STUDIA MATHEMATICA

Online First version

Integer part independent polynomial averages and applications along primes

by

DIMITRIS KARAGEORGOS (Athens) and
ANDREAS KOUTSOGIANNIS (Columbus, OH)

Abstract. Exploiting the equidistribution properties of polynomial sequences, following the methods developed by Leibman (2005) and Frantzikinakis (2009, 2010), we show that the ergodic averages with iterates given by the integer parts of strongly independent real valued polynomials converge in the mean to the expected limit. These results have, via Furstenberg's correspondence principle, immediate combinatorial applications, while combining these results with methods of Frantzikinakis et al. (2013) and Koutsogiannis (2018) we get the expected limits and combinatorial results for multiple averages for a single sequence, as well as for several sequences along prime numbers.

1. Introduction. The study of the limiting behavior in $L^2(\mu)$ of multiple ergodic averages of the form

$$\frac{1}{N}\sum_{n=1}^{N} T^{a_1(n)}f_1\cdot\ldots\cdot T^{a_\ell(n)}f_\ell \tag{1.1}$$

where $(a_1(n))_n,\ldots,(a_\ell(n))_n$ are sequences of integers, $T : X \to X$ is an invertible measure preserving transformation on the probability space $(X,\mathcal{B},\mu)$ (1) and $f_1,\ldots,f_\ell \in L^\infty(\mu)$ has been of great importance in ergodic theory. It was pioneered by Furstenberg [Fur] who studied averages as in (1.1) with $a_i(n) = in$, $1 \le i \le \ell$, in order to provide an ergodic-theoretical proof of Szemerédi's theorem on arithmetic progressions.

Bergelson and Liebman [BL] studied the case where the a_i's are integer polynomials with no constant term, proving a polynomial extension of Szemerédi's theorem.

2010 *Mathematics Subject Classification*: Primary 37A45; Secondary 37A05, 37A30.
Key words and phrases: ergodic average, recurrence, equidistribution, nilmanifold, prime numbers.
Received 2 November 2017; revised 19 August 2018.
Published online 24 June 2019.

(1) We shall call the quadruple $(X,\mathcal{B},\mu,T)$ a *system*.

DOI: 10.4064/sm171102-18-9

Another question that arises when studying the limiting behavior of (1.1) is: if the limit exists, what can we say about it? In this direction and for the case of weakly mixing systems Furstenberg, Katznelson and Ornstein [FKO] proved that if $a_i(n) = in$, $1 \leq i \leq \ell$, then the multiple ergodic averages in (1.1) converge to the product of the integrals of the f_i's. Bergelson [B2] extended this result to the case where a_i's are essentially distinct integer polynomials (i.e., all polynomials and their differences are non-constant). Furthermore, Frantzikinakis and Kra [FK] established the same result under the total ergodicity assumption of the system for a_i's independent integer polynomials (i.e., every non-trivial integral combination of a_i's is non-constant). Precise knowledge of the limit for a general system would a priori imply "nice" recurrence and combinatorial results.

Convergence of multiple ergodic averages with several commuting transformations has been studied as well. Under weaker assumptions than weak mixing, convergence to the expected limit in the case of iterates of integer polynomials (i.e., polynomials with rational coefficients that take integer values at integers) with different degrees and in the case of the integer parts of real polynomials with different degrees is treated in [CFH] and [K1] respectively, where by *expected limit* we mean that under the ergodicity assumption, the limit is equal to the product of the integrals of the f_i's.

There are also results in other, non-polynomial classes of iterates. Bergelson and Håland-Knutson [BK] considered the case of iterates determined by integer parts of tempered functions on a weakly mixing system, while Frantzikinakis dealt with the case of iterates determined by integer parts of logarithmico-exponential Hardy field functions of different growth with polynomial degrees (in [F1] for a single T and in [F4] for multiple commuting T_i's where he also showed that in the sublinear, 0-degree case, the commutativity assumption on T_i's can be removed). All these results give convergence to the expected limit. We highlight at this point that the last results of Frantzikinakis are very strong and hold for general systems, a behavior that we have not had so far for any class of polynomial iterates with multiple polynomials of degree greater than 1 (for a single polynomial see Theorem 2.10 of Frantzikinakis below).

In this paper we study the convergence of multiple ergodic averages with integer-part-of-real-polynomial iterates of several sequences of the form

$$\frac{1}{N}\sum_{n=1}^{N} T^{[p_1(n)]}f_1 \cdot \ldots \cdot T^{[p_\ell(n)]}f_\ell. \tag{1.2}$$

In our main theorem, Theorem 2.1, we show that for polynomial sequences $(p_1(n))_n, \ldots, (p_\ell(n))_n$ where every non-trivial linear combination of p_i's has at least one non-constant irrational coefficient, under the assumption of ergodicity on T (i.e., there is no non-trivial set invariant under T), the limit of (1.2) as $N \to \infty$ in $L^2(\mu)$ is the expected one, $\prod_{i=1}^{\ell} \int f_i \, d\mu$. So, by ergodic decomposition, for a general system the above limit is the product of the conditional expectations of f_i's, i.e., $\prod_{i=1}^{\ell} \mathbb{E}(f_i|\mathcal{I}(T))$. This is arguably the first time that we get existence and precise expression of the limit for general multiple polynomial expressions. As we make no assumptions on our system, even though it is in ergodic theory language, our result is of combinatorial nature. Measure-theoretical and hence, via Furstenberg's correspondence principle, combinatorial applications of Theorem 2.1 are given in Theorems 2.2–2.4. The strong nature of our results is also reflected in Theorem 2.5, Corollary 2.6, Theorem 2.8 and Corollary 2.9, where we obtain some additional applications in topological dynamics and combinatorics.

To prove Theorem 2.1 we follow Frantzikinakis' approach (from [F1] and [F2]) and we also use some results of Leibman [L1] and Host–Kra [HK]. More specifically, via Lemma 5.1 (that is, [F1, Lemma 4.7]), stating that the nilfactor of our system is also the characteristic factor for the family of polynomials of interest, and the structure theorem (Theorem 3.2) of Host and Kra (which follows from [HK, Theorem 10.1]), it suffices to prove Theorem 2.1 when our system is a nilsystem. To complete the proof, we use Theorem 4.1, an equidistribution result first proved by Frantzikinakis [F2] for logarithmico-exponential Hardy field functions with polynomial degree of different growth. In order to derive Theorem 4.1, we use Theorem 3.1, a central equidistribution result due to Leibman [L1, Theorems B and C] for a polynomial sequence in a connected and simply connected Lie group.

Combining Theorem 2.10, which is a result from [F1] on multiple convergence of a single real polynomial sequence, and Theorem 2.1 with some results from [K2], we get the analogous results in Theorems 2.12 and 2.14 respectively, together with their implications, for averages along prime numbers.

Lastly, in Theorems 2.18 and 2.19 we show the corresponding recurrence results along shifted primes ($\mathbb{P}-1$ and $\mathbb{P}+1$), together with their combinatorial implications via Furstenberg's correspondence principle in Theorems 2.20 and 2.21.

Throughout the article we highlight the fact that one cannot expect nice convergence and recurrence results for other classical families of polynomials, say, integer polynomials, or even for very special families of them. This forces one to deal with real-valued polynomial families which satisfy some Weyl-type assumptions (see next section).

Notation. We denote by $\mathbb{P}$ the set of prime numbers. For $N \in \mathbb{N} = \{1, 2, \ldots\}$ we write $[1, N]$ for $\{1, \ldots, N\}$. For a measurable function f on a measure space X with a transformation $T : X \to X$, we denote by Tf

the composition $f \circ T$. Moreover, $\mathbb{T}^s = \mathbb{R}^s/\mathbb{Z}^s$ denotes the s-dimensional torus, $e(t) = e^{2\pi i t}$ the exponential map, $(a(n))_n$ a sequence indexed over the natural numbers (i.e., $(a(n))_{n\in\mathbb{N}}$), and $[\cdot]$ the integer part (floor) function.

2. Main results. It is as a special case of [K1, Theorem 1.3] that for any $\ell \in \mathbb{N}$, any system $(X, \mathcal{B}, \mu, T)$, and any real polynomials $p_1, \dots, p_\ell$ and functions $f_1, \dots, f_\ell \in L^\infty(\mu)$ the limit

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T^{[p_1(n)]} f_1 \cdot \ldots \cdot T^{[p_\ell(n)]} f_\ell \tag{2.1}$$

exists in $L^2(\mu)$.

We now define a notion of independence of real polynomials that will ensure convergence of (2.1) to the expected limit (see Theorem 2.1 below).

Definition. For $\ell \in \mathbb{N}$, let $p_1, \dots, p_\ell$ be real polynomials. We say they are *strongly independent* if any non-trivial real linear combination of p_i's has a non-constant irrational coefficient.

Note that a family with one element, $\{p\}$, where $p \in \mathbb{R}[t]$, is strongly independent iff $p(t)$ is not of the form $cq(t) + d$ for all $c, d \in \mathbb{R}$ and $q \in \mathbb{Q}[t]$ (or $\mathbb{Z}[t]$ equivalently).

Examples. The family $\{\sqrt{2}\,t^3 + t^2, \sqrt{3}\,t^3 - t\}$ is strongly independent while the families $\{\sqrt{5}\,t^3 + t^2 + \sqrt{6}\,t, t^2, \sqrt{7}\,t\}$ and $\{\sqrt{2}\,t^2 + t, \sqrt{5}\,t^2 - t\}$ are not.

We also remark that our definition coincides with the definition of the *good* family of polynomials given in [F3, Problem 10], in the special case where the polynomial sequences are constant.

From our method of proof, it will become clear that our assumptions on the real polynomials are in a sense optimal, since they are exactly what one has to assume in order to obtain the Weyl-type equidistribution results we have to show in order to prove our main result, that the limit of the ergodic averages over strongly independent real polynomials exists and is as expected.

Theorem 2.1. *Let $\ell \in \mathbb{N}$, and let $p_1, \dots, p_\ell \in \mathbb{R}[t]$ be strongly independent real polynomials, $(X, \mathcal{B}, \mu, T)$ an ergodic system and $f_1, \dots, f_\ell \in L^\infty(\mu)$. Then*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T^{[p_1(n)]} f_1 \cdot \ldots \cdot T^{[p_\ell(n)]} f_\ell = \prod_{i=1}^{\ell} \int f_i \, d\mu \tag{2.2}$$

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

Remark. The assumption that the polynomials are strongly independent is necessary, since even for $\ell = 1$, $p(t) = \sqrt{2}\,t$ and ergodic rotations on the torus, (2.2) typically fails.

In the study of ergodic averages with polynomial iterates, of importance are families of finitely many *independent, integer polynomials* (i.e., any non-trivial linear combination with integer coefficients is non-constant). We note that even for such families it is not true in general that one has convergence as in (2.2), i.e., to the expected limit for a general ergodic system (see remark after Theorem 2.2). Such a result requires more assumptions on the system, like total ergodicity (see [FK]). Hence, one is forced to work with real polynomials in order to have this nice convergence behavior.

We now state a principle due to Furstenberg which allows one to obtain combinatorial results from ergodic-theoretical ones. We present it in a formulation from [B1].

THEOREM (Furstenberg correspondence principle, [B1], [Fur, Theorem 1.1]). *Let E be a subset of positive integers. There exists a system $(X,\mathcal{B},\mu,T)$ and a set $A\in\mathcal{B}$ with $\mu(A)=\bar{d}(E)$ (2) such that*

$$\bar{d}(E\cap(E-n_1)\cap\cdots\cap(E-n_\ell))\geq\mu(A\cap T^{-n_1}A\cap\cdots\cap T^{-n_\ell}A)$$

for every $\ell\in\mathbb{N}$ and $n_1,\ldots,n_\ell\in\mathbb{Z}$.

As a consequence of Theorem 2.1 we get the following recurrence result (for a proof of this implication see for example [F1, Theorem 2.8]).

THEOREM 2.2. *Let $\ell\in\mathbb{N}$ and let $p_1,\ldots,p_\ell\in\mathbb{R}[t]$ be strongly independent real polynomials. Then for every system $(X,\mathcal{B},\mu,T)$ and $A\in\mathcal{B}$ we have*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mu(A\cap T^{-[p_1(n)]}A\cap\cdots\cap T^{-[p_\ell(n)]}A)\geq(\mu(A))^{\ell+1}. \tag{2.3}$$

REMARK. The assumption that the polynomials are strongly independent is necessary since even for $\ell=1$ and $p(t)=t^2$, (2.3) typically fails.

This remark shows that (2.3) typically fails even for families of independent, integer polynomials. Hence, Theorem 2.2 is another indication that one has to work with real polynomials in order to have nice lower bounds as in (2.3) for general systems.

Note at this point that our arguments show that the uniform version of Theorem 2.1, and hence its implications, holds, meaning that one can replace the standard Cesàro averages, $\lim_{N\to\infty}N^{-1}\sum_{n=1}^{N}$, with the respective uniform ones, $\lim_{N-M\to\infty}(N-M)^{-1}\sum_{n=M+1}^{N}$, and the natural upper density, $\bar{d}$, with the respective upper Banach density, d^* (3).

(2) For a set $A\subseteq\mathbb{N}$ we define its *upper density* to be $\bar{d}(A)=\limsup_{N\to\infty}|A\cap\{1,\ldots,N\}|/N$. If the limit as $N\to\infty$ of the previous expression exists, it is denoted by $d(A)$ and called the *density* of A.

(3) For a set $A\subseteq\mathbb{Z}$, we call $d^*(A)=\limsup_{N-M\to\infty}|A\cap\{M+1,\ldots,N\}|/(N-M)$ the *upper Banach density* of A.

Then the uniform version of Theorem 2.2 implies that for any $A \in \mathcal{B}$ with $\mu(A) > 0$, and every $\varepsilon > 0$, the set

$$R_\varepsilon(A) = \{n \in \mathbb{Z} : \mu(A \cap T^{-[p_1(n)]}A \cap \cdots \cap T^{-[p_\ell(n)]}A) > (\mu(A))^{\ell+1} - \varepsilon\}$$

is syndetic (i.e., it has bounded gaps).

We note that this general result, which holds under no assumption on the system, implies that a family of strongly independent real polynomials has a much different behavior than a family of linear integer polynomials, since its stands in contrast with the Bergelson–Host–Kra–Ruzsa counterexample to the higher-order Khinchin recurrence theorem. Indeed, in [BHK], those authors found an ergodic system $(X, \mathcal{B}, \mu, T)$ and a set $A \in \mathcal{B}$ with $\mu(A) > 0$ such that

$$\mu(A \cap T^{-n}A \cap T^{-2n}A \cap T^{-3n}A \cap T^{-4n}A) \leq \mu(A)^5/2 \quad \text{for all } n \neq 0$$

(so, for $p_i(t) = it$ the syndeticity of $R_\varepsilon(A)$ fails for certain ergodic systems when $\ell \geq 4$, while for certain non-ergodic systems it fails even when $\ell \geq 2$; for examples covering both cases, see [BHK]).

Using Theorem 2.2 and Furstenberg's correspondence principle, we obtain:

Theorem 2.3. *Let $\ell \in \mathbb{N}$ and let $p_1, \ldots, p_\ell \in \mathbb{R}[t]$ be strongly independent real polynomials. Then for every $E \subseteq \mathbb{N}$ we have*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \bar{d}\big(E \cap (E - [p_1(n)]) \cap \cdots \cap (E - [p_\ell(n)])\big) \geq (\bar{d}(E))^{\ell+1}.$$

An immediate implication is:

Theorem 2.4. *Let $\ell \in \mathbb{N}$ and let $p_1, \ldots, p_\ell \in \mathbb{R}[t]$ be strongly independent real polynomials. Then every $E \subseteq \mathbb{N}$ with $\bar{d}(E) > 0$ contains arithmetic configurations of the form*

$$\{m, m + [p_1(n)], m + [p_2(n)], \ldots, m + [p_\ell(n)]\}$$

for some $m \in \mathbb{Z}$ and $n \in \mathbb{N}$ with $[p_i(n)] \neq 0$ for all $1 \leq i \leq \ell$.

We note that one can get the conclusion of Theorem 2.3 for integer polynomials with no constant term from the polynomial Szemerédi theorem [BL, Theorem A_0], but in the generality that we present it here it is not clear to us at all if the theorem follows from previous results in the literature.

In the next two applications of Theorem 2.1 we follow Subsections 2.3 and 2.4 of [F4] respectively, getting similar results for sequences of strongly independent polynomials instead of sequences of Hardy field functions.

2.0.1. *An application in topological dynamics.* Let (X, T) be a (topological) dynamical system, where (X, d) is a compact metric space and

$T : X \to X$ an invertible continuous transformation. Suppose T is minimal (i.e., $\overline{\{T^n x : n \in \mathbb{N}\}} = X$ for all $x \in X$, hence for every $x \in X$ and non-empty open set U the set $\{n \in \mathbb{N} : T^n x \in U\}$ is syndetic). There exists a T-invariant Borel measure which gives positive value to every non-empty open set. So, due to syndeticity, for every $x \in X$ and every non-empty open set U we have

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_U(T^n x) > 0 \tag{2.4}$$

(actually, the limit exists). Note that from Theorem 2.1, by ergodic decomposition, it follows that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} T^{[p_1(n)]} f_1 \cdot \ldots \cdot T^{[p_\ell(n)]} f_\ell = \prod_{i=1}^{\ell} \mathbb{E}(f_i|\mathcal{I}(T)) \tag{2.5}$$

with convergence in $L^2(\mu)$, where $p_1, \ldots, p_\ell$ are strongly independent real polynomials, $f_1, \ldots, f_\ell \in L^\infty(\mu)$, $\mathcal{I}(T)$ denotes the σ-algebra of T-invariant sets and $\mathbb{E}(f|\mathcal{I}(T))$ is the conditional expectation with respect to $\mathcal{I}(T)$.

Indeed, if $\mu = \int \mu_t \, d\lambda(t)$ denotes the ergodic decomposition of μ, it suffices to show that if $\mathbb{E}(f_i|\mathcal{I}(T)) = 0$ for some i then the averages converge to 0. Since $\mathbb{E}(f_i|\mathcal{I}(T)) = 0$, we have $\int f_i \, d\mu_t = 0$ for λ-a.e. t. By (2.2), the averages go to 0 in $L^2(\mu_t)$ for λ-a.e. t, hence the limit is equal to 0 in $L^2(\mu)$.

Since $\mathbb{E}(f_i|\mathcal{I}(T)) = \lim_{N\to\infty} N^{-1} \sum_{n=1}^{N} T^n f_i$, combining (2.5) with (2.4) we deduce that for almost every $x \in X$ (and hence for a dense set) and any $U_1, \ldots, U_\ell$ from a given countable basis of non-empty open sets,

$$\limsup_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{U_1}(T^{[p_1(n)]} x) \cdot \ldots \cdot \mathbf{1}_{U_\ell}(T^{[p_\ell(n)]} x) > 0. \tag{2.6}$$

Using this we get

THEOREM 2.5. *Let $\ell \in \mathbb{N}$, and let $p_1, \ldots, p_\ell \in \mathbb{R}[t]$ be strongly independent real polynomials and (X, T) a minimal dynamical system. Then for a residual and T-invariant set of $x \in X$ we have*

$$\overline{\{(T^{[p_1(n)]} x, \ldots, T^{[p_\ell(n)]} x) : n \in \mathbb{N}\}} = X \times \cdots \times X. \tag{2.7}$$

Proof. Relation (2.6) immediately implies that the set of points that satisfy (2.7), say R, is dense. To see that it is G_δ, take $\ell = 1$ (the general case is analogous). Then

$$R = \{x \in X : \text{for all } m, r \in \mathbb{N} \text{ there is } n \in \mathbb{N} \text{ with } T^{[p_1(n)]} x \in B(x_m, 1/r)\},$$

where $\{x_m : m \in \mathbb{N}\}$ is a countable, dense subset of X and $B(x_m, 1/r)$ denotes the open ball centered at x_m with radius $1/r$. The claim now follows since

$$R=\bigcap_{m,r\in\mathbb{N}}\bigcup_{n\in\mathbb{N}}T^{-[p_1(n)]}B(x_m,1/r).$$

With a standard convergence argument, we can also show that R is T-invariant. ■

Remark. Even for $\ell=1$ examples of minimal rotations on finite cyclic groups show that if $p\in\mathbb{Z}[t]$ is any polynomial different from $\pm t+d$, then (2.7) may fail for every $x\in X$.

Using Zorn's lemma we can easily show that every dynamical system has a minimal subsystem. Using this and Theorem 2.5 we get:

Corollary 2.6. *Let $\ell\in\mathbb{N}$, and let $p_1,\ldots,p_\ell\in\mathbb{R}[t]$ be strongly independent real polynomials and (X,T) a dynamical system. Then for a non-empty T-invariant set of $x\in X$ we have*

$$\overline{\{(T^{[p_1(n)]}x,\ldots,T^{[p_\ell(n)]}x):n\in\mathbb{N}\}}=\overline{o_T(x)}\times\cdots\times\overline{o_T(x)}, \tag{2.8}$$

where $o_T(x):=\{T^nx:n\in\mathbb{N}\}$ is the (forward) orbit of x under T.

Remark. As in the previous remark, examples for $\ell=1$ and $p\in\mathbb{Z}[t]$ with $p(t)$ different from $\pm t+d$ show that (2.8) typically fails.

2.0.2. *An application in combinatorics.* Using Theorem 2.1, we can obtain the following recurrence result (for a proof use an argument similar to the one in [F4, Theorem 2.4]):

Theorem 2.7. *Let $\ell\in\mathbb{N}$, and let $p_1,\ldots,p_\ell\in\mathbb{R}[t]$ be strongly independent real polynomials, $(X,\mathcal{B},\mu,T)$ a system and $A_0,A_1,\ldots,A_\ell\in\mathcal{B}$ such that*

$$\mu(A_0\cap T^{k_1}A_1\cap\cdots\cap T^{k_\ell}A_\ell)=\alpha>0$$

for some $k_1,\ldots,k_\ell\in\mathbb{Z}$. Then

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\mu(A_0\cap T^{-[p_1(n)]}A_1\cap\cdots\cap T^{-[p_\ell(n)]}A_\ell)\geq\alpha^{\ell+1}.$$

Using this result and a variant of Furstenberg's correspondence principle for several sets A_i (see [F5, Proposition 3.3]) we get (see [F4, the $d=1$ case of Theorem 2.8]):

Theorem 2.8. *Let $\ell\in\mathbb{N}$, and let $p_1,\ldots,p_\ell\in\mathbb{R}[t]$ be strongly independent real polynomials and $E_0,E_1,\ldots,E_\ell\subseteq\mathbb{N}$ such that*

$$\bar{d}(E_0\cap(E_1+k_1)\cap\cdots\cap(E_\ell+k_\ell))=\alpha>0$$

for some $k_1,\ldots,k_\ell\in\mathbb{Z}$. Then

$$\liminf_{N\to\infty}\frac{1}{N}\sum_{n=1}^{N}\bar{d}\big(E_0\cap(E_1-[p_1(n)])\cap\cdots\cap(E_\ell-[p_\ell(n)])\big)\geq\alpha^{\ell+1}.$$

We will sketch the proof of this last result. First we recall a definition:

Definition ([F5, Definition 5]). We say that sequences $a_1, \ldots, a_\ell \in \ell^\infty(\mathbb{Z})$ *admit correlations along a sequence* $([1, N_k])_k$ *of intervals* with $N_k \to \infty$ as $k \to \infty$ if the limit

$$\lim_{k\to\infty} \frac{1}{N_k} \sum_{n=1}^{N_k} b_1(n+m_1) \cdot \ldots \cdot b_s(n+m_s)$$

exists for all $s \in \mathbb{N}$, $m_1, \ldots, m_s \in \mathbb{Z}$ and all sequences $b_1, \ldots, b_s \in \{a_1, \ldots, a_\ell, \overline{a}_1, \ldots, \overline{a}_\ell\}$.

We remark that for $a_1, \ldots, a_\ell \in \ell^\infty(\mathbb{Z})$, using a diagonal argument, for any sequence $(N_k)_k \subseteq \mathbb{N}$ with $N_k \to \infty$ as $k \to \infty$, we can find a subsequence $(N'_k)_k$ such that $a_1, \ldots, a_\ell$ admit correlations along $([1, N'_k])_k$.

Proof of Theorem 2.8. Find a sequence $\mathbf{N} := ([1, N_k])_k$ of intervals along which the upper density of the intersection in the assumption is attained. Let $d_{\mathbf{N}}$ denote the corresponding density. Passing to a subsequence, if needed, which we denote again by $([1, N_k])_k$, we can assume that the functions $\mathbf{1}_{E_0}, \ldots, \mathbf{1}_{E_\ell}$ admit correlations along $([1, N_k])_k$. By [F5, Proposition 3.3], there exists a system $(X, \mathcal{B}, \mu, T)$ and $A_0, \ldots, A_\ell \in \mathcal{B}$ such that

$$d_{\mathbf{N}}(E_0 \cap (E_1 - m_1) \cap \cdots \cap (E_\ell - m_\ell)) = \mu(A_0 \cap T^{m_1} A_1 \cap \cdots \cap T^{m_\ell} A_\ell)$$

for all $m_1, \ldots, m_\ell \in \mathbb{Z}$. The result now follows from Theorem 2.7. ■

This result can be applied to several syndetic sets $E_0, E_1, \ldots, E_\ell \subseteq \mathbb{N}$ with constant $\alpha = (\prod_{i=0}^{\ell} r_i)^{-1}$, where r_i is the syndeticity constant of E_i, $0 \le i \le \ell$. So, one immediately gets:

Corollary 2.9. *Let $\ell \in \mathbb{N}$, and let $p_1, \ldots, p_\ell \in \mathbb{R}[t]$ be strongly independent real polynomials and $E_0, E_1, \ldots, E_\ell \subseteq \mathbb{N}$ be syndetic sets. Then there exist $m, n \in \mathbb{N}$ such that*

$$m \in E_0, \quad m + [p_1(n)] \in E_1, \ \ldots, \ m + [p_\ell(n)] \in E_\ell.$$

Via this last result, for a syndetic set $E \subseteq \mathbb{N}$, real strongly independent polynomials $p_1, \ldots, p_\ell \in \mathbb{R}[t]$ and $c_0, c_1, \ldots, c_\ell \in \mathbb{N}$, setting $E_i = c_i E$, (4) $0 \le i \le \ell$, we can find $x_0, x_1, \ldots, x_\ell \in E$ and $n \in \mathbb{N}$ solving the following system of equations:

$$\begin{aligned} c_1 x_1 - c_0 x_0 &= [p_1(n)], \\ c_2 x_2 - c_0 x_0 &= [p_2(n)], \\ &\vdots \\ c_\ell x_\ell - c_0 x_0 &= [p_\ell(n)]. \end{aligned}$$

(4) Here $cE := \{cn : n \in E\}$.

Note at this point that similar results fail even for $\ell = 1$, i.e., a single polynomial sequence, and also fail when E is only assumed to be piecewise syndetic. Easy examples show that if $p \in \mathbb{Z}[t]$ is any polynomial different from $\pm t + d$ and $k \in \mathbb{N} \setminus \{1\}$, then the equation $kx - y = p(n)$ has no solution with x, y belonging to some set E that is an arithmetic progression.

2.1. Convergence along primes. Using Theorems 2.10 (see below) and 2.1 and some results from [FHK1], [FHK2] and [K2], we can prove integer part polynomial multiple convergence along primes to the expected limit for strongly independent polynomial families.

2.1.1. *Single sequence.* The next result tells us that the limit of the ergodic averages with integer-part-of-real-polynomial iterates of a single sequence is equal to the limit of the "Furstenberg averages", and it follows from [F1, Theorem 2.2].

THEOREM 2.10 ([F1]). *Let $p \in \mathbb{R}[t]$ be different from $cq+d$ for all $c, d \in \mathbb{R}$ and $q \in \mathbb{Q}[t]$. Then for all $\ell \in \mathbb{N}$, all systems $(X, \mathcal{B}, \mu, T)$ and all $f_1, \dots, f_\ell \in L^\infty(\mu)$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{\ell} T^{i[p(n)]} f_i = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{\ell} T^{in} f_i \tag{2.9}$$

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

This theorem, via Furstenberg's correspondence principle and Furstenberg's ergodic Szemerédi theorem, immediately implies the following Szemerédi type result:

THEOREM 2.11 ([F1]). *Let $p \in \mathbb{R}[t]$ be different from $cq+d$ for all $c, d \in \mathbb{R}$ and $q \in \mathbb{Q}[t]$. Then for each $\ell \in \mathbb{N}$, every set $E \subseteq \mathbb{N}$ with $\bar{d}(E) > 0$ contains arithmetic progressions of the form*

$$\{m, m + [p(n)], m + 2[p(n)], \dots, m + \ell[p(n)]\}$$

for some $m \in \mathbb{Z}$ and $n \in \mathbb{N}$ with $[p(n)] \neq 0$.

We will show the respective versions of these two last results along primes.

THEOREM 2.12. *Let $q \in \mathbb{R}[t]$ be different from $c\tilde{q} + d$ for all $c, d \in \mathbb{R}$ and $\tilde{q} \in \mathbb{Q}[t]$. Then for all $\ell \in \mathbb{N}$, all systems $(X, \mathcal{B}, \mu, T)$ and all $f_1, \dots, f_\ell \in L^\infty(\mu)$, we have*

$$\lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p \in \mathbb{P}\cap[1,N]} \prod_{i=1}^{\ell} T^{i[q(p)]} f_i = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{i=1}^{\ell} T^{in} f_i$$

with convergence in $L^2(\mu)$, where $\pi(N) = |\mathbb{P} \cap [1, N]|$ denotes the number of primes up to N.

THEOREM 2.13. *Let $q \in \mathbb{R}[t]$ be different from $c\tilde{q} + d$ for all $c, d \in \mathbb{R}$ and $\tilde{q} \in \mathbb{Q}[t]$. Then for each $\ell \in \mathbb{N}$, every set $E \subseteq \mathbb{N}$ with $\bar{d}(E) > 0$ contains arithmetic progressions of the form*

$$\{m, m + [q(p)], m + 2[q(p)], \ldots, m + \ell[q(p)]\}$$

for some $m \in \mathbb{Z}$ and $p \in \mathbb{P}$ with $[q(p)] \neq 0$.

2.1.2. *Several sequences.* We also get the corresponding result and applications of Theorem 2.1 along primes:

THEOREM 2.14. *Let $\ell \in \mathbb{N}$, and let $p_1, \ldots, p_\ell \in \mathbb{R}[t]$ be strongly independent real polynomials, $(X, \mathcal{B}, \mu, T)$ an ergodic system and $f_1, \ldots, f_\ell \in L^\infty(\mu)$. Then*

$$\lim_{N \to \infty} \frac{1}{\pi(N)} \sum_{p \in \mathbb{P} \cap [1,N]} T^{[p_1(p)]} f_1 \cdot \ldots \cdot T^{[p_\ell(p)]} f_\ell = \prod_{i=1}^{\ell} \int f_i \, d\mu$$

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

Theorem 2.14 has the following implications, analogous to the ones that Theorem 2.1 has.

THEOREM 2.15. *Let $\ell \in \mathbb{N}$ and let $p_1, \ldots, p_\ell \in \mathbb{R}[t]$ be strongly independent real polynomials. Then for every system $(X, \mathcal{B}, \mu, T)$ and $A \in \mathcal{B}$ we have*

$$\lim_{N \to \infty} \frac{1}{\pi(N)} \sum_{p \in \mathbb{P} \cap [1,N]} \mu(A \cap T^{-[p_1(p)]} A \cap \cdots \cap T^{-[p_\ell(p)]} A) \geq (\mu(A))^{\ell+1}.$$

THEOREM 2.16. *Let $\ell \in \mathbb{N}$ and let $p_1, \ldots, p_\ell \in \mathbb{R}[t]$ be strongly independent real polynomials. Then for every $E \subseteq \mathbb{N}$ we have*

$$\liminf_{N \to \infty} \frac{1}{\pi(N)} \sum_{p \in \mathbb{P} \cap [1,N]} \bar{d}\big(E \cap (E - [p_1(p)]) \cap \cdots \cap (E - [p_\ell(p)])\big) \geq (\bar{d}(E))^{\ell+1}.$$

THEOREM 2.17. *Let $\ell \in \mathbb{N}$ and let $p_1, \ldots, p_\ell \in \mathbb{R}[t]$ be strongly independent real polynomials. Then every set $E \subseteq \mathbb{N}$ with $\bar{d}(E) > 0$ contains arithmetic configurations of the form*

$$\{m, m + [p_1(p)], m + [p_2(p)], \ldots, m + [p_\ell(p)]\}$$

for some $m \in \mathbb{Z}$ and $p \in \mathbb{P}$ with $[p_i(p)] \neq 0$ for all $1 \leq i \leq \ell$.

2.2. Recurrence along shifted primes. In this last subsection we get some recurrence results along shifted primes for the polynomial families of interest. Namely, we show the following:

THEOREM 2.18. *Let $\ell \in \mathbb{N}$ and $p \in \mathbb{R}[t]$ different from $cq + d$ for all $c, d \in \mathbb{R}$ and $q \in \mathbb{Q}[t]$. Then, for every system $(X, \mathcal{B}, \mu, T)$ and $A \in \mathcal{B}$ with*

$\mu(A) > 0$, the set of integers n such that

$$\mu(A \cap T^{-[p(n)]}A \cap T^{-2[p(n)]}A \cap \cdots \cap T^{-\ell[p(n)]}A) > 0$$

has non-empty intersection with $\mathbb{P} - 1$ and with $\mathbb{P} + 1$.

THEOREM 2.19. *Let $\ell \in \mathbb{N}$ and let $p_1, \ldots, p_\ell \in \mathbb{R}[t]$ be strongly independent real polynomials. Then, for every system $(X, \mathcal{B}, \mu, T)$ and $A \in \mathcal{B}$ with $\mu(A) > 0$, the set of integers n such that*

$$\mu(A \cap T^{-[p_1(n)]}A \cap \cdots \cap T^{-[p_\ell(n)]}A) > 0$$

has non-empty intersection with $\mathbb{P} - 1$ and with $\mathbb{P} + 1$.

Via Furstenberg's correspondence principle the previous results imply:

THEOREM 2.20. *Let $\ell \in \mathbb{N}$, let $p \in \mathbb{R}[t]$ be different from $cq + d$ for all $c, d \in \mathbb{R}$ and $q \in \mathbb{Q}[t]$, and let $E \subseteq \mathbb{N}$ with $\bar{d}(E) > 0$. Then the set of integers n such that*

$$\bar{d}\big(E \cap (E - [p(n)]) \cap (E - 2[p(n)]) \cap \cdots \cap (E - \ell[p(n)])\big) > 0$$

has non-empty intersection with $\mathbb{P} - 1$ and with $\mathbb{P} + 1$.

THEOREM 2.21. *Let $\ell \in \mathbb{N}$, let $p_1, \ldots, p_\ell \in \mathbb{R}[t]$ be strongly independent real polynomials and let $E \subseteq \mathbb{N}$ with $\bar{d}(E) > 0$. Then the set of integers n such that*

$$\bar{d}\big(E \cap (E - [p_1(n)]) \cap \cdots \cap (E - [p_\ell(n)])\big) > 0$$

has non-empty intersection with $\mathbb{P} - 1$ and with $\mathbb{P} + 1$.

REMARK. The arguments will show that the intersections of Theorems 2.18 and 2.19 have positive measure for a set of positive relative density in the shifted primes (and so the same holds for the conclusions of Theorems 2.20 and 2.21 as well).

We close this section with the remark that we believe all the reformulations of the results stated in this section for a single transformation hold for several commuting transformations but the method we use does not allow us to prove any of them in this more general setting.

3. Background material

3.1. Nilmanifolds. In this subsection we recall some basic facts concerning nilmanifolds and equidistribution results on them.

3.1.1. *Definitions and basic properties.* Let G be a k-step nilpotent Lie group, meaning $G_{k+1} = \{e\}$ for some $k \in \mathbb{N}$, where $G_k = [G, G_{k-1}]$ denotes the kth commutator subgroup, and let Γ be a discrete cocompact subgroup of G. The compact homogeneous space $X = G/\Gamma$ is called a *k-step nilmanifold* (or just *nilmanifold*).

The group G acts on G/Γ by left translations where the translation by an element $b \in G$ is given by $T_b(g\Gamma) = (bg)\Gamma$. We denote by m_X the normalized *Haar measure* on X, the unique probability measure that is invariant under the action of G by left translations. Let $\mathcal{G}/\Gamma$ denote the Borel σ-algebra of G/Γ. If $b \in G$, we call the system $(G/\Gamma, \mathcal{G}/\Gamma, m_X, T_b)$ a *k-step nilsystem* (or just *nilsystem*) and the elements of G *nilrotations*.

3.1.2. *Equidistribution on nilmanifolds.* Let $\exp : \mathfrak{g} \to G$ be the exponential map, where $\mathfrak{g}$ is the Lie algebra of G for a connected and simply connected Lie group G. For $b \in G$ and $s \in \mathbb{R}$ we define the element b^s of G as follows: If $Z \in \mathfrak{g}$ is such that $\exp(Z) = b$, then $b^s = \exp(sZ)$ (this is well defined since exp is a bijection).

If $(a(n))_n$ is a sequence of real numbers and $X = G/\Gamma$ is a nilmanifold with G connected and simply connected, we say that the sequence $(b^{a(n)}x)_n$ is *equidistributed* in a sub-nilmanifold Y of X if for every $F \in C(Y)$ we have

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} F(b^{a(n)}x) = \int F\, dm_Y.$$

If the sequence $(a(n))_n$ takes only integer values, we need not assume that G is connected and simply connected.

A nilrotation $b \in G$ is *ergodic*, or *acts ergodically* on X, if the sequence $(b^n\Gamma)_n$ is dense in X. If $b \in G$ is ergodic, then for every $x \in X$ the sequence $(b^n x)_n$ is equidistributed in X (a non-trivial fact which follows from unique ergodicity).

Let $X = G/\Gamma$ be a nilmanifold and $b \in G$. Then the orbit closure $\overline{(b^n\Gamma)_n}$ has the structure of a nilmanifold. Furthermore, the sequence $(b^n\Gamma)_n$ is equidistributed in $\overline{(b^n\Gamma)_n}$. If G is connected and simply connected and $b \in G$, then $\overline{(b^s\Gamma)_{s\in\mathbb{R}}}$ is a nilmanifold. Furthermore, the nilflow $(b^s\Gamma)_{s\in\mathbb{R}}$ is equidistributed in $\overline{(b^s\Gamma)_{s\in\mathbb{R}}}$.

If G is a nilpotent group, then a sequence $g : \mathbb{N} \to G$ of the form $g(n) = b_1^{p_1(n)} \cdots b_k^{p_k(n)}$, where $b_i \in G$ and p_i are polynomials taking integer values at the integers for every $1 \le i \le k$, is called a *polynomial sequence* in G. A polynomial sequence on the nilmanifold $X = G/\Gamma$ is a sequence of the form $(g(n)\Gamma)_n$ where $g : \mathbb{N} \to G$ is a polynomial sequence in G.

The following qualitative equidistribution result was established by Leibman [L1]:

THEOREM 3.1 ([L1, Theorems B, C]). *Suppose that $X = G/\Gamma$ is a nilmanifold with G connected and simply connected and $(g(n))_n$ is a polynomial sequence in G. Let $Z = G/([G,G]\Gamma)$ and $\pi : X \to Z$ be the natural projection. Then:*

(i) *For every* $x \in X$ *the sequence* $(g(n)x)_n$ *is equidistributed in a finite union of subnilmanifolds of* X.
(ii) *For every* $x \in X$ *the sequence* $(g(n)x)_n$ *is equidistributed in* X *if and only if* $(g(n)\pi(x))_n$ *is equidistributed in* Z.

If $X = G/\Gamma$ is a nilmanifold with G connected and simply connected, then Z is a connected compact abelian Lie group, hence a torus $\mathbb{T}^s$ for some $s \in \mathbb{N}$, and as a consequence every nilrotation in Z is isomorphic to a rotation on $\mathbb{T}^s$.

3.2. Ergodic theory. Below we gather some basic notions and facts from ergodic theory that we use throughout the paper.

3.2.1. *Factors.* A *homomorphism* from a system $(X, \mathcal{X}, \mu, T)$ onto a system $(Y, \mathcal{Y}, \nu, S)$ is a measurable map $\pi : X \to Y$ such that $\mu \circ \pi^{-1} = \nu$ and $S \circ \pi(x) = \pi \circ T(x)$ for $x \in X$. When we have such a homomorphism, we say that the system $(Y, \mathcal{Y}, \nu, S)$ is a *factor* of $(X, \mathcal{X}, \mu, T)$. If the factor map $\pi : X \to Y$ can be chosen to be injective, then we say that the systems $(X, \mathcal{X}, \mu, T)$ and $(Y, \mathcal{Y}, \nu, S)$ are *isomorphic*. A factor can also be characterized by $\pi^{-1}(\mathcal{Y})$, which is a T-invariant sub-σ-algebra of $\mathcal{X}$. By a classical abuse of notation we denote by the same letter the σ-algebras $\mathcal{Y}$ and $\pi^{-1}(\mathcal{Y})$.

3.2.2. *Characteristic factors.* Let $(X, \mathcal{X}, \mu, T)$ be a system. We say that the σ-algebra $\mathcal{Y}$ of $\mathcal{X}$ is a *characteristic factor* for the family $\{(a_1(n))_n, \ldots, (a_\ell(n))_n\}$ of integer sequences if $\mathcal{Y}$ is T-invariant and the difference

$$\frac{1}{N}\sum_{n=1}^{N} T^{a_1(n)} f_1 \cdot \ldots \cdot T^{a_\ell(n)} f_\ell - \frac{1}{N}\sum_{n=1}^{N} T^{a_1(n)} \tilde{f}_1 \cdot \ldots \cdot T^{a_\ell(n)} \tilde{f}_\ell$$

goes to 0 in $L^2(\mu)$ norm as $N \to \infty$, where $\tilde{f}_i = \mathbb{E}(f_i|\mathcal{Y})$, $f_i \in L^\infty(\mu)$ for all $1 \le i \le \ell$ (5).

3.2.3. *Seminorms and nilfactors.* Following [HK] and [CFH] we inductively define the seminorms $\|\!|\cdot|\!\|_k$. More specifically, our definition comes from [HK] (in the ergodic case), [CFH] (in the general case) and the use of von Neumann's ergodic theorem.

Let $(X, \mathcal{B}, \mu, T)$ be a system and $f \in L^\infty(\mu)$. We define inductively the seminorms $\|\!|f|\!\|_k$ as follows: For $k = 1$ we set

$$\|\!|f|\!\|_1 := \|\mathbb{E}(f|\mathcal{I}(T))\|_{L^2(\mu)}.$$

(5) Equivalently, $\lim_{N\to\infty} \|N^{-1}\sum_{n=1}^{N} T^{a_1(n)} f_1 \cdot \ldots \cdot T^{a_\ell(n)} f_\ell\|_{L^2(\mu)} = 0$ if $\mathbb{E}(f_i|\mathcal{Y}) = 0$ for some $1 \le i \le \ell$.

For $k \geq 1$, we let

$$\|\!|f|\!\|_{k+1}^{2^{k+1}} := \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \|\!|\bar{f}\cdot T^n f|\!\|_k^{2^k}.$$

It was shown in [HK] that for every $k \geq 1$ all these limits exist and $\|\!|\cdot|\!\|_k$ defines a seminorm on $L^\infty(\mu)$.

Using these seminorms we can construct factors $\mathcal{Z}_k = \mathcal{Z}_k(T)$ of X characterized by the property that

$$\text{for } f \in L^\infty(\mu),\ \mathbb{E}(f|\mathcal{Z}_{k-1}) = 0 \text{ if and only if } \|\!|f|\!\|_k = 0.$$

It was also shown in [HK] that for every $k \in \mathbb{N}$ the factor $\mathcal{Z}_k$ has an algebraic structure, in fact we can assume that it is a k-step nilsystem. This is the content of the following structure theorem, which we recall in the ergodic case and which follows from [HK, Theorem 10.1]:

THEOREM 3.2 (Host & Kra, [HK]). *Let $(X, \mathcal{B}, \mu, T)$ be an ergodic system and $k \in \mathbb{N}$. Then the factor $\mathcal{Z}_k(T)$ is an inverse limit of k-step nilsystems* ([6]).

Because of this result we call $\mathcal{Z}_k$ the *k-step nilfactor* of the system. The smallest factor that is an extension of all finite step nilfactors is denoted by $\mathcal{Z} = \mathcal{Z}(T)$, that is, $\mathcal{Z} = \bigvee_{k\in\mathbb{N}} \mathcal{Z}_k$, and is called the *nilfactor* of the system.

4. Equidistribution results. In this section we establish some equidistribution results in order to prove the convergence and recurrence results stated in Section 2. To obtain the equidistribution results we follow the main strategy introduced in [F2, Sections 4 and 5].

First, we give an equidistribution result involving nil-orbits of several sequences of strongly independent polynomials (first proved for Hardy field functions [F2, Theorem 1.3]).

THEOREM 4.1. *Let $\ell \in \mathbb{N}$ and let $p_1, \ldots, p_\ell \in \mathbb{R}[t]$ be strongly independent real polynomials.*

(i) *If $X_i = G_i/\Gamma_i$, $1 \leq i \leq \ell$, are nilmanifolds with G_i connected and simply connected, then for any $b_i \in G_i$ and $x_i \in X_i$ the sequence*

$$(b_1^{p_1(n)}x_1, \ldots, b_\ell^{p_\ell(n)}x_\ell)_n$$

is equidistributed in the nilmanifold

$$\overline{(b_1^s x_1)}_{s\in\mathbb{R}} \times \cdots \times \overline{(b_\ell^s x_\ell)}_{s\in\mathbb{R}}.$$

([6]) By this we mean that there exist T-invariant sub-σ-algebras $\mathcal{Z}_{k,i}$, $i \in \mathbb{N}$, of $\mathcal{B}$ such that $\mathcal{Z}_k = \bigcup_{i\in\mathbb{N}} \mathcal{Z}_{k,i}$ and for every $i \in \mathbb{N}$ the factors induced by the σ-algebras $\mathcal{Z}_{k,i}$ are isomorphic to k-step nilsystems.

(ii) *If $X_i = G_i/\Gamma_i$, $1 \le i \le \ell$, are nilmanifolds, then for any $b_i \in G_i$ and $x_i \in X_i$ the sequence*

$$(b_1^{[p_1(n)]}x_1, \dots, b_\ell^{[p_\ell(n)]}x_\ell)_n$$

is equidistributed in the nilmanifold

$$\overline{(b_1^n x_1)_n} \times \dots \times \overline{(b_\ell^n x_\ell)_n}.$$

REMARKS. (i) To prove Theorem 4.1, we can assume that $X_1 = \dots = X_\ell = X$.

Indeed, in the general case we consider the nilmanifold $\tilde{X} = X_1 \times \dots \times X_\ell$. Then $\tilde{X} = \tilde{G}/\tilde{\Gamma}$, where $\tilde{G} = G_1 \times \dots \times G_\ell$ is connected and simply connected and $\tilde{\Gamma} = \Gamma_1 \times \dots \times \Gamma_\ell$ is a discrete cocompact subgroup of $\tilde{G}$. Each b_i can be considered as an element of $\tilde{G}$ and each x_i as an element of $\tilde{X}$.

(ii) If $X = G/\Gamma$ is a nilmanifold, since for every $b \in G$ the nil-orbit $(b^n\Gamma)_n$ is equidistributed in $X_b = \overline{\{b^n\Gamma : n \in \mathbb{N}\}}$, using the fact that $b^n g = g(g^{-1}bg)^n$ we deduce that $(b^n g\Gamma)_n$ is equidistributed in $g \cdot X_{g^{-1}bg}$ (in case G is connected and simply connected a similar formula holds where $n \in \mathbb{N}$ is replaced by $s \in \mathbb{R}$ and the nilmanifold X_b with $Y_b = \overline{\{b^s\Gamma : s \in \mathbb{R}\}}$). Because of this (which is called the *change of base point formula*), changing the base point we can assume that $x = \Gamma$ in the previous statement.

The following lemma shows that part (ii) of Theorem 4.1 follows from (i).

LEMMA 4.2 ([F2, Lemma 5.1]). *Let $\ell \in \mathbb{N}$ and let $(a_1(n))_n, \dots, (a_\ell(n))_n$ be sequences of real numbers. Suppose that for every nilmanifold $X = G/\Gamma$ with G connected and simply connected and all $b_1, \dots, b_\ell \in G$ the sequence*

$$(b_1^{a_1(n)}\Gamma, \dots, b_\ell^{a_\ell(n)}\Gamma)_n$$

is equidistributed in the nilmanifold

$$\overline{(b_1^s\Gamma)_{s\in\mathbb{R}}} \times \dots \times \overline{(b_\ell^s\Gamma)_{s\in\mathbb{R}}}.$$

Then for every nilmanifold $X = G/\Gamma$ and all $b_1, \dots, b_\ell \in G$ and $x_1, \dots, x_\ell \in X$ the sequence

$$(b_1^{[a_1(n)]}x_1, \dots, b_\ell^{[a_\ell(n)]}x_\ell)_n$$

is equidistributed in the nilmanifold

$$\overline{(b_1^n x_1)_n} \times \dots \times \overline{(b_\ell^n x_\ell)_n}.$$

Next we give a result needed to prove Theorem 4.1(i).

PROPOSITION 4.3. *Let $\ell \in \mathbb{N}$ and let $p_1, \dots, p_\ell \in \mathbb{R}[t]$ be strongly independent real polynomials. Let $X = G/\Gamma$ be a nilmanifold with G connected and simply connected and assume that $b_1, \dots, b_\ell \in G$ act ergodically on X. Then the sequence*

$$(b_1^{p_1(n)}\Gamma, \dots, b_\ell^{p_\ell(n)}\Gamma)_n$$

is equidistributed in the nilmanifold X^ℓ.

Proof. First notice that $(b_1^{p_1(n)},\dots,b_\ell^{p_\ell(n)})_n$ is a polynomial sequence in G^ℓ. Indeed, if $p(t)=a_dt^d+\dots+a_1t+a_0$, then $b^{p(n)}$ can be written as $(b^{a_d})^{n^d}\cdots(b^{a_1})^n b^{a_0}$. Since $X^\ell=G^\ell/\Gamma^\ell$ with G^ℓ connected and simply connected, we can apply Theorem 3.1, so in order to prove that $(b_1^{p_1(n)}\Gamma,\dots,b_\ell^{p_\ell(n)}\Gamma)_n$ is equidistributed in G^ℓ it suffices to show that $(\pi(b_1^{p_1(n)}\Gamma),\dots,\pi(b_\ell^{p_\ell(n)}\Gamma))_n$ is equidistributed in Z^ℓ, where $Z=G/[G,G]\Gamma$ and $\pi:X\to Z$ is the natural projection.

Since G is connected and simply connected, Z is isomorphic to some finite-dimensional torus $\mathbb{T}^s$ and every nilrotation in Z is isomorphic to a rotation on $\mathbb{T}^s$. Hence, for every $1\le i\le\ell$ we have $\pi(b_i\Gamma)=(\beta_{i,1}\mathbb{Z},\dots,\beta_{i,s}\mathbb{Z})$, where $\beta_{i,j}\in\mathbb{R}$ and $(\beta_{i,1},\dots,\beta_{i,s})$ is the projection of b_i on $\mathbb{T}^s$ (note that the s is bounded by the dimension of X). Since every b_i acts ergodically on X, for each $i\in[1,\ell]$ the set $\{1,\beta_{i,1},\dots,\beta_{i,s}\}$ of real numbers is rationally independent. Also, for every $t\in\mathbb{R}$ and $1\le i\le\ell$ we have $\pi(b_i^t\Gamma)=(t\beta'_{i,1}\mathbb{Z},\dots,t\beta'_{i,s}\mathbb{Z})$ for some $\beta'_{i,j}\in\mathbb{R}$ with $\beta'_{i,j}\mathbb{Z}=\beta_{i,j}\mathbb{Z}$, and so $\pi(b_i^{p_i(n)}\Gamma)=(p_i(n)\beta'_{i,1}\mathbb{Z},\dots,p_i(n)\beta'_{i,s}\mathbb{Z})$. Note that the set $\{1,\beta'_{i,1},\dots,\beta'_{i,s}\}$ is also rationally independent for each i.

Our objective now is to establish the equidistribution of the sequence

$$\big((p_1(n)\beta'_{1,1}\mathbb{Z},\dots,p_1(n)\beta'_{1,s}\mathbb{Z},\dots,p_\ell(n)\beta'_{\ell,1}\mathbb{Z},\dots,p_\ell(n)\beta'_{\ell,s}\mathbb{Z})\big)_n$$

on $\mathbb{T}^{\ell s}$. To verify this we use Weyl's criterion (see [W]; for a reference in English and an extensive study of uniform distribution see also [KN]).

Let $\mathbf{h}=(h_{1,1},\dots,h_{1,s},\dots,h_{\ell,1},\dots,h_{\ell,s})\in\mathbb{Z}^{\ell s}\setminus\{(0,\dots,0)\}$. By rational independence, not all the sums $\sum_{j=1}^s h_{i,j}\beta'_{i,j}$, $1\le i\le\ell$, are zero, so by the strong independence of $\{p_1,\dots,p_\ell\}$ the polynomial $\sum_{i=1}^\ell(\sum_{j=1}^s h_{i,j}\beta'_{i,j})p_i(n)$ has at least one non-constant irrational coefficient. Hence

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N \mathrm{e}\big(\mathbf{h}\cdot(p_1(n)\beta'_{1,1},\dots,p_1(n)\beta'_{1,s},\dots,p_\ell(n)\beta'_{\ell,1},\dots,p_\ell(n)\beta'_{\ell,s})\big)$$
$$=\lim_{N\to\infty}\frac{1}{N}\sum_{n=1}^N \mathrm{e}\Big(\sum_{i=1}^\ell\Big(\sum_{j=1}^s h_{i,j}\beta'_{i,j}\Big)p_i(n)\Big)=0.$$

By Weyl's equidistribution criterion, the result follows. ∎

The last ingredient in proving Theorem 4.1(i) is:

LEMMA 4.4 ([F2, Lemma 5.2]). *Let $X=G/\Gamma$ be a nilmanifold with G connected and simply connected. Then for any $b_1,\dots,b_\ell\in G$ there exists $s_0\in\mathbb{R}$ such that $b_i^{s_0}$ acts ergodically on the nilmanifold $\overline{(b_i^s\Gamma)}_{s\in\mathbb{R}}$ for all $1\le i\le\ell$.*

We are now ready to prove Theorem 4.1.

Proof of Theorem 4.1. Using Lemma 4.2 we see that (ii) follows from (i). To establish (i) let $b_1, \ldots, b_\ell \in G$. By Lemma 4.4 there exists a non-zero $s_0 \in \mathbb{R}$ such that for every $1 \leq i \leq \ell$ the element $b_i^{s_0}$ acts ergodically on $\overline{(b_i^s \Gamma)}_{s\in\mathbb{R}}$. Using Proposition 4.3 for the elements $b_i^{s_0}$ and the polynomials $p_i(s)/s_0$ (which trivially are still strongly independent) we find that the sequence $(b_1^{p_1(n)}\Gamma, \ldots, b_\ell^{p_\ell(n)}\Gamma)_n$ is equidistributed in $\overline{(b_1^s\Gamma)}_{s\in\mathbb{R}} \times \cdots \times \overline{(b_\ell^s\Gamma)}_{s\in\mathbb{R}}$ and the conclusion follows. ∎

5. Proof of main results. In this last section we prove our main results, Theorems 2.1, 2.12 and 2.14.

We first give the proof of Theorem 2.1. In order to do so we recall from [F1] that the nilfactor $\mathcal{Z}$ of a system is characteristic for the family $\{p_1, \ldots, p_\ell\}$, where $p_1, \ldots, p_\ell \in \mathbb{R}[t]$ are strongly independent real polynomials. Actually, the statement in [F1] is about *nice* families of polynomials (see definition below), a notion more general than the strong independence that we use here.

Definition ([F1]). Let $\ell \in \mathbb{N}$ and for $N \in \mathbb{N}$, let $\mathcal{P}_N = \{p_{1,N}, \ldots, p_{\ell,N}\}$ be a family of polynomials with real coefficients. We say that the collection $(P_N)_N$ is *nice* if for every $N \in \mathbb{N}$ the polynomials $p_{i,N}$ and $p_{i,N} - p_{j,N}$, $i \neq j$, are non-constant and their leading coefficients are independent of N.

The following lemma shows that for a nice collection of polynomial families the nilfactor is the characteristic factor as well (a different proof of this fact is also given in [L2]).

Lemma 5.1 ([F1, Lemma 4.7]). *Let $(\{p_{1,N}, \ldots, p_{\ell,N}\})_N$ be a nice collection of polynomial families and let $(X, \mathcal{B}, \mu, T)$ be a system, and suppose that one of the functions $f_1, \ldots, f_\ell \in L^\infty(\mu)$ is orthogonal to the nilfactor $\mathcal{Z}$. Then for any Følner sequence $(\Phi_N)_N$ in $\mathbb{Z}$ (7) and any bounded two-parameter sequence $(c_{N,n})_{N,n}$ of real numbers we have*

$$\lim_{N\to\infty} \frac{1}{|\Phi_N|} \sum_{n\in\Phi_N} c_{N,n} T^{[p_{1,N}(n)]} f_1 \cdot \ldots \cdot T^{[p_{\ell,N}(n)]} f_\ell = 0 \tag{5.1}$$

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

Remark. For $\ell \in \mathbb{N}$, let $\{p_1, \ldots, p_\ell\}$ be a strongly independent family of polynomials. Then trivially this collection is nice, so we have (5.1), hence the nilfactor $\mathcal{Z}$ is the characteristic factor for this family.

For the sake of completeness, and since the lemma is crucial for our study, we will present most of the details of the proof in the special setting where

(7) A *Følner sequence in* $\mathbb{Z}$ is a sequence $(\Phi_n)_n$ of finite subsets of $\mathbb{Z}$ such that for any $m \in \mathbb{Z}$ we have $\lim_{n\to\infty} |(\Phi_n + m) \cap \Phi_n|/|\Phi_n| = 1$.

$p_{i,N} = p_i$, $c_{N,n} = 1$, and $\Phi_N = \{1, \ldots, N\}$ for $N \in \mathbb{N}$ (which is a Følner sequence in $\mathbb{N}$). Note that in this case that the sequences of polynomials are constant the notion of "nice" polynomials coincides with the notion of "essentially distinct" polynomials [B2].

Sketch of proof of special case of Lemma 5.1. We follow the argument in [F1, Lemma 4.7]. Without loss of generality we assume that f_1 is orthogonal to $\mathcal{Z}$, $\|f_i\|_\infty \le 1$ for all $i \in [1, \ell]$, and p_1 is a polynomial of maximum degree in $\mathcal{P} = \{p_1, \ldots, p_\ell\}$. Under these assumptions it suffices to show that

$$\lim_{N\to\infty} \sup_{\|f_0\|_\infty, \|f_2\|_\infty, \ldots, \|f_\ell\|_\infty \le 1} \frac{1}{N}\sum_{n=1}^{N} \Big| \int f_0 \cdot \prod_{i=1}^{\ell} T^{[p_i(n)]} f_i \, d\mu \Big| = 0. \tag{5.2}$$

To any family of (finitely many) polynomials, we assign the vector $(d, w_d, \ldots, w_1)$, where d is the maximum degree of the polynomials and w_i the number of distinct coefficients of polynomials of degree i in the family. We call this vector the *type* of the family. We order the types lexicographically. We will show (5.2) by induction on the type of $\{p_1, \ldots, p_\ell\}$.

The case where $d = \deg p_1 = 1$ can be treated as in [F1, Proposition 5.3] (or by using [F4, Lemmas 4.11 and 4.13]).

Let $d = \deg p_1 \ge 2$ and suppose that the statement holds for nice polynomial families of type smaller than $(d, w_d, \ldots, w_1)$. Using the Cauchy–Schwarz inequality, (5.2) will follow if we show

$$\lim_{N\to\infty} \sup_{\|f_0\|_\infty, \|f_2\|_\infty, \ldots, \|f_\ell\|_\infty \le 1} \frac{1}{N}\sum_{n=1}^{N} \Big| \int f_0 \cdot T^{[p_1(n)]} f_1 \cdot \ldots \cdot T^{[p_\ell(n)]} f_\ell \, d\mu \Big|^2 = 0.$$

So, if $\tilde{T} = T \times T$, $\tilde{f} = f \otimes \bar{f}$ and $\tilde{\mu} = \mu \times \mu$, using Cauchy–Schwarz it suffices to show

$$\lim_{N\to\infty} \sup_{\|\tilde{f}_2\|_\infty, \ldots, \|\tilde{f}_\ell\|_\infty \le 1} \Big\| \frac{1}{N}\sum_{n=1}^{N} \tilde{T}^{[p_1(n)]} \tilde{f}_1 \cdot \ldots \cdot \tilde{T}^{[p_\ell(n)]} \tilde{f}_\ell \Big\|_{L^2(\tilde{\mu})} = 0. \tag{5.3}$$

Using a variation of the classical van der Corput lemma (see [F1, Lemma 4.6]), setting $b(n_1, \ldots, n_\ell) = \tilde{T}^{n_1} \tilde{f}_1 \cdot \ldots \cdot \tilde{T}^{n_\ell} \tilde{f}_\ell$, we get (5.3) once we show that for large enough h the supremum over $\|\tilde{f}_2\|_\infty, \ldots, \|\tilde{f}_\ell\|_\infty \le 1$ of

$$\Big| \frac{1}{N}\sum_{n=1}^{N} \int b([p_1(n+h)], \ldots, [p_\ell(n+h)]) \overline{b([p_1(n)], \ldots, [p_\ell(n)])} \, d\tilde{\mu} \Big|$$

goes to 0 as $N \to \infty$. We factor out the transformation $\tilde{T}^{[p(n)]}$, where $p = p_{i_0}$ for some $1 \le i_0 \le \ell$ is chosen so that for every large h the polynomial subfamily $\mathcal{P}(p, h)$ which is obtained from $\{p_1(n+h) - p(n), \ldots, p_\ell(n+h) - p(n), p_1(n) - p(n), \ldots, p_\ell(n) - p(n)\}$ after successfully removing the smallest number of polynomials so that the resulting family consists of non-constant, essentially distinct polynomials, has type smaller than $\{p_1(n),\ldots,p_\ell(n)\}$ (see [F1, Lemma 4.5]). We write $[p_i(n+h)]-[p(n)] = [p_i(n+h)-p(n)]+e_{1,i}(h,n)$ and $[p_i(n)]-[p(n)] = [p_i(n)-p(n)]+e_{2,i}(h,n)$, for all $1\le i\le \ell$, where $e_{j,i}(h,n)$ are error terms with values in $\{0,1\}$. For every fixed h we partition the integers into a finite number of sets, that depend only on ℓ, where all the error terms are constant. Therefore, it suffices to show that

(5.4)
$$\lim_{N\to\infty}\sup_{\|\tilde f_2\|_\infty,\ldots,\|\tilde f_\ell\|_\infty\le 1}\frac{1}{N}\sum_{n=1}^{N}\Big|\int b([p_1(n+h)-p(n)],\ldots,[p_\ell(n+h)-p(n)]) \cdot \overline{b([p_1(n)-p(n)],\ldots,[p_\ell(n)-p(n)])}\,d\tilde\mu\Big| = 0.$$

If some p_i happens to have degree 1 (hence $i\neq 1$), we write $p_i(n+h) = p_i(n)+c(h)$, where $c(h)=p_i(h)-p_i(0)\in\mathbb{R}$. Hence, in this case,

$$\tilde T^{[p_i(n+h)-p(n)]}\tilde f_i\cdot \overline{\tilde T^{[p_i(n)-p(n)]}\tilde f_i} = \tilde T^{[p_i(n)-p(n)]}(\tilde T^{c(h)+e(h,n)}\tilde f_i\cdot \overline{\tilde f_i})$$

for some $e(h,n)\in\{0,1\}$. Since the error terms take finitely many values, to show (5.4) we can assume that $e(h,n)=0$. Therefore, it suffices to show that for large h the supremum over $\|\tilde f_0\|_\infty,\|\tilde f_2\|_\infty,\ldots,\|\tilde f_r\|_\infty\le 1$ of

$$\frac{1}{N}\sum_{n=1}^{N}\Big|\int \tilde f_0\cdot \tilde T^{[p_1(n+h)-p(n)]}\tilde f_1\cdot\prod_{i=2}^{r}\tilde T^{[\tilde p_{h,i}(n)]}\tilde f_i\,d\tilde\mu\Big| \tag{5.5}$$

goes to 0 as $N\to\infty$, for some $r\in\mathbb{N}$, where all the polynomials $\tilde p_{h,i}$ belong to the family $\mathcal{P}(p,h)$. For sufficiently large h, this family consist of essentially distinct polynomials with type less than that of $\mathcal{P}$. Since $p_1(n+h)-p(n)$ is the polynomial of maximum degree in $\mathcal{P}(p,h)$, and since f_1 is orthogonal to $\mathcal{Z}(T)$, we infer that $\tilde f_1$ is orthogonal to $\mathcal{Z}(\tilde T)$ and the result now follows by the induction hypothesis. ∎

Proof of Theorem 2.1. We start by using Lemma 5.1 to deduce that the nilfactor $\mathcal{Z}$ is characteristic for the corresponding multiple ergodic average. Via Theorem 3.2 we can assume without loss of generality that our system is an inverse limit of nilsystems. By a standard approximation argument, we can further assume that our system is a nilsystem.

Let $(X=G/\Gamma,\mathcal{G}/\Gamma,m_X,T_b)$ be a nilsystem, where $b\in G$ is ergodic, and $F_1,\ldots,F_\ell\in L^\infty(m_X)$. Our objective now is to show that if $\{p_1,\ldots,p_\ell\}$ is a strongly independent family of polynomials then

$$\lim_{N\to\infty}\sum_{n=1}^{N}F_1(b^{[p_1(n)]}x)\cdot\ldots\cdot F_\ell(b^{[p_\ell(n)]}x) = \int F_1\,dm_X\cdot\ldots\cdot\int F_\ell\,dm_X \tag{5.6}$$

with convergence in $L^2(m_X)$. By density, we can assume that $F_1,\ldots,F_\ell$

are continuous. Then we can apply Theorem 4.1 to the nilmanifold X^ℓ, the nilrotation $\tilde{b} = (b, \dots, b) \in G^\ell$, the point $\tilde{x} = (x, \dots, x) \in X^\ell$, and the continuous function $\tilde{F}(x_1, \dots, x_\ell) = F_1(x_1) \cdot \ldots \cdot F_\ell(x_\ell)$, to get

$$\lim_{N\to\infty} \sum_{n=1}^{N} \tilde{F}(b^{[p_1(n)]}x, \dots, b^{[p_\ell(n)]}x) = \int \tilde{F}\, dm_{X^\ell};$$

this gives the desired limit in (5.6), completing the proof. ∎

5.1. Convergence along primes. We first give the definitions and the main ideas in order to prove Theorems 2.12 and 2.14.

We start by recalling the definition of the *von Mangoldt function* $\Lambda : \mathbb{N} \to \mathbb{R}$:

$$\Lambda(n) = \begin{cases} \log(p) & \text{if } n = p^k \text{ for some } p \in \mathbb{P} \text{ and some } k \in \mathbb{N}, \\ 0 & \text{otherwise.} \end{cases}$$

As in [FHK2] and [K2] it is more natural for us to work instead with the function $\Lambda' : \mathbb{N} \to \mathbb{R}$, where $\Lambda'(n) = \mathbf{1}_{\mathbb{P}}(n) \cdot \Lambda(n) = \mathbf{1}_{\mathbb{P}}(n) \cdot \log(n)$.

The function Λ', according to the following lemma, will allow us to relate averages along primes to weighted averages over the integers.

LEMMA 5.2 ([FHK1]). *If $a : \mathbb{N} \to \mathbb{C}$ is bounded, then*

$$\lim_{N\to\infty} \left| \frac{1}{\pi(N)} \sum_{p\in\mathbb{P}\cap[1,N]} a(p) - \frac{1}{N}\sum_{n=1}^{N} \Lambda'(n) \cdot a(n) \right| = 0.$$

For $w > 2$, let

$$W = \prod_{p\in\mathbb{P}\cap[1,w-1]} p$$

be the product of primes bounded above by w. For $r \in \mathbb{N}$, let

$$\Lambda'_{w,r}(n) = \frac{\phi(W)}{W} \cdot \Lambda'(Wn + r)$$

be the *modified von Mangoldt function*; here ϕ is the Euler function.

DEFINITION. For $\ell \in \mathbb{N}$, we call the setting $(X, \mathcal{B}, \mu, T_1, \dots, T_\ell)$ a *system*, where $T_1, \dots, T_\ell : X \to X$ are invertible commuting measure preserving transformations on the probability space $(X, \mathcal{B}, \mu)$.

The proposition below, the proof of which relies on a deep result due to Green and Tao [GT] on the inverse conjecture for the Gowers norms, will provide us with a crucial intermediate step in order to prove Theorems 2.12 and 2.14, as well as Theorems 2.18 and 2.19 (we will actually use a very weak version of it for all these results).

Proposition 5.3 ([K2, Proposition 3.2]). *Let* $\ell, m \in \mathbb{N}$, *and let* $(X, \mathcal{B}, \mu, T_1, \ldots, T_m)$ *be a system,* $p_{i,j} \in \mathbb{R}[t]$ *for* $1 \le i \le m$ *and* $1 \le j \le \ell$, *and* $f_1, \ldots, f_\ell \in L^\infty(\mu)$. *Then*

$$\max_{1\le r\le W,\,(r,W)=1} \Big\| \frac{1}{N}\sum_{n=1}^{N} (\Lambda'_{w,r}(n)-1)\cdot \prod_{i=1}^{\ell}\Big(\prod_{j=1}^{m} T_j^{[p_{j,i}(Wn+r)]}\Big) f_i \Big\|_{L^2(\mu)}$$

converges to 0 *as* $N \to \infty$ *and then* $w \to \infty$.

Proof of Theorem 2.12. We borrow the arguments from [FHK2, proof of Theorem 1.3] (see also [K2, Theorem 1.3]). By Lemma 5.2 it suffices to show that the sequence

$$A(N) := \frac{1}{N}\sum_{n=1}^{N} \Lambda'(n)\cdot T^{[q(n)]}f_1 \cdot T^{2[q(n)]}f_2 \cdot \ldots \cdot T^{\ell[q(n)]}f_\ell$$

converges in $L^2(\mu)$ to the same limit as $N^{-1}\sum_{n=1}^{N}\prod_{i=1}^{\ell} T^{in}f_i$ as $N \to \infty$. For $w \in \mathbb{N}$ (which gives a corresponding W) and $r \in \mathbb{N}$, we define

$$B_{w,r}(N) := \frac{1}{N}\sum_{n=1}^{N} T^{[q(Wn+r)]}f_1 \cdot T^{2[q(Wn+r)]}f_2 \cdot \ldots \cdot T^{\ell[q(Wn+r)]}f_\ell.$$

For any $\varepsilon > 0$, using Proposition 5.3 with $m = \ell$, $T_i = T$, $1 \le i \le \ell$, and

$$p_{i,j} = \begin{cases} 0 & \text{if } i \le \ell - j, \\ q & \text{otherwise,} \end{cases}$$

for sufficiently large N and some w_0 we have

$$\Big\| A(W_0N) - \frac{1}{\phi(W_0)} \sum_{1\le r\le W_0,\,(r,W_0)=1} B_{w_0,r}(N) \Big\|_{L^2(\mu)} < \varepsilon.$$

Note at this point that for all $W, r \in \mathbb{N}$ we have $q(Wt + r) \notin c\mathbb{Q}[t] + d$ for $c, d \in \mathbb{R}$, for otherwise q would have the same property, contradicting our assumption.

By Theorem 2.10, for any $r \in [1, W_0]$ the sequence $(B_{w_0,r}(N))_N$ converges to the same limit as $N^{-1}\sum_{n=1}^{N}\prod_{i=1}^{\ell} T^{in}f_i$, and since

$$\lim_{N\to\infty} \|A(W_0N + r) - A(W_0N)\|_{L^2(\mu)} = 0$$

for every $r \in [1, W_0]$, we get the result. ∎

Proof of Theorem 2.14. The proof is analogous to the previous one. In this case we define

$$A(N) := \frac{1}{N}\sum_{n=1}^{N} \Lambda'(n)\cdot T^{[p_1(n)]}f_1 \cdot \ldots \cdot T^{[p_\ell(n)]}f_\ell$$

and for $w, r \in \mathbb{N}$,

$$B_{w,r}(N) := \frac{1}{N} \sum_{n=1}^{N} T^{[p_1(Wn+r)]} f_1 \cdot \ldots \cdot T^{[p_\ell(Wn+r)]} f_\ell.$$

We use Proposition 5.3 with $m = \ell$, $T_i = T$, $1 \le i \le \ell$, $p_{i,j} = \delta_{i,j} p_i$ where $\delta_{i,j}$ is Kronecker's delta, and we note that the family $\{\tilde{p}_1, \ldots, \tilde{p}_\ell\}$, where $\tilde{p}_i(t) = p_i(Wt + r)$, is strongly independent for all $W, r \in \mathbb{N}$. (Indeed, if for some $(\lambda_1, \ldots, \lambda_\ell) \in \mathbb{R}^\ell \setminus \{\vec{0}\}$, $d \in \mathbb{R}$, $q \in \mathbb{Q}[t]$ and $W, r \in \mathbb{N}$ we had $\sum_{i=1}^{\ell} \lambda_i p_i(Wt + r) = q(t) + d$, then $\sum_{i=1}^{\ell} \lambda_i p_i(t) = \tilde{q}(t) + d$, where $\tilde{q}(t) = q((t - r)/W) \in \mathbb{Q}[t]$, a contradiction to the strong independence assumption.) The result now follows similarly to the previous proof since by Theorem 2.1, the sequence $(B_{w_0,r}(N))_N$ converges in $L^2(\mu)$ to $\prod_{i=1}^{\ell} \int f_i \, d\mu$ for any $r \in [1, W_0]$. ∎

5.2. Recurrence along shifted primes. In this last subsection we prove Theorems 2.18–2.21. To show Theorem 2.18 we use the following result.

THEOREM 5.4 ([BHMP, Theorem 2.1]). *Let $\ell \in \mathbb{N}$ and let $(X, \mathcal{B}, \mu, T)$ be a system. Then for every $A \in \mathcal{B}$ with $\mu(A) > 0$ there exists a constant $c \equiv c_{\ell,\mu(A)} > 0$ such that*

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-n} A \cap T^{-2n} A \cap \cdots \cap T^{-\ell n} A) \ge c. \tag{5.7}$$

REMARK. Actually, in (5.7) the limit exists by [HK].

Proof of Theorem 2.18. Using Proposition 5.3 with $m = \ell$, $T_i = T$, $1 \le i \le \ell$, $r = 1$ and

$$p_{i,j}(n) = \begin{cases} 0 & \text{if } i \le \ell - j, \\ p(n-1) & \text{otherwise,} \end{cases}$$

and combining Theorem 5.4 with Theorem 2.10 we deduce, for sufficiently large $\omega \in \mathbb{N}$, that

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \Lambda'_{\omega,1}(n) \cdot \mu\Big(A \cap \bigcap_{i=1}^{\ell} T^{-i[p(Wn)]} A\Big) > 0,$$

from which we get the required non-empty intersection with $\mathbb{P} - 1$. ∎

Proof of Theorem 2.19. The proof is analogous to the preceding one. More specifically, we use Theorem 2.10 instead of Theorem 5.4 and Proposition 5.3 with $m = \ell$, $T_i = T$, $1 \le i \le \ell$, $r = 1$ and $p_{i,j}(n) = \delta_{i,j} p_i(n-1)$ to get, for some sufficiently large $\omega \in \mathbb{N}$,

$$\liminf_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \Lambda'_{\omega,1}(n) \cdot \mu(A \cap T^{-[p_1(Wn)]} A \cap \cdots \cap T^{-[p_\ell(Wn)]} A) > 0,$$

from which we get the result. ∎

Proof of Theorems 2.20 and 2.21. Both theorems follow immediately by using Theorems 2.18 and 2.19 together with Furstenberg's correspondence principle. ∎

REMARKS. (1) According to Lemma 5.2, we see that the conclusions of Theorems 2.18 and 2.19, and so of Theorems 2.20 and 2.21, are satisfied for a set of integers n with positive relative density in the shifted primes $\mathbb{P} - 1$ (the analogous results, by a similar argument, hold for $\mathbb{P} + 1$ as well).

(2) In the special case where $p_i(0) = 1/2$ for all $i \in [1, \ell]$, the results of Theorems 2.18–2.21 follow from [K2, Theorem 1.2].

Acknowledgements. We are greatly indebted to N. Frantzikinakis who introduced us to this problem, for his in-depth suggestions and the fruitful discussions we had during the writing of this article. The first author thanks his advisor V. Farmaki for her support during the writing of this paper, and the second author thanks V. Bergelson for his constant support. We also thank D. Glasscock for his detailed corrections. Finally, we would like to thank the anonymous referee for the suggestions and remarks.

References

[B1] V. Bergelson, *Ergodic Ramsey theory*, in: Logic and Combinatorics (Arcata, CA, 1985), Contemp. Math. 65, Amer. Math. Soc., Providence, RI, 1987, 63–87.

[B2] V. Bergelson, *Weakly mixing PET*, Ergodic Theory Dynam. Systems 7 (1987), 337–349.

[BK] V. Bergelson and I. Håland-Knutson, *Weak mixing implies mixing of higher orders along tempered functions*, Ergodic Theory Dynam. Systems 29 (2009), 1375–1416.

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

[BHMP] V. Bergelson, B. Host, R. McCutcheon and F. Parreau, *Aspects of uniformity in recurrence*, Colloq. Math. 84/85 (2000), 549–576.

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

[CFH] Q. Chu, N. Frantzikinakis and B. Host, *Ergodic averages of commuting transformations with distinct degree polynomial iterates*, Proc. London Math. Soc. (3) 102 (2011), 801–842.

[F1] N. Frantzikinakis, *Multiple recurrence and convergence for Hardy field sequences of polynomial growth*, J. Anal. Math. 112 (2010), 79–135.

[F2] N. Frantzikinakis, *Equidistribution of sparse sequences on nilmanifolds*, J. Anal. Math. 109 (2009), 353–395.

[F3] N. Frantzikinakis, *Some open problems on multiple ergodic averages*, Bull. Hellenic Math. Soc. 60 (2016), 41–90.

[F4] N. Frantzikinakis, *A multidimensional Szemerédi theorem for Hardy sequences of different growth*, Trans. Amer. Math. Soc. 367 (2015), 5653–5692.

[F5] N. Frantzikinakis, *An averaged Chowla and Elliott conjecture along independent polynomials*, Int. Math. Res. Notices 2018, 3721–3743.

[FHK1] N. Frantzikinakis, B. Host and B. Kra, *Multiple recurrence and convergence for sequences related to the prime numbers*, J. Reine Angew. Math. 611 (2007), 131–144.

[FHK2] N. Frantzikinakis, B. Host and B. Kra, *The polynomial multidimensional Szemerédi theorem along shifted primes*, Israel J. Math. 194 (2013), 331–348.

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

[Fur] H. Furstenberg, *Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions*, J. Anal. Math. 31 (1977), 204–256.

[FKO] H. Furstenberg, Y. Katznelson and D. Ornstein, *The ergodic theoretical proof of Szemerédi's theorem*, Bull. Amer. Math. Soc. 7 (1982), 527–552.

[GT] B. Green and T. Tao, *Linear equations in primes*, Ann. of Math. (2) 171 (2010), 1753–1850.

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

[K1] A. Koutsogiannis, *Integer part polynomial correlation sequences*, Ergodic Theory Dynam. Systems 38 (2018), 1525–1542.

[K2] A. Koutsogiannis, *Closest integer polynomial multiple recurrence along shifted primes*, Ergodic Theory Dynam. Systems 38 (2018), 666–685.

[KN] L. Kuipers and H. Niederreiter, *Uniform Distribution of Sequences*, Wiley, 1974.

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

[L2] A. Leibman, *Convergence of multiple ergodic averages along polynomials of several variables*, Israel J. Math. 146 (2005), 303–316.

[W] H. Weyl, *Über die Gleichverteilung von Zahlen mod Eins*, Math. Ann. 77 (1916), 313–352.

Dimitris Karageorgos
Department of Mathematics
National and Kapodistrian University of Athens
Panepistemioupolis, Athens, Greece
E-mail: dimitris_kar@math.uoa.gr

Andreas Koutsogiannis
Department of Mathematics
The Ohio State University
Columbus, OH 43210, U.S.A.
E-mail: koutsogiannis.1@osu.edu