0, \ldots, 0) \in H/\Gamma$. This completes the proof. □

We illustrate the construction of this proof with an example:

Example. Suppose that $k = 2$ and $p_1(n) = 2n + 1$, $p_2(n) = n^2/2 - n/2$, $G = \mathbb{T}$, and $T\colon \mathbb{T} \to \mathbb{T}$ is given by $T(t) = t + \alpha \pmod 1$ for some irrational $\alpha \in \mathbb{T}$. Then

$$a_n = \int f_0\big(t\big) \cdot f_1\big(t + (2n+1)\alpha\big) \cdot f_2\Big(t + (\binom{n}{2} - \binom{n}{1})\alpha\Big) \ dt \ ,$$

$S\colon \mathbb{T}^2 \to \mathbb{T}^2$ is defined by

$$S(t_1, t_2) = (t_1 + \alpha, t_2 + t_1) \ ,$$

and $\phi\colon \mathbb{T}^2 \to \mathbb{C}$ is defined by

$$\phi(t_1, t_2) = \int f_0(t) \cdot f_1(t + \alpha + 2t_1) \cdot f_2(t - t_1 + t_2) \ dt \ .$$

References

[1] V. Bergelson, B. Host and B. Kra, with an appendix by I. Ruzsa. Multiple recurrence and nilsequences, *Inventiones Math.*, **160**, (2005), 261–303.

[2] V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems, *J. Amer. Math. Soc.* **9**, (1996) 725–753.

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

[4] H. Furstenberg. Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. d'Analyse Math*, **71**, (1977), 204–256.

[5] H. Furstenberg. Recurrence in ergodic theory and combinatorial number theory, *Princeton University Press*, Princeton, (1981).

[6] B. Green. A Szemerédi-type regularity lemma in abelian groups. *GAFA*, **15**, (2005), no. 2, 340–376.

[7] B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Annals. Math.* **161**, (2005), 397–488.

[8] B. Host and B. Kra. Convergence of Polynomial ergodic averages, *Isr. J. Math.*, **149**, (2005), 1-19.

[9] A. Khintchine. Eine Verschärfung des Poincaréschen "Wiederkehrsatzes"' *Comp. Math.*, **1**, (1934), 177–179.

[10] A. Leibman. Convergence of multiple ergodic averages along polynomials of several variables, *Isr. J. Math.*, **146**, (2005), 303–316.

[11] A. Sárközy. On difference sets of integers III, *Acta Math. Acad. Sci. Hungar.*, **31**, (1978), 125–149.

[12] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression, *Acta Arith.*, **27**, (1975), 299–345.

Department of Mathematics, Institute for Advanced Study, 1 Einstein Drive, Princeton, NJ 08540

Department of Mathematics, Northwestern University, 2033 Sheridan Road, Evanston, IL 60208-2730

E-mail address: nikos@math.ias.edu

E-mail address: kra@math.northwestern.edu