Ergod. Th. & Dynam. Sys. (2021), **41**, 606–621 © *Cambridge University Press, 2019*
doi:10.1017/etds.2019.67

Weak ergodic averages over dilated measures

WENBO SUN

Department of Mathematics, The Ohio State University, 231 West 18th Avenue, Columbus OH, 43210-1174, USA
(e-mail: sun.1991@osu.edu)

(*Received* 18 *October* 2018 *and accepted in revised form* 27 *June* 2019)

Abstract. Let $m \in \mathbb{N}$ and $\mathbf{X} = (X, \mathcal{X}, \mu, (T_\alpha)_{\alpha \in \mathbb{R}^m})$ be a measure-preserving system with an $\mathbb{R}^m$-action. We say that a Borel measure ν on $\mathbb{R}^m$ is weakly equidistributed for $\mathbf{X}$ if there exists $A \subseteq \mathbb{R}$ of density 1 such that, for all $f \in L^\infty(\mu)$, we have

$$\lim_{t \in A, t \to \infty} \int_{\mathbb{R}^m} f(T_{t\alpha}x)\, d\nu(\alpha) = \int_X f\, d\mu$$

for μ-almost every $x \in X$. Let $W(\mathbf{X})$ denote the collection of all $\alpha \in \mathbb{R}^m$ such that the $\mathbb{R}$-action $(T_{t\alpha})_{t \in \mathbb{R}}$ is not ergodic. Under the assumption of the pointwise convergence of the double Birkhoff ergodic average, we show that a Borel measure ν on $\mathbb{R}^m$ is weakly equidistributed for an ergodic system $\mathbf{X}$ if and only if $\nu(W(\mathbf{X}) + \beta) = 0$ for every $\beta \in \mathbb{R}^m$. Under the same assumption, we also show that ν is weakly equidistributed for all ergodic measure-preserving systems with $\mathbb{R}^m$-actions if and only if $\nu(\ell) = 0$ for all hyperplanes ℓ of $\mathbb{R}^m$. Unlike many equidistribution results in literature whose proofs use methods from harmonic analysis, our results adopt a purely ergodic-theoretic approach.

Key words: ergodic theory, weak equidistribution, characteristic factor, nilsystem
2010 Mathematics Subject Classification: 37A05 (Primary); 37A15, 37A17, 37A25 (Secondary)

1. *Introduction*

1.1. *Strong equidistribution over dilated measures.* Let G be a locally compact Hausdorff topological group. A *measure-preserving G-system* (or a *G-system*) is a tuple $\mathbf{X} = (X, \mathcal{X}, \mu, (T_g)_{g \in G})$, where $(X, \mathcal{X}, \mu)$ is a separable probability space and $T_g : X \to X$, $g \in G$ are measurable and measure-preserving transformations such that $T_g \circ T_h = T_{gh}$, $T_{e_G} = \mathrm{id}$ for all $g, h \in G$. We also require that, for all $x \in X$, the map $G \to X$, $g \to T_g x$ is measurable. We say that $\mathbf{X}$ is *ergodic* if $A \in \mathcal{X}$, $T_g A = A$ for all $g \in G$ implies that $\mu(A) = 0$ or 1.

Let $m \in \mathbb{N}$, ν be a Borel measure on $\mathbb{R}^m$ and $\mathbf{X} = (X, \mathcal{X}, \mu, (T_g)_{g \in \mathbb{R}^m})$ be an ergodic $\mathbb{R}^m$-system. Throughout this paper, we assume that $\nu(\mathbb{R}^m) = 1$. We say that ν is *(strongly)* *equidistributed* for **X** if, for all $f \in L^\infty(\mu)$, we have

$$\lim_{t\to\infty} \int_{\mathbb{R}^m} f(T_{t\alpha}x)\, d\nu(\alpha) = \int_X f\, d\mu$$

for μ-almost every (μ-a.e.) $x \in X$. The Birkhoff ergodic theorem for $\mathbb{R}$-systems (see, for example, [**9**, Corollary 8.15]) states that for every ergodic $\mathbb{R}$-system $(X, \mathcal{X}, \mu, (T_t)_{t\in\mathbb{R}})$ and every $f \in L^\infty(\mu)$,

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f(T_t x) = \int_X f\, d\mu$$

for μ-a.e. $x \in X$, which is equivalent to saying that the Lebesgue measure restricted to the interval $[-1, 1]$ is equidistributed for every ergodic $\mathbb{R}$-system. Similarly, the Birkhoff ergodic theorem holds for $\mathbb{R}^m$-systems for all $m \in \mathbb{N}$: the Lebesgue measure restricted to the unit cube or ball in $\mathbb{R}^m$ is equidistributed for every ergodic $\mathbb{R}^m$-system (see, for example, [**9**, Theorem 8.19]).

It is an interesting question to ask if similar results hold for the Lebesgue measure restricted to the boundary of the unit cube or ball. The motivation for this question was from a result of Stein [**23**] in 1976, who showed that for $\phi \in L^p(\mathbb{R}^m)$, $p > m/(m-1)$, $m \geq 3$, and for Lebesgue-a.e. $x \in \mathbb{R}^m$, we have that

$$\lim_{t\to 0} \int_{S_t} \phi(x-u)\, d\sigma_{m,t}(u) = \phi(x),$$

where $\sigma_{m,t}$ is the Lebesgue measure on $\mathbb{R}^m$ restricted to S_t, the sphere of radius t centered at the origin. Later an analog of this result was proved in the ergodic-theoretic setting. It was proved by Jones [**16**] (for $m \geq 3$) and Lacey [**18**] (for $m = 2$) that the Lebesgue measure restricted to the boundary of the unit ball $\sigma_{m,1}$ is equidistributed for all ergodic $\mathbb{R}^m$-systems. We remark that, on the other hand, it is not hard to see that the Lebesgue measure restricted to the boundary of the unit cube is not equidistributed for some ergodic $\mathbb{R}^m$-systems. It is then natural to ask which measure ν on $\mathbb{R}^m$ is equidistributed for all ergodic $\mathbb{R}^m$-systems. It was proved by Björklund [**5**] that if ν has *Fourier dimension* $a > 1$, meaning that a is the supremum over all $0 \leq a \leq d$ such that $\lim_{\zeta\to\infty} |\widehat{\nu}(\zeta)| \cdot |\zeta|^{a/2} < \infty$, then ν is equidistributed for all ergodic $\mathbb{R}^m$-systems. It is worth noting that strong equidistribution for polynomial maps on special homogeneous systems have also been studied in recent years (see [**17**, **22**], for example).

1.2. *Weak equidistribution over dilated measures.* In contrast to the strong equidistribution, a notion called 'weak equidistribution' has recently been studied, and various results have been obtained in the settings of translation surfaces [**7**] and nilmanifolds [**17**]. To be more precise, we say that a Borel measure ν on $\mathbb{R}^m$ is *weakly equidistributed* for an $\mathbb{R}^m$-system **X** if there exists $A \subseteq \mathbb{R}$ of density 1 such that, for all $f \in L^\infty(\mu)$, we have

$$\lim_{t\in A, t\to\infty} \int_{\mathbb{R}^m} f(T_{t\alpha}x)\, d\nu(\alpha) = \int_X f\, d\mu$$

for μ-a.e. $x \in X$. By the examples in [**17**, §5], strong and weak equidistributions are not equivalent conditions.

It is natural to ask which measures ν on $\mathbb{R}^m$ are weakly equidistributed for all ergodic $\mathbb{R}^m$-systems. In this paper, we provide a necessary and sufficient condition for such ν under the assumption of the pointwise convergence of the double Birkhoff ergodic average. We say that an $\mathbb{R}^m$-system $(X, \mathcal{X}, \mu, (T_g)_{g\in\mathbb{R}^m})$ is *good for double Birkhoff averages* if, for all $f_1, f_2 \in L^\infty(\mu)$ and $\alpha_1, \alpha_2 \in \mathbb{R}^m$, the limit

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f_1(T_{\alpha_1 t}x) f_2(T_{\alpha_2 t}x)\, dt$$

exists for μ-a.e. $x \in X$.

We now state our first theorem. A *hyperplane* of $\mathbb{R}^m$ is $V + \beta$ for some subspace V of $\mathbb{R}^m$ of codimension 1 and $\beta \in \mathbb{R}^m$.

THEOREM 1.1. *Let $m \in \mathbb{N}$. A Borel measure ν on $\mathbb{R}^m$ is weakly equidistributed for all ergodic $\mathbb{R}^m$-systems which are good for double Birkhoff averages if and only if $\nu(\ell) = 0$ for all hyperplanes ℓ of $\mathbb{R}^m$.*

Theorem 1.1 will provide a complete answer for the weak equidistribution problem if the following conjecture holds.

CONJECTURE 1.2. *Every $\mathbb{R}^m$-system is good for double Birkhoff averages.*

Conjecture 1.2 is still an open question in ergodic theory. Nevertheless, various partial results on Conjecture 1.2 have been obtained in recent years for some special type of systems, and so this paper can be viewed as an application of these results. We defer the discussion of this topic to §5.

Another question we study in this paper is the necessary and sufficient conditions for a Borel measure ν on $\mathbb{R}^m$ to be weakly equidistributed on a particular $\mathbb{R}^m$-system. Let $m \in \mathbb{N}$ and $\mathbf{X} = (X, \mathcal{X}, \mu, (T_g)_{g\in\mathbb{R}^m})$ be an $\mathbb{R}^m$-system. We use $I(H)$ to denote the σ-algebra of $\mathcal{X}$ consisting of all the H-invariant sets for every subgroup H of $\mathbb{R}^m$. Let $W(\mathbf{X})$ denote the collection of all $\alpha \in \mathbb{R}^m$ such that $I((t\alpha)_{t\in\mathbb{R}}) \neq I(\mathbb{R}^m)$. If $\mathbf{X}$ is an ergodic $\mathbb{R}^m$-system, then $W(\mathbf{X})$ is the collection of all $\alpha \in \mathbb{R}^m$ such that the $\mathbb{R}$-action $(T_{t\alpha})_{t\in\mathbb{R}}$ is not ergodic on $\mathbf{X}$. We have the following result.

THEOREM 1.3. *Let $m \in \mathbb{N}$ and ν be a Borel measure on $\mathbb{R}^m$. If $\mathbf{X}$ is an ergodic $\mathbb{R}^m$-system which is good for double Birkhoff averages such that $\nu(W(\mathbf{X}) + \beta) = 0$ for every $\beta \in \mathbb{R}^m$, then ν is weakly equidistributed for $\mathbf{X}$. Conversely, if $\mathbf{X}$ is an ergodic $\mathbb{R}^m$-system such that $\nu(W(\mathbf{X}) + \beta) \neq 0$ for some $\beta \in \mathbb{R}^m$, then ν is not weakly equidistributed for $\mathbf{X}$.*

We remark that the second part of Theorem 1.3 holds for every ergodic $\mathbb{R}^m$-system. We give an example to illustrate Theorem 1.3.

Example 1.4. Let $m \in \mathbb{N}$ and $(X = \mathbb{T}^m, \mathcal{X}, \mu)$ be an m-dimensional torus endowed with the Lebesgue measure μ. For all $\alpha \in \mathbb{R}^m$, denote $T_\alpha\beta = \alpha + \beta \mod \mathbb{Z}^m$ for all $\beta \in \mathbb{T}^m$. Then $\mathbf{X} = (X, \mathcal{X}, \mu, (T_\alpha)_{\alpha\in\mathbb{R}^m})$ is an ergodic $\mathbb{R}^m$-system and is good for double Birkhoff averages. In this case, $W(\mathbf{X})$ consists of all the $(m-1)$-dimensional rational subspaces of $\mathbb{R}^m$. By Theorem 1.3, a Borel measure ν on $\mathbb{R}^m$ is weakly equidistributed for $\mathbf{X}$ if and only if the ν-measure of any translation of a rational subspace of $\mathbb{R}^m$ is equal to zero. This recovers a special case of [**17**, Theorem 1.1].

In the case $m = 1$, the assumption of goodness for double Birkhoff averages can be dropped by using Bourgain's result [**6**] (see §5).

PROPOSITION 1.5. *Let ν be a Borel measure on $\mathbb{R}$ and $\mathbf{X}$ be an ergodic $\mathbb{R}$-system. Then ν is weakly equidistributed for $\mathbf{X}$ if and only if ν is atomless (meaning that $\nu(\{\beta\}) = 0$ for all $\beta \in \mathbb{R}$).*

It is an interesting question to understand the algebraic structure of $W(\mathbf{X})$. Let $W'(\mathbf{X})$ denote the collection of all $\alpha \in \mathbb{R}^m$ such that $I((n\alpha)_{n\in\mathbb{Z}}) \neq I(\mathbb{R}^m)$. Then $W(\mathbf{X}) \subseteq W'(\mathbf{X})$. By a result of Pugh and Shub [**21**] (see also Theorem 2.1), $W'(\mathbf{X})$ is contained in the union of at most countably many hyperplanes of $\mathbb{R}^m$. We show in §2 an analog of this result for $W(\mathbf{X})$.

THEOREM 1.6. *Let $m \in \mathbb{N}$ and $\mathbf{X}$ be an $\mathbb{R}^m$-system. Then $W(\mathbf{X})$ is the union of at most countably many proper subspaces of $\mathbb{R}^m$.*

In other words, $W(\mathbf{X})$ is contained in the union of at most countably many hyperplanes of $\mathbb{R}^m$ passing through the origin.

While all the previous mentioned results on the strong equidistribution rely heavily on tools from harmonic analysis, in this paper we provide purely ergodic-theoretic proofs for Theorems 1.1 and 1.3 and Proposition 1.5. An advantage of considering the weak equidistribution problem is that while the conditions in Theorems 1.1 and 1.3 and Proposition 1.5 are almost necessary and sufficient, the conditions imposed in all the previously mentioned results for strong equidistribution seem to be far from necessary. Moreover, we make no smoothness assumption for the Borel measure ν in the main results of this paper, as we do not apply Fourier analysis in the proofs.

1.3. *Organization of the paper.* In §2 we provide for later use two variations of the result of Pugh and Shub [**21**] on the ergodic directions of $\mathbb{R}^m$-systems. In §3 we introduce Host–Kra characteristic factors, which are the main tool of this paper. For the convenience of our purpose and future work, we develop the existing results on this topic into a more general setting. The proofs of the main results (Theorems 1.1 and 1.3) are in §4. In §5, we review systems which are good for double Birkhoff averages, and discuss applications of the main theorems of this paper to such systems (including the proof of Proposition 1.5).

2. *Ergodic elements in ergodic systems*

Let $m \in \mathbb{N}$, $\mathbf{X}$ be an $\mathbb{R}^m$-system and H be a subgroup of G. We say that $(T_h)_{h\in H}$ is *ergodic* for $\mathbf{X}$ if all the H-invariant subsets of $\mathbf{X}$ are of measure either 0 or 1.

A key ingredient connecting $\mathbb{R}^m$-systems and $\mathbb{Z}^m$-systems is the following theorem.

THEOREM 2.1. (Pugh and Shub [**21**, Theorem 1.1]) *Let $m \in \mathbb{N}$ and $\mathbf{X}$ be an ergodic $\mathbb{R}^m$-system. Then, for all $\alpha \in \mathbb{R}^m$ except at most a countable family of hyperplanes of $\mathbb{R}^m$, the $\mathbb{Z}$-action $(T_{n\alpha})_{n\in\mathbb{Z}}$ is ergodic for $\mathbf{X}$.*

In this section, we provide two generalizations of Pugh and Shub's theorem. The first is a relative version of Theorem 2.1.

LEMMA 2.2. *Let $m \in \mathbb{N}$ and $\mathbf{X}$ be a (not necessarily ergodic) $\mathbb{R}^m$-system. Then, for all $\alpha \in \mathbb{R}^m$ except at most a countable family of hyperplanes of $\mathbb{R}^m$, we have that $I((n\alpha)_{n\in\mathbb{Z}}) = I(\mathbb{R}^m)$.*

The proof of this lemma is almost identical to that of Theorem 2.1, so we only provide a sketch.

Sketch of the proof. Let G be an abelian, Hausdorff, locally compact and separable group, and $\mathbf{X} = (X, \mathcal{X}, \mu, (T_g)_{g\in G})$ be a G-system. Using Zorn's lemma, and the fact that $\mathbf{X}$ is separable, we may decompose $L^2(\mu)$ as a countable direct sum of orthogonal closed subspaces

$$L^2(\mu) = H \bigoplus_i H_i,$$

where H consists of all the G-invariant functions and, for each i, H_i is the smallest closed subspace of $L^2(\mu)$ containing the G-orbit of some $f_i \in L^2(\mu)$. To each i there corresponds a unique normalized Borel measure β_i on the dual group $\widehat{G} = \mathrm{Hom}(G, \mathbb{T}^1)$ such that $(T_g)_{g\in G}$ restricted to H_i is unitarily equivalent to the 'direct integral' representation $m_i : G \to Un(L^2(\widehat{G}, \beta_i))$,

$$g \to \langle \cdot, g\rangle f(\cdot), \quad f \in L^2(\widehat{G}, \beta_i).$$

For $g \in G$, denote

$$\ker(g) = \{\chi \in \widehat{G} : \langle \chi, g\rangle = 1\}.$$

Following the proof in [**21**], we can deduce the following claims.

Claim 1. The identity element of $\widehat{G}$ has zero β_i measure for all i.

Claim 2. If $I((g^n)_{n\in\mathbb{Z}}) \neq I(G)$ for some $g \in G$, then there exists i such that $\beta_i(\ker(g)) > 0$.

For Claim 1, if the identity element of $\widehat{G}$ has positive β_i measure for some i, by the argument of [**21**, proof of Lemma 1], one can construct a non-trivial G-invariant function lying in H_i, a contradiction. For Claim 2, if $I((g^n)_{n\in\mathbb{Z}}) \neq I(G)$ for some $g \in G$, then there exists a g-invariant function which does not belong to H, and the rest of the proof is identical to [**21**, Lemma 2].

We now return to the case where $G = \mathbb{R}^m$. By using Claims 1 and 2 to replace [**21**, Lemmas 1 and 2], and following the same argument as in [**21**, §5], we finish the proof. □

We now prove Theorem 1.6, which is a variation of Theorem 2.1 for the ergodicity of $\mathbb{R}$-actions. This result is of interest in its own right.

Proof of Theorem 1.6. We first claim that for every subspace V of $\mathbb{R}^m$, either $V \subseteq W(\mathbf{X})$ or there exists a family of at most countably many proper subspaces $(V_j)_{j\in J}$ of V such that $W(\mathbf{X}) \cap V \subseteq \bigcup_{j\in J} V_j$.

Let $\mathbf{X} = (X, \mathcal{X}, \mu, (T_\alpha)_{\alpha\in\mathbb{R}^m})$ and suppose that $V \not\subseteq W(\mathbf{X})$. Then there exists $\alpha \in V\backslash W(\mathbf{X})$ such that $I((t\alpha)_{t\in\mathbb{R}}) = I(\mathbb{R}^m) \subseteq I(V) \subseteq I((t\alpha)_{t\in\mathbb{R}})$. Therefore $I(\mathbb{R}^m) = I(V)$. Now consider the V-system $\mathbf{Y} = (X, \mathcal{X}, \mu, (T_\alpha)_{\alpha\in V})$. Since $I(\mathbb{R}^m) = I(V)$, we have that $W(\mathbf{X}) \cap V = W(\mathbf{Y})$.

Suppose that, for every family of at most countably many proper subspaces $(V_j)_{j\in J}$ of V, we have that $W(\mathbf{Y}) = W(\mathbf{X}) \cap V \not\subseteq \bigcup_{j\in J} V_j$. Since $I((t\alpha)_{t\in\mathbb{R}}) \subseteq I((n\alpha)_{n\in\mathbb{Z}})$, applying Lemma 2.2 to $\mathbf{Y}$, there exist at most countably many proper subspaces $(V_j)_{j\in J}$ of V, and at most countably many hyperplanes $(V_j)_{j\in J'}$ of V not passing through the origin such that $W(\mathbf{Y}) \subseteq \bigcup_{j\in J\cup J'} V_j$. By assumption, $W(\mathbf{Y}) \nsubseteq \bigcup_{j\in J} V_j$. So there exists $\alpha \in W(\mathbf{Y})\setminus \bigcup_{j\in J} V_j$. By the definition of $W(\mathbf{Y})$, it is easy to see that $\alpha \in W(\mathbf{Y})$ implies that $t\alpha \in W(\mathbf{Y})$ for all $t \in \mathbb{R}$. Since $\alpha \notin \bigcup_{j\in J} V_j$ implies that $t\alpha \notin \bigcup_{j\in J} V_j$ for all $t \neq 0$, we must have that $\{t\alpha : t \in \mathbb{R}\} \subseteq \bigcup_{j\in J'} V_j$. However, since V_j does not pass through the origin for all $j \in J'$, $\{t\alpha : t \in \mathbb{R}\} \cap \bigcup_{j\in J'} V_j$ is a countable set, which leads to a contradiction. This proves the claim.

Now we return to the proof of the theorem. By Lemma 2.2, $W(\mathbf{X}) \neq \mathbb{R}^m$. By the claim, there exists a family of at most countably many subspaces $(V_j)_{j\in J_1\cup L_1}$ of $\mathbb{R}^m$ of codimension at least 1 such that $W(\mathbf{X}) \subseteq \bigcup_{j\in J_1\cup L_1} V_j$, where $V_j \subseteq W(\mathbf{X})$ if $j \in L_1$ and $V_j \nsubseteq W(\mathbf{X})$ if $j \in J_1$. Applying the claim to each subspace in J_1, there exists a family of at most countably many subspaces $(V_j)_{j\in J_2\cup L_2}$ of $\mathbb{R}^m$ such that $W(\mathbf{X}) \subseteq \bigcup_{j\in J_2\cup L_2} V_j$, where all V_j, $j \in J$ are of codimension at least 2, $V_j \subseteq W(\mathbf{X})$ if $j \in L_2$, and $V_j \nsubseteq W(\mathbf{X})$ if $j \in J_2$. Using the claim repeatedly, there exists a family of at most countably many subspaces $(V_j)_{j\in J_m\cup L_m}$ of $\mathbb{R}^m$ such that $W(\mathbf{X}) \subseteq \bigcup_{j\in J_m\cup L_m} V_j$, where all V_j, $j \in J$ are of codimension at least m, $V_j \subseteq W(\mathbf{X})$ if $j \in L_m$, and $V_j \nsubseteq W(\mathbf{X})$ if $j \in J_m$. Since J_m is an empty set, we have that $W(\mathbf{X}) = \bigcup_{j\in J_m} V_j$, which finishes the proof. □

3. *Characteristic factors and structure theorem*

3.1. *Host–Kra characteristic factors.* Let G be an abelian locally compact Hausdorff topological group and $H_1, \ldots, H_d$ be subgroups of G. Let $\mathbf{X} = (X, \mathcal{X}, \mu, (T_g)_{g\in G})$ be a G-system. For convenience we denote $X^{[d]} = X^{2^d}$, $\mathcal{X}^{[d]} = \mathcal{X}^{2^d}$ and $T_g^{[d]} = T_g^{2^d}$. For any subgroup H of G, let $I(H)$ denote the σ-algebra of $\mathcal{X}$ consisting of all the H-invariant sets. For $1 \leq j \leq d-1$, let $I_\Delta(H_{j+1}^{[j]})$ denote the sub-σ-algebra of $\mathcal{X}^{[j]}$ consisting of all the sets which are invariant under $T_g^{[j]}$ for all $g \in H_{j+1}$. We inductively define the *Host–Kra measures* $\mu_{H_1,\ldots,H_j}$ on $X^{[j]}$ by setting $\mu_{H_1} = \mu \times_{I(H_1)} \mu$, meaning that

$$\int_{X^2} f \otimes g \, d\mu_{H_1} = \int_X \mathbb{E}(f|I(H_1)) \cdot \mathbb{E}(g|I(H_1)) \, d\mu$$

for all $f, g \in L^\infty(\mu)$; and for all $1 \leq j \leq d-1$, define $\mu_{H_1,\ldots,H_{j+1}} = \mu_{H_1,\ldots,H_j} \times_{I_\Delta(H_{j+1}^{[j]})} \mu_{H_1,\ldots,H_j}$, meaning that

$$\int_{X^{[j+1]}} F \otimes G \, d\mu_{H_1,\ldots,H_{j+1}} = \int_{X^{[j]}} \mathbb{E}(F|I_\Delta(H_{j+1}^{[j]})) \cdot \mathbb{E}(G|I_\Delta(H_{j+1}^{[j]})) \, d\mu_{H_1,\ldots,H_j}$$

for all $F, G \in L^\infty(\mu^{[j]})$. We define the *Host–Kra seminorm* by

$$\|f\|_{\mathbf{X},H_1,\ldots,H_d} := \left(\int_{X^{[d]}} f^{\otimes 2^d} \, d\mu_{H_1,\ldots,H_d}\right)^{1/2^d}$$

for all $f \in L^\infty(\mu)$. Let $Z_{H_1,\ldots,H_d}(\mathbf{X})$ (or $Z_{H_1,\ldots,H_d}$ when there is no confusion) be the sub-σ-algebra of $\mathcal{X}$ such that, for all $f \in L^\infty(\mu)$,

$$\mathbb{E}(f|Z_{H_1,\ldots,H_d}(\mathbf{X})) = 0 \quad \text{if and only if } \|f\|_{\mathbf{X},H_1,\ldots,H_d} = 0.$$

Similar to [**13**, proof of Lemma 4] (or [**14**, Lemma 4.3]), one can show that $Z_{H_1,\ldots,H_d}$ is well defined and we call it a *Host–Kra characteristic factor*. Sometimes we will slightly abuse the notation and say that '$Z_{H_1,\ldots,H_d}$ is a factor X', meaning that the system $(X, Z_{H_1,\ldots,H_d}, \mu, G)$ is a factor of $(X, \mathcal{X}, \mu, G)$.

https://doi.org/10.1017/etds.2019.67 Published online by Cambridge University Press

The following lemma is useful in many circumstances.

LEMMA 3.1. *Let G be an abelian locally compact Hausdorff topological group and $\mathbf{X}$ be a G-system. Let $H_1, \ldots, H_d, H_j'$ be subgroups of G for some $1 \le j \le d$.*

(i) *For every permutation $\sigma : \{1, \ldots, d\} \to \{1, \ldots, d\}$, we have that $Z_{H_1,\ldots,H_d}(\mathbf{X}) = Z_{H_{\sigma(1)},\ldots,H_{\sigma(d)}}(\mathbf{X})$;*

(ii) *If $I(H_j) = I(H_j')$, then $Z_{H_1,H_2,\ldots,H_j,\ldots,H_d}(\mathbf{X}) = Z_{H_1,H_2,\ldots,H_j',\ldots,H_d}(\mathbf{X})$.*

Proof. (i) The proof is similar to [**13**] and so we only provide a sketch. It suffices to show that, for all subgroups $H_1, \ldots, H_d$ of G and $1 \le i \le d-1$, we have that $Z_{H_1,\ldots,H_i,H_{i+1},\ldots,H_d}(\mathbf{X}) = Z_{H_1,\ldots,H_{i+1},H_i,\ldots,H_d}(\mathbf{X})$, or

$$\|f\|_{\mathbf{X},H_1,\ldots,H_i,H_{i+1},\ldots,H_d} = 0 \Leftrightarrow \|f\|_{\mathbf{X},H_1,\ldots,H_{i+1},H_i,\ldots,H_d} = 0$$

for all $f \in L^\infty(\mu)$†. By the definition of the Host–Kra measure, it suffices to show that $\|f\|_{\mathbf{X},H_1,\ldots,H_i,H_{i+1}} = 0 \Leftrightarrow \|f\|_{\mathbf{X},H_1,\ldots,H_{i+1},H_i} = 0$. Replacing the system $\mathbf{X} = (X, \mathcal{X}, \mu, (T_g)_{g\in G})$ with $(X^{[i-1]}, \mathcal{X}^{[i-1]}, \mu_{H_1,\ldots,H_{i-1}}, (T_g^{[i-1]})_{g\in G})$, it suffices to show that for all G-systems $\mathbf{X}$ and subgroups H_1, H_2 of G, we have that

$$\|f\|_{\mathbf{X},H_1,H_2} = 0 \Leftrightarrow \|f\|_{\mathbf{X},H_2,H_1} = 0.$$

Suppose first that $\|f\|_{\mathbf{X},H_1,H_2} = 0$. We may assume that $\|f\|_{L^\infty(\mu)} \le 1$. Let $(F_{1,n})_{n\in\mathbb{N}}$ and $(F_{2,n})_{n\in\mathbb{N}}$ be any Følner sequences of H_1 and H_2, respectively. Similarly to [**13**, Lemma 2], it is not hard to show that

$$\left|\lim_{N\to\infty} \frac{1}{|F_{1,N}|\cdot|F_{2,N}|} \sum_{g_1\in F_{1,N}, g_2\in F_{2,N}} \int_X f \cdot T_{g_1} f \cdot T_{g_2} f \cdot T_{g_1g_2} f \, d\mu\right| \le \|f\|_{\mathbf{X},H_1,H_2} = 0.$$

So the limit $\lim_{N\to\infty} (1/(|F_{1,N}|\cdot|F_{2,N}|)) \sum_{g_1\in F_{1,N}, g_2\in F_{2,N}} \int_X f\cdot T_{g_1} f\cdot T_{g_2} f\cdot T_{g_1g_2} f \, d\mu$ exists and is equal to 0. On the other hand, similarly to [**13**, (11)] (and invoking [**9**, Theorem 8.13], the Birkhoff ergodic theorem for G-systems),

$$\|f\|^4_{\mathbf{X},H_2,H_1} = \lim_{N\to\infty} \frac{1}{|F_{2,N}|} \sum_{g_2\in F_{2,N}} \lim_{N\to\infty} \frac{1}{|F_{1,N}|} \sum_{g_1\in F_{1,N}} \int_X f \cdot T_{g_1} f \cdot T_{g_2} f \cdot T_{g_1g_2} f \, d\mu,$$

where the limit

$$\lim_{N\to\infty} \frac{1}{|F_{1,N}|} \sum_{g_1\in F_{1,N}} \int_X f \cdot T_{g_1} f \cdot T_{g_2} f \cdot T_{g_1g_2} f \, d\mu$$

exists for all $g_2 \in H_2$. By [**3**, Lemmas 1.1 and 1.2],

$$\begin{aligned}\lim_{N\to\infty} \frac{1}{|F_{2,N}|} \sum_{g_2\in F_{2,N}} \lim_{N\to\infty} \frac{1}{|F_{1,N}|} \sum_{g_1\in F_{1,N}} \int_X f \cdot T_{g_1} f \cdot T_{g_2} f \cdot T_{g_1g_2} f \, d\mu \\ = \lim_{N\to\infty} \frac{1}{|F_{1,N}|\cdot|F_{2,N}|} \sum_{g_1\in F_{1,N}, g_2\in F_{2,N}} \int_X f \cdot T_{g_1} f \cdot T_{g_2} f \cdot T_{g_1g_2} f \, d\mu = 0,\end{aligned}$$

and so $\|f\|_{\mathbf{X},H_2,H_1} = 0$. Similarly, $\|f\|_{\mathbf{X},H_2,H_1} = 0$ implies that $\|f\|_{\mathbf{X},H_1,H_2} = 0$.

† It seems that [**13**, proof of Proposition 3] can be adapted to proving that $\|f\|_{\mathbf{X},H_1,\ldots,H_i,H_{i+1},\ldots,H_d} = \|f\|_{\mathbf{X},H_1,\ldots,H_{i+1},H_i,\ldots,H_d}$ for all $f \in L^\infty(X)$. But we do not need this property in this paper.

We now prove (ii). By (i), we may assume without loss of generality that $j=1$. Note that

$$\mu_{H_1}=\mu\times_{I(H_1)}\mu=\mu\times_{I(H_1')}\mu=\mu_{H_1'}.$$

By induction, $\mu_{H_1,H_2,\ldots,H_d}=\mu_{H_1',H_2,\ldots,H_d}$ and so $Z_{H_1,H_2,\ldots,H_d}(\mathbf{X})=Z_{H_1',H_2,\ldots,H_d}(\mathbf{X})$, which finishes the proof. □

The following is an immediate corollary of Lemma 3.1.

LEMMA 3.2. *Let G be an abelian locally compact Hausdorff topological group and $\mathbf{X}$ be a G-system. Let $H_1,\ldots,H_d$, $H_1',\ldots,H_d'$ be subgroups of G. If $I(H_i)=I(H_i')$ for all $1\le i\le d$. Then $Z_{H_1,H_2,\ldots,H_d}(\mathbf{X})=Z_{H_1',H_2',\ldots,H_d'}(\mathbf{X})$.*

3.2. *Structure theorems for $\mathbb{R}^m$-systems.* In this section, we establish structure theorems for $\mathbb{R}^m$-systems. These questions have been studied in various papers; see, for example, [**2**, **4**, **20**, **24**]. As none of the existing results can be applied directly to our problem, we need to develop the past results into a more general setting. In this paper, we only use some special cases of the theorems developed in this section. But we still write all the results in full generality for the convenience of future researchers.

CONVENTION 3.3. *Let $m\in\mathbb{N}$, $\mathbf{X}$ be an $\mathbb{R}^m$-system and $H_1,\ldots,H_d$ be subgroups of $\mathbb{R}^m$. In the notation $\mu_{H_1,H_2,\ldots,H_d}$, $Z_{H_1,H_2,\ldots,H_d}$ and $\|\cdot\|_{\mathbf{X},H_1,H_2,\ldots,H_d}$, if $H_i=(t\alpha_i)_{t\in\mathbb{R}}$ for some $\alpha_i\in\mathbb{R}^m$, we abbreviate H_i by α_i. If $H_i=(n\alpha_i)_{n\in\mathbb{Z}}$ for some $\alpha_i\in\mathbb{R}^m$, we abbreviate H_i by $\widehat{\alpha}_i$. For example, the notion $Z_{\alpha_1,\alpha_2,\widehat{\alpha}_3}$ represents $Z_{(t\alpha_1)_{t\in\mathbb{R}},(t\alpha_2)_{t\in\mathbb{R}},(n\alpha_3)_{n\in\mathbb{Z}}}$, and $\mu_{\alpha_1,\alpha_2,\widehat{\alpha}_3}$ represents $\mu_{(t\alpha_1)_{t\in\mathbb{R}},(t\alpha_2)_{t\in\mathbb{R}},(n\alpha_3)_{n\in\mathbb{Z}}}$.*

The Host–Kra characteristic factor is an important tool in the study of problems related to multiple averages. For example, certain Host–Kra characteristic factors control the L^2 limit of multiple averages for $\mathbb{Z}^m$-systems.

THEOREM 3.4. *Let $m\in\mathbb{N}$, $\mathbf{X}=(X,\mathcal{X},\mu,(T_g)_{g\in\mathbb{R}^m})$ be an $\mathbb{R}^m$-system and let $\alpha_1,\ldots,\alpha_d\in\mathbb{R}^m$. Denote $\widehat{Z}_i:=Z_{\widehat{\alpha_i},\widehat{\alpha_1-\alpha_i},\ldots,\widehat{\alpha_d-\alpha_i}}(\mathbf{X})$ for all $1\le i\le d$. Then, for all $f_1,\ldots,f_d\in L^\infty(\mu)$, both the $L^2(\mu)$ limits of*

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}f_1(T_{n\alpha_1}x)\cdot\ldots\cdot f_d(T_{n\alpha_d}x)$$

and

$$\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N-1}\mathbb{E}(f_1|\widehat{Z}_1)(T_{n\alpha_1}x)\cdot\ldots\cdot\mathbb{E}(f_d|\widehat{Z}_d)(T_{n\alpha_d}x)$$

exist and coincide (as $L^2(\mu)$ functions). Moreover, if both limit exist for μ-a.e. $x\in X$, then they coincide for μ-a.e. $x\in X$.

Proof. The existence and coincidence of the $L^2(\mu)$ limits is a result of Host [**13**, Proposition 1]. The existence and coincidence of the pointwise limit follows from the fact that if a sequence of bounded functions converge both as $L^2(\mu)$ functions and almost everywhere, then both limits are the same. □

The following lemma illustrates the connection between Host–Kra measures, seminorms and characteristic factors for $\mathbb{R}^m$-systems and $\mathbb{Z}^m$-systems.

LEMMA 3.5. *Let* $\mathbf{X}$ *be an* $\mathbb{R}^m$*-system and* $\alpha_1, \ldots, \alpha_d \in \mathbb{R}^m$. *Then for Lebesgue almost every* $s \in \mathbb{R}$, $\mu_{\widehat{s\alpha_1},\ldots,\widehat{s\alpha_d}} = \mu_{\alpha_1,\ldots,\alpha_d}$, $\|\cdot\|_{\mathbf{X},\widehat{s\alpha_1},\ldots,\widehat{s\alpha_d}} = \|\cdot\|_{\mathbf{X},\alpha_1,\ldots,\alpha_d}$ *and* $Z_{\widehat{s\alpha_1},\ldots,\widehat{s\alpha_d}} = Z_{\alpha_1,\ldots,\alpha_d}$.

Proof. Let H_i denote the $\mathbb{R}$-span of α_i and $\widehat{H_i}$ denote the $\mathbb{Z}$-span of $\widehat{\alpha_i}$. Applying Lemma 2.2 to each H_i, for Lebesgue almost every $s \in \mathbb{R}$, we have that $I(H_i) = I(s\widehat{H_i})$ for all $1 \le i \le d$. By Corollary 3.2, $Z_{s\widehat{H}_1,\ldots,s\widehat{H}_d} = Z_{H_1,\ldots,H_d}$. By definition, $\mu_{s\widehat{H}_1,\ldots,s\widehat{H}_d} = \mu_{H_1,\ldots,H_d}$ and $\|\cdot\|_{\mathbf{X},s\widehat{H}_1,\ldots,s\widehat{H}_d} = \|\cdot\|_{\mathbf{X},H_1,\ldots,H_d}$. This finishes the proof. □

We can now prove the following analog of Theorem 3.4 for $\mathbb{R}^m$-systems.

PROPOSITION 3.6. *Let* $m \in \mathbb{N}$, $\mathbf{X} = (X, \mathcal{X}, \mu, (T_g)_{g\in\mathbb{R}^m})$ *be an* $\mathbb{R}^m$*-system and let* $\alpha_1, \ldots, \alpha_d \in \mathbb{R}^m$. *Denote* $Z_i := Z_{\alpha_i,\alpha_1-\alpha_i,\ldots,\alpha_d-\alpha_i}(\mathbf{X})$ *for all* $1 \le i \le d$. *Then, for all* $f_1, \ldots, f_d \in L^\infty(\mu)$*, both the limits*

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T f_1(T_{t\alpha_1}x) \cdot \ldots \cdot f_d(T_{t\alpha_d}x)\, dt$$

and

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \mathbb{E}(f_1|Z_1)(T_{t\alpha_1}x) \cdot \ldots \cdot \mathbb{E}(f_d|Z_d)(T_{t\alpha_d}x)\, dt$$

exist and coincide (as $L^2(\mu)$ *functions). Moreover, if both limit exists for* μ*-a.e.* $x \in X$*, then they coincide for* μ*-a.e.* $x \in X$.

Proof. By Lemma 3.5, there exists $s \in \mathbb{R}$ such that $Z_i = Z_{\alpha_i,\alpha_1-\alpha_i,\ldots,\alpha_d-\alpha_i}(\mathbf{X}) = Z_{\widehat{s\alpha_i},\widehat{s(\alpha_1-\alpha_i)},\ldots,\widehat{s(\alpha_d-\alpha_i)}}(\mathbf{X})$ for all $1 \le i \le d$. For convenience we may assume without loss of generality that $s = 1$. By Theorem 3.4, for all $f_1, \ldots, f_d \in L^\infty(\mu)$,

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T_{n\alpha_1} f_1 \cdot \ldots \cdot T_{n\alpha_d} f_d = \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} T_{n\alpha_1}\mathbb{E}(f_1|Z_1) \cdot \ldots \cdot T_{n\alpha_d}\mathbb{E}(f_d|Z_d),$$

where the limits are taken in $L^2(\mu)$. Using the fact that every Z_i is G-invariant, we have that, as $L^2(\mu)$ functions,

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{T} \int_0^T f_1(T_{t\alpha_1}x) \cdot \ldots \cdot f_d((T_{t\alpha_d}x))\, dt \\
&= \int_0^1 \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} (T_{r\alpha_1} f_1)(T_{n\alpha_1}x) \cdot \ldots \cdot (T_{r\alpha_d} f_d)(T_{n\alpha_d}x)\, dr \\
&= \int_0^1 \lim_{N\to\infty} \frac{1}{N} \sum_{n=0}^{N-1} \mathbb{E}(T_{r\alpha_1} f_1|Z_1)(T_{n\alpha_1}x) \cdot \ldots \cdot \mathbb{E}(T_{r\alpha_d} f_d|Z_d)(T_{n\alpha_d}x)\, dr \\
&= \lim_{T\to\infty} \frac{1}{T} \int_0^T \mathbb{E}(f_1|Z_1)(T_{t\alpha_1}x) \cdot \ldots \cdot \mathbb{E}(f_d|Z_d)(T_{t\alpha_d}x)\, dt. \qquad (1)
\end{aligned}$$

Note that if a sequence of bounded functions converge both as $L^2(\mu)$ functions and almost everywhere, then both limits are the same. So (1) also holds for μ-a.e. $x \in X$ if all the limits in (1) exist for μ-a.e. $x \in X$. □

Let $X = N/\Gamma$, where N is a (k-step) nilpotent group and Γ is a discrete cocompact subgroup of N. Let $\mathcal{X}$ and μ be the Borel σ-algebra and Haar measure of X. Let $T_g \colon X \to X$, $T_g x = b_g \cdot x$, $g \in G$, for some group homomorphism $g \to b_g$ from G to N. We say that $\mathbf{X} = (X, \mathcal{X}, \mu, (T_g)_{g\in G})$ is a *(k-step) G-nilsystem*. It is classical that we can choose N to be simply connected, and we make this assumption throughout this paper. We remark that if G is connected, then we may also assume that N is connected. The following theorem is a combination of Theorem 3.4 in this paper and [**24**, Theorem 3.7]. We omit the proof.

THEOREM 3.7. *Let $m \in \mathbb{N}$, $\mathbf{X}$ be an ergodic $\mathbb{R}^m$-system and let $\alpha_1, \ldots, \alpha_d \in \mathbb{R}^m$. If the $\mathbb{Z}$-action $(T_{n\alpha_i})_{n\in\mathbb{Z}}$ is ergodic for $\mathbf{X}$ for all $1 \le i \le d$, then $Z_{\widehat{\alpha_1},\ldots,\widehat{\alpha_d}}(\mathbf{X})$ is an inverse limit of $(d-1)$-step $\mathbb{R}^m$-nilsystems.*

We have the following structure theorem for $\mathbb{R}^m$-actions, which should be viewed as an analog of the Host–Kra structure theorem [**14**].

PROPOSITION 3.8. *Let $m \in \mathbb{N}$ and $\mathbf{X}$ be an ergodic $\mathbb{R}^m$-system. Then $Z_{\mathbb{R}^m,\ldots,\mathbb{R}^m}(\mathbf{X})$ with d-copies of $\mathbb{R}^m$ is an inverse limit of $(d-1)$-step $\mathbb{R}^m$-nilsystems. Moreover, if $\alpha_1, \ldots, \alpha_d \in \mathbb{R}^m$ are such that the $\mathbb{R}$-action $(T_{t\alpha_i})_{t\in\mathbb{R}}$ is ergodic for $\mathbf{X}$ for all $1 \le i \le d$, then $Z_{\alpha_1,\ldots,\alpha_d}(\mathbf{X}) = Z_{\mathbb{R}^m,\ldots,\mathbb{R}^m}(\mathbf{X})$ with d copies of $\mathbb{R}^m$.*

Proof. By Lemma 2.2, it is not hard to show that there exists a $\mathbb{Z}$-action $(T_{n\alpha_i})_{n\in\mathbb{Z}}$ ergodic for $\mathbf{X}$ for all $1 \le i \le d$. By Lemma 3.2, we have that $Z_{\mathbb{R}^m,\ldots,\mathbb{R}^m} = Z_{\widehat{\alpha_i},\widehat{\alpha_1-\alpha_i},\ldots,\widehat{\alpha_d-\alpha_i}}$ with d copies of $\mathbb{R}^m$, which is an inverse limit of $(d-1)$-step $\mathbb{R}^m$-nilsystems by Theorem 3.7.

If $\alpha_1, \ldots, \alpha_d \in \mathbb{R}^m$ are such that the $\mathbb{R}$-action $(T_{t\alpha_i})_{t\in\mathbb{R}}$ is ergodic for $\mathbf{X}$ for all $1 \le i \le d$, then $I((t\alpha_i)_{t\in\mathbb{R}}) = I(\mathbb{R}^m)$ for all $1 \le i \le d$. By Lemma 3.2, we have that $Z_{\alpha_1,\ldots,\alpha_d}(\mathbf{X}) = Z_{\mathbb{R}^m,\ldots,\mathbb{R}^m}(\mathbf{X})$. □

4. *Proof of the main theorems*

We prove Theorems 1.1 and 1.3 in this section.

LEMMA 4.1. *Let $m \in \mathbb{N}$, ν be a Borel measure on $\mathbb{R}^m$ and $\mathbf{X}$ be an $\mathbb{R}^m$-system. If $\nu(W(\mathbf{X}) + \beta) = 0$ for all hyperplanes $\beta \in \mathbb{R}^m$, then the set of all $(\alpha, \beta) \in \mathbb{R}^{2m}$ such that $Z_{\alpha,\alpha-\beta} = Z_{\beta,\alpha-\beta} = Z_{\mathbb{R}^m,\mathbb{R}^m}$ is of $\nu \times \nu$ measure 1.*

Proof. By Proposition 3.8, if all the three $\mathbb{R}$-actions $(T_{t\alpha})_{t\in\mathbb{R}}$, $(T_{t\beta})_{t\in\mathbb{R}}$ and $(T_{t(\alpha-\beta)})_{t\in\mathbb{R}}$ are ergodic for $\mathbf{X}$, then $Z_{\alpha,\alpha-\beta} = Z_{\beta,\alpha-\beta} = Z_{\mathbb{R}^m,\mathbb{R}^m}$. So it suffices to show that the sets

$$E_1 = \{(\alpha, \beta) \in \mathbb{R}^{2m} : \alpha \in W(\mathbf{X})\}, \quad E_2 = \{(\alpha, \beta) \in \mathbb{R}^{2m} : \beta \in W(\mathbf{X})\}$$

and

$$E_3 = \{(\alpha, \beta) \in \mathbb{R}^{2m} : \alpha - \beta \in W(\mathbf{X})\},$$

have zero $\nu \times \nu$ measure. Obviously,

$$\nu \times \nu(E_1) = \nu \times \nu(E_2) = \nu(W(\mathbf{X})) = 0.$$

On the other hand,

$$\nu \times \nu(E_3) = \int_{\mathbb{R}^m} \nu(\{\alpha \in \mathbb{R}^m : \alpha \in W(\mathbf{X}) + \beta\})\, d\nu(\beta)$$
$$= \int_{\mathbb{R}^m} \nu(W(\mathbf{X}) + \beta)\, d\nu(\beta) = \int_{\mathbb{R}^m} 0\, d\nu(\beta) = 0.$$

This finishes proof. □

We start with a special case of Theorem 1.3.

PROPOSITION 4.2. *Let $d, m \in \mathbb{N}$, ν be a Borel measure on $\mathbb{R}^m$ and $\mathbf{X}$ be an ergodic $\mathbb{R}^m$-system such that $Z_{\mathbb{R}^m,\ldots,\mathbb{R}^m}(\mathbf{X}) = \mathbf{X}$ with d copies of $\mathbb{R}^m$. If $\nu(W(\mathbf{X}) + \beta) = 0$ for all hyperplanes $\beta \in \mathbb{R}^m$, then ν is weakly equidistributed for $\mathbf{X}$.*

Proof. It suffices to show that, for all $f \in L^\infty(\mu)$ with $\int_X f\, d\mu = 0$,

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \left| \int_{\mathbb{R}^m} f(T_{t\alpha}x)\, d\nu(\alpha) \right|^2 dt = 0$$

for μ-a.e. $x \in X$. Let J_0 denote the set of all $(\alpha, \beta) \in \mathbb{R}^{2m}$ such that $Z_{\alpha,\alpha-\beta} = Z_{\beta,\alpha-\beta} = Z_{\mathbb{R}^m,\mathbb{R}^m}$. By Lemma 4.1, $\nu \times \nu(J_0) = 1$. So

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \left| \int_{\mathbb{R}^m} f(T_{t\alpha}x)\, d\nu(\alpha) \right|^2 dt$$
$$= \lim_{T\to\infty} \frac{1}{T} \int_0^T \int_{\mathbb{R}^{2m}} f(T_{t\alpha}x)\overline{f}(T_{t\beta}x)\, d\nu(\alpha)\, d\nu(\beta)\, dt$$
$$= \int_{\mathbb{R}^{2m}} \left(\lim_{T\to\infty} \frac{1}{T} \int_0^T f(T_{t\alpha}x)\overline{f}(T_{t\beta}x)\, dt \right) d\nu(\alpha)\, d\nu(\beta)$$
$$= \int_{J_0} \left(\lim_{T\to\infty} \frac{1}{T} \int_0^T f(T_{t\alpha}x)\overline{f}(T_{t\beta}x)\, dt \right) d\nu(\alpha)\, d\nu(\beta). \tag{2}$$

Since $Z_{\mathbb{R}^m,\ldots,\mathbb{R}^m}(\mathbf{X}) = \mathbf{X}$, by Proposition 3.8 and an approximation argument, we may assume without loss of generality that $\mathbf{X}$ is an $\mathbb{R}^m$-nilsystem. We may assume without loss of generality that $\mathbf{X}$ is connected.

Suppose that $X = N/\Gamma$, where N is a (k-step) nilpotent group and Γ is a discrete cocompact subgroup of N. Let $\mathcal{X}$ and μ be the Borel σ-algebra and Haar measure of X. Assume that $T_g : X \to X$, $T_g x = b_g \cdot x$, $g \in G$, for some group homomorphism $g \to b_g$ from G to N. Let J denote the set of all $(\alpha, \beta) \in \mathbb{R}^{2m}$ such that $((b_{t\alpha}\Gamma, b_{t\beta}\Gamma))_{t\in\mathbb{R}}$ is equidistributed on $X \times X$. If $\nu \times \nu(J) = 1$, then

$$(2) = \int_J \int_{X\times X} f \otimes \overline{f}\, d\mu \times \mu\, d\nu(\alpha)\, d\nu(\beta) = \int_J \left| \int_X f\, d\mu \right|^2 d\nu(\alpha)\, d\nu(\beta) = 0,$$

which finishes the proof.

We now prove that $\nu \times \nu(J) = 1$. Since $\mathbb{R}^m$ is a connected group, we may assume that $X = N/\Gamma$ with N being connected and simply connected. Note that for all non-trivial horizontal characters χ of X (a *horizontal character* on $X = N/\Gamma$ is a continuous group homomorphism χ from N to $\mathbb{T}$ such that $\chi(\Gamma) = 1$), the complement of the set

$$A_\chi := \{\alpha \in \mathbb{R}^m : \chi(b_\alpha) \neq 1\}$$

is contained in $W(\mathbf{X})$. So $\nu(A_\chi) = 1$.

Let

$$A = \{\alpha \in \mathbb{R}^m : \chi(b_\alpha) \neq 1 \text{ for all non-trivial horizontal characters } \chi\}.$$

Since there are only countably many horizontal characters, $\nu(A) = 1$.

Fix $\alpha \in A$. Let B_α denote the set of β such that $((b_{t\alpha}\Gamma, b_{t\beta}\Gamma))_{t\in\mathbb{R}}$ is not equidistributed on $X \times X$. Then, for $\beta \in B_\alpha$, by Leibman's theorem [**19**], there exists a non-trivial horizontal character $\chi_{\alpha,\beta}$ of $X \times X$ such that $\chi_{\alpha,\beta}(b_\alpha, b_\beta) = 1$. Since $\alpha \in A$, there exist horizontal characters χ and χ' of X such that $\chi(b_\alpha) = \chi'(b_\beta) \neq 1$.

For all horizontal characters χ and χ' of X, let

$$B_{\chi,\chi'} = \{\gamma \in \mathbb{R}^m : \chi(b_\alpha) = \chi'(b_\gamma)\}.$$

$B_{\chi,\chi'}$ is obviously non-empty. Pick any $\gamma_0 \in B_{\chi,\chi'}$. Then

$$B_{\chi,\chi'} = \{\gamma \in \mathbb{R}^m : \chi'(b_{\gamma-\gamma_0}) = 1\}$$

which is contained in $W(\mathbf{X}) + \gamma_0$. By assumption, $\nu(B_{\chi,\chi'}) = 0$.

Since $B_\alpha = \bigcup_{\chi\neq 1}\bigcup_{\chi'} B_{\chi,\chi'}$, we have that $\nu(B_\alpha) = 0$. So $\nu \times \nu(J^c) \leq \nu \times \nu(\{(\alpha, \beta) : \alpha \in A, \beta \in B_\alpha\}) = 0$. This finishes the proof. □

Proof of Theorem 1.3. We start with the first part. Let $\mathbf{X} = (X, \mathcal{X}, \mu, (T_\alpha)_{\alpha\in\mathbb{R}^m})$ be an ergodic $\mathbb{R}^m$-system which is good for double Birkhoff averages such that $\nu(W(\mathbf{X}) + \beta) = 0$ for all $\beta \in \mathbb{R}^m$. Since $\mathbf{X}$ is separable, it suffices to show that, for all $f \in L^\infty(\mu)$, there exists $A \subseteq \mathbb{R}$ of density 1 such that

$$\lim_{t\in A, t\to\infty} \int_{\mathbb{R}^m} f(T_{t\alpha}x)\, d\nu(\alpha) = \int_X f\, d\mu \tag{3}$$

for μ-a.e. $x \in X$. Suppose first that f is measurable with respect to $Z_{\mathbb{R}^m,\mathbb{R}^m}$. Consider the factor system $\mathbf{Y} = (X, Z_{\mathbb{R}^m,\mathbb{R}^m}, \mu, (T_\alpha)_{\alpha\in\mathbb{R}^m})$ of $\mathbf{X}$. Since $W(\mathbf{Y}) \subseteq W(\mathbf{X})$, we have that $\nu(W(\mathbf{Y}) + \beta) = 0$ for all $\beta \in \mathbb{R}^m$. So (3) follows from Propositions 4.2.

We now assume that $\mathbb{E}(f|Z_{\mathbb{R}^m,\mathbb{R}^m}) = 0$. To show (3), it suffices to show that

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \left|\int_{\mathbb{R}^m} f(T_{t\alpha}x)\, d\nu(\alpha)\right|^2 dt = 0$$

for μ-a.e. $x \in X$. By Proposition 3.6 and the assumption that $\mathbf{X}$ is good for double Birkhoff averages,

$$\begin{aligned}
&\lim_{T\to\infty} \frac{1}{T}\int_0^T \left|\int_{\mathbb{R}^m} f(T_{t\alpha}x)\, d\nu(\alpha)\right|^2 dt \\
&= \lim_{T\to\infty} \frac{1}{T}\int_0^T \int_{\mathbb{R}^{2m}} f(T_{t\alpha}x)\overline{f}(T_{t\beta}x)\, d\nu(\alpha)\, d\nu(\beta)\, dt \\
&= \int_{\mathbb{R}^{2m}} \left(\lim_{T\to\infty} \frac{1}{T}\int_0^T f(T_{t\alpha}x)\overline{f}(T_{t\beta}x)\, dt\right) d\nu(\alpha)\, d\nu(\beta) \\
&= \int_{\mathbb{R}^{2m}} \left(\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathbb{E}(f|Z_{\alpha,\alpha-\beta})(T_{t\alpha}x)\mathbb{E}(\overline{f}|Z_{\beta,\alpha-\beta})(T_{t\beta}x)\, dt\right) d\nu(\alpha)\, d\nu(\beta).
\end{aligned}$$

Let J denote the set of all $(\alpha, \beta) \in \mathbb{R}^{2m}$ such that $Z_{\alpha,\alpha-\beta} = Z_{\beta,\alpha-\beta} = Z_{\mathbb{R}^m,\mathbb{R}^m}$. By Lemma 4.1, $\nu \times \nu(J) = 1$. So

$$\int_{\mathbb{R}^{2m}} \bigg(\lim_{T\to\infty} \frac{1}{T} \int_0^T \mathbb{E}(f|Z_{\alpha,\alpha-\beta})(T_{t\alpha}x)\mathbb{E}(\overline{f}|Z_{\beta,\alpha-\beta})(T_{t\beta}x)\, dt \bigg)\, d\nu(\alpha)\, d\nu(\beta)$$
$$= \int_J \bigg(\lim_{T\to\infty} \frac{1}{T} \int_0^T \mathbb{E}(f|Z_{\mathbb{R}^m,\mathbb{R}^m})(T_{t\alpha}x)\mathbb{E}(\overline{f}|Z_{\mathbb{R}^m,\mathbb{R}^m})(T_{t\beta}x)\, dt \bigg)\, d\nu(\alpha)\, d\nu(\beta) = 0$$

and we are done.

We now prove the second part. Let $\mathbf{X} = (X, \mathcal{X}, \mu, (T_\alpha)_{\alpha\in\mathbb{R}^m})$ be an ergodic $\mathbb{R}^m$-system and suppose that $\nu(W(\mathbf{X}) + \beta) > 0$ for some $\beta \in \mathbb{R}^m$. We wish to show that ν is not weakly equidistributed for $\mathbf{X}$. By Theorem 1.6, there exist $\beta \in \mathbb{R}^m$ and a subspace V of $\mathbb{R}^m$ contained in $W(\mathbf{X})$ such that $\nu(V + \beta) > 0$ and $\nu(V' + \beta) = 0$ for every proper subspace V' of V. Again by Theorem 1.6, it is not hard to show that there exists an $(m-1)$-dimensional subspace V_0 of $\mathbb{R}^m$ which contains V such that, for every subspace V'' of $\mathbb{R}^m$ which is contained in $W(\mathbf{X})$ but not contained in V, we have $V'' + V_0 = \mathbb{R}^m$. Let U be the one-dimensional subspace of $\mathbb{R}^m$ which is the orthogonal complement of V_0, and let $\pi : \mathbb{R}^m \to U$ be the natural projection. Let $\mathbf{Y} = (X, I(V_0), \mu, (T_\alpha)_{\alpha\in U})$. Since, for all V_0-invariant functions f, we have

$$\int_{\mathbb{R}^m} f(T_{t\alpha}x)\, d\nu(\alpha) = \int_{\mathbb{R}^m} f(T_{t\pi(\alpha)}x)\, d\nu(\alpha) = \int_U f(T_{t\alpha}x)\, d\pi_*\nu(\alpha),$$

where $\pi_*\nu$ is the push-forward of ν under π, in order to show that ν is not weakly equidistributed for the ergodic $\mathbb{R}^m$-system $\mathbf{X}$, it suffices to show that $\pi_*\nu$ is not weakly equidistributed for the ergodic U-system $\mathbf{Y}$.

We may decompose $\pi_*\nu$ as the sum of two (unnormalized) measures $\pi_*\nu = \nu_c + \nu_d$, where $\nu_c(\{\beta\}) = 0$ for all $\beta \in U$, and ν_d is supported on at most countably many points on U. Since $\pi_*\nu(\{\pi(\beta)\}) = \nu(V_0 + \beta) \geq \nu(V + \beta) > 0$, we have that $\nu_d \neq 0$. Since U is isomorphic to $\mathbb{R}$, applying the conclusion of the first part, we have that (the normalization of) ν_c is weakly equidistributed for $\mathbf{Y}$. So it suffices to show that (the normalization of) ν_d is not weakly equidistributed for $\mathbf{Y}$.

Suppose that (the normalization of) ν_d is weakly equidistributed for $\mathbf{Y}$. We may assume that $\nu_d = \sum_{j\in J} c_j \delta_{\alpha_j}$ for some non-empty countable index set J, $c_j > 0$, $\alpha_j \in U$, where $\alpha_j \neq \alpha_{j'}$ for $j \neq j'$. We assume without loss of generality that $\sum_{j\in J} c_j = 1$. Then, for all $f \in L^\infty(\mu)$ with $\int_X f\, d\mu = 0$,

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \bigg| \int_{\mathbb{R}^m} f(T_{t\alpha}x)\, d\nu(\alpha) \bigg|^2 dt = \lim_{T\to\infty} \frac{1}{T} \int_0^T \bigg| \sum_{j\in J} c_j f(T_{t\alpha_j}x) \bigg|^2 dt = 0$$

for μ-a.e. $x \in X$. Since $f \in L^\infty(\mu)$, we have that the $L^1(\mu)$ limit of

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \bigg| \sum_{j\in J} c_j f(T_{t\alpha_j}x) \bigg|^2 dt$$

is also equal to 0. Since U is of dimension 1, by Theorem 1.6, for all $\alpha \in U\backslash\{0\}$, $(T_{t\alpha})_{t\in\mathbb{R}}$ is ergodic for $\mathbf{Y}$. So

$$\lim_{T\to\infty} \frac{1}{T} \int_0^T \bigg(\int_X f(T_{t\alpha}x)\overline{f}(x)\, d\mu \bigg)\, dt = \mathbf{1}_{\alpha=0} \cdot \|f\|^2_{L^2(\mu)}.$$

So

$$\begin{aligned}\int_X \lim_{T\to\infty} \frac{1}{T}\int_0^T \left|\sum_{j\in J} c_j f(T_{t\alpha_j}x)\right|^2 dt\, d\mu(x) \\ = \sum_{j,j'\in J} c_j c'_j \lim_{T\to\infty} \frac{1}{T}\int_0^T \left(\int_X f(T_{t\alpha_j}x)\overline{f}(T_{t\alpha_{j'}}x)\, d\mu\right) dt \\ = \sum_{j,j'\in J} c_j c'_j \lim_{T\to\infty} \frac{1}{T}\int_0^T \left(\int_X f(T_{t(\alpha_j-\alpha'_j)}x)\overline{f}(x)\, d\mu\right) dt \\ = \sum_{j\in J} c_j^2 \cdot \|f\|^2_{L^2(\mu)} > 0\end{aligned}$$

whenever $\|f\|_{L^2(\mu)} > 0$ (since J is non-empty), a contradiction. This proves the second part of the theorem. □

We are now ready to prove Theorem 1.1.

Proof of Theorem 1.1. Suppose first that ν is a Borel measure on $\mathbb{R}^m$ such that $\nu(\ell) = 0$ for all hyperplanes ℓ of $\mathbb{R}^m$. Let $\mathbf{X}$ be an ergodic $\mathbb{R}^m$-system which is good for double Birkhoff ergodic averages. By Theorem 1.6, for every $\beta \in \mathbb{R}^m$, $W(\mathbf{X}) + \beta$ is contained in an at most countable union of hyperplanes of $\mathbb{R}^m$. So $\nu(W(\mathbf{X}) + \beta) = 0$. By Theorem 1.3, ν is weakly equidistributed for $\mathbf{X}$. This proofs the 'if' part.

We now prove the 'only if' part. Suppose that there exists a hyperplane of $\mathbb{R}^m$,

$$\ell = \{\alpha \in \mathbb{R}^m : \alpha \cdot \beta = c\},$$

such that $\nu(\ell) \neq 0$, where $\beta \in \mathbb{R}^m$ and $c \in \mathbb{R}$. Let $(X, \mathcal{X}, \mu)$ be the one-dimensional torus. Let $(S_s)_{s\in\mathbb{R}}$ be the ergodic $\mathbb{R}$-action on X given by $S_s x = x + s \mod 1$, $x \in [0, 1)$. We now consider the $\mathbb{R}^m$-system $\mathbf{X}_0 = (X, \mathcal{X}, \mu, (T_\alpha)_{\alpha\in\mathbb{R}^m})$, where $T_\alpha = S_{\pi(\alpha)}$ for all $\alpha \in \mathbb{R}^m$ with $\pi : \mathbb{R}^m \to \mathbb{R}$ being the linear map given by $\pi(\alpha) = \alpha \cdot \beta$, $\alpha \in \mathbb{R}^m$. This system is obviously good for Birkhoff double averages.

Note that $W(\mathbf{X}_0) = \{\alpha \in \mathbb{R}^m : \alpha \cdot \beta = 0\}$ and so $\nu(\ell) = \nu(W(\mathbf{X}_0) + c) = 0$. By Theorem 1.3, ν is not weakly equidistributed for $\mathbf{X}_0$. □

5. *Systems good for double Birkhoff averages*

In this section we discuss to what extent do the main theorems of this paper apply, that is, which systems are good for double Birkhoff averages. Using an argument similar to the proof of Proposition 3.6 (or use [**4**, Theorem 3.1]), it is not hard to show the following lemma.

LEMMA 5.1. *Let* $\mathbf{X} = (X, \mathcal{X}, \mu, (T_\alpha)_{\alpha\in\mathbb{R}^m})$ *be an* $\mathbb{R}^m$*-system. If, for all* $f_1, f_2 \in L^\infty(\mu)$ *and* $\alpha_1, \alpha_2 \in \mathbb{R}^m$*, the limit*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=0}^{N-1} f_1(T_{\alpha_1 n}x) f_2(T_{\alpha_2 n}x) \tag{4}$$

exists for μ*-a.e.* $x \in X$*, then* $\mathbf{X}$ *is good for double Birkhoff averages.*

https://doi.org/10.1017/etds.2019.67 Published online by Cambridge University Press

Combining Lemma 5.1 with past results in the literature, we have that the following $\mathbb{R}^m$-systems $\mathbf{X} = (X, \mathcal{X}, \mu, (T_\alpha)_{\alpha\in\mathbb{R}^m})$ are good for double Birkhoff averages.

- Assani [**1**]: every weakly mixing $\mathbb{R}^m$-system $\mathbf{X}$ such that, for every $g \in \mathbb{R}^m$, the restriction of T_g to the Pinsker algebra of $\mathbf{X}$ (the maximal sub-σ-algebra on which T_g has zero entropy) has singular spectrum with respect to the Lebesgue measure†.
- Bourgain [**6**]: every $\mathbb{R}$-system (or every $\mathbb{R}^m$-system $\mathbf{X} = (X, \mathcal{X}, \mu, (T_\alpha)_{\alpha\in\mathbb{R}^m})$ for which there exist an $\mathbb{R}$-action $(S_t)_{t\in\mathbb{R}}$ on $\mathbf{X}$ and $\beta \in \mathbb{R}^m$ such that $T_\alpha = S_{\alpha\cdot\beta}$ for all $\alpha \in \mathbb{R}^m$).
- Donoso and Sun [**8**]: every distal $\mathbb{R}^m$-system $\mathbf{X}$ (the case $m = 1$ was proved by Huang, Shao and Ye [**15**]).

For completeness we recall the definition of distal systems (and refer readers to [**10**, Ch. 10] for further details). Let G be a group and $\pi : \mathbf{X} = (X, \mathcal{X}, \mu, (T_g)_{g\in G}) \to \mathbf{Y} = (Y, \mathcal{Y}, \nu, (S_g)_{g\in G})$ be a factor map between two G-systems. We say π is an *isometric extension* (or $\mathbf{X}$ is an *isometric extension* of $\mathbf{Y}$) if there exist a compact group H, a closed subgroup Γ of H, and a cocycle $\rho : G \times Y \to H$ such that $(X, \mathcal{X}, \mu, (T_g)_{g\in G}) \cong (Y \times H/\Gamma, \mathcal{Y} \times \mathcal{H}, \nu \times m, (T_g)_{g\in G})$, where m is the Haar measure on H/Γ, $\mathcal{H}$ is the Borel σ-algebra on H/Γ, and that for all $g \in G$ and $(y, a\Gamma) \in Y \times H/\Gamma$, we have

$$T_g(y, a\Gamma) = (S_g y, \rho(g, y)a\Gamma).$$

Definition 5.2. Let $\mathbf{X}$ be a G-system. We say that $\mathbf{X}$ is *distal* if there exist a countable ordinal η and a directed family of factors $\mathbf{X}_\theta$, $\theta \leq \eta$ of $\mathbf{X}$ such that:

(1) $\mathbf{X}_0$ is the trivial system, and $\mathbf{X}_\eta = \mathbf{X}$;
(2) for $\theta < \eta$, the extension $\pi_\theta : \mathbf{X}_{\theta+1} \to \mathbf{X}_\theta$ is isometric and is not an isomorphism;
(3) for a limit ordinal $l \leq \eta$, $\mathbf{X}_l = \lim_{\leftarrow\theta<l} \mathbf{X}_\theta$.

As a result of [**8**], we have the following applications of Theorems 1.1 and 1.3.

PROPOSITION 5.3. *Let $m \in \mathbb{N}$ and ν be a Borel measure on $\mathbb{R}^m$.*

(i) *ν is weakly equidistributed for an ergodic distal $\mathbb{R}^m$-system $\mathbf{X}$ if and only if $\nu(W(\mathbf{X}) + \beta) = 0$ for every $\beta \in \mathbb{R}^m$.*
(ii) *ν is weakly equidistributed for all ergodic distal $\mathbb{R}^m$-systems if and only if $\nu(\ell) = 0$ for all hyperplanes ℓ of $\mathbb{R}^m$.*

Proof. Fix $\alpha_1, \alpha_2 \in \mathbb{R}^m$ and let G' denote the $\mathbb{Z}$-span of α_1, α_2. Then it is easy to see by definition that if $\mathbf{X} = (X, \mathcal{X}, \mu, (T_\alpha)_{\alpha\in\mathbb{R}^m})$ is a distal $\mathbb{R}^m$-system, then $(X, \mathcal{X}, \mu, (T_\alpha)_{\alpha\in G'})$ is a distal $\mathbb{Z}^2$-system. By [**8**], the limit (4) exists for all $f_1, f_2 \in L^\infty(\mu)$ and μ-a.e. $x \in X$. By Lemma 5.1, $\mathbf{X}$ is good for double Birkhoff averages, and so the 'if' parts of (i) and (ii) follow from the first part of Theorem 1.3 and the 'if' part of Theorem 1.1, respectively. The 'only if' part of (i) follows from the second part of Theorem 1.3 (which is valid for every $\mathbb{R}^m$-system). The 'only if' part of (ii) follows from the 'only if' of Theorem 1.1 as the system $\mathbf{X}_0$ constructed in the proof of Theorem 1.1 is distal. □

† Ref. [**1**] only covered the case where $m = 1$, but the general case can be deduced by a similar argument combined with results in [**12**]. It is worth noting that the results in [**1**] were recently improved by Gutman *et al* [**11**].

Using the result of [**6**], we can deduce Proposition 1.5 from Theorem 1.3.

Proof of Proposition 1.5. Let ν be a Borel measure on $\mathbb{R}$ and $\mathbf{X}$ be an ergodic $\mathbb{R}$-system. By [**6**], $\mathbf{X}$ is good for double Birkhoff averages. Then, by Theorem 1.3, ν is weakly equidistributed for $\mathbf{X}$ if and only if $\nu(W(\mathbf{X})+\beta)=0$ for all $\beta\in\mathbb{R}$. Since $\mathbf{X}$ is an ergodic $\mathbb{R}$-system, it is easy to see that $W(\mathbf{X})=\{0\}$. This finishes the proof. □

REFERENCES

[**1**] I. Assani. Multiple recurrence and almost sure convergence for weakly mixing dynamical systems. *Israel J. Math.* **103** (1998), 111–124.

[**2**] T. Austin. Norm convergence of continuous-time polynomial multiple ergodic averages. *Ergod. Th. & Dynam. Sys.* **32**(2) (2012), 361–382.

[**3**] V. Bergelson and A. Leibman. Cubic averages and large intersections. *Recent Trends in Ergodic Theory and Dynamical Systems (Contemporary Mathematics, 631).* American Mathematical Society, Providence, RI, 2015, pp. 5–19.

[**4**] V. Bergelson, A. Leibman and J. Moreira. From discrete- to continuous-time ergodic theorems. *Ergod. Th. & Dynam. Sys.* **32**(2) (2012), 383–426.

[**5**] M. Björklund. Ergodic theorems for homogeneous dilations. *Random Walks, Boundaries and Spectra (Progress in Probability, 64).* Eds. D. Lenz, F. Sobieczky and W. Woess. Springer, Basel, 2011.

[**6**] J. Bourgain. Double recurrence and almost sure convergence. *J. Reine Angew. Math.* **404** (1990), 140–161.

[**7**] J. Chaika and P. Hubert. Circle averages and disjointness in typical flat surfaces on every Teichmüller disc. *Bull. Lond. Math. Soc.* **49**(5) (2017), 755–769.

[**8**] S. Donoso and W. Sun. Pointwise convergence of some multiple ergodic averages. *Adv. Math.* **330** (2018), 946–996.

[**9**] M. Einsiedler and T. Ward. *Ergodic Theory with a View towards Number Theory*. Springer, London, 2011, 259.

[**10**] E. Glasner. *Ergodic Theory via Joinings (Mathematical Surveys and Monographs, 101).* American Mathematical Society, Providence, RI, 2003.

[**11**] Y. Gutman, W. Huang, S. Shao and X. Ye. Almost sure convergence of the multiple ergodic average for certain weakly mixing systems. *Acta Math. Sin. (Engl. Ser.) (1)* **34** (2018), 79–90.

[**12**] B. Host. Mixing of all orders and pairwise independent joining. *Israel J. Math.* **76** (1991), 289–298.

[**13**] B. Host. Ergodic seminorms for commuting transformations and applications. *Studia Math.* **195** (2009), 31–49.

[**14**] B. Host and B. Kra. Nonconventional ergodic averages and nilmanifolds. *Ann. of Math. (2)* **161**(1) (2005), 397–488.

[**15**] W. Huang, S. Shao and X. Ye. Pointwise convergence of multiple ergodic averages and strictly ergodic models. *J. Anal. Math.*, to appear. *Preprint*, 2014, arXiv:1406.5930.

[**16**] R. L. Jones. Ergodic averages on spheres. *J. Anal. Math.* **61** (1993), 29–45.

[**17**] B. Kra, N. Shah and W. Sun. Equidistribution of dilated curves on nilmanifolds. *J. Lond. Math. Soc.* **98**(3) (2018), 708–732.

[**18**] M. T. Lacey. Ergodic averages on circles. *J. Anal. Math.* **67** (1995), 199–206.

[**19**] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold. *Ergod. Th. & Dynam. Sys.* **25**(1) (2005), 201–213.

[**20**] A. Potts. Multiple ergodic averages for flows and an application. *Illinois J. Math.* **55**(2) (2011), 589–621.

[**21**] C. Pugh and M. Shub. Ergodic elements of ergodic actions. *Compos. Math.* **23** (1971), 115–122.

[**22**] N. Shah and P. Yang. Stretching translates of shrinking curves and Dirichlet's simultaneous approximation. *Preprint*, 2018, arXiv:1809.05570.

[**23**] E. M. Stein. Maximal functions: spherical means. *Proc. Natl Acad. Sci. USA* **73** (1976), 2174–2175.

[**24**] T. Ziegler. Nilfactors of $\mathbb{R}^m$-actions and configurations in sets of positive upper density in $\mathbb{R}^m$. *J. Anal. Math.* **99** (2006), 249–266.