

ERGODIC AVERAGES AND THE LARGE INTERSECTION PROPERTY ALONG IP SETS.

BRYNA KRA AND OR SHALOM

ABSTRACT. We study multiple ergodic averages along IP sets, meaning we restrict iterates in the averages to all finite sums of some infinite sequence of natural numbers. We give criteria for convergence and divergence in mean of these multiple averages and derive sufficient conditions for convergence to the projection onto the space of invariant functions. For a class of sequences that, roughly speaking, only have rational obstructions to such a limit, we show that the behavior is controlled by nilsystems. We also consider pointwise convergence, obtaining convergence and a formula for a set of functions on nilsystems that are dense in L^2 . Finally, we show that certain correlations have optimally large intersections along an IP set.

1. INTRODUCTION

1.1. IP convergence. Since Furstenberg's proof [20] of Szemerédi's Theorem [39] using ergodic theory, dynamical methods have been used to prove numerous generalizations (some examples include [22, 9, 7, 23, 24]). Among the many generalizations, we focus on the IP version of Szemerédi's theorem. An IP, short for *infinite parallelepiped* or *idempotent* depending on the context, is defined as all the finite sums of some sequence $(n_j)_{j \in \mathbb{N}}$ and an IP-set is a set containing an IP (for technical reasons, we deviate from the standard definition and allow an IP to be a multiset). Furstenberg and Katznelson [23] showed that for any IP-set, any set of natural numbers with positive upper Banach density contains arbitrary long arithmetic progressions whose common step lies in that IP. To prove this refinement, they change the type of convergence studied, introducing taking a limit along

Date: June 30, 2025.

2020 Mathematics Subject Classification. Primary 37A15, 11B30; Secondary 28D15, 37A35.

Key words and phrases. multiple ergodic average, IP-set, nilsystem.

IP-sets. This type of limit has been used in other contexts to obtain further combinatorial and ergodic generalizations, including for example in [10].

In using ergodic methods to prove these results, one is naturally led to study *multicorrelation sequences*, meaning sequences of the form

$$(1.1) \quad a(n) = \int_{\mathbf{X}} \prod_{i=0}^k f_i(T^{in}x) \, d\mu(x),$$

where $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ is an invertible measure preserving system and $f_0, \dots, f_k \in L^\infty(\mu)$ (we defer precise definitions and further background to Section 2). Taking all functions to be the indicator function of some set with positive measure, showing that the \liminf of a multicorrelation sequence is positive corresponds to proving a recurrence statement, and such results have led to generalizations of Szemerédi's theorem along subsequences. Convergence in mean has been shown for many multicorrelation sequences, including [30, 43, 25, 40]. Further refinements have been given, such as large intersections for the indicator function of a set in a multicorrelation [7, 16, 19, 15] (see also [2, 38, 37, 3, 1] for extensions to abelian groups beyond \mathbb{Z}).

For each of these types of results (recurrence, convergence, and large intersections), it is natural to ask if there are IP versions, and this is our focus. We start in Section 3 by giving necessary and sufficient conditions for mean convergence of an average along an IP. Using spectral criteria, in Theorem 3.3, we show when such an average for a function converges to its projection onto the invariant σ -algebra. This allows us to deduce sufficient conditions for convergence along an IP in Section 3.2 and an example of divergence along an IP in Section 3.3. In particular, we show that for IP sequences with rational spectrum (see Definition 3.7) we can describe the limit (Theorem 3.8), and this is the class of IPs for which we can describe more detailed behavior (Theorem 5.2).

In Section 4, we turn our attention to multiple ergodic averages along an IP. Proving convergence for these multiple averages requires the development of IP versions of the tools used for linear averages, including a version of the van der Corput Lemma (Lemma 4.3), a version of Gowers-Host-Kra seminorms (Definition 4.5), and IP cubic measures (Definition 4.6). This allows us to show (Theorem 4.2) that for sequences with rational spectrum,

the behavior of the multiple ergodic average along an IP is controlled by the same factors that control the well understood case of linear averages.

We then use the tools developed to derive several applications. In Theorem 5.2, we give a limit formula for certain multiple averages along IPs, and this in turn is used to derive a large intersection property for progressions of length three and four along an IP with rational spectrum. The ergodic version of these results are given in Theorems 6.2 and 6.4, and via variant of Furstenberg's correspondence principle for ergodic systems (see [7, Proposition 3.1]), we immediately deduce an application in the integers (see Definition 6.1 for the notion of IP density). To state the result, recall that if $E \subseteq \mathbb{Z}$, we define the *upper Banach density* $d^*(E)$ by

$$d^*(E) = \lim_{N \rightarrow \infty} \sup_{M \in \mathbb{Z}} \frac{|E \cap \{M+1, \dots, M+N\}|}{N}.$$

Theorem 1.1. *Let $(n_j)_{j \in \mathbb{N}}$ be a sequence of natural numbers with rational spectrum and let $E \subseteq \mathbb{Z}$ be a set with $d^*(E) > 0$. Then for all coprime integers $\ell_1, \ell_2 \in \mathbb{Z}$ and all $\varepsilon > 0$, we have that both*

$$d_{\text{IP}_\Phi} \left\{ n \in \text{IP}((n_j)_{j \in \mathbb{N}}) : d^*(E \cap (E - \ell_1 n) \cap (E - \ell_2 n)) > d^*(E)^3 - \varepsilon \right\}$$

and

$$d_{\text{IP}_\Phi} \left\{ n \in \text{IP}((n_j)_{j \in \mathbb{N}}) : d^*(E \cap (E - \ell_1 n) \cap (E - \ell_2 n) \cap (E - (\ell_1 + \ell_2)n)) > d^*(E)^4 - \varepsilon \right\}$$

are positive for any increasing Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$.

As a sample application (see Exmaple 2.4), this applies to the IP generated by $1, 10, 100, \dots$, which generates the set of all numbers in base 10 that can be written only using the digits 0 and 1.

1.2. Open questions. It is natural to ask if the sets of large intersections in Theorem 1.1 are large with respect to other measurements of largeness. One phrasing of this question is the following syndetic version of this question.

Question 1.3. Let $(n_j)_{j \in \mathbb{N}}$ be a sequence of natural numbers with rational spectrum, let $E \subseteq \mathbb{Z}$ be a set with positive upper Banach density $d^*(E) > 0$, and let $\varepsilon > 0$. Do finitely many translates of the set

$$\{n \in \text{IP}((n_j)_{j \in \mathbb{N}}) : d^*(E \cap (E - n) \cap (E - 2n)) > d^*(E)^3 - \varepsilon\}$$

cover $\text{IP}((n_j)_{j \in \mathbb{N}})$? Does the analogous statement hold for four term arithmetic progressions?

From the example in Theorem B.1, we know that there exist IPs for which the set of large intersections is trivial, and this leads to the next question.

Question 1.4. Are there IPs other than those with rational spectrum for which the set of large intersections in Theorem 1.1 for three (or four) sets is nonempty?

A related question on these intersections is if there is some other bound for which the large intersection property holds for all IPs. It follows from Furstenberg and Katznelson [23] that there exists a constant $C > 0$ such that

$$(1.2) \quad \{n \in \mathbb{IP}((n_j)_{j \in \mathbb{N}}) : d^*(E \cap (E - n) \cap (E - 2n)) > C\}$$

is nonempty, and it further follows from [8] that $C = C(d^*(E))$ is a constant depending only on $d^*(E)$. This leads to the next question.

Question 1.5. Let $E \subseteq \mathbb{N}$ with $d^*(E) = \delta$ for some $\delta > 0$. What is the optimal value of $C = C(\delta)$ in (1.2)?

We note that we phrase this as the optimal value, as in Theorem B.1 we show that $C(\delta) \leq \delta^{-c \log \delta}$ for some absolute constant $c > 0$.

A last question that we pose is on IP pointwise convergence.

Question 1.6. Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ be an ergodic invertible measure preserving system, let $(n_j)_{j \in \mathbb{N}}$ be a sequence with rational spectrum, and let $f \in L^\infty(\mu)$. Does the limit

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \mathbb{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} T^n f(x)$$

exist for μ -almost every $x \in X$?

While such a result is of interest on its own, it would also simplify the computations of the limit of an average along an IP in nilsystems, leading to simplifications of the proofs in Section 5.

Acknowledgements. The first author was partially supported by the National Science Foundation grant DMS-2348315. The second author was partially supported by the National Science Foundation grant DMS-1926686. We thank Ethan Ackelsberg for useful discussions during the preparation of this article and for pointing out useful references.

2. BACKGROUND AND ADAPTATIONS FOR IP SEQUENCES

2.1. Measure preserving systems. An *invertible measure preserving system* $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ is a quadruple in which (X, \mathcal{B}, μ) is a probability space and $T: X \rightarrow X$ is an invertible, measurable, measure preserving transformation (meaning that $\mu(A) = \mu(T^{-1}(A))$ for all measurable $A \subseteq X$). When the context is clear, we shorten this and refer to an invertible measure preserving system as a *system*. Throughout, we assume that all systems $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ are regular, meaning that X is a compact metric space, \mathcal{B} is the Borel- σ algebra, and μ is a regular measure.

Given a system $\mathbf{X} = (X, \mathcal{B}, \mu, T)$, let $L^2(\mu)$ denote the Hilbert space of all complex valued square integrable functions on X modulo μ -almost everywhere equivalence. We make the usual abuse of notation, using T both for the transformation on X and for the associated unitary operator $T: L^2(\mu) \rightarrow L^2(\mu)$ defined by $f \mapsto f \circ T$ for $f \in L^2(\mu)$. The measure preserving system is *ergodic* if the only T -invariant functions are constant μ -almost everywhere.

Let S^1 denote the one dimensional circle. For $t \in S^1$, the t -*eigenspace* is the closed Hilbert space consisting of all $f \in L^2(\mu)$ satisfying $Tf = t \cdot f$, and any nonzero function $f \in L^2(\mu)$ in the t -eigenspace is an *eigenfunction with eigenvalue* t . For any $t \in S^1$, we let $P_t: L^2(\mu) \rightarrow L^2(\mu)$ denote orthogonal projection onto the t -eigenspace.

2.2. Factors and cocycles. If $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ and $\mathbf{Y} = (Y, \mathcal{D}, \nu, S)$ are systems, we say that \mathbf{Y} is a *factor* of \mathbf{X} (or equivalently \mathbf{X} is an *extension* of \mathbf{Y}) if there exists a measurable map $\pi: X \rightarrow Y$ such that $\pi\mu = \nu$ and $S \circ \pi = \pi \circ T$ holds μ -almost everywhere. When the map π is also invertible with a measurable inverse, we say that the systems \mathbf{X} and \mathbf{Y} are *isomorphic*.

If $\mathbf{Y} = (Y, \mathcal{Y}, \nu, S)$ is a factor of the system $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ with factor map $\pi: X \rightarrow Y$ and $f \in L^2(\mu)$, let $\mathbb{E}(f \mid \mathbf{Y})$ denote the *conditional expectation* of f onto Y , defined as the unique function in $L^2(\nu)$ such that $\int_{\pi^{-1}(A)} f d\mu = \int_A \mathbb{E}(f \mid \mathbf{Y}) d\nu$ for all measurable set $A \subset \mathcal{Y}$. To avoid confusion, we use $\mathbb{E}(f \mid \mathcal{Y})$ to denote the lift of $\mathbb{E}(f \mid \mathbf{Y})$ to $L^2(\mu)$.

Let $\mathbf{X} = (X, \mathcal{B}_X, \mu_X, T)$ be a system and $K = (K, \mathcal{B}_K, \mu_K)$ be a compact abelian group equipped with the Borel σ -algebra and Haar measure. A *cocycle* on \mathbf{X} with values in K is a measurable map $\rho: X \rightarrow K$. The *extension defined by ρ* is the system $\mathbf{X} \times_\rho K = (X \times K, \mathcal{B}_X \times \mathcal{B}_K, \mu_X \times \mu_K, T_\rho)$, where

$$T_\rho(x, u) = (Tx, \rho(x) + u).$$

A cocycle $\rho: X \rightarrow K$ gives rise to an extended map $\rho: \mathbb{Z} \times X \rightarrow K$ defined by (Id_K denotes the identity in K)

$$\rho(n, x) = \begin{cases} \sum_{i=0}^{n-1} \rho(T^i x) & n \geq 1 \\ \text{Id}_K & n = 0 \\ -\sum_{i=n+1}^0 \rho(T^i x) & n \leq -1 \end{cases}$$

and

$$T_\rho^n(x, u) = (T^n x, \rho(n, x) + u).$$

Two cocycles $\rho, \rho': X \rightarrow K$ are *cohomologous* if there exists a measurable map $F: X \rightarrow K$ such that $\rho(x) = \rho'(x) + F(Tx) - F(x)$. We note that if ρ, ρ' are cohomologous then the systems $\mathbf{X} \times_\rho K$ and $\mathbf{X} \times_{\rho'} K$ are isomorphic and the isomorphism is given by $(x, t) \mapsto (x, t - F(x))$.

Definition 2.1 (Continuous vertical characters). Let $\mathbf{X} = (X, \mathcal{B}_X, \mu_X, T)$ be a system, let K be a compact abelian group, and let $\rho: X \rightarrow K$ be a cocycle. A function $f: X \times_\rho K \rightarrow \mathbb{C}$ is a *continuous vertical character* if there exists a character $\chi: K \rightarrow \mathcal{S}^1$ and a continuous function $g: X \rightarrow \mathbb{C}$ such that $f(x, k) = g(x) \cdot \chi(k)$.

We note for use in the sequel that it follows from the Stone-Weierstrass theorem that the space spanned by the continuous vertical characters is dense in the continuous functions $C(X \times_\rho K)$.

Lemma 2.2. *Let $\mathbf{X} = (X, \mathcal{B}_X, \mu_X, T)$ be a system, let K be a compact abelian group and $\rho: X \rightarrow K$ a cocycle. The space spanned by the continuous vertical characters on $X \times_\rho K$ is dense in $C(X \times_\rho K)$.*

2.3. Definition of an IP.

Definition 2.3. If $(n_j)_{j \in \mathbb{N}}$ is a sequence of natural numbers, define $\text{IP}((n_j)_{j \in \mathbb{N}})$ to be the multiset containing all finite sums of the sequence $(n_j)_{j \in \mathbb{N}}$, meaning

$$\text{IP}((n_j)_{j \in \mathbb{N}}) = \{n_{i_1} + \cdots + n_{i_\ell} : i_1, \dots, i_\ell \text{ are distinct}\},$$

with the convention that the empty sum 0 is included. We refer to the numbers $(n_j)_{j \in \mathbb{N}}$ as the *generators*, and when the generators are clear from the context we refer to this simply as an IP.

Our definition differs slightly from some uses in the literature, where there is no assumption that 0 lies in the IP and an IP is defined as a set rather, than a multiset. The assumption that an IP contains 0 is merely for notational convenience, but the use of a multiset is necessary for our analysis. We could avoid this by taking the generators to be a sequence $(n_j)_{j \in \mathbb{N}}$ that grows sufficiently fast such that $n_j > \sum_{i=1}^{j-1} n_i$ for all $j \in \mathbb{N}$, but we prefer to work in the more general setting. We note that in the literature, a set is sometimes called an IP-set if it contains an IP. However, as we also prove convergence results, we can not adopt that convention.

Example 2.4. Set $n_j = 10^{j-1}$ for all $j \geq 1$. Then $\text{IP}((n_j)_{j \in \mathbb{N}})$ is the set of all natural numbers that can be written in base 10 only using the digits 0 and 1, meaning that

$$\text{IP}((n_j)_{j \in \mathbb{N}}) = \{0, 1, 10, 11, 100, 101, 110, 111, 1000, \dots\}.$$

Note that this IP is a set, not a multiset, as all finite sums are distinct. If instead we take the generators $n_j = 2^{j-1}$ for $j \geq 1$, then $\text{IP}((n_j)_{j \in \mathbb{N}}) = \mathbb{N}$.

It is easy to see that there are many subsets of integers that are not IPs. For example, it is easy to check that an IP contains infinitely many multiples of a given positive integer.

2.4. Averages along an IP. A *Følner sequence* $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ in \mathbb{N} is a sequence of finite subsets of natural numbers satisfying

$$(2.1) \quad \lim_{N \rightarrow \infty} \frac{|(a + \Phi_N) \Delta \Phi_N|}{|\Phi_N|} = 0$$

for all $a \in \mathbb{N}$, where Δ denotes the symmetric difference. A standard example of a Følner sequence is taking intervals, such as $\Phi_N = [1, N]$ or $\Phi_N = [N^2, N^2 + N]$. We say that a Følner sequence Φ_N is *increasing* if $\Phi_N = [M, a_N]$ for some fixed number M and a sequence a_N tending to infinity as $N \rightarrow \infty$.

To define averages along an IP, we adapt Følner sequences to our context.

Definition 2.5. Let $(n_j)_{j \in \mathbb{N}}$ be a sequence of natural numbers and $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ be a Følner sequence in \mathbb{N} . The *IP-Følner sequence associated to $(\Phi_N)_{N \in \mathbb{N}}$* is the sequence

$$\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}}) = \{n_{i_1} + \cdots + n_{i_\ell} : i_1, \dots, i_\ell \in \Phi_N \text{ are distinct}\}.$$

If the Følner sequence $(\Phi_N)_{N \in \mathbb{N}}$ is increasing, we say that the IP-Følner sequence $\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})$ is *increasing*. When the generating sequence $(n_j)_{j \in \mathbb{N}}$ is clear from the context, we omit it from the notation.

Given a Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ in \mathbb{N} and sequence $(a_n)_{n \in \mathbb{N}}$, we use the shorthand notation

$$\mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} a_n = \frac{1}{|\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})|} \sum_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} a_n,$$

and when the generators are clear from the context, we write $\mathbb{E}_{n \in \text{IP}_{\Phi_N}} a_n$.

This allows us to define ergodic averages along IPs.

Definition 2.6 (The ergodic average). Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ be an invertible measure preserving system, let $(n_j)_{j \in \mathbb{N}}$ be a sequence of natural numbers, let $(\Phi_N)_{N \in \mathbb{N}}$ be a Følner sequence, and let $f \in L^2(\mu)$. The *ergodic average of f along the IP-Følner sequence $\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})$* is defined to be the limit

$$(2.2) \quad \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} T^n f$$

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

Considering the ergodic average along $\text{IP}_{[1, N]}((2^{j-1})_{j \in \mathbb{N}})$, von Neumann's mean ergodic theorem states that this average exists for all $f \in L^2(\mu)$ and converges to the projection onto the T -invariant functions. In contrast, we give examples of IP-Følner sequences where the ergodic averages do not converge, and examples where the averages converge but the limit is not projection onto the T -invariant functions.

3. ERGODIC AVERAGES ALONG AN IP-SEQUENCE

3.1. Convergence along an IP-Følner sequence. In this section, we give necessary and sufficient conditions in Theorem 3.3 for mean convergence of ergodic averages along an IP-Følner sequence, with criteria for when the

limit is the projection onto the T -invariant functions. We start by recalling some standard definitions and setting notation.

Let δ_1 denote the *Dirac function* on the circle \mathcal{S}^1 , meaning

$$(3.1) \quad \delta_1(\alpha) = \begin{cases} 1 & \alpha = 1 \\ 0 & \alpha \neq 1, \end{cases}$$

Our first condition for convergence is a consequence of the spectral theorem for unitary operators.

Theorem 3.1. *Let $(n_j)_{j \in \mathbb{N}}$ be a sequence of natural numbers and let $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ be a Følner sequence. The following are equivalent:*

- (1) *For all invertible measure preserving systems $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ and all $f \in L^2(\mu)$, the ergodic averages (2.2) along $\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})$ exists,*
- (2) *The limit*

$$(3.2) \quad \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \alpha^n$$

exists for all $\alpha \in \mathcal{S}^1$.

Furthermore, the following are equivalent:

- (3) *For all invertible measure preserving systems $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ and all $f \in L^2(\mu)$, the ergodic averages (2.2) along $\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})$ converges to the projection onto the T -invariant functions.*
- (4) *The limit in (3.2) is equal to $\delta_1(\alpha)$.*

Proof. We start by proving the equivalence of (1) and (2). First suppose that for every invertible measure preserving system $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ and all $f \in L^2(\mu)$, the ergodic averages along $\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})$ exist. Given $\alpha \in \mathcal{S}^1$, take \mathbf{X} to be the ergodic rotation on \mathcal{S}^1 by α and let $f: X \rightarrow \mathcal{S}^1$ be the natural embedding. Then,

$$(3.3) \quad \mathbb{E}_{n \in \text{IP}_{\Phi_N}} T^n f = \left(\mathbb{E}_{n \in \text{IP}_{\Phi_N}} \alpha^n \right) \cdot f$$

and it follows immediately that the limit (3.2) exists. Conversely, suppose that (3.2) exists for all $\alpha \in \mathcal{S}^1$ and let $d\rho$ denote the projection valued measure for T , meaning that

$$T^n = \int_{\mathcal{S}^1} t^n d\rho(t)$$

for all $n \in \mathbb{N}$. Then by the Lebesgue dominated convergence theorem,

$$(3.4) \quad \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} T^n = \int_{S^1} \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} t^n d\rho(t) = \int_{S^1} \phi(t) d\rho(t),$$

where $\phi(t) = \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} t^n$ is the pointwise limit which exists by assumption. In particular, we see that the ergodic average converges for all $f \in L^2(\mu)$.

To check the equivalence of (3) and (4), note that iff $\alpha = 1$, it is immediate that the limit (3.2) equals 1. Supposing that (3) holds, assume that $\alpha \neq 1$ and again take \mathbf{X} to be the ergodic rotation by α and $f: X \rightarrow S^1$ to be the natural embedding. Since $\int_X f d\mu = 0$, it follows from (3.3) that the limit (3.2) is 0. Conversely, suppose that (4) holds, meaning that the map ϕ in (3.4) satisfies $\phi(t) = \delta_1(t)$. Using the same argument as in the first part, the ergodic average converges to the projection $\rho(\{1\})$. Observe that

$$T\rho(\{1\}) = \int_{S^1} t \cdot \delta_1(t) d\rho(t) = \int_{S^1} \delta_1(t) d\rho(t)$$

and similarly $\rho(\{1\})T = \rho(\{1\})$. It follows that $\rho(\{1\})$ is the projection onto the subspace of T -invariant functions. \square

To refine this characterization of convergence, we use the next result to rewrite averages as infinite products. This use also justifies our conventions on the definition of an IP-set, defining it as a multiset and always including 0 in the set.

Proposition 3.2. *Let $(n_j)_{j \in \mathbb{N}}$ be a sequence of natural numbers, let $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ be a Følner sequence in \mathbb{N} , let $\alpha \in S^1$, and set $z_j = \frac{1+\alpha^{n_j}}{2}$ for $j \in \mathbb{N}$. For every $N \in \mathbb{N}$, we have*

$$\mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \alpha^n = \prod_{j \in \Phi_N} z_j.$$

Proof. We have

$$\prod_{j \in \Phi_N} z_j = \prod_{j \in \Phi_N} \left(\frac{1 + \alpha^{n_j}}{2} \right) = \frac{\prod_{j \in \Phi_N} (1 + \alpha^{n_j})}{2^{|\Phi_N|}}$$

and by expanding the terms, we obtain

$$\prod_{j \in \Phi_N} z_j = \frac{\sum_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \alpha^n}{2^{|\Phi_N|}} = \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \alpha^n. \quad \square$$

Let $\arg: \mathcal{S}^1 \rightarrow \mathbb{R}/2\pi\mathbb{Z}$ denote the inverse of the isomorphism $t \mapsto e^{it}$. We have the following criteria for convergence.

Theorem 3.3 (Criteria for convergence). *Let $(n_j)_{j \in \mathbb{N}}$ be a sequence of natural numbers and let $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ be a Følner sequence in \mathbb{N} .*

- (1) *For all invertible measure preserving system $\mathbf{X} = (X, \mathcal{B}, \mu, T)$, the ergodic average (2.2) along $\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})$ converges to the projection onto the T -invariant functions if and only if*

$$\lim_{N \rightarrow \infty} \prod_{j \in \Phi_N} \frac{(1 + \cos(\arg(\alpha^{n_j}))}{2} = \delta_1(\alpha)$$

for all $\alpha \in \mathcal{S}^1$.

- (2) *For all invertible measure preserving system $\mathbf{X} = (X, \mathcal{B}, \mu, T)$, the ergodic average (2.2) along $\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})$ converges if and only if each $\alpha \in \mathcal{S}^1$ satisfies either*

$$\lim_{N \rightarrow \infty} \prod_{j \in \Phi_N} \frac{(1 + \cos(\arg(\alpha^{n_j}))}{2} = 0$$

or

$$\lim_{N \rightarrow \infty} \sum_{j \in \Phi_N} \arctan\left(\frac{\sin(\arg(\alpha^{n_j}))}{1 + \cos(\arg(\alpha^{n_j}))}\right)$$

exists.

Proof. Fixing $j \in \mathbb{N}$ and $\alpha \in \mathcal{S}^1$, set $z_j = z_j(\alpha) = \frac{1 + \alpha^{n_j}}{2}$. Rewriting, we have that $z_j = r_j e(i\theta_j)$ for some $r_j \geq 0$ and $\theta_j \in (-\frac{\pi}{2}, \frac{\pi}{2})$. Using the identity

$$z_j(\alpha) = \frac{1 + \cos(\arg(\alpha^{n_j}))}{2} + i \frac{\sin(\arg(\alpha^{n_j}))}{2},$$

we can rewrite $r_j = \sqrt{\frac{1 + \cos(\arg(\alpha^{n_j}))}{2}}$ and $\theta_j = \arctan\left(\frac{\sin(\arg(\alpha^{n_j}))}{1 + \cos(\arg(\alpha^{n_j}))}\right)$. In particular, it follows from Proposition 3.2 that

$$\mathbb{E}_{n \in \text{IP}_{\Phi_N}} \alpha^n = \left(\prod_{j \in \Phi_N} r_j \right) \cdot e(i \sum_{j \in \Phi_N} \theta_j).$$

By Theorem 3.1, the ergodic average along $\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})$ exists if and only if this quantity converges as $N \rightarrow \infty$. We use the general criterion that a sequence of complex numbers converges if and only if either the norms converge to zero or the norms converge and the arguments converge. In the first case, if the norms converge to zero for all $\alpha \neq 1$, then $\mathbb{E}_{n \in \text{IP}_{\Phi_N}} \alpha^n = \delta_1(\alpha)$ and so Part (1) follows from Theorem 3.1. Otherwise, since z_j is the average

of two numbers in the unit circle, it follows that $|z_j| \leq 1$ and so the norm of $\mathbb{E}_{n \in \text{IP}_{\Phi_N}} \alpha^n$ always converges. If the limit is nonzero, then by Theorem 3.1 the limit exists if and only if the arguments converge, proving Part (2). \square

3.2. Sufficient conditions for convergence. We start with a simple proposition showing that for the average along an IP to converge to the projection onto the T -invariant functions, it suffices that the average along the generators does so.

Proposition 3.4. *Let $(n_j)_{j \in \mathbb{N}}$ be a sequence of natural numbers and let $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ be a Følner sequence in \mathbb{N} . Assume that for all ergodic rotations $\mathbf{Z} = (Z, C, \nu, R_a)$ and all $f \in L^2(\nu)$, we have*

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \Phi_N} f(a^n x) = \int_Z f d\nu$$

in $L^2(\nu)$. Then for every ergodic system $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ and all $f \in L^2(\mu)$, we have

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} T^n f = \int_{\mathbf{X}} f d\mu.$$

Proof. The hypothesis of convergence in the proposition holds if and only if $\lim_{N \rightarrow \infty} \mathbb{E}_{j \in \Phi_N} t^{n_j} = \delta_1(t)$ for all $t \in \mathcal{S}^1$. Fix $t \neq 1$, and consider

$$A_N := \{j \in \Phi_N : |t^{n_j} - 1| < 1/10\}.$$

Using the triangle inequality, we have that

$$\left| \frac{1}{|\Phi_N|} \sum_{j \in A_N} t^{n_j} \right| \leq \left| \mathbb{E}_{j \in \Phi_N} t^{n_j} \right| + \left| \frac{1}{|\Phi_N|} \sum_{j \notin A_N} t^{n_j} \right|.$$

Since $|t^{n_j} - 1| < 1/10$ for all $j \in A_N$, the left hand side is bounded from below by $\frac{9|A_N|}{10|\Phi_N|}$. Fixing arbitrary $\varepsilon > 0$, it follows that for sufficiently large N we have that

$$\frac{9|A_N|}{10|\Phi_N|} \leq \varepsilon + \left(1 - \frac{|A_N|}{|\Phi_N|}\right).$$

Choosing any $0 < \varepsilon < 9/10$, it follows that $\limsup_{N \rightarrow \infty} \frac{|A_N|}{|\Phi_N|} < 1$, and so for all sufficiently large N , we have $\frac{|A_N|}{|\Phi_N|} < 1 - c$ for some sufficiently small $c > 0$. In particular, the number of $j \in \Phi_N$ satisfying $|t^{n_j} - 1| > 1/10$ is arbitrarily large as $N \rightarrow \infty$. By Proposition 3.2, it follows that

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} t^n = \lim_{N \rightarrow \infty} \prod_{j \in \Phi_N} \left(\frac{1 + t^{n_j}}{2} \right) = 0.$$

As this holds for all $1 \neq t \in \mathcal{S}^1$, convergence follows from Theorem 3.1. \square

This proposition applies to many sequences, including polynomial sequences (see for example [40]), the primes and polynomials in primes [41, 18], and many other sequences such as those in [11, 17].

A particular class of dynamical systems in which we can describe convergence is when the eigenspaces generate $L^2(\mu)$ and additionally all nontrivial eigenspaces are associated to rational eigenvalues.

Definition 3.5. If $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ is an invertible measure preserving system, let $\mathcal{K}_{\text{rat}}(\mathbf{X})$ denote the minimal factor of \mathbf{X} such that any eigenfunction of \mathbf{X} with rational eigenvalue is measurable with respect to $\mathcal{K}_{\text{rat}}(\mathbf{X})$. The system \mathbf{X} is *rational* if $\mathbf{X} = \mathcal{K}_{\text{rat}}(\mathbf{X})$.

Thus \mathcal{K}_{rat} is the factor generated by the eigenfunctions that are associated to the rational eigenvalues. It is immediate that any finite system is rational. A more interesting example is obtained by taking $X = \mathbb{Z}_p$ is the p -adic numbers for some fixed prime p , equipped with the Borel σ -algebra, Haar measure, and the transformation is translation by some fixed $t \in \mathbb{Z}_p$. Any character χ of \mathbb{Z}_p is an eigenfunction whose eigenvalue $\chi(t)$ is a p^m root of unity for some $m \in \mathbb{N}$, and these characters form an orthonormal basis for \mathbb{Z}_p . Note that if the group generated by t is dense, then the system is also ergodic.

Recall that a function $f: X \rightarrow \mathbb{C}$ is *1-bounded* if $|f(x)| \leq 1$ for all $x \in X$ and we let $P_t: L^2(\mu) \rightarrow L^2(\mu)$ denote the projection onto the t -eigenspace of T .

Theorem 3.6 (Convergence in $\mathcal{K}_{\text{rat}}(\mathbf{X})$). *Let $(n_j)_{j \in \mathbb{N}}$ be a sequence of natural numbers and let $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ be an increasing Følner sequence. Then there exists a 1-bounded function $\omega_\Phi: e^{2\pi i \mathbb{Q}} \rightarrow \mathbb{C}$ such that for every invertible measure preserving system $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ and $f \in L^2(\mu)$ measurable with respect to $\mathcal{K}_{\text{rat}}(\mathbf{X})$, we have*

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} T^n f = \sum_{t \in e^{2\pi i \mathbb{Q}}} \omega_\Phi(t) \cdot P_t(f)$$

in $L^2(\mu)$.

Proof. Let $(n_j)_{j \in \mathbb{N}}$ and $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ be as in the statement. We first check that the limit

$$(3.5) \quad \omega_\Phi(t) := \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} t^n$$

exists for all $t \in e^{2\pi i \mathbb{Q}}$. By Proposition 3.2, this limit is equal to

$$\lim_{N \rightarrow \infty} \prod_{j \in \Phi_N} \left(\frac{1 + t^{n_j}}{2} \right).$$

We consider two cases. If $t^{n_j} = 1$ for all but finitely many values of j , then this infinite product is eventually equal to

$$\prod_{\{j : t^{n_j} \neq 1\}} \left(\frac{1 + t^{n_j}}{2} \right)$$

and so the limit exists. Otherwise, there are infinitely many j such that $t^{n_j} \neq 1$. Then we can choose some $\varepsilon_0 > 0$, depending only on t , such that $\left| \left(\frac{1 + t^{n_j}}{2} \right) \right| < 1 - \varepsilon_0$. Since the Følner sequence Φ is increasing, it eventually contains arbitrarily many such j and

$$\lim_{N \rightarrow \infty} \prod_{j \in \Phi_N} \left(\frac{1 + t^{n_j}}{2} \right) = 0.$$

Thus the limit (3.5) exists.

Since f is measurable with respect to $\mathcal{K}_{\text{rat}}(\mathbf{X})$, we can write

$$f = \sum_{t \in e^{2\pi i \mathbb{Q}}} P_t(f).$$

Since $T(P_t(f)) = t \cdot P_t(f)$, we conclude that

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} ((n_j)_{j \in \mathbb{N}})^{T^n} f = \sum_{t \in e^{2\pi i \mathbb{Q}}} \omega_\Phi(t) P_t(f),$$

and the statement follows. \square

This leads to the following definition.

Definition 3.7 (Spectrum of a sequence). The *spectrum* $\sigma((n_j)_{j \in \mathbb{N}})$ of a sequence of natural numbers $(n_j)_{j \in \mathbb{N}}$ is defined to be the group generated by the complement of the set

$$\bigcap_{m \in \mathbb{N}} \left\{ \alpha \in \mathcal{S}^1 : \lim_{d \rightarrow \infty} \prod_{j=m}^d \left(\frac{1 + \alpha^{n_j}}{2} \right) = 0 \right\}.$$

We say that a sequence has *rational spectrum* if $\sigma((n_j)_{j \in \mathbb{N}}) \subseteq e^{2\pi i \mathbb{Q}}$.

The intersection over m is taken to ensure that for every $\alpha \notin \sigma((n_j)_{j \in \mathbb{N}})$, and every increasing Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$, the quantity

$$(3.6) \quad \omega_\Phi(\alpha) := \lim_{N \rightarrow \infty} \prod_{j \in \Phi_N} \left(\frac{1 + \alpha^{n_j}}{2} \right)$$

is zero.

The next result follows quickly from Theorem 3.6.

Corollary 3.8. *Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ be an invertible measure preserving system and let $(n_j)_{j \in \mathbb{N}}$ be a sequence with rational spectrum. Then for any increasing Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$, we have*

$$\mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} T^n = \sum_{t \in \sigma((n_j)_{j \in \mathbb{N}})} \omega_\Phi(t) \cdot P_t$$

in the strong operator topology.

For example, for every integer $a \geq 2$, the sequence $n_j = a^{j-1}$ has rational spectrum. To check this it suffices to show that for any irrational α , the sequence α^{n_j} does not converges to 1 as $j \rightarrow \infty$. Considering the expansion of $\arg(\alpha)$ in base a , taking a power α^{n_j} shifts the digits $(j-1)$ times to the left. If $\alpha^{n_j} \rightarrow 1$, this implies that for sufficiently large j all entries in the decimal expansion of $\arg(\alpha)$ are 0, contradicting that α is irrational. On the other hand, any sequence $(n_j)_{j \in \mathbb{N}}$ satisfying $\lim_{j \rightarrow \infty} n_{j+1}/n_j$ has uncountable spectrum (see [6, Proposition 3.5]) and so in particular is not rational.

The next corollary is a direct application of Corollary 3.8.

Corollary 3.9. *Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ be an invertible measure preserving system, let $a \geq 2$ be an integer, and set $n_j = a^{j-1}$.*

- (1) *For all $f \in L^2(\mu)$ and any increasing Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$, the ergodic average (2.2) of f along $\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})$ exists in $L^2(\mu)$.*
- (2) *Furthermore, if \mathbf{X} is ergodic and the points $e(m/p^n)$ for $m, n \in \mathbb{N}$ and p any prime divisor of a do not lie in the point spectrum of \mathbf{X} , then the ergodic average (2.2) of f along $\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})$ converges to $\int_X f \, d\mu$.*

Note that the hypothesis in Part (2) holds whenever the system \mathbf{X} is totally ergodic. We use this to deduce an equidistribution result along $\text{IP}((a^{j-1})_{j \in \mathbb{N}})$.

Corollary 3.10 (Equidistribution). *Let $a \geq 2$ be an integer and let $\alpha \in \mathcal{S}^1$ be either an irrational or be a root of unity whose order is coprime to a . Then for every continuous function $f: \mathcal{S}^1 \rightarrow \mathbb{C}$, we have*

$$(3.7) \quad \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N((a^{j-1})_{j \in \mathbb{N}})}} f(\alpha^n x) = \int_{\mathcal{S}^1} f \, d\mu$$

for all $x \in \mathcal{S}^1$.

Proof. By the Stone-Weierstrass theorem, any continuous function on a compact abelian group is a uniform limit of linear combinations of characters. Thus it suffices to prove (3.7) when f is a character. If $f = 1$, this is immediate. Otherwise, choose $0 \neq \ell \in \mathbb{Z}$ such that $f(x) = x^\ell$. It follows that (recall that ω_Φ is defined in (3.6))

$$\begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} f(\alpha^n x) &= \left(\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} (\alpha^\ell)^n \right) f(x) \\ &= \omega_\Phi(\alpha^\ell) \cdot f(x) = 0, \end{aligned}$$

and since $\int_{\mathcal{S}^1} f(x) \, d\mu = 0$, the statement follows. \square

3.3. An example of non-convergence along an IP. To produce a sequence such that convergence fails, fix some irrational $\alpha \in \mathcal{S}^1$ and the standard Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ with $\Phi_N = [1, N]$ for all $N \in \mathbb{N}$. We use that the orbit of α is dense to produce a sequence $(n_j)_{j \in \mathbb{N}}$ such that

$$(3.8) \quad \lim_{N \rightarrow \infty} \sum_{j=1}^N \arctan \left(\frac{\sin(\arg(\alpha^{n_j}))}{1 + \cos(\arg(\alpha^{n_j}))} \right)$$

diverges while

$$(3.9) \quad \lim_{N \rightarrow \infty} \prod_{j=1}^N \frac{1 + \cos(\arg(\alpha^{n_j}))}{2}$$

converges to a nonzero value. It then follows from Theorem 3.3 the ergodic average along IP generated by $(n_j)_{j \in \mathbb{N}}$ diverges.

To do so, set $z_j = \frac{1 + \alpha^{n_j}}{2}$ and write $z_j = r_j e(i \cdot \theta_j)$ for some $r_j > 0$ and $\theta_j \in (-\frac{\pi}{2}, \frac{\pi}{2})$. For each $j \in \mathbb{N}$, choose n_j such that $\frac{1}{j} < \theta_j < \frac{1}{j} + \frac{1}{j^2}$. Comparing with the harmonic series we see that $\sum_{j=1}^\infty \theta_j$ diverges, and it follows that (3.8) diverges. Considering (3.9), we have that $r_j = \sqrt{\cos(\theta_j)}$ and taking the absolute value of the product we obtain $\lim_{N \rightarrow \infty} \prod_{j=1}^N \frac{1 + \cos(\arg(\alpha^{n_j}))}{2} = \prod_{j=1}^\infty r_j$.

Combining these facts, we have that

$$\lim_{N \rightarrow \infty} \left(\prod_{j=1}^N r_j \right)^2 = \exp \left(\lim_{N \rightarrow \infty} \sum_{j=0}^N \ln(\cos(\theta_j)) \right).$$

Approximating via the Taylor expansion $\ln \cos(\theta_j) = O(\theta_j^2) = O(1/j^2)$ and comparing to the convergent series $\sum_{j=0}^N \frac{1}{j^2}$, the limit $\lim_{N \rightarrow \infty} \sum_{j=0}^N \ln(\cos(\theta_j))$ exists. It follows that (3.9) is nonzero.

4. CHARACTERISTIC FACTORS FOR MULTIPLE ERGODIC AVERAGES ALONG AN IP

4.1. Preliminaries on characteristic factors. Characteristic factors were implicit in the work of Furstenberg [20], and were given this name by Furstenberg and Weiss [26] in their study of some double averages. They have been used to prove the convergence of multiple ergodic averages [30, 43], and though there are other methods to prove the convergence (such as [40]), characteristic factors give more information that can be used to derive limit formulas and combinatorial corollaries. We adapt these techniques for IP averages, and start by recalling the precise definition.

Definition 4.1 (Characteristic factors). Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$, let $(A_N)_{N \in \mathbb{N}} \subseteq \mathbb{N}$ be finite subsets, and let $k \geq 1$. A factor $\mathbf{Y} = (Y, \mathcal{Y}, \nu, S)$ of \mathbf{X} is *characteristic for the average*

$$(4.1) \quad \mathbb{E}_{n \in A_N} T^n f_1 \cdot \dots \cdot T^{kn} f_k$$

if the difference between this average and the average

$$\mathbb{E}_{n \in A_N} T^n \mathbb{E}(f_1 \mid \mathcal{Y}) \cdot \dots \cdot T^{kn} \mathbb{E}(f_k \mid \mathcal{Y})$$

where each function is replaced by its expectation on \mathbf{Y} converges to 0 in $L^2(\mu)$ for all $f_1, \dots, f_k \in L^\infty(\mu)$.

Equivalently, the factor \mathbf{Y} is characteristic for (4.1) if this average converges to 0 in $L^2(\mu)$ whenever there is some $i \in \{1, \dots, k\}$ such that expectation $\mathbb{E}(f_i \mid \mathcal{Y}) = 0$.

In [30] (and later in [43]), characteristic factors for the average (4.1) with $(A_N)_{N \in \mathbb{N}}$ being a Følner sequence in \mathbb{N} are described, defining a factor $Z_{k-1} = Z_{k-1}(X)$ (known as the *Host-Kra factor*) of the system \mathbf{X} for each $k \geq 1$. Our next result shows that the same factors Z_{k-1} are characteristic along IPs.

Theorem 4.2 (Characteristic factors for multiple ergodic averages along IPs.). *Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ be an invertible measure preserving system, let $(n_j)_{j \in \mathbb{N}}$ be a sequence with rational spectrum, let ℓ_1, \dots, ℓ_k be distinct and let $k \geq 2$. For any increasing Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$, the factor $Z_{k-1}(\mathbf{X})$ is characteristic for the average $\mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \prod_{i=1}^k T^{\ell_i n} f_i$.*

When passing to a subsequence, there is no reason, apriori, that the factors Z_k that are characteristic for the multiple average (4.1) taken along a Følner in \mathbb{N} are also characteristic for the multiple ergodic averages along the subsequence. In particular, for an IP, this fails for even the simplest case $k = 1$ that corresponds to the mean ergodic theorem, explaining the assumption that $k \geq 2$. To prove Theorem 4.2 in this case, we develop IP versions of the tools used in the proof in [30], including IP versions of the cubic systems, the Gowers-Host-Kra seminorms, and the associated factors. A posteriori, we show that for $k \geq 2$, the factors agree with the factors Z_k .

4.2. An IP-version of the van der Corput lemma. We start with an IP version of a standard tool for identifying characteristic factors for multiple ergodic averages, the van der Corput Lemma. For a Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ and fixed $M \in \mathbb{N}$, define the Følner sequence $\Phi^M = (\Phi_N^M)_{N \in \mathbb{N}}$ by setting $\Phi_N^M = \Phi_N \setminus \Phi_M$. It follows immediately from the definitions that Φ^M is a Følner sequence for every $M \in \mathbb{N}$ and that if Φ is an increasing Følner sequence then so is Φ^M .

Lemma 4.3 (van der Corput lemma for increasing IP-Følner sequences). *Let \mathcal{H} be a Hilbert space with norm $\|\cdot\|$ and inner product $\langle \cdot, \cdot \rangle$, and let $(x_n)_{n \in \mathbb{N}}$ be a sequence in \mathcal{H} . For a sequence $(n_j)_{j \in \mathbb{N}}$ and increasing Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$, we have*

$$\begin{aligned} & \limsup_{N \rightarrow \infty} \|\mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} x_n\|^2 \\ & \leq \limsup_{M \rightarrow \infty} \mathbb{E}_{m_1, m_2 \in \text{IP}_{\Phi_M}((n_j)_{j \in \mathbb{N}})} \limsup_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N^M}((n_j)_{j \in \mathbb{N}})} \langle x_{n+m_1}, x_{n+m_2} \rangle. \end{aligned}$$

Proof. Setting $\beta(n, m) = n + m$, we have a bijection

$$\beta: \text{IP}_{\Phi_M}((n_j)_{j \in \mathbb{N}}) \times \text{IP}_{\Phi_N^M}((n_j)_{j \in \mathbb{N}}) \rightarrow \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})$$

for all $N, M \in \mathbb{N}$. It follows that for all $M \in \mathbb{N}$, we have

$$\mathbb{E}_{n \in \text{IP}_{\Phi_N}} x_n = \mathbb{E}_{n \in \text{IP}_{\Phi_N^M}} \mathbb{E}_{m \in \text{IP}_{\Phi_M}} x_{n+m}.$$

By Jensen's inequality, we have that

$$\begin{aligned} \left\| \mathbb{E}_{n \in \text{IP}_{\Phi_N^M}} \mathbb{E}_{m \in \text{IP}_{\Phi_M}} x_{n+m} \right\|^2 &\leq \mathbb{E}_{n \in \text{IP}_{\Phi_N^M}} \left\| \mathbb{E}_{m \in \text{IP}_{\Phi_M}} x_{n+m} \right\|^2 \\ &= \mathbb{E}_{n \in \text{IP}_{\Phi_N^M}} \mathbb{E}_{m_1, m_2 \in \text{IP}_{\Phi_M}} \langle x_{n+m_1}, x_{n+m_2} \rangle. \end{aligned}$$

Interchanging the order of summation and taking $N \rightarrow \infty$, we obtain that for all $M \in \mathbb{N}$,

$$\limsup_{N \rightarrow \infty} \left\| \mathbb{E}_{n \in \text{IP}_{\Phi_N}} x_n \right\|^2 \leq \mathbb{E}_{m_1, m_2 \in \text{IP}_{\Phi_M}} \limsup_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N^M}} \langle x_{n+m_1}, x_{n+m_2} \rangle.$$

The statement follows as $M \rightarrow \infty$. \square

Combining this lemma and induction, we have the generalization of Corollary 3.10 to polynomial sequences.

Corollary 4.4 (Equidistribution along polynomials). *Let $a \geq 2$ be an integer and let $\alpha \in \mathcal{S}^1$ be either an irrational or be a root of unity whose order is coprime to a . Then for every $k \geq 1$, integer valued (non-constant) polynomial $p: \mathbb{N} \rightarrow \mathbb{N}$ of degree k , and continuous function $f: \mathcal{S}^1 \rightarrow \mathbb{C}$, we have*

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N((a^{j-1})_{j \in \mathbb{N}})}} f(\alpha^{p(n)} x) = \int_{\mathcal{S}^1} f \, d\mu$$

for all $x \in \mathcal{S}^1$.

To control multiple ergodic averages along an IP associated to a sequence with rational spectrum, we introduce the relevant version of the uniformity seminorms, defined inductively analogous to the construction in [30].

Definition 4.5 (The (rational) IP Gowers-Host-Kra seminorms). Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ be an invertible measure preserving system and $f \in L^\infty(\mu)$. For $k = 1$, define the *first seminorm* $\| \cdot \|_{\text{IP}^1(\mathbf{X}, T)}$ by setting

$$\|f\|_{\text{IP}^1(\mathbf{X}, T)} := \sup_{(n_j)_{j \in \mathbb{N}}} \sup_{\Phi} \lim_{N \rightarrow \infty} \left\| \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} T^n f \right\|_{L^2(\mu)},$$

where the supremums are taken over all sequences $(n_j)_{j \in \mathbb{N}}$ with rational spectrum and all increasing IP-Følner sequences Φ . For $k \geq 1$, define the $(k+1)^{\text{st}}$ *seminorm* $\| \cdot \|_{\text{IP}^{k+1}(\mathbf{X}, T)}$ by setting

$$\|f\|_{\text{IP}^{k+1}(\mathbf{X}, T)} := \sup_{(n_j)_{j \in \mathbb{N}}} \sup_{\Phi} \lim_{N \rightarrow \infty} \left(\mathbb{E}_{m_1, m_2 \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \left\| T^{m_1} f \cdot T^{m_2} \bar{f} \right\|_{\text{IP}^k(\mathbf{X}, T)}^{2^k} \right)^{1/2^{k+1}},$$

again taking the supremums over all sequences $(n_j)_{j \in \mathbb{N}}$ with rational spectrum and all increasing IP-Følner sequences Φ .

As in the case of a Følner sequence in \mathbb{N} , these seminorms are characterized in terms of certain measures, inductively constructed analogously to how the cubic measures are constructed in [30]. Using notation of [30], set $V_k = \{0, 1\}^k$ to be the k -dimensional cube. We write a vertex $\epsilon \in V_k$ without commas and set $|\epsilon| = \sum_{i=1}^k \epsilon_i$. If $\epsilon \in V_k$ and $\eta \in V_\ell$, we concatenate them to obtain an element $\epsilon\eta \in V_{k+\ell}$. Let $C: \mathbb{C} \rightarrow \mathbb{C}$ denote complex conjugation. In [30], joinings are built inductively over the invariant factor, and in our inductive construction the invariant factor is replaced by the rational Kronecker factor.

Definition 4.6 (The (rational) IP cubic measures). Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ be an invertible measure preserving system, set $X^{[0]} = X$, and set $\tilde{\mu}^{[0]} = \mu$. For $k \geq 1$, identifying $X^{[k+1]}$ with $X^{[k]} \times X^{[k]}$, we define the *cubic measure* $\tilde{\mu}^{[k+1]}$ on $\mathbf{X}^{[k+1]}$ to be the relatively independent joining of $\mathbf{X}^{[k]}$ with itself over the factor $\mathcal{K}_{\text{rat}}(\mathbf{X}^{[k]})$.

This means that if $(f_\epsilon)_{\epsilon \in V_{k+1}}$ are 2^{k+1} bounded functions on X , the measure $\tilde{\mu}^{[k+1]}$ is defined by

$$\int_{X^{[k+1]}} \bigotimes_{\epsilon \in V_{k+1}} f_\epsilon d\tilde{\mu}^{[k+1]} = \int_{X^{[k]}} \mathbb{E}\left(\bigotimes_{\eta \in V_k} f_{\eta 0} \mid \mathcal{K}_{\text{rat}}\right) \cdot \mathbb{E}\left(\bigotimes_{\eta \in V_k} f_{\eta 1} \mid \mathcal{K}_{\text{rat}}\right) d\tilde{\mu}^{[k]}.$$

We check that rational IP-cubic measures control IP-seminorms.

Proposition 4.7. *Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ be an invertible measure preserving system and let $k \geq 1$. For every $f \in L^\infty(\mu)$, we have*

$$\|f\|_{\mathcal{U}_{\text{IP}(\mathbf{X}, T)}^k}^{2^k} \leq \int_{X^{[k]}} \bigotimes_{\epsilon \in V_k} C^{|\epsilon|} f d\tilde{\mu}^{[k]}.$$

Proof. We proceed by induction on k . For $k = 1$, Corollary 3.8 implies that

$$(4.2) \quad \|f\|_{\mathcal{U}_{\text{IP}(\mathbf{X}, T)}^1}^2 = \sup_{(n_j)_{j \in \mathbb{N}}} \sup_{\Phi} \left\| \sum_{t \in \sigma((n_j)_{j \in \mathbb{N}})} \omega_\Phi(t) \cdot P_t(f) \right\|_{L^2(\mu)}^2,$$

where $\omega_\Phi(t) = \lim_{N \rightarrow \infty} \prod_{j \in \Phi_N} \frac{1+t^{n_j}}{2}$. Since eigenspaces corresponding to different eigenvalues are orthogonal and $|\omega_\Phi(t)| \leq 1$ for all t independently

both of the sequence $(n_j)_{j \in \mathbb{N}}$ and of the increasing Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$, the quantity in (4.2) is bounded by

$$\left\| \sum_{\{t=e^{2\pi i r}: r \in \mathbb{Q}\}} P_t(f) \right\|_{L^2(\mu)}^2 = \|\mathbb{E}(f \mid \mathcal{K}_{\text{rat}}(\mathbf{X}))\|_{L^2(\mu)}^2.$$

On the other hand,

$$\int_{X^{[1]}} f \otimes \bar{f} \, d\tilde{\mu}^{[1]} = \int_X \mathbb{E}(f \mid \mathcal{K}_{\text{rat}}(\mathbf{X})) \cdot \overline{\mathbb{E}(f \mid \mathcal{K}_{\text{rat}}(\mathbf{X}))} \, d\mu = \|\mathbb{E}(f \mid \mathcal{K}_{\text{rat}}(\mathbf{X}))\|_{L^2(\mu)}^2,$$

completing the base case.

For $k \geq 2$, by the inductive hypothesis, Corollary 3.8 and the same argument used in the base case, we have

$$\begin{aligned} \|f\|_{\mathcal{U}_{\text{IP}}^k(\mathbf{X}, T)}^k &= \sup_{(n_j)_{j \in \mathbb{N}}} \sup_{\Phi} \lim_{N \rightarrow \infty} \left(\mathbb{E}_{m_1, m_2 \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \left\| T^{m_1} f \cdot \overline{T^{m_2} f} \right\|_{\mathcal{U}_{\text{IP}}^{k-1}(\mathbf{X}, T)}^{2^{k-1}} \right) \\ &\leq \sup_{(n_j)_{j \in \mathbb{N}}} \sup_{\Phi} \lim_{N \rightarrow \infty} \mathbb{E}_{m_1, m_2 \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \int_{X^{[k-1]}} \bigotimes_{\epsilon \in V_{k-1}} C^{|\epsilon|} T^{m_1} f \cdot \overline{T^{m_2} f} \, d\tilde{\mu}^{[k-1]} \\ &= \sup_{(n_j)_{j \in \mathbb{N}}} \sup_{\Phi} \lim_{N \rightarrow \infty} \int_{X^{[k-1]}} \left| \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \bigotimes_{\epsilon \in V_{k-1}} C^{|\epsilon|} T^n f \right|^2 d\tilde{\mu}^{[k-1]} \\ &= \sup_{(n_j)_{j \in \mathbb{N}}} \sup_{\Phi} \lim_{N \rightarrow \infty} \int_{X^{[k-1]}} \left| \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} (T^{[k-1]})^n \left(\bigotimes_{\epsilon \in V_{k-1}} C^{|\epsilon|} f \right) \right|^2 d\tilde{\mu}^{[k-1]} \\ &\leq \int_{X^{[k-1]}} \mathbb{E} \left(\bigotimes_{\epsilon \in V_{k-1}} C^{|\epsilon|} f \mid \mathcal{K}_{\text{rat}}(X^{[k-1]}) \right) \cdot \overline{\mathbb{E} \left(\bigotimes_{\epsilon \in V_{k-1}} C^{|\epsilon|} f \mid \mathcal{K}_{\text{rat}}(X^{[k-1]}) \right)} d\tilde{\mu}^{[k-1]} \\ &= \int_{X^{[k]}} \bigotimes_{\epsilon \in V_k} C^{|\epsilon|} f \, d\tilde{\mu}^{[k]}, \end{aligned}$$

completing the argument. \square

We state a simple lemma on rational spectrum.

Lemma 4.8. *If the spectrum of a sequence $(n_j)_{j \in \mathbb{N}}$ is rational, then for every integer $a \in \mathbb{N}$, the spectrum of the sequence $(a \cdot n_j)_{j \in \mathbb{N}}$ is also rational.*

Proof. If $\alpha \in \sigma(a \cdot (n_j)_{j \in \mathbb{N}})$, then $\alpha^a \in \sigma((n_j)_{j \in \mathbb{N}})$ and so α is rational. \square

We check that IP-seminorms control multiple ergodic averages along an IP, and note that it then follows immediately from Proposition 4.7 that rational IP cubic measures control multiple ergodic averages along an IP.

Proposition 4.9. *Let $(n_j)_{j \in \mathbb{N}}$ be a sequence with rational spectrum. Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ be an invertible measure preserving system, let $k \geq 1$, let $f_1, \dots, f_k \in L^\infty(\mu)$ with $\|f_i\|_{L^\infty(\mu)} \leq 1$ for $i = 1, \dots, k$, and let ℓ_1, \dots, ℓ_k be distinct integers. For any increasing Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$, we have*

$$(4.3) \quad \limsup_{N \rightarrow \infty} \left\| \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} T^{\ell_1 n} f_1 \cdot \dots \cdot T^{\ell_k n} f_k \right\|_{L^2(\mu)} \leq \min_{1 \leq i \leq k} \|f_i\|_{\text{U}_{\text{IP}}^k(\mathbf{X}, T)}.$$

Proof. We proceed by induction on k , noting that the case $k = 1$ follows from Lemma 4.8 and Theorem 3.8. Fixing $k \geq 2$ and changing the order of the functions f_1, \dots, f_k if needed, it suffices to show that the left hand side of (4.3) is bounded by $\|f_1\|_{\text{U}_{\text{IP}}^{k+1}(\mathbf{X}, T)}$. Setting

$$x_n = T^{\ell_1 n} f_1 \cdot \dots \cdot T^{\ell_k n} f_k$$

and applying Lemma 4.3, we have that

$$\begin{aligned} & \limsup_{N \rightarrow \infty} \left\| \mathbb{E}_{n \in \text{IP}_{\Phi_N}} x_n \right\|_{L^2(\mu)}^2 \\ & \leq \limsup_{M \rightarrow \infty} \mathbb{E}_{m_1, m_2 \in \text{IP}_{\Phi_M}} \limsup_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N^M}} \int_X \prod_{i=1}^k T^{\ell_i(n+m_1)} f_i \cdot \prod_{i=1}^k \overline{T^{\ell_i(n+m_2)} f_i} \, d\mu \\ & = \limsup_{M \rightarrow \infty} \mathbb{E}_{m_1, m_2 \in \text{IP}_{\Phi_M}} \limsup_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N^M}} \int_X \left(T^{\ell_k m_1} f_k \overline{T^{\ell_k m_2} f_k} \right) \prod_{i=1}^{k-1} T^{(\ell_i - \ell_k)n} \left(T^{\ell_i m_1} f_i \cdot \overline{T^{\ell_i m_2} f_i} \right) \, d\mu, \end{aligned}$$

where the last equality follows from T -invariance of μ . Applying the Cauchy-Schwartz Inequality and using that $\|f_k\|_{L^\infty(\mu)} \leq 1$, this last quantity is bounded by

$$\limsup_{M \rightarrow \infty} \mathbb{E}_{m_1, m_2 \in \text{IP}_{\Phi_M}} \limsup_{N \rightarrow \infty} \left\| \mathbb{E}_{n \in \text{IP}_{\Phi_N^M}} \prod_{i=1}^{k-1} T^{(\ell_i - \ell_k)n} \left(T^{\ell_i m_1} f_i \cdot \overline{T^{\ell_i m_2} f_i} \right) \right\|_{L^2(\mu)}.$$

By the induction hypothesis, we bound this by

$$\limsup_{M \rightarrow \infty} \mathbb{E}_{m_1, m_2 \in \text{IP}_{\Phi_M}} \|T^{\ell_1 m_1} f_1 \cdot \overline{T^{\ell_1 m_2} f_1}\|_{\text{U}_{\text{IP}}^{k-1}(\mathbf{X}, T)}$$

and by convexity of the map $x \mapsto x^{2^{k-1}}$, this is bounded by

$$\limsup_{M \rightarrow \infty} (\mathbb{E}_{m_1, m_2 \in \text{IP}_{\Phi_M}} \|T^{\ell_1 m_1} f_1 \cdot \overline{T^{\ell_1 m_2} f_1}\|_{\text{U}_{\text{IP}}^{k-1}(\mathbf{X}, T)}^{2^{k-1}})^{1/2^{k-1}} \leq \|f_1\|_{\text{U}_{\text{IP}}^k(\mathbf{X}, T)}^2.$$

Thus we shown that

$$\limsup_{N \rightarrow \infty} \left\| \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} T^{\ell_1 n} f_1 \cdot \dots \cdot T^{\ell_k n} f_k \right\|_{L^2(\mu)}^2 \leq \min_{1 \leq i \leq k} \|f_i\|_{\text{U}_{\text{IP}}^k(\mathbf{X}, T)}^2,$$

and taking square roots, the statement follows. \square

Lemma 4.10. *Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ be an invertible measure preserving system, let $k \geq 2$, and let $(f_\epsilon)_{\epsilon \in V_k} \in L^\infty(\mu)$ be 2^k functions. If $\|f_\epsilon\|_{U^k(\mathbf{X}, T)} = 0$ for some $\epsilon \in V_k$, then*

$$\int_{X^{[k]}} \bigotimes_{\epsilon \in V_k} C^{|\epsilon|} f_\epsilon d\tilde{\mu}^{[k]} = 0.$$

Proof. Let $\mathbf{Y} = (Y, C, \nu, S)$ be an invertible measure preserving system. For each eigenfunction $f \in L^2(\mathbf{Y})$ with rational eigenvalue, there exists some ℓ such that f is S^ℓ -invariant. Letting \mathcal{I}_ℓ denote the σ -algebra of ℓ -invariant functions, the sequence of σ -algebras $\mathcal{I}_{\ell!}$ is increasing and $\mathcal{K}_{\text{rat}}(\mathbf{Y})$ is a sub- σ -algebra of the join $\bigvee_{\ell \in \mathbb{N}} \mathcal{I}_{\ell!}$.

Conversely, for each $\ell \in \mathbb{N}$, taking f to be an S^ℓ -invariant function and letting ξ_ℓ be the first root of unity of order ℓ , we have that the functions

$$f_j = \sum_{i=0}^{\ell-1} \xi_\ell^{i \cdot j} \cdot T^i f$$

are eigenfunctions for all $1 \leq j \leq \ell$. Since $\sum_{i=0}^{\ell-1} \xi_\ell^{i \cdot j} = 0$ for all $1 \leq j \leq \ell-1$, we deduce that $f = \frac{1}{\ell} \sum_{j=1}^{\ell} f_j$. Thus the σ -algebra \mathcal{I}_ℓ of ℓ -invariant functions is a sub- σ -algebra of $\mathcal{K}_{\text{rat}}(\mathbf{Y})$, and $\mathcal{K}_{\text{rat}}(\mathbf{Y}) = \bigvee_{\ell \in \mathbb{N}} \mathcal{I}_{\ell!}$. In particular, for any function $f \in L^\infty(\mathbf{Y})$, we have that

$$\mathbb{E}(f \mid \mathcal{K}_{\text{rat}}(\mathbf{Y})) = \lim_{\ell \rightarrow \infty} \mathbb{E}(f \mid \mathcal{I}_{\ell!}).$$

Taking the Host-Kra measure $\mu_\ell^{[k]}$ on $\mathbf{X} = (X, \mathcal{B}, \mu, T^{\ell!})$ (see [30, Section 3]), then it follows from this and the definition of $\tilde{\mu}^{[k]}$ that for any function $f \in L^\infty(\mu)$ we have

$$\int_{X^{[k]}} \bigotimes_{\epsilon \in V_k} C^{|\epsilon|} f d\tilde{\mu}^{[k]} = \lim_{\ell \rightarrow \infty} \int_{X^{[k]}} \bigotimes_{\epsilon \in V_k} C^{|\epsilon|} f d\mu_\ell^{[k]}.$$

By a theorem of Leibman [35, Theorem 2], it follows that $\|f\|_{U^k(\mathbf{X}, T)} = 0$ is equivalent to

$$\int_{X^{[k]}} \bigotimes_{\epsilon \in V_k} C^{|\epsilon|} f d\mu_\ell^{[k]} = 0$$

for all $\ell \in \mathbb{N}$. It follows that

$$\int_{X^{[k]}} \bigotimes_{\epsilon \in V_k} C^{|\epsilon|} f d\tilde{\mu}^{[k]} = 0$$

\square

We have now assembled the tools to complete the proof of Theorem 4.2.

Proof of Theorem 4.2. Let $f_1, \dots, f_k \in L^\infty(\mu)$ and assume $\mathbb{E}(f_j \mid Z_{k-1}(\mathbf{X})) = 0$ for some $1 \leq j \leq k$. It follows from this assumption that $\|f_j\|_{\cup^k(\mathbf{X}, T)} = 0$ (see for example [30, Lemma 4.3]). By Propositions 4.7 and 4.9, we have that

$$\lim_{N \rightarrow \infty} \left\| \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \prod_{i=1}^k T^{\ell_i n} f_i \right\|_{L^2(\mu)}^{2k} \leq \int_{X^{[k]}} \bigotimes_{\epsilon \in V_k} f_j \, d\tilde{\mu}^{[k]}.$$

By Lemma 4.10, we deduce that

$$\lim_{N \rightarrow \infty} \left\| \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \prod_{i=1}^k T^{\ell_i n} f_i \right\|_{L^2(\mu)}^{2k} = 0$$

and therefore

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \prod_{i=1}^k T^{\ell_i n} f_i = 0. \quad \square$$

5. LIMIT FORMULA IN NILSYSTEMS

5.1. Limits for systems in which the spectrum has no nontrivial eigenvalues. Nilsystems are the natural class that arise in understanding uniformity norms and cubes in [30] and generalizations have been to understand combinatorial structures associated to arithmetic progressions and cubes, including work in [12, 13, 14, 28, 27, 29, 32, 33].

Recall that $\mathbf{X} = (X, \mathcal{B}, \mu, T_\tau)$ is a k -step nilsystem if $X = G/\Gamma$ for a k -step nilpotent Lie group G with discrete, cocompact subgroup Γ , \mathcal{B} is the Borel σ -algebra, μ is the unique Borel probability measure on X invariant under the action of G on X by left translation, and $T_\tau: X \rightarrow X$ is the action $x \mapsto \tau \cdot x$ for some fixed element $\tau \in G$. In case of ambiguity, we include the space on the measure and write $\mu = \mu_X = \mu_{G/\Gamma}$.

As with the ergodic average, the limit formula in a nilsystem depends crucially on the spectrum of T_τ . We start with the special case of a 1-step nilsystem whose spectrum is disjoint from $\sigma((n_j)_{j \in \mathbb{N}})$. Recall that $\mathbf{Z} = (Z, \mathcal{B}, \mu, R_a)$ is a 1-step nilsystem means that Z is a compact abelian group, \mathcal{B} is the Borel σ -algebra, μ is the Haar measure, and R_a is the rotation by a fixed element $a \in Z$. This system is ergodic if and only if the orbit of the element a is dense.

Theorem 5.1 (Limit formula for two terms). *Let $\mathbf{Z} = (Z, \mathcal{B}, \mu, R_a)$ be an ergodic 1-step nilsystem, let $(n_j)_{j \in \mathbb{N}}$ be a sequence whose spectrum contains no nontrivial eigenvalue of Z , and let $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ be an increasing Følner sequence. Then for all $f_1, \dots, f_k \in L^\infty(\mu)$ and integers $\ell_1, \dots, \ell_k \in \mathbb{N}$, we have*

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \prod_{i=1}^k f_i(T^{\ell_i n} x) = \int_Z \prod_{i=1}^k f_i(x + \ell_i t) d\mu(t)$$

in $L^2(\mu)$.

Proof. Since linear combinations of characters are dense in $L^2(\mu)$, it suffices to establish the formula when f_1, \dots, f_k are characters of Z . Furthermore, since Z is an abelian group, it suffices to assume that they are eigenfunctions with eigenvalues $\alpha_1, \dots, \alpha_k$ respectively. Recall that ω_Φ is defined in (3.6) and δ_1 in (3.1). Since the spectrum of $(n_j)_{j \in \mathbb{N}}$ contains no nontrivial eigenvalue of Z , we have that $\omega_\Phi(\alpha) = \delta_1(\alpha)$ for all eigenvalues of the system. By Theorem 3.8, and Fourier analysis it follows that

$$\begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \prod_{i=1}^k f_i(T^{\ell_i n} x) &= \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \left(\prod_{i=1}^k \alpha_i^{\ell_i n} \right) \cdot \left(\prod_{i=1}^k f_i(x) \right) \\ &= \omega_\Phi \left(\prod_{i=1}^k \alpha_i^{\ell_i} \right) \cdot \left(\prod_{i=1}^k f_i(x) \right) = \delta_1 \left(\prod_{i=1}^k \alpha_i^{\ell_i} \right) \cdot \left(\prod_{i=1}^k f_i(x) \right) \\ &= \int_Z \prod_{i=1}^k f_i(x + \ell_i t) d\mu(t), \end{aligned}$$

completing the proof. \square

For a nilsystem $\mathbf{X} = (G/\Gamma, \mathcal{B}, \mu, R_a)$ and integers ℓ_1, \dots, ℓ_k , set

$$\tilde{G} = \tilde{G}_{\ell_1, \dots, \ell_k} = \left\{ \left(\prod_{i=1}^{k-1} g_i^{\binom{\ell_j}{i}} u_j \right)_{j \in \{1, \dots, k\}} : g_i \in G_i, u_j \in G_k, 1 \leq i, j \leq k \right\} \leq G^k,$$

where $\binom{n}{m} = 0$ if $m > n$. (As ℓ_1, \dots, ℓ_k are fixed, we simplify the notation and write $\tilde{G} = \tilde{G}_{\ell_1, \dots, \ell_k}$.) Setting $\tilde{\Gamma} = \tilde{G} \cap \Gamma^k$, it is shown in [34] that \tilde{G} is a closed subgroup of G^k . The discrete subgroup $\tilde{\Gamma}$ is cocompact and so $\tilde{X} = \tilde{G}/\tilde{\Gamma}$ is a nilmanifold endowed with a Haar measure $\tilde{\mu}$.

The main goal of this section is prove the following limit formula.

Theorem 5.2. *Let $k \geq 1$, let $\mathbf{X} = (G/\Gamma, \mathcal{B}, \mu_{G/\Gamma}, T_\tau)$ be an ergodic k -step nilsystem, let $(n_j)_{j \in \mathbb{N}}$ be a sequence whose spectrum contains no nontrivial eigenvalue of T_τ , and let $(\Phi_N)_{N \in \mathbb{N}}$ be an increasing Følner sequence. For all $f_1, \dots, f_k \in L^\infty(\mu)$ and distinct integers $\ell_1, \dots, \ell_k \in \mathbb{N}$, we have*

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \prod_{i=1}^k f_i(T_\tau^{\ell_i n} x) = \int_{\tilde{X}} \prod_{i=1}^k f_i(xy_i) d\tilde{\mu}(y\tilde{\Gamma})$$

in $L^2(\mu)$, where $y = (y_1, \dots, y_k) \in \tilde{G}$.

For general (non-IP) multiple averages Ziegler [42] derived a limit formula, and a different method for doing so was given in [7]. Both proofs rely on the unique ergodicity of nilsystems and the existence of the limit of the mean ergodic average everywhere for continuous functions on nilsystems. In our context of evaluating the averages along IP-Følner sequences, we need a different approach. Instead, we identify a class of functions that is dense and for which we can show convergence everywhere, and this class suffices to prove for convergence in $L^2(\mu)$ for all bounded functions.

We start with the case where the set of eigenvalues of the nilsystem is disjoint from the spectrum of $(n_j)_{j \in \mathbb{N}}$.

Theorem 5.3 (Pointwise convergence along IPs for nilsystems with disjoint spectrum). *Let $k \geq 1$ and let $\mathbf{X} = (G/\Gamma, \mathcal{B}, \mu, T_\tau)$ be an ergodic k -step nilsystem. There exists an ergodic system $\mathbf{Y} = (Y, \mathcal{C}, \nu, T)$ on a compact metric space Y such that the set of T -periodic points has ν -measure zero and there exists a measure theoretic isomorphism $\pi: X \rightarrow Y$ such that for any sequence $(n_j)_{j \in \mathbb{N}}$ whose spectrum contain no nontrivial eigenvalue of T_τ and every increasing Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$, we have that*

$$(5.1) \quad \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} f(T_\tau^n y) = \int_Y f d\mu$$

for any continuous function $f \in C(Y)$ and $y \in Y$.

In general, the isomorphism $\pi: X \rightarrow Y$ need not be continuous. The hypotheses, and in particular the assumption that the T -periodic points have ν -measure zero, is used to apply a theorem of Lindenstrauss [36, Theorem 3.1].

Proof. We proceed by induction on the step k . When $k = 1$, \mathbf{X} is an ergodic rotation on a compact abelian group $Z = G/\Gamma$ and we can take $\mathbf{Y} = \mathbf{X}$ and $\pi: X \rightarrow Y$ to be the identity map. In this case, there are at most finitely many T -periodic points in Y . Let $p: Z \rightarrow \mathbb{C}$ be a linear combination of characters in Z and write

$$p = \sum_{i=1}^n b_i \xi_i$$

for some $b_i \in \mathbb{C}$ and $\xi_i \in \widehat{Z}$. Letting $a \in Z$ denote the rotation induced by T_τ , we then have that for all $y \in Y$,

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} p(T^n y) = \sum_{i=0}^n b_i \left(\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \xi_i(a)^n \right) \xi_i(y).$$

By assumption, no nontrivial eigenvalue of Z lies in the spectrum of $(n_j)_{j \in \mathbb{N}}$, and so for each i , the limit on the right hand side is zero unless $\xi_i = 0$ and in this case the limit is 1. In both cases, the limit exists and is equal to $b_0 = \int_Y p(y) d\mu$. Taking $f \in C(Y)$ and $\varepsilon > 0$, there exists a finite linear combination of characters $p: Z \rightarrow \mathbb{C}$ such that $\|f - p\|_\infty < \varepsilon/2$. Fixing some $y \in Y$, using this approximation of f by p , and noting that by choice we have $X = Y$ and T is the rotation induced by $a \in Z$, we have

$$\begin{aligned} & \limsup_{N \rightarrow \infty} \left| \mathbb{E}_{n \in \text{IP}_{\Phi_N}} f(T^n y) - \int_Y f d\mu \right| \\ & \leq \limsup_{N \rightarrow \infty} \left| \mathbb{E}_{n \in \text{IP}_{\Phi_N}} f(a^n y) - \mathbb{E}_{n \in \text{IP}_{\Phi_N}} p(a^n y) \right| \\ & \quad + \limsup_{N \rightarrow \infty} \left| \mathbb{E}_{n \in \text{IP}_{\Phi_N}} p(a^n y) - \int_Z p d\mu \right| + \left| \int_Y f d\mu - \int_Z p d\mu \right| \\ & \leq 2\|f - p\|_\infty < \varepsilon. \end{aligned}$$

Since $\varepsilon > 0$ is arbitrary, we deduce that

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} f(T^n y) = \int_Y f d\mu,$$

completing the case $k = 1$.

Fix some $k \geq 2$. Combining [30, Theorem 10.3 and Section 11], since \mathbf{X} is an ergodic k -step nilsystem, it is measure theoretically isomorphic to a k -step toral system, meaning that we can assume that the Kronecker factor $Z_1(\mathbf{X})$ is a compact abelian Lie group and each nilfactor $Z_{j+1}(\mathbf{X})$ is an extension of $Z_j(\mathbf{X})$ by a torus for $1 \leq j < k$. Thus we can write $X = Z_{k-1}(\mathbf{X}) \times_\rho U_k$,

for some torus U_k and cocycle $\rho: Z_{k-1}(\mathbf{X}) \rightarrow U_k$. Applying the induction hypothesis to $Z_{k-1}(\mathbf{X})$, there exists a system $\mathbf{Y}_{k-1} = (Y_{k-1}, C_{k-1}, \nu_{k-1}, T_{k-1})$ that is isomorphic to $Z_{k-1}(\mathbf{X})$ satisfying the properties of the theorem, meaning that \mathbf{X} is measure theoretically isomorphic to an abelian extension of \mathbf{Y}_{k-1} by a torus U_k via a cocycle $\rho: Y_{k-1} \rightarrow U_k$. Using a theorem of Lindenstrauss [36, Theorem 3.1], ρ is cohomologous to a continuous cocycle. As cohomologous cocycles give rise to isomorphic extensions, without loss of generality we may assume that ρ is continuous and write $Y = Y_{k-1} \times_\rho U_k$ and let $T: Y \rightarrow Y$ denote the action induced by the cocycle ρ . Letting P_{k-1} denote the set of periodic points in the system \mathbf{Y}_{k-1} , then the set of periodic points in the system \mathbf{Y} is a subset of $P_{k-1} \times U_k$ and is therefore of measure 0. We are left with showing that \mathbf{Y} satisfies the remaining properties in the statement.

By Lemma 2.2, the space of continuous vertical characters on \mathbf{Y} is dense with respect to the uniform norm. Thus it suffices to prove the limit formula for functions of the form $f_\chi(y, u) = g(y)\chi(u)$ for some character $\chi: U_k \rightarrow \mathcal{S}^1$ and continuous function $g: Y_{k-1} \rightarrow \mathbb{C}$. If χ is trivial, then the convergence follows from the induction hypothesis. Otherwise, we have

$$\limsup_{N \rightarrow \infty} |\mathbb{E}_{n \in \text{IP}_{\Phi_N}} f_\chi(T^n(y, t))| = \limsup_{N \rightarrow \infty} |\mathbb{E}_{n \in \text{IP}_{\Phi_N}} g(T_{k-1}^n y) \chi \circ \rho(n, y) f_\chi(y, t)|.$$

By the van der Corput inequality (Lemma 4.3), we can bound the square of the right hand side by

$$(5.2) \quad \limsup_{M \rightarrow \infty} \mathbb{E}_{m_1, m_2 \in \text{IP}_{\Phi_M}} \limsup_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N^M}} g(T_{k-1}^{n+m_1} y) \cdot \overline{g(T_{k-1}^{n+m_2} y)} \cdot \overline{\rho(m_1, T^n y)} \cdot \overline{\rho(m_2, T^n y)},$$

where $\Phi_N^M = \Phi_N \setminus \Phi_M$ and we have used the cocycle identity $\rho(n+m, y) = \rho(n, y)\rho(m, T^n y)$ for all $m, n \in \mathbb{N}$ and $y \in Y$. Since ρ and g are continuous, using the inductive assumption applied to the convergence in (5.1), we can rewrite the quantity in (5.2) as

$$\begin{aligned} \limsup_{M \rightarrow \infty} \mathbb{E}_{m_1, m_2 \in \text{IP}_{\Phi_M}} \int_{Y_{k-1}} g(T_{k-1}^{m_1}(y)) \overline{g(T_{k-1}^{m_2} y)} \overline{\rho(m_1, y)} \cdot \overline{\rho(m_2, y)} d\mu(y) \\ = \limsup_{M \rightarrow \infty} \left\| \mathbb{E}_{m \in \text{IP}_{\Phi_M}} g(T_{k-1}^m \cdot) \rho(m, \cdot) \right\|_{L^2(Y_{k-1})}^2. \end{aligned}$$

Applying the mean ergodic theorem in \mathbf{Y} to the function f_χ , the right hand side of this equation converges to 0 as $M \rightarrow \infty$. We conclude that whenever $\chi \neq 1$,

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} f_\chi(T^n(y, t)) = 0$$

for all $(y, t) \in Y$, completing the proof. \square

We are set to prove Theorem 5.2.

Proof. With the preliminaries adapted to the setting of IPs, we now follow the argument in [7]. For every $x \in G/\Gamma$, let $g_x \in G$ be a lift of x and set $\tilde{\Gamma}_x = \{(g_x \gamma_1 g_x^{-1}, \dots, g_x \gamma_k g_x^{-1}) : (\gamma_1, \dots, \gamma_k) \in \tilde{\Gamma}\}$. We consider the nilmanifold $\tilde{X}_x = \tilde{G}/\tilde{\Gamma}_x$ endowed with the Haar measure $\tilde{\mu}_x$. By [7, Corollary 5.5], the action of $T_\tau^{\ell_1} \times \dots \times T_\tau^{\ell_k}$ on \tilde{X}_x is ergodic for μ -almost every $x \in X$.

Let \mathbf{Y} be as in Theorem 5.3 and let $f_1, \dots, f_k: X \rightarrow \mathbb{C}$ be the lifts of some continuous functions $g_1, \dots, g_k: Y \rightarrow \mathbb{C}$ (thus $f_i = g_i \circ \pi$ where $\pi: \mathbf{X} \rightarrow \mathbf{Y}$ is an isomorphism). By Theorem 5.3 applied to the point (x, x, \dots, x) , we deduce that

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \prod_{i=1}^k f_i(T_\tau^{\ell_i n} x)$$

converges everywhere. By ergodicity, it must converge to the integral with respect to the Haar measure $\tilde{\mu}_x$. This holds for all f_1, \dots, f_k that arise as such lifts, and since the continuous functions are dense in $L^\infty(\mu_Y)$, the lifts of such functions are dense in $L^\infty(\mu_X)$. Thus, we have convergence in $L^2(\mu_X)$ to the desired limit for all $f_1, \dots, f_k \in L^\infty(\mu_X)$. \square

5.2. Limit formulae for weighted averages. In Theorems 5.1 and 5.2, we assumed that the spectrum of $(n_j)_{j \in \mathbb{N}}$ contains no nontrivial eigenvalues. To address settings where this does not hold, we study certain weighted averages.

Lemma 5.4 (Characteristic factors for multiple weighted averages.). *Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ be an ergodic system, $(n_j)_{j \in \mathbb{N}}$ be a sequence with rational spectrum, $k \geq 2$, $f_1, \dots, f_k \in L^\infty(\mathbf{X})$, $\ell_1, \dots, \ell_k \in \mathbb{N}$, and let $a \in Z_1(\mathbf{X})$ denote the rotation induced on the Kronecker factor $Z_1(\mathbf{X})$ by T . If $\ell_1, \ell_2 \in \mathbb{N}$ are coprime, $\eta: Z_1(\mathbf{X}) \rightarrow \mathbb{R}$ is a non-negative continuous function with*

$\mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \eta(na) = 1$, and $\delta > 0$, there exists a $(k-1)$ -step nilsystem factor $\mathbf{Y} = (Y, \mathcal{Y}, \nu, T)$ such that

$$\left\| \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \eta(na) \prod_{i=1}^k f_i(T^{\ell_i n} x) - \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \eta(na) \prod_{i=1}^k \mathbb{E}(f_i \mid \mathcal{Y})(T^{\ell_i n} x) \right\|_{L^2(\mu)} < \delta$$

for all increasing Følner sequences and sufficiently large N .

Proof. We start with the case that η is a character. Since ℓ_1, ℓ_2 are coprime, we can choose $s, t \in \mathbb{Z}$ such that $s \cdot \ell_1 + t \cdot \ell_2 = 1$. Letting $\pi: X \rightarrow Z_1(\mathbf{X})$ denote the factor map, set $\tilde{\eta} = \eta \circ \pi$ and write $f'_1 = f_1 \cdot \tilde{\eta}^s$, $f'_2 = f_2 \cdot \tilde{\eta}^t$, and $f'_i = f_i$ for all $i \geq 3$. Then

$$\mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \eta(na) \prod_{i=1}^k f_i(T^{\ell_i n} x) = \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \prod_{i=1}^k f'_i(T^{\ell_i n} x).$$

By Theorem 4.2, this limit is zero if at least one function f'_i is orthogonal to $Z_{k-1}(\mathbf{X})$. Since η is measurable with respect to $Z_1(\mathbf{X})$, the function f'_i is orthogonal to the factor $Z_{k-1}(\mathbf{X})$ if and only if f_i is orthogonal to $Z_1(\mathbf{X})$. For general η , using the Stone-Weierstrass Theorem we can approximate η uniformly by linear combinations of characters of $Z_1(\mathbf{X})$. Thus, we may assume without loss of generality that all f_i are all measurable with respect to $Z_{k-1}(\mathbf{X})$.

Let $\delta > 0$. Since the factor $Z_{k-1}(\mathbf{X})$ is an inverse limit of nilsystems [30, 43], there is a $(k-1)$ -step nilsystem \mathbf{Y} such that $\|f_i - \mathbb{E}(f_i \mid \mathcal{Y})\| < \delta/2k$. By the Cauchy-Schwartz and triangle inequalities, we have that

$$\left\| \prod_{i=1}^k f_i \circ T^{\ell_i n} - \prod_{i=1}^k \mathbb{E}(f_i \mid \mathcal{Y}) \circ T^{\ell_i n} \right\|_{L^2(\mu)} < \delta/2.$$

Since $\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta(na) = 1$, it follows that for N sufficiently large we have that

$$\begin{aligned} \left\| \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta(na) \prod_{i=1}^k f_i(T^{\ell_i n}(x)) - \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta(na) \prod_{i=1}^k \mathbb{E}(f_i \mid \mathcal{Y})(T^{\ell_i n}(x)) \right\|_{L^2(\mu)} \\ < \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta(na) \cdot \delta/2 \leq \delta. \quad \square \end{aligned}$$

We make use of various results about nilmanifolds, and refer to [31] for further details and background. For an ergodic nilsystem $\mathbf{X} = (G/\Gamma, \mathcal{B}, \mu, T_\tau)$, we let G^0 denote the connected component of G and set $\Gamma^0 = G^0 \cap \Gamma$. The nilmanifold G^0/Γ^0 can be embedded as a subnilmanifold of G/Γ , and since the quotient $G/G^0\Gamma$ is finite, there is some $r \in \mathbb{N}$ such that τ^r maps G^0/Γ^0 to itself and we have

$$G/\Gamma = \bigsqcup_{i=0}^{r-1} \tau^i G^0/\Gamma^0.$$

Let \mathbf{X}^0 denote the associated system (endowed with the Borel σ -algebra, Haar measure μ^0 , and transformation $T_\tau^r = T_{\tau^r}$).

Definition 5.5. An ergodic nilsystem $\mathbf{X} = (G/\Gamma, \mathcal{B}, \mu, T_\tau)$ is *synchronized* on G^0/Γ^0 , or *synchronized* if the context is clear, if there exist some $b \in G$ and $r \in \mathbb{N}$ such that

- (1) $T_b: G/\Gamma \rightarrow G/\Gamma$ maps G^0/Γ^0 to itself;
- (2) The action of T_b on G^0/Γ^0 is ergodic;
- (3) The elements $b, \tau \in G$ satisfy $b^r = \tau^r$.

We note that for general nilsystems, property (1) does not imply that $b \in G^0$. For instance, any totally ergodic nilsystem \mathbf{X} is synchronized, as then $X = G^0/\Gamma^0$ and we can take $b = \tau$ and $r = 1$. Moreover, all ergodic 1-step nilsystems are synchronized, as seen by projecting the rotation τ of a compact abelian Lie group Z to the torus part of Z . We show that this extends to 2-step nilsystems.

Lemma 5.6. *Let $k = 1$ or $k = 2$. Then every ergodic k -step nilsystem $\mathbf{X} = (G/\Gamma, \mathcal{B}, \mu, T_\tau)$ is synchronized on G^0/Γ^0 for some choice of G, Γ .*

The proof is technical and relies on results from [30] not directly relevant to this paper, and so we defer it to Appendix C. While it is not needed for our current work, it would be interesting to know if all nilsystems are synchronized.

For $r \in \mathbb{N}$, let $\mathbf{1}_{r\mathbb{N}}: \mathbb{N} \rightarrow \{0, 1\}$ denote the function which assigns 1 to any integer divisible by r and 0 otherwise. Given a Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$, set

$$\mathbf{1}_{r,\Phi}(n) = \frac{\mathbf{1}_{r\mathbb{N}}(n)}{\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \mathbf{1}_{r\mathbb{N}}(n)}.$$

This is well defined: the limit in the denominator exists because $\mathbf{1}_{r\mathbb{N}}(n) = \frac{1}{r} \sum_{i=0}^{r-1} \xi_r^{in}$, where ξ_r is the first root of unity of degree r , and so is a periodic function.

Proposition 5.7. *Let $\mathbf{X} = (G/\Gamma, \mathcal{B}, \mu, T_\tau)$ be a nilsystem and suppose that \mathbf{X} is synchronized on G^0/Γ^0 . Let $(n_j)_{j \in \mathbb{N}}$ be a sequence with rational spectrum and let $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ be an increasing Følner sequence. Then there exists some $r \in \mathbb{N}$ and a set of bounded functions \mathcal{F} that are dense in $L^2(\mu)$ such that for all $f \in \mathcal{F}$, we have*

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \mathbf{1}_{r, \Phi}(n) f(\tau^n x) = \int_{X^0} f(xy) d\mu^0(y \Gamma^0)$$

for μ_G -almost every $x \in G$, where $X^0 = G^0/\Gamma^0$ is the connected component of the identity in X and is endowed with the Haar measure μ^0 .

Proof. Since the system \mathbf{X} is synchronized, we can find some r and $\tau \in G$ such that $\tau^r = b^r$ and b maps the connected component of X to itself. Since $\mathbf{1}_{r, \Phi}(n) \neq 0$ if and only if $\tau^n = b^n$, it suffices to compute the limit

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \mathbf{1}_{r, \Phi}(n) f(b^n x).$$

Decompose $X = \bigsqcup_{i=0}^{r'-1} X_i$ into connected components and note that by increasing r if necessary we may assume without loss of generality that r' divides r . The translation by b induces an ergodic action on each X_i . Thus the function f can be written as a sum of functions on each of these components. Therefore, without loss of generality, it suffices to find a dense set of bounded functions $f: X^0 \rightarrow \mathbb{C}$ such that

$$(5.3) \quad \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \mathbf{1}_{r, \Phi}(n) f(b^n x) = \int_{X^0} f(x) d\mu^0(x).$$

We proceed by induction on k . For $k = 1$, the system $\mathbf{X}^0 = (X^0, \mathcal{B}_{X^0}, \mu^0, R_b)$ is an ergodic rotation on a connected compact abelian group Z . Linear combinations of characters are dense in $C(Z)$, and so it suffices to prove the result for a character f . If $f = 1$, then it follows from the definition of $\mathbf{1}_{r, \Phi}$ that the limit in (5.3) is 1. Otherwise, since X^0 is connected, $f(b^n x) = \alpha^n f(x)$ for some irrational α . On the other hand $\mathbf{1}_{r, \Phi}(n) = c \cdot \sum_{i=0}^{r-1} \xi_r^{in}$ where ξ_r is the first root of unity of degree r and $c \in \mathbb{R}$ is a constant. We have

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \mathbf{1}_{r, \Phi}(n) f(b^n x) = c \sum_{i=0}^{r-1} \omega_\Phi(\xi_r^i \cdot \alpha) \cdot f(x)$$

where ω_Φ is defined as in (3.6). However, since $\xi_r^i \alpha$ is irrational, and $(n_j)_{j \in \mathbb{N}}$ has rational spectrum this quantity is zero, completing the case $k = 1$.

To proceed inductively, we use a similar argument to that in the proof of Theorem 5.3. Namely, the system \mathbf{X}^0 is isomorphic to a system \mathbf{Y} and $Y = Y_{k-1} \times_\rho U$ for some connected group U and a continuous cocycle ρ . We then consider a nilcharacter $f_\chi(y, t) = g(y)\chi(t)$, where $\chi: U \rightarrow \mathcal{S}^1$ is a nontrivial character and $g: Y_{k-1} \rightarrow \mathbb{C}$ is continuous. Once again write $\mathbf{1}_{r,\Phi}(n) = c \cdot \sum_{i=0}^{r-1} \xi_r^{in}$, fix some $0 \leq i \leq r-1$, and set $s = \xi_r^i$. Using the same notation as in the proof of Theorem 5.3 (so T_{k-1} denotes the action on \mathbf{Y}_{k-1} and T the action on \mathbf{Y}), we have

$$\begin{aligned} \limsup_{N \rightarrow \infty} \left| \mathbb{E}_{n \in \text{IP}_{\Phi_N}} s^n \cdot f_\chi(T^n(y, t)) \right| \\ = \limsup_{N \rightarrow \infty} \left| \mathbb{E}_{n \in \text{IP}_{\Phi_N}} s^n g(T_{k-1}^n y) \chi \circ \rho(n, y) f_\chi(y, t) \right|. \end{aligned}$$

By the van der Corput inequality (Lemma 4.3) and the induction hypothesis, this is bounded by

$$\limsup_{M \rightarrow \infty} \left\| \mathbb{E}_{m \in \text{IP}_{\Phi_M}} s^m g(T_{k-1}^m y) \rho(m, y) \right\|_{L^2(\mu_{k-1})}^2 = 0$$

Since X^0 is connected, it is totally ergodic and so the spectrum is supported on the irrational elements in \mathcal{S}^1 . Therefore, we have the L^2 -convergence of the average $\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} s^n f_\chi(T^n(y, t)) = 0$ for all rational $s \in \mathcal{S}^1$. Th, completing the proof. \square

As a corollary, we obtain a formula for weighted averages on synchronized nilsystems.

Corollary 5.8. *Let $\mathbf{X} = (G/\Gamma, \mathcal{B}, \mu, T_\tau)$ be an ergodic k -step nilsystem that is synchronized on G^0/Γ^0 . Let $\eta: Z_1(\mathbf{X}) \rightarrow \mathbb{R}$ be a non-negative continuous function such that $\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \eta(na) = 1$ where a denotes the rotation induced on the Kronecker factor $Z_1(\mathbf{X})$, let $f_1, \dots, f_k \in L^\infty(\mu)$ be bounded functions, and let $\ell_1, \dots, \ell_k \in \mathbb{N}$ be distinct integers such that ℓ_1, ℓ_2 are coprime. Then there exists $r \in \mathbb{N}$ such that for every increasing Følner*

sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$, we have

$$\begin{aligned} \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})} \eta(na) \mathbf{1}_{r, \Phi}(n) \prod_{i=1}^k f_i(T^{\ell_i n} x) \\ = \int_{\tilde{X}^0} \eta(y_1) \prod_{i=1}^k f_i(xy_i) \, d\tilde{\mu}(y \tilde{\Gamma}_{\ell_1, \dots, \ell_k}), \end{aligned}$$

where \tilde{X}^0 denotes the connected component of \tilde{X} and $\tilde{\mu}^0$ the Haar measure on \tilde{X}^0 .

Proof. As in the proof of Lemma 5.4, we can absorb η into f_1, f_2 and so without loss of generality we assume that $\eta = 1$. As in the proof of Proposition 5.7, we can reduce to the case that the functions are defined on the connected component of the identity (note that all of these component are isomorphic). Using the proof of Theorem 5.2, the limit is obtained by applying a pointwise convergence result for an ergodic average of a single function on a particular nilsystem. We therefore obtain the required formula using the same proof as in Theorem 5.2, only replacing the use of Theorem 5.3 by Proposition 5.7. \square

6. THE LARGE INTERSECTION PROPERTY

6.1. Notions of largeness. Depending on the choice of sequence, different notions of size are relevant.

Definition 6.1 (IP-density). Let $(n_j)_{j \in \mathbb{N}}$ be a sequence of natural numbers, let $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ be an increasing sequence, and let $A \subseteq \text{IP}((n_j)_{j \in \mathbb{N}})$ be a subset. We define the *lower IP-density* of A with respect to Φ as the quantity

$$d_{\text{IP}\Phi}(A) := \liminf_{N \rightarrow \infty} \frac{|A \cap \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})|}{|\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})|}.$$

Other notions of largeness for subsets of \mathbb{N} also have counterparts for IPs, and we include a discussion in Appendix A.

6.2. Lower bounds for three terms.

Theorem 6.2. Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ be an ergodic invertible measure preserving system, let $(n_j)_{j \in \mathbb{N}}$ be a sequence with rational spectrum, and let

$A \subseteq X$ be a measurable set. Then for all $\varepsilon > 0$, increasing Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$, and coprime and nonzero integers $\ell_1, \ell_2 \in \mathbb{Z}$, we have

$$\underline{d}_{\text{IP}_\Phi} \{n \in \text{IP}((n_j)_{j \in \mathbb{N}}) : \mu(A \cap T^{-\ell_1 n} A \cap T^{-\ell_2 n} A) > \mu(A)^3 - \varepsilon\} > 0.$$

Proof. The factor $Z_1(\mathbf{X})$ is (measurably isomorphic to) a rotation on a compact abelian group and we let $a \in Z_1(\mathbf{X})$ denote the rotation. Let $\eta : Z_1(\mathbf{X}) \rightarrow \mathbb{R}$ be a non-negative continuous function (yet to be specified) such that $\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta(na) = 1$. For contradiction, suppose that the statement is false for some $0 < \varepsilon < 1$. Then there exists an increasing Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$ so that

$$(6.1) \quad \lim_{N \rightarrow \infty} \frac{|R_{\varepsilon, N}|}{|\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})|} = 1,$$

where

$$R_{\varepsilon, N} = \{n \in \text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}}) : \mu(A \cap T^{-\ell_1 n} A \cap T^{-\ell_2 n} A) \leq \mu(A)^3 - \varepsilon\}.$$

Since η is continuous and $Z_1(\mathbf{X})$ is compact, it is bounded by a constant (not depending on N), and so we have

$$\begin{aligned} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta(na) \mu(A \cap T^{-\ell_1 n} A \cap T^{-\ell_2 n} A) \\ \leq \frac{1}{|\text{IP}_{\Phi_N}|} \sum_{n \in R_{\varepsilon, N}} \eta(na) (\mu(A)^3 - \varepsilon) + \|\eta\|_\infty \cdot o_{N \rightarrow \infty}(1) \\ \leq \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta(na) \cdot (\mu(A)^3 - \varepsilon) + \|\eta\|_\infty \cdot o_{N \rightarrow \infty}(1). \end{aligned}$$

Passing to a subsequence of Φ_N if necessary and taking $N \rightarrow \infty$, we conclude that

$$(6.2) \quad \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta(na) \mu(A \cap T^{-\ell_1 n} A \cap T^{-\ell_2 n} A) < \mu(A)^3 - \varepsilon.$$

Choosing $\delta > 0$ sufficiently small with respect to ε , it follows from Lemma 5.4 that there exists a 1-step nilsystem factor \mathbf{Y} such that if f is the projection of $\mathbf{1}_A$ onto \mathbf{Y} , then

$$(6.3) \quad \left| \lim_{N \rightarrow \infty} \left(\mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta(na) \mu(A \cap T^{-\ell_1 n} A \cap T^{-\ell_2 n} A) - \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta(na) \int_{\mathbf{Y}} f(x) f(T^{\ell_1 n} x) f(T^{\ell_2 n} x) d\mu_{\mathbf{Y}}(x) \right) \right| < \varepsilon/2.$$

As a 1-step nilsystem is a rotation on a compact abelian Lie group, we can write $Y = \mathbb{T}^m \times D$ for some finite dimensional torus $\mathbb{T}^m = (\mathbb{R}/\mathbb{Z})^m$ for some $m \geq 1$ and finite cyclic group D , with the rotation by the element $a = (a_1, a_2)$

for some $a_1 \in \mathbb{T}^m$ and $a_2 \in D$. Let $\eta: Y \rightarrow \mathbb{R}$ be a function of the form $\eta(t, d) = \eta_1(t) \cdot c^{-1} \mathbf{1}_{\{0\}}(d)$, where $\eta_1: \mathbb{T}^m \rightarrow \mathbb{R}$ is a non-negative continuous function still to be chosen, $\mathbf{1}_{\{0\}}(d)$ takes the value 1 if $d = 0$ and the value 0 otherwise, and $c := \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \mathbf{1}_{\{0\}}(na_2)$. Then equation (6.3) holds for this η . Note that $\mathbf{1}_{\{0\}}(na_2) = \mathbf{1}_r(n)$ where $r = |D|$ and so c is nonzero and $c^{-1} \cdot \mathbf{1}_{\{0\}}(na_2) = \mathbf{1}_{r, \Phi}(n)$. Lifting η to the Kronecker factor of \mathbf{X} , by equation (6.3) and Corollary 5.8, we have that the left hand side of (6.2) is $\varepsilon/2$ -close to the quantity

$$\int_Y \int_{\mathbb{T}} \eta(t) f(x) f(x + \ell_1 t) f(x + \ell_2 t) d\mu_{\mathbb{T}}(t) d\mu_Y(x).$$

Let V be small neighborhood of identity in \mathbb{T}^m and assume that η_1 is a continuous function supported on V such that $\int_{\mathbb{T}^m} \eta_1 d\mu = 1$. Since the action of a_1 on \mathbb{T}^m is totally ergodic, we have that

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta_1(na_1) \cdot c^{-1} \mathbf{1}_{\{0\}}(na_2) = \int_{\mathbb{T}} \eta(t) dt = 1.$$

It follows that by taking the neighborhood V to be sufficiently small, the left hand side of Equation (6.2) is ε -close to

$$\int_Z f(x) f(x) f(x) d\mu_Z(x) \geq \left(\int_Z f(x) d\mu_Z(x) \right)^3 = \mu(A)^3.$$

In other words,

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta(na) \mu(A \cap T^{-\ell_1 n} A \cap T^{-\ell_2 n} A) > \mu(A)^3 - \varepsilon,$$

a contradiction of the inequality in (6.2). \square

6.3. Four terms. Before extending the lower bounds to four terms, we derive a result that follows from Frantzikinakis [16, Proof of Theorem C].

Lemma 6.3. *Let $\mathbf{X} = (G/\Gamma, \mathcal{B}, \mu, T)$ be an ergodic 2-step nilsystem, $f \in L^\infty(\mu)$ be a bounded function, ℓ_1, ℓ_2 be coprime and nonzero integers, and set $\ell_0 = 0$ and $\ell_3 = \ell_1 + \ell_2$. Then*

$$\int_X \int_{G_2} \int_{G_2} \prod_{i=0}^3 f(xy_1^{\ell_i} y_2^{\binom{\ell_i}{2}}) d\mu_{G_2}(y_2) d\mu_{G_2}(y_1) d\mu(x) \geq \left(\int_X f d\mu \right)^4.$$

Proof. Since G_2 commutes with all elements in G , we can rewrite the integral on the left hand side as

$$\int_X \int_{G_2 \times G_2} \prod_{i=0}^3 f(y_1^{\ell_i} y_2^{\binom{\ell_i}{2}} x) d\mu_{G_2 \times G_2}(y_1, y_2) d\mu_X(x).$$

Taking the integral over G_2 , this becomes

$$(6.4) \quad \int_X \int_{G_2 \times G_2 \times G_2} \prod_{i=0}^3 f(y y_1^{\ell_i} y_2^{\binom{\ell_i}{2}} x) d\mu_{G_2 \times G_2 \times G_2}(y, y_1, y_2) d\mu_X(x).$$

Reparameterizing the set

$$\{(y, y y_1^{\ell_1} y_2^{\binom{\ell_1}{2}}, y y_1^{\ell_2} y_2^{\binom{\ell_2}{2}}, y y_1^{\ell_1 + \ell_2} y_2^{\binom{\ell_1 + \ell_2}{2}} : y, y_1, y_2 \in G_2\}$$

as

$$\{(h_1, h_2, h_3, h_4) \in G_2^4 : h_1^{\ell_2 - \ell_1} h_3^{\ell_1 + \ell_2} = h_4^{\ell_2 - \ell_1} h_2^{\ell_1 + \ell_2}\},$$

we can rewrite (6.4) as

$$\int_X \int_{G_2} \left(\int_{\{h_1^{\ell_2 - \ell_1} h_3^{\ell_1 + \ell_2} = h\}} f(h_1 x) f(h_3 x) d\mu_{G_2 \times G_2}(h_1, h_3) \right)^2 d\mu_{G_2}(h) d\mu_X(x).$$

By the Cauchy-Schwartz Inequality and a change of variables, this last expression is greater than or equal to

$$\int_X \left(\int_{G_2} f(hx) d\mu_{G_2}(h) \right)^4 d\mu_X(x).$$

This, in turn, is greater or equal than

$$\left(\int_X \int_{G_2} f(hx) d\mu_{G_2}(h) d\mu_X(x) \right)^4 = \left(\int_X f(x) d\mu_X(x) \right)^4,$$

as required. \square

Theorem 6.4. *Let $\mathbf{X} = (X, \mathcal{B}, \mu, T)$ be an ergodic invertible measure preserving system, let $(n_j)_{j \in \mathbb{N}}$ be a sequence with rational spectrum, and let $A \subseteq X$ be a measurable set. Then for all $\varepsilon > 0$, increasing Følner sequence $\Phi = (\Phi_N)_{N \in \mathbb{N}}$, and coprime and nonzero integers $\ell_1, \ell_2 \in \mathbb{Z}$, we have*

$$d_{\text{IP}_\Phi} \{n \in \text{IP}((n_j)_{j \in \mathbb{N}}) : \mu(A \cap T^{-\ell_1 n} A \cap T^{-\ell_2 n} A \cap T^{-(\ell_1 + \ell_2)n} A) > \mu(A)^4 - \varepsilon\} > 0.$$

Proof. We start as in the proof of Theorem 6.2, identifying $Z_1(\mathbf{X})$ with a compact abelian group and letting $a \in Z_1(\mathbf{X})$ denote the rotation. Again let $\eta: Z_1(\mathbf{X}) \rightarrow \mathbb{R}$ be a continuous non-negative function such that

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta(na) = 1$$

and suppose by contradiction that the statement does not hold. Then, as in the proof of Theorem 6.2, for can find some $0 < \varepsilon < 1$ such that

$$(6.5) \quad \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta(na) \mu(A \cap T^{-\ell_1 n} A \cap T^{-\ell_2 n} A \cap T^{-(\ell_1 + \ell_2)n}) \leq \mu(A)^4 - \varepsilon.$$

Taking $\delta > 0$ sufficiently small with respect to $\varepsilon > 0$, it follows from Lemma 5.4 that there exists a factor isomorphic to an ergodic 2-step nil-system \mathbf{Y} such that letting f denote the projection of $\mathbf{1}_A$ onto this factor, we have that the left hand side of (6.5) is $\varepsilon/2$ -close to

$$(6.6) \quad \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta(na) \int_Y f(x) f(T^{\ell_1 n} x) f(T^{\ell_2 n} x) f(T^{(\ell_1 + \ell_2)n} x) d\mu(x).$$

By Lemma 5.6, the system \mathbf{Y} is synchronized. Let r be as in Lemma 5.6 and let $Z(\mathbf{Y})$ denote the Kronecker factor of \mathbf{Y} . Again write $Z = \mathbb{T}^m \times D$ for some finite dimensional torus \mathbb{T}^m and a finite cyclic group D , and we write the translation by T as $a = (a_1, a_2) \in Z(\mathbf{Y})$ with $a_1 \in \mathbb{T}^m$ and $a_2 \in D$. For convenience, we assume that r divides $|D|$, and otherwise we can pass to an extension such that this holds. Take $\eta: Z(\mathbf{Y}) \rightarrow \mathbb{R}$ to be in the proof of Theorem 6.2, meaning that it is a function the form $\eta(t, d) = \eta_1(t) \cdot c^{-1} \mathbf{1}_{\{0\}}(d)$ where $\eta_1: \mathbb{T}^m \rightarrow \mathbb{R}$ is a non-negative continuous function yet to be specified and $c := \lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \mathbf{1}_{\{0\}}(na_2)$, and lift this function to the Kronecker factor of \mathbf{X} . Let $Y^0 = N/\Lambda$ denote the connected component of Y . For notational simplicity, set $\ell_0 = 0$ and $\ell_3 = \ell_1 + \ell_2$. Then, by (6.6) and Proposition 5.8, the left hand side of (6.5) is $\varepsilon/2$ close to

$$(6.7) \quad \int_Y \int_{N/\Lambda} \int_{N_2/\Lambda_2} \eta_1(y_1 \Lambda) \prod_{i=0}^3 f(xy_1^{\ell_i} y_2^{\binom{\ell_i}{2}}) d\mu_{N_2/\Gamma_2}(y_2) d\mu_{N/\Gamma}(y_1) d\mu_Y(x).$$

Let $V = B(G_2, r)$ be an open ball with radius $r > 0$ and choose r to be sufficiently small with respect to $\varepsilon > 0$ and such that V has trivial intersection with the group D . Let $\eta_1: \mathbb{T}^m \rightarrow \mathbb{R}$ be a function supported on the projection of V onto \mathbb{T}^m and such that the lift to Z satisfies $\int_{\mathbb{T}} \eta_1(t) d\mu(t) = 1$. Since

the action of a_1 on \mathbb{T}^m is totally ergodic, it follows that

$$\lim_{N \rightarrow \infty} \mathbb{E}_{n \in \text{IP}_{\Phi_N}} \eta_1(na_1) \cdot c^{-1} \mathbf{1}_{\{0\}}(d) = \int_{\mathbb{T}} \eta_1(t) \, d\mu_{\mathbb{T}}(t) = 1.$$

Using (6.7), it follows that the left hand side of equation (6.5) is ε -close to

$$\int_Y \int_{N_2/\Lambda_2} \int_{N_2/\Lambda_2} \prod_{i=0}^3 f(xy_1^{\ell_i} y_2^{\binom{\ell_i}{2}}) \, d\mu_{N_2/\Lambda_2}(y_1, y_2) \, d\mu_Y(x).$$

By Lemma 6.3, this is bounded below by $\mu(A)^4$, a contradiction of (6.5). \square

APPENDIX A. ON DIFFERENT NOTIONS OF LARGENESS

There are several other notions of largeness that are appropriate for IP-sequences, and we discuss relations among some of them. Let $(n_j)_{j \in \mathbb{N}}$ be a sequence of natural numbers and let $A \subseteq \text{IP}((n_j)_{j \in \mathbb{N}})$.

- (1) The set A has *positive lower IP-density* if for any increasing Følner sequence $(\Phi_N)_{N \in \mathbb{N}}$, we have $\underline{d}_{\text{IP}_{\Phi}}(A) > 0$. Replacing the \liminf in the definition of $\underline{d}_{\text{IP}_{\Phi}}$ with the limit or with \limsup , we define *positive IP-density* and *positive upper IP-density* of the set A .
- (2) The set A is *syndetic in* $\text{IP}((n_j)_{j \in \mathbb{N}})$ if finitely many translations of A cover $\text{IP}((n_j)_{j \in \mathbb{N}})$.
- (3) The set A is *IP-syndetic* if it has nontrivial intersection with every IP-Følner sequence.

We check that none of these density notions are equivalent.

Lemma A.1. *For each of the density conditions (1), (2), and (3), there exists a set satisfying that condition but not the other two.*

Proof. Let $(n_j)_{j \in \mathbb{N}}$ be a sequence such that all of its finite sums are distinct (for example, we can take $n_j = 10^{j-1}$).

To construct a set that has positive IP-density but is neither syndetic in $\text{IP}((n_j)_{j \in \mathbb{N}})$ nor is IP-syndetic, set $\Phi_N = [N^2, N^2 + N]$, let $\Phi = \bigcup_{N \in \mathbb{N}} \Phi_N$, and take $A = \text{IP}((n_j)_{j \in \mathbb{N}}) \setminus \text{IP}_{\Phi}((n_j)_{j \in \mathbb{N}})$. The set A is not IP-syndetic because it has trivial intersection with $\text{IP}_{\Phi_N}((n_j)_{j \in \mathbb{N}})$ and is not syndetic as infinitely many translates are needed to cover $\text{IP}((n_j)_{j \in \mathbb{N}})$. By construction of the sequence Φ_N , the set A has positive IP-density.

The set $A = \{n_j : j \in \mathbb{N}\}$ is easily checked to be IP-syndetic, but does not have positive lower IP-density and is not syndetic in $\text{IP}((n_j)_{j \in \mathbb{N}})$.

Taking the set $A \subseteq \text{IP}((n_j)_{j \in \mathbb{N}})$ to be the sequence which contains all summands which include n_1 , then we have that this set avoids the IP-Følner sequence $\Phi_N = [2, N]$ and is therefore not IP-large nor IP-syndetic. However, since $A \cup A - n_1 = \text{IP}((n_j)_{j \in \mathbb{N}})$, this set is syndetic. \square

We note that it follows immediately from the definitions that for any sequence $(n_j)_{j \in \mathbb{N}}$ of natural numbers, if a set A intersects every IP set, then $A \cap \text{IP}((n_j)_{j \in \mathbb{N}})$ is IP-syndetic in $\text{IP}((n_j)_{j \in \mathbb{N}})$.

APPENDIX B. FAILURE OF THE LARGE INTERSECTION PROPERTY FOR NON-RATIONAL SEQUENCES

We show that Theorem 6.2 does not hold for arbitrary sequences, even when the notion of IP-largeness is replaced by some other notions of largeness, including stronger ones. A set of integers is *central* if it is an element in a minimal idempotent in the Stone-Čech compactification $\beta\mathbb{N}$ of the integers (see [21] for more on central sets and [4] for this equivalent definition).

It was established in [4] that a central set I contains all finite sums of a sequences $(n_j)_{j \in \mathbb{N}}$ however, without multiplicities. Throughout we assume that all central sets contains 0 by default. Though our notion of IP differs from that used in [21], given IP generated by a sequence of distinct elements $(n_j)_{j \in \mathbb{N}}$, we can always find a sub-IP by passing to a subsequence $(n_{j_i})_{i \in \mathbb{N}}$ satisfying that $n_{j_i} > \sum_{l=1}^{i-1} n_{j_l}$ for all $i > 1$. Thus the IP generated by this subsequence is a set, rather than a multiset, and a central set contains an IP.

Theorem B.1. *There exists an ergodic invertible measure preserving system $\mathbf{X} = (X, \mathcal{B}, \mu, T)$, a constant $c > 0$, and a set $A \subseteq X$ of positive measure such that any central set I contains a central subset $I' \subseteq I$ with the property that*

$$\{0 \neq n \in I' : \mu(A \cap T^{-n}A \cap T^{-2n}A) \geq \mu(A)^{-c \log \mu(A)}\} = \emptyset.$$

Proof. Consider the space $X = \mathbb{T}^2$, where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$, equipped with the action $T(x, y) = (\alpha + x, y + x)$ for some irrational $\alpha \in \mathbb{T}$. Since α is irrational, this system is ergodic and it is easy to check that the iterates of a point (x, y) are given by

$$T^n(x, y) = (n\alpha + x, y + nx + \binom{n}{2}\alpha).$$

Set $A = \mathbb{T} \times B$ for a subset B yet to be chosen. Writing $\mathbf{1}_A(x, y) = \sum_{n \in \mathbb{Z}} a_n y^n$ in terms of the Fourier series for $\mathbf{1}_A$ (note that $\mathbf{1}_A$ does not depend on the first coordinate), we have that

$$\begin{aligned} \mu(A \cap T^{-n}A \cap T^{-2n}A) &= \int_{\mathbf{X}} \mathbf{1}_A \cdot T^n \mathbf{1}_A \cdot T^{2n} \mathbf{1}_A \, d\mu \\ &= \sum_{m_0, m_1, m_2 \in \mathbb{Z}} a_{m_0} a_{m_1} a_{m_2} \alpha^{m_1 \binom{n}{2} + m_2 \binom{2n}{2}} \int_X y^{m_0 + m_1 + m_2} x^{n(m_1 + 2m_2)} \, d\mu(x, y). \end{aligned}$$

The last integral is zero unless $m_0 + m_1 + m_2 = 0$ and $m_1 + 2m_2 = 0$ (note that $n \neq 0$). In this case, the integral is 1. It follows that $m_0 = m_1 = -2m_2$, and so

$$(B.1) \quad \mu(A \cap T^{-n}A \cap T^{-2n}A) = \sum_{m \in \mathbb{Z}} a_m^2 a_{2m} \alpha^{mn}.$$

Let I be a central set and let p be a minimal idempotent containing I . For an ultrafilter $p \in \beta\mathbb{N}$, we write $p\text{-}\lim_{n \in \mathbb{N}} x_n = x$ to mean that for every neighborhood U of x we have $\{n : x_n \in U\} \in p$. Bergelson [5] showed that $p\text{-}\lim_{n \in \mathbb{N}} T^n = P_K$ converges weakly to the projection onto the Kronecker factor. In particular, we have that $p\text{-}\lim_{n \in \mathbb{N}} (\alpha^{mn}) = 1$ for all $m \in \mathbb{Z}$. Therefore, applying $p\text{-}\lim$ to (B.1), we deduce that

$$\begin{aligned} (B.2) \quad p\text{-}\lim_{n \in \mathbb{N}} \mu(A \cap T^{-n}A \cap T^{-2n}A) &= \sum_{m \in \mathbb{Z}} a_m^2 a_{2m} \\ &= \int_{\mathbb{R}/\mathbb{Z} \times \mathbb{R}/\mathbb{Z}} \mathbf{1}_B(t) \mathbf{1}_B(t+s) \mathbf{1}_B(t+2s) \, ds \, dt. \end{aligned}$$

In [7, Proof of Theorem 2.1], it is shown that for every integer $L \geq 1$ there is a set B_L of measure $\frac{\exp(-c\sqrt{\log L})}{4}$, where c is the constant from Behrend's construction of a set of integers of size L containing no three term arithmetic progressions, for which the quantity in (B.2) is bounded by $\frac{\exp(-c\sqrt{\log L})}{16L}$. Taking L sufficiently large, this implies that

$$p\text{-}\lim_{n \in \mathbb{N}} \mu(A \cap T^{-n}A \cap T^{-2n}A) < \mu(A)^{-c \log \mu(A)}.$$

It follows that there exists a central set $I' \subseteq I$ such that $\mu(A \cap T^{-n}A \cap T^{-2n}A) \leq \mu(A)^{-c \log \mu(A)}$ for all $n \in I'$, completing the proof. \square

APPENDIX C. PROOF OF LEMMA 5.6

To prove Lemma 5.6, we freely make use of notation and results from [30], giving precise references but omitting the details.

We start with the following structure theorem for cocycles of type 2 (see [30, Section 7] for definitions).

Lemma C.1. *Assume that U_1 is a finite dimensional torus, K is a finite cyclic group, and let $Z = U_1 \times K$ be a 1-step nilsystem with the rotation given by some $a = (a_1, b) \in Z$ for some generator $b \in K$ and irrational $a_1 \in U_1$. For a finite dimensional torus U , any cocycle $\rho: Z \rightarrow U$ of type 2 is cohomologous to a sum of a K -invariant cocycle ρ^0 and a homomorphism $\phi: K \rightarrow U$.*

Proof. Set $r = |K|$. For $z \in Z$, write $z = (x, t)$ for $x \in U_1$ and $t \in K$. For $s \in Z$ and $f: Z \rightarrow U$, we write $\partial_s f(x) = f(x + s) - f(x)$. By [30, Lemma 8.1], we have that

$$\partial_b \rho = c - \partial_a F$$

for some $c \in U$ and measurable function $F: Z \rightarrow U$. Combining this equation with the telescoping identity $\sum_{i=0}^{r-1} \partial_b \rho(x, t + ib) = 0$, it follows that

$$r \cdot c = \partial_a \left(\sum_{i=0}^{r-1} F(x, t + ib) \right).$$

Thus rc is an eigenvalue of a b -invariant eigenfunction. Since any eigenfunction is of the form $\psi: U_1 \rightarrow U$ for some affine map ψ (meaning that $\psi(x) - \psi(0)$ is a homomorphism from U_1 to U), we can write

$$(C.1) \quad \psi(x) = \sum_{i=0}^{r-1} F(x, t + ib).$$

We claim that ψ determines F up to a coboundary with respect to the action of K . More precisely, the claim is that

$$F(x, t) = \partial_b F'(x, t) + \binom{|t|}{r-1} \psi(x)$$

for some measurable map $F': Z \rightarrow U$, where $\binom{|t|}{r-1}$ maps all cosets of $r\mathbb{Z}$ to 0 other than the coset $(r-1) + r\mathbb{Z}$, which is mapped to 1.

To check the claim, note that $\sum_{i=0}^{r-1} \binom{|t+ib|}{r-1} \psi(x) = \psi(x)$ and so by (C.1) it follows that

$$(C.2) \quad \sum_{i=0}^{r-1} F(x, t+ib) - \binom{|t+ib|}{r-1} \psi(x) = 0.$$

Defining $F'(x, jb) = \sum_{i=0}^{j-1} F(x, 0+ib) - \binom{|jb|}{r-1} \psi(x)$ for all $0 \leq j \leq r-1$, we have that (C.2) implies that $\partial_b F'(x, t) = F(x, t) - \binom{|t|}{r-1} \psi(x)$, proving the claim.

Thus by the claim, we may modify ρ by the cohomologous cocycle $\rho - \partial_a F'$. Therefore, we can assume without loss of generality that $F(x, t) = \binom{|t|}{r-1} \psi(x)$ for some affine map $\psi: Z \rightarrow U$. In this case

$$\partial_b \rho = c - \partial_b \left(\binom{|t|}{r-1} \psi(x) \right) = c - \binom{|t+b|}{r-1} \psi(x+a_1) + \binom{|t|}{r-1} \psi(x)$$

and $rc = \partial_{a_1} \psi(x)$. This equation determines ρ up to a b -invariant function. Indeed, let $\sigma: Z \rightarrow U$ be the cocycle defined by

$$\sigma(x, jb) := \partial_{jb} \rho(x, 0) = j \cdot c - \binom{|jb|}{r-1} \psi(x+a_1).$$

Then $\partial_b \sigma(x, t) = \partial_b \rho(x, t)$ and so there is a map $f: U_1 \rightarrow U$ such that $\rho(x, t) = \sigma(x, t) + f(x)$. Since $\rho(x, t)$ is of type 2, for every $s \in U_1$ there is a constant $c_s \in U$ and a measurable map $F_s: Z \rightarrow U$ such that

$$\partial_s \sigma(x, t) + \partial_s f(x) = \partial_s \rho(x, t) = c_s - \partial_a F_s(x, t).$$

When $t = 0$, we have that $\sigma(x, 0) = 0$ and so

$$\partial_s f(x) = c_s - \partial_a F_s(x, t).$$

Choosing t_0 such that $t_0 b = r-1$, we have $\sigma(x, t_0 b) = t_0 \cdot c - \psi(x+a_1)$. From this it follows that,

$$\partial_s \psi(x+a_1) + \partial_s f(x) = c_s - \partial_a F_s(x, t_0).$$

It follows that $\partial_s \psi(x+a_1) = \partial_s \psi(x)$ is a coboundary for all $s \in U_1$, but this can not happen unless $\psi = 0$.

In this case, $\partial_b \rho = c$, and again using the telescoping identity $\sum_{i=0}^{r-1} \partial_b \rho(x, t+ib) = 0$, we deduce that $rc = 0$. Hence, we can define a homomorphism $\phi: K \rightarrow U$ taking b to c , and then $\rho^0 := \rho - \phi$ is a b -invariant cocycle and ρ is cohomologous to $\rho^0 + \phi$. \square

We use this to complete the proof of the lemma.

Proof of Lemma 5.6. Let $k = 1$ or $k = 2$ and let $\mathbf{X} = (G/\Gamma, \mathcal{B}, \mu, T_\tau)$ be an ergodic k -step nilsystem.

Suppose first that $k = 1$. Then X is isomorphic to a rotation on a compact abelian Lie group Z and we can write $Z = U_1 \times K$ for some torus U_1 and finite abelian cyclic group K , with the rotation $a = (a_1, b)$, where $a_1 \in U_1$ and $b \in K$. Taking $r = |K|$, we have that $ra = r(a_1, 0)$ and so X is synchronized.

Now suppose that $k = 2$, and write $X = \mathcal{G}(\mathbf{X})/\Gamma$ where $\mathcal{G}(\mathbf{X})$ is the Host-Kra group of \mathbf{X} (see [30, Section 5]). By [30, Section 8], we have that $X = Z \times_\rho U$ where Z , U , and ρ are as in Lemma C.1. Thus, without loss of generality we may assume that $\rho = \rho^0 + \phi$, where ρ^0 is a K -invariant cocycle for a finite cyclic group K and $\phi: K \rightarrow U$ is a homomorphism. Since K has some finite order, it follows from the cocycle identity that there exists some $r \in \mathbb{N}$ such that $\rho(r, x) = \rho^0(r, x)$. The translation T_τ on \mathbf{X} is given by the element $\tau = S_{(a, \rho)} \in \mathcal{G}(\mathbf{X})$. On the other hand we also have $b := S_{(a_1, 0, \rho^0)} \in \mathcal{G}(\mathbf{X})$ since $\partial_{a_1}\rho = \partial_a\rho^0$ (see [30, Lemma 8.7]) and the latter maps the connected component of \mathbf{X} to itself. Furthermore, it follows from the construction that $b^r = S_{((a_1, 0), \rho^0)}^r = S_{(a, \rho)}^r = \tau^r$, as required. \square

REFERENCES

- [1] E. Ackelsberg. Khintchine-type double recurrence in abelian groups. *Ergodic Theory and Dynamical Systems*, 45(1):1–33, 2024.
- [2] E. Ackelsberg, V. Bergelson, and A. Best. Multiple recurrence and large intersections for abelian group actions. *Discrete Anal.*, pages Paper No. 18, 91, 2021.
- [3] E. Ackelsberg, V. Bergelson, and O. Shalom. Khintchine-type recurrence for 3-point configurations. *Forum Math. Sigma*, 10:Paper No. e107, 57, 2022.
- [4] V. Bergelson. Minimal idempotents and ergodic ramsey theory. *Topics in dynamics and ergodic theory, London Math. Soc. Lecture Note Ser.*, 310:8–39, 2003.
- [5] V. Bergelson. Minimal idempotents and ergodic Ramsey theory. In *Topics in Dynamics and Ergodic Theory, London Math. Soc. Lecture Note Ser.*, volume 310. Cambridge Univ. Press, 2003.
- [6] V. Bergelson, A. del Junco, M. Lemańczyk, and J. Rosenblatt. Rigidity and non-recurrence along sequences. *Ergodic Theory Dynam. Systems*, 34(5):1464–1502, 2014.
- [7] V. Bergelson, B. Host, and B. Kra. Multiple recurrence and nilsequences. *Invent. Math.*, 160(2):261–303, 2005. With an appendix by Imre Ruzsa.
- [8] V. Bergelson, B. Host, R. McCutcheon, and F. Parreau. aspects of uniformity in recurrence. *Colloquium Mathematicae*, 84/85(2):549–576, 2000.

- [9] V. Bergelson and A. Leibman. Polynomial extensions of van der Waerden's and Szemerédi's theorems. *J. Amer. Math. Soc.*, 9(3):725–753, 1996.
- [10] V. Bergelson and R. McCutcheon. An ergodic IP polynomial Szemerédi theorem. *Memoirs of the AMS*, 146, 2000.
- [11] M. Boshernitzan, G. Kolesnik, A. Quas, and M. Wierdl. Ergodic averaging sequences. *J. Anal. Math.*, 95:63–103, 2005.
- [12] P. Candela, D. González-Sánchez, and B. Szegedy. Lie-fibered nilspaces, double-coset nilspaces, and gowers norms. *Preprint*, 2023.
- [13] P. Candela and B. Szegedy. Regularity and inverse theorems for uniformity norms on compact abelian groups and nilmanifolds. *J. für die Reine und Angew. Math.*, 789:1–42, 2022.
- [14] P. Candela and B. Szegedy. Nilspace factors for general uniformity seminorms, cubic exchangeability and limits. *Mem. Amer. Math. Soc.*, 287(1425):v+101, 2023.
- [15] A. Ferré Moragues. Properties of multicorrelation sequences and large returns under some ergodicity assumptions. *Discrete Contin. Dyn. Syst.*, 41(6):2809–2828, 2021.
- [16] N. Frantzikinakis. Multiple ergodic averages for three polynomials and applications. *Trans. Amer. Math. Soc.*, 360(10):5435–5475, 2008.
- [17] N. Frantzikinakis. Multiple recurrence and convergence for Hardy sequences of polynomial growth. *J. Anal. Math.*, 112:79–135, 2010.
- [18] N. Frantzikinakis, B. Host, and B. Kra. The polynomial multidimensional Szemerédi theorem along shifted primes. *Israel J. Math.*, 194(1):331–348, 2013.
- [19] N. Frantzikinakis and B. Kuca. Joint ergodicity for commuting transformations and applications to polynomial sequences. *Inventiones Mathematicae*, 239(2):621–706, 2025.
- [20] H. Furstenberg. Ergodic behaviour of diagonal measures and a theorem of Szemerédi on arithmetic progressions. *J. Anal. Math.*, 31:204–256, 1977.
- [21] H. Furstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory*. Princeton Legacy Library. Princeton University Press, 2014.
- [22] H. Furstenberg and Y. Katznelson. An ergodic Szemerédi theorem for commuting transformations. *J. Anal. Math.*, 34:275–291, 1978.
- [23] H. Furstenberg and Y. Katznelson. An ergodic Szemerédi theorem for IP-systems and combinatorial theory. *J. Analyse Math.*, 45:117–168, 1985.
- [24] H. Furstenberg and Y. Katznelson. A density version of the Hales-Jewett theorem. *J. Anal. Math.*, 57:64–119, 1991.
- [25] H. Furstenberg and B. Weiss. A mean ergodic theorem for $(1/N) \sum_{n=1}^N f(T^n x)g(T^{n^2} x)$. In *Convergence in ergodic theory and probability (Columbus, OH, 1993)*, volume 5 of *Ohio State Univ. Math. Res. Inst. Publ.*, pages 193–227. de Gruyter, Berlin, 1996.
- [26] H. Furstenberg and B. Weiss. A mean ergodic theorem for $(1/N) \sum_{n=1}^N f(T^n x)g(T^{n^2} x)$. In *Convergence in ergodic theory and probability (Columbus, OH, 1993)*, volume 5 of *Ohio State Univ. Math. Res. Inst. Publ.*, pages 193–227. de Gruyter, Berlin, 1996.
- [27] Y. Gutman, F. Manners, and P. P. Varjú. The structure theory of nilspaces II: Representation as nilmanifolds. *Trans. Amer. Math. Soc.*, 371(7):4951–4992, 2019.
- [28] Y. Gutman, F. Manners, and P. P. Varjú. The structure theory of nilspaces I. *J. Anal. Math.*, 140(1):299–369, 2020.
- [29] Y. Gutman, F. Manners, and P. P. Varjú. The structure theory of nilspaces III: Inverse limit representations and topological dynamics. *Adv. Math.*, 365:107059, 53, 2020.
- [30] B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. Math.*, 161(1):397–488, 2005.

- [31] B. Host and B. Kra. *Nilpotent structures in ergodic theory*, volume 236 of *Mathematical Surveys and Monographs*. American Mathematical Society, Providence, RI, 2018.
- [32] A. Jamneshan, O. Shalom, and T. Tao. The structure of arbitrary Conze-Lesigne systems. *Commun. Am. Math. Soc.*, 4:182–229, 2024.
- [33] A. Jamneshan and T. Tao. The inverse theorem for the U^3 Gowers uniformity norm on arbitrary finite abelian groups: Fourier-analytic and ergodic approaches. *Discrete Anal.*, pages Paper No. 11, 48, 2023.
- [34] A. Leibman. Polynomial mappings of groups. *Israel Journal of Mathematics*, 129:29–60, 2002.
- [35] A. Leibman. Host–Kra and Ziegler factors and convergence of multiple averages, (appendix to Combinatorial and Diophantine applications of ergodic theory). In *Handbook of dynamical systems. Vol. 1B*, pages 745–869. Elsevier B. V., Amsterdam, 2006.
- [36] E. Lindenstrauss. Measurable distal and topological distal systems. *Ergodic Theory and Dynamical Systems.*, 19(4):1063–1076, 1999.
- [37] O. Shalom. Multiple ergodic averages in abelian groups and Khintchine type recurrence. *Trans. Amer. Math. Soc.*, 375(4):2729–2761, 2022.
- [38] O. Shalom. Host-Kra theory for $\bigoplus_{p \in P} \mathbb{F}_p$ -systems and multiple recurrence. *Ergodic Theory Dynam. Systems*, 43(1):299–360, 2023.
- [39] E. Szemerédi. On sets of integers containing no k elements in arithmetic progression. *Acta Arith.*, 27:199–245, 1975.
- [40] N. M. Walsh. Norm convergence of nilpotent ergodic averages. *Ann. of Math.*, 175(3):1667–1688, 2012.
- [41] T. D. Wooley and T. D. Ziegler. Multiple recurrence and convergence along the primes. *Amer. J. Math.*, 134(6):1705–1732, 2012.
- [42] T. Ziegler. A non-conventional ergodic theorem for a nilsystem. *Ergodic Theory and Dynamical Systems*, 25 no. 4:1357–1370., 2005.
- [43] T. Ziegler. Universal characteristic factors and Furstenberg averages. *J. Amer. Math. Soc.*, 20:53–97, 2007.

DEPARTMENT OF MATHEMATICS, NORTHWESTERN UNIVERSITY, EVANSTON, IL 60208
 Email address: kra@math.northwestern.edu

DEPARTMENT OF MATHEMATICS, BAR ILAN UNIVERSITY, RAMAT GAN, 5290002, ISRAEL
 Email address: Or.Shalom@math.biu.ac.il