HIGHER-DIMENSIONAL MOVING AVERAGES AND SUBMANIFOLD GENERICITY

JIAJUN CHENG, REYNOLD FREGOLI, AND BEINUO GUO

Abstract. We generalize results of Bellow, Jones, and Rosenblatt on moving ergodic averages to measure-preserving actions of $\mathbb{Z}^d$ and $\mathbb{R}^d$ for $d \geq 1$. In particular, we give necessary and sufficient conditions for the pointwise convergence of certain sequences of functions defined by averaging over families of boxes in $\mathbb{Z}^d$ and $\mathbb{R}^d$. As an application of our characterization, we show that averages along dilates of "locally flat" submanifolds in $\mathbb{R}^d$ do not necessarily converge point-wise for bounded measurable functions. This is closely related to the concept of submanifold-genericity recently introduced in [BFK25].

1. Introduction

Let T be a measure-preserving transformation defined on a non-atomic probability space (X, μ). Let $\mathbb{N}$ denote the set of strictly positive integers and let Ω be a sequence of integer pairs $\{(n_k, l_k)\} \subset \mathbb{N} \times \mathbb{N}$. For each $k \in \mathbb{N}$ consider the averaging operator

$$A_k f := \frac{1}{l_k} \sum_{j=0}^{l_k - 1} f \circ T^{n_k + j},$$

where $f \in \mathrm{L}^1(X)$, and let

$$M_\Omega f := \sup_k \frac{1}{l_k} \sum_{j=0}^{l_k - 1} \left| f \circ T^{n_k + j} \right|.$$

In [BJR90], Bellow, Jones, and Rosenblatt gave necessary and sufficient conditions on the sequence Ω for M_Ω to satisfy a maximal inequality. Let us briefly recall their characterization. For $\alpha > 0$ define

$$\Omega^{(\alpha)} := \left\{ (x, y) \in \mathbb{N}^2 : |x - n_k| \leq \alpha (y - l_k) \text{ for some } (n_k, l_k) \in \Omega \right\}$$

and for $\lambda \in \mathbb{N}$ set

$$\Omega^{(\alpha)}(\lambda) := \left\{ x \in \mathbb{N} : (x, \lambda) \in \Omega^{(\alpha)} \right\}.$$

Then, according to [BJR90, Theorem 1], M_Ω is of weak type $(1,1)$ and of strong type (p,p) for any $p > 1$ if and only if the following condition on the sequence Ω is satisfied

$$(C_1) \qquad \exists\, A, \alpha > 0 : \quad \#\Omega^{(\alpha)}(\lambda) \leq A\lambda \quad \text{for all } \lambda \in \mathbb{N}.$$

2020 *Mathematics Subject Classification.* 37A25, 37A30, 28D05, 28D15.

This paper was written during an REU project organized by the department of mathematics of the University of Michigan, Ann Arbor, whose support is gratefully acknowledged. The third author is sponsored by the Derksen Research Incentive fund.

Condition (C_1) asserts in other words that the horizontal cross-section of all the cones of aperture α above points in the sequence Ω at fixed height must grow linearly as the height increases. For example, the sequence (k, rk) for any $r \in \mathbb{N}$ has this property, while the sequence $(k, \sqrt{k})$ does not (see also picture below).

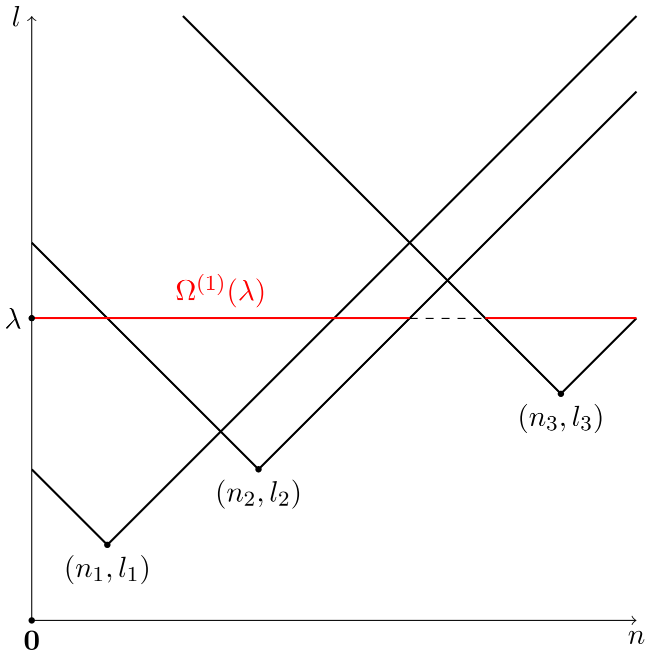

FIGURE 1. Set $\Omega^{(1)}(\lambda)$ for some sequence (n_k, l_k).

It is a well-known fact that, when the operator M_Ω is of weak type $(1, 1)$, the sequence of functions $A_k f$ converges point-wise for any $f \in \mathrm{L}^1(X)$. [BJR90, Theorem 4] shows additionally that this is not the case if Condition (C_1) fails.

In this paper, we generalize [BJR90, Theorem 1, Part a)] and [BJR90, Theorem 4] to measurable actions of $\mathbb{Z}^d$ and $\mathbb{R}^d$ for $d \geq 1$. Based on these extensions, we study the property of submanifold genericity, recently introduced in [BFK25]. We now proceed to lay out our results in greater detail.

1.1. **The Discrete Case.** Let (X, μ) be a non-atomic probability space and let $T_1, \ldots,$ T_d be commuting measure-preserving transformations on the space X (which is our standing assumption throughout this section). Given sequences $\{\boldsymbol{n}_k\} \subset \mathbb{Z}^d$ and $\{\boldsymbol{l}_k\} \subset \mathbb{N}^d$ $(k \in \mathbb{N})$, put

$$B_k := [n_{k1}, n_{k1} + l_{k1}) \times \cdots \times [n_{kd}, n_{kd} + l_{kd})$$

and let $\mathcal{B} := \{B_k\}$. To any box $B_k \in \mathcal{B}$ associate an averaging operator A_k defined by

$$A_k f(x) := \frac{1}{l_{k1} \cdots l_{kd}} \sum_{\boldsymbol{j} \in B_k \cap \mathbb{Z}^d} f\left(T_1^{j_1} \cdots T_d^{j_d} x\right)$$

for $f \in \mathrm{L}^1(X)$, and denote by $M_{\mathcal{B}}$ the maximal operator

$$M_{\mathcal{B}} f(x) := \sup_{B_k \in \mathcal{B}} \frac{1}{l_{k1} \cdots l_{kd}} \sum_{\boldsymbol{j} \in B_k \cap \mathbb{Z}^d} \left| f\left(T_1^{j_1} \cdots T_d^{j_d} x\right)\right|.$$

Our first result shows, roughly speaking, that, if Condition (C_1) holds in each coordinate, then the operator $M_{\mathcal{B}}$ is of strong type (p,p) for any $p > 1$. More precisely, for $i = 1, \dots, d$ let

$$\Omega_i := \{(n_{ki}, l_{k,i}) : k \in \mathbb{N}\}$$

and for $\alpha > 0$ define

$$\Omega_i^{(\alpha)} := \{(x,y) \in \mathbb{Z} \times \mathbb{N} : |x - n| \le \alpha(y - l) \text{ for some } (n,l) \in \Omega_i\}.$$

Finally, given $\lambda \in \mathbb{N}$ put

$$\Omega_i^{(\alpha)}(\lambda) := \left\{x \in \mathbb{Z} : (x, \lambda) \in \Omega_i^{(\alpha)}\right\}.$$

Then the following holds.

Theorem 1.1. *Assume that for all $i = 1, \dots, d$*

$$(C_i) \qquad \exists\, A, \alpha > 0 : \quad \#\Omega_i^{(\alpha)}(\lambda) \le A\lambda \quad \text{for all } \lambda \in \mathbb{N}.$$

Then the operator $M_{\mathcal{B}}$ is of strong type (p,p) for any $p > 1$.

Remark 1.2. It does not follow from our proof that $M_{\mathcal{B}}$ is of weak type $(1,1)$. In particular, it appears that extending the ideas of Bellow, Jones, and Rosenblatt to prove a weak type-$(1,1)$ inequality for $\mathbb{Z}^d$- or $\mathbb{R}^d$-actions with $d \ge 2$ is a non trivial task. This is mainly due to obstructions in the use of the Hardy-Littlewood Maximal Inequality in higher-dimension (see [BJR90, pages 45 and 46]).

As a corollary to Theorem 1.1 we deduce that, under Condition (C_i) for $i = 1, \dots, d$, the sequence of functions $A_k f$ converges point-wise for any $f \in \mathrm{L}^p(X)$ with $p > 1$.

Corollary 1.3. *Assume that the action generated by the transformations $T_1, \dots, T_d$ on the space X is ergodic and that Condition (C_i) is satisfied for all $i = 1, \dots, d$. Then for any $f \in \mathrm{L}^p(X)$ with $p > 1$ and μ-a.e. $x \in X$ we have that*

$$A_k f(x) \to \mu(f) \quad \text{as } k \to \infty.$$

If Condition (C_i) is not satisfied for at least one $1 \le i \le d$, on the other hand, and the action of $\mathbb{Z}^d$ on X is aperiodic, the conclusion of Corollary (1.3) fails in a strong sense. Recall that a measurable action of $\mathbb{Z}^d$ on X is said to be *aperiodic* if $\mu(x \in X : \boldsymbol{a}.x = x) = 0$ for all $\boldsymbol{a} \in \mathbb{Z}^d \setminus \{\boldsymbol{0}\}$.

Before formally stating the converse of Corollary 1.3, let us also recall the definition of *mixing family* introduced by Sawyer. Following [Saw66], we say that a family of measure-preserving transformations $\{S_h\}$ on X is *mixing* if for any pair of measurable subsets A, B of X and any $\rho > 1$

$$(1.1) \qquad \exists\, h : \quad \mu(A \cap S_h^{-1}(B)) < \rho \cdot \mu(A)\mu(B).$$

Remark 1.4. Note that the notion of "mixing family" differs from that of "mixing action". More precisely, suppose that $\{S_h\}$ is a locally compact group of measure-preserving transformations on X. Then Condition (1.1) is equivalent to

$$\inf_h \mu\left(A \cap S_h(B)\right) \leq \mu(A)\mu(B).$$

Compare this to

$$\lim_{S_h \to \infty} \mu\left(A \cap S_h(B)\right) = \mu(A)\mu(B),$$

which corresponds the usual definition of mixing action. In fact, it is possible to show that any group of measure-preserving transformations that acts ergodically on X satisfies (1.1) (see, e.g., [Saw66, Lemma 1, page 163]). This is an easy exercise if $\{S_h\}$ is just the group generated by one ergodic transformation on X.

We are now in a position to state our next result.

Theorem 1.5. *Assume that the action generated by the transformations $T_1, \ldots, T_d$ on the space X is aperiodic and that $T_1, \ldots, T_d$ commute with a mixing family $\{S_h\}$ on X. Then, if for some $1 \leq i \leq d$ it holds that*

$$(!C_i) \qquad \forall\, A, \alpha > 0 \quad \#\Omega_i^{(\alpha)}(\lambda) \geq A \cdot \lambda \quad \textit{for an unbounded set of } \lambda > 0,$$

the operators A_k have the "strong sweeping out" property, that is, for any $\varepsilon > 0$ there exists a set $B_\varepsilon \subset X$, with $\mu(B_\varepsilon) < \varepsilon$, such that for μ-a.e. $x \in X$ it holds that

$$\liminf_k A_k \chi_{B_\varepsilon}(x) = 0 \quad \textit{and} \quad \limsup_k A_k \chi_{B_\varepsilon}(x) = 1.$$

Remark 1.6. It is worth pointing out that, in view of Remark 1.4, any subgroup of an Abelian group acting ergodically on the space (X, μ) commutes with a mixing family, even if the subgroup itself does not act ergodically.

1.2. **The Continuous Case.** We will now state the continuous version of the results of the previous subsection.

Let (X, μ) be a non-atomic probability space and let $U_{\boldsymbol{t}}$ for $\boldsymbol{t} \in \mathbb{R}^d$ denote a measurable and measure-preserving d-dimensional flow on X (which is our standing assumption throughout this section). Given sequences $\{\boldsymbol{w}_k\} \subset \mathbb{R}^d$ and $\{\boldsymbol{s}_k\} \subset (0, +\infty)^d$, put

$$B_k := [w_{k1}, w_{k1} + s_{k1}) \times \cdots \times [w_{kd}, w_{kd} + s_{kd})$$

and let $\mathcal{B} := \{B_k\}$. To any box $B_k \in \mathcal{B}$ associate a continuous averaging operator R_k defined by

$$R_k f(x) := \frac{1}{s_{k1} \cdots s_{kd}} \int_{B_k} f\left(U_{\boldsymbol{t}} x\right) \, \mathrm{d}\boldsymbol{t}$$

for $f \in \mathrm{L}^1(X)$, and let $N_{\mathcal{B}}$ denote the maximal operator

$$N_{\mathcal{B}} f(x) := \sup_{B_k \in \mathcal{B}} \frac{1}{s_{k1} \cdots s_{kd}} \int_{B_k} \left| f\left(U_{\boldsymbol{t}} x\right) \right| \, \mathrm{d}\boldsymbol{t}.$$

For $i = 1, \ldots, d$ let

$$\Omega_i := \{(w_{ki}, s_{k,i}) : k \in \mathbb{N}\}$$

and for $\alpha > 0$ define

$$\Omega_i^{(\alpha)} := \{(x, y) \in \mathbb{R} \times (0, +\infty) : |x - t| \le \alpha(y - s) \text{ for some } (t, s) \in \Omega_i\}.$$

Finally, for any real $\lambda > 0$ put

$$\Omega_i^{(\alpha)}(\lambda) := \left\{x \in \mathbb{R} : (x, \lambda) \in \Omega_i^{(\alpha)}\right\}.$$

Then, in analogy to Theorem 1.1, we have the following.

Theorem 1.7. *Assume that for all* $i = 1, \dots, d$

$$(\tilde{C}_i) \qquad \exists\, A, \alpha > 0 : \quad \operatorname{Leb}\left(\Omega_i^{(\alpha)}(\lambda)\right) \le A\lambda \quad \textit{for all } \lambda > 0,$$

where Leb *denotes the Lebesgue measure on* $\mathbb{R}$. *Then the operator* $N_{\mathcal{B}}$ *is of strong type* (p, p) *for any* $p > 1$.

Once again, the analog of Corollary 1.3 holds.

Corollary 1.8. *Assume that the flow* $U_{\boldsymbol{t}}$ *acts ergodically on the space* (X, μ) *and that Condition* $(\tilde{C}_i)$ *is satisfied for all* $i = 1, \dots, d$. *Then for any* $f \in \mathrm{L}^p(X)$ *with* $p > 1$ *and* μ*-a.e.* $x \in X$ *we have that*

$$R_k f(x) \to \mu(f) \quad \textit{as } k \to \infty.$$

We now proceed to state the following converse to Corollary 1.8. Recall that a d-dimensional flow $U_{\boldsymbol{t}}$ on X is said to be *aperiodic* if there exists a set $N \subset X$, with $\mu(N) = 0$, such that $U_{\boldsymbol{t}} x \neq x$ for all $x \in X \setminus N$ and all $\boldsymbol{t} \in \mathbb{R}^d \setminus \{\boldsymbol{0}\}$.

Theorem 1.9. *Assume the flow* $U_{\boldsymbol{t}}$ *is aperiodic and that it commutes with a mixing family of transformations* $\{S_h\}$ *on* X. *Then, if for some* $1 \le i \le d$ *it holds that*

$$(!\tilde{C}_i) \qquad \forall\, A, \alpha > 0 \quad \operatorname{Leb}\left(\Omega_i^{(\alpha)}(\lambda)\right) \ge A \cdot \lambda \quad \textit{for an unbounded set of } \lambda > 0,$$

the operators R_k *have the "strong sweeping out" property (see Theorem 1.5).*

In the next subsection, we present an application of Theorem 1.9.

1.3. Failure of Submanifold Genericity for Essentially Bounded Functions.

Let (X, μ) be a probability space equipped with a measure preserving action of $\mathbb{R}^d$, which we denote by $\boldsymbol{a}.x$ for all $\boldsymbol{a} \in \mathbb{R}^d$ and $x \in X$. In [BFK25], the notion of submanifold genericity for a measure ν on X was introduced, in connection with certain problems in Diophantine approximation. We recall it here, for the convenience of the reader.

Let $M \subset \mathbb{R}^d$ be a compact m-dimensional $\mathscr{C}^1$ submanifold of $\mathbb{R}^d$ and let $\mathcal{F} \subset \mathrm{L}^1(X)$ be a collection of functions. Let vol_m denote the m-dimensional volume measure induced by the Euclidean metric on $\mathbb{R}^d$. We say that a measure ν on X (potentially equal to μ) is $(M, \mathcal{F})$-generic if for ν-a.e. $x \in X$ it holds that

$$\frac{1}{t^m \cdot \mathrm{vol}_m(M)} \int_{tM} f(\boldsymbol{a}.x)\, \mathrm{dvol}_m(\boldsymbol{a}) \to \mu(f) \quad \text{as } t \to \infty. \tag{1.2}$$

We also say that a measure ν is $(m, \mathcal{F})$-generic (for fixed $1 \le m \le d$) if for all compact m-dimensional $\mathscr{C}^1$ submanifolds M in $\mathbb{R}^d$ it holds that ν is $(M, \mathcal{F})$-generic.

When M is a d-dimensional bounded positive-measure set, $(M, \mathrm{L}^1(X))$-genericity is equivalent to ergodicity of the $\mathbb{R}^d$-action (see [Lin01, Theorem 1.2]). However, if M is a proper submanifold of $\mathbb{R}^d$, the convergence in (1.2) is more delicate and may not occur for all $f \in \mathrm{L}^1(X)$. For example, when $M = \mathbb{S}^{d-1}$ (i.e., the $(d-1)$-dimensional sphere in $\mathbb{R}^d$) and $d \geq 3$, Jones [Jon93, Theorem 2.1] showed that μ is $(M, \mathrm{L}^p(X))$-generic for any $p > d/(d-1)$ and that the bound on p is sharp (see [Jon93, Theorem 2.3] for a constructive counterexample). [Lac95] extended this result to the case $d = 2$. Such analytical arguments, however, do not apply to different manifolds M such as, e.g., the boundary of a hypercube.

If one restricts to very special families of functions $\mathcal{F}$, on the other hand, there are actions for which point-wise convergence in (1.2) occurs for essentially any sufficiently regular submanifold of $\mathbb{R}^d$. For example, let $X = G/\Gamma$, where G is a semisimple Lie group (with no factors of rank 1) and Γ is a lattice in G. Let μ denote the canonical left-invariant measure on G/Γ and let $A < G$ be a Cartan subgroup, whose Lie algebra will be denoted by $\mathfrak{a}$. Consider the action of $\mathfrak{a}$ on X given by $x \mapsto \exp(a)x$ (which is identifiable with an $\mathbb{R}^d$ action on X, if G is of rank d). Then [BFK25, Theorem 1.3] asserts that the measure μ on X is $(k, \mathscr{C}_c^\infty(X))$-generic for all $1 \leq k \leq d$. Here, $\mathscr{C}_c^\infty(X)$ stands for the space of smooth compactly supported functions on X. In addition, any invariant measure supported on a closed unipotent orbit in X is also $(k, \mathscr{C}_c^\infty(X))$-generic for all $k = 1, \ldots, d$ (see for instance [BFK25, Theorem 1.4]). More generally, any measure-preserving and exponentially mixing action of $\mathbb{R}^d$ on a probability space (X, μ) (under some minor assumptions) enjoys the same property (see [BFK25, Theorem 2.2]).

In view of the above discussion, it is natural to ask whether there exist ergodic actions of $\mathbb{R}^d$ on some probability space (X, μ) such that the measure μ is $(k, \mathcal{F})$-generic for some much larger class of functions $\mathcal{F}$, e.g., $L^\infty(X)$. As a consequence of Theorem 1.9, we are able to show the following.

Theorem 1.10. *Let (X, μ) be a probability space, equipped with an ergodic and aperiodic measure-preserving action of $\mathbb{R}^d$. Let M be a compact m-dimensional $\mathscr{C}^1$ submanifold of $\mathbb{R}^d$ such that $M \cap \pi$ is a non-empty open set in M for a given m-dimensional affine subspace π of $\mathbb{R}^d$, which does not contain the origin. Then the measure μ is not $(M, \mathrm{L}^\infty(X))$-generic. In particular, there exist no ergodic and aperiodic measure-preserving actions of $\mathbb{R}^d$ on X that are $(m, \mathrm{L}^\infty(X))$-generic for any $m = 1, \ldots, d-1$.*

Remark 1.11. The left-multiplication action of a Cartan subgroup A of a Lie group G on the quotient $X = G/\Gamma$ is always ergodic and aperiodic (see explanation below). Thus, Theorem 1.10 implies that the G-invariant measure μ on X is not $(m, \mathrm{L}^\infty(X))$-generic for any $m = 1, \ldots, d-1$. In other words, a point-wise ergodic theorem for dilates of submanifolds M such as those in Theorem 1.10 cannot hold. However, it is interesting to observe that for the same submanifolds the convergence in (1.2) still occurs in the L^p-norm for any $p \geq 1$, i.e., a mean ergodic theorem holds. This follows easily from the fact that the measure μ is $(m, \mathscr{C}_c^\infty(X))$-generic, by a density argument.

Let us briefly explain why the left-multiplication action of a Cartan subgroup A on G is aperiodic. If $ag\Gamma = g\Gamma$ for some $a \in A$ and $g \in G$, then there exists $\gamma \in \Gamma$ such that $g^{-1}ag = \gamma$. Put $S(A, \gamma) := \{g \in G : g\gamma g^{-1} \in A\}$. Hence, any periodic point $g\Gamma$ for the group A belongs to the set

$$\bigcup_{\gamma\in\Gamma} S(A,\gamma)\Gamma.$$

This is a countable union of algebraic varieties and therefore has measure 0.

We conclude with the following observation. If M is a polynomial curve in $\mathbb{R}^d$ not entirely contained in a proper affine subspace, then, according to [BF09, Theorem 2] (with a slight modification to account for the density induced by the measure vol_m), the convergence in (1.2) occurs in norm (more precisely, in the L^2 norm). As suggested in [BF09], it would be natural to study point-wise convergence as in (1.2) for such curves. However, polynomial curves fall outside the scope of Theorem 1.10, since any smooth submanifold M of $\mathbb{R}^d$ not entirely contained in, but intersecting an affine subspace in an open set, cannot be analytically parametrized.

1.4. **Acknowledgments.** The authors are indebted to Alexander Gorodnik for suggesting to read [BJR90], to Dmitry Kleinbock for pointing out Remark 1.11, and to Amos Nevo for many fruitful conversations which eventually lead to this paper.

2. Proof of Theorems 1.1 and 1.7 and of Corollaries 1.3 and 1.8

2.1. **Proof of Theorem 1.1.** First, we observe that, if a box B_k contains both integer vectors whose i-th component is negative and integer vectors whose i-th component is positive for some k and i, then it must be $n_{ki} < 0$ and $l_{ki} > |n_{ki}|$. In this case, we may write

$$B_k = B_k^{i-} \cup B_k^{i+},$$

where

$$B_k^{i-} := B_k \cap \{x_i < 0\} \quad \text{and} \quad B_k^{i+} = B_k \cap \{x_i \geq 0\}.$$

If $\mathcal{B}'$ is the collection of boxes where B_k is replaced by B_k^{i-} and B_k^{i+}, it is obvious that $M_{\mathcal{B}} f \leq M_{\mathcal{B}'} f$ for any $f \in \mathrm{L}^p(X)$. Moreover, for any k and i such that $l_{ki} > |n_{ki}|$ and any $\lambda \in \mathbb{N}$ we have that

$$\#\{x \in \mathbb{Z} : |x - n_{ki}| \leq \alpha(\lambda - l_{ki})\} \leq (1+\alpha)\lambda,$$

so that the boxes B_k that are broken in two or more parts do not contribute to the validity of Condition (C_i). Hence, by working in each orthant separately and by replacing T_i with T_i^{-1} if necessary, we may always assume that $n_{ki} \geq 0$ for all k and i.

Take $f \in \mathrm{L}^p(X)$ for fixed $p > 1$ and observe that

$$\begin{aligned} M_{\mathcal{B}} f(x) &= \sup_{B_k \in \mathcal{B}} \frac{1}{l_{k1} \cdots l_{kd}} \sum_{\boldsymbol{j} \in B_k \cap \mathbb{Z}^d} \left| f\left(T_1^{j_1} \cdots T_d^{j_d} x\right)\right| \\ &\leq \sup_{(n_{k1}, l_{k1}) \in \Omega_1} \cdots \sup_{(n_{kd}, l_{kd}) \in \Omega_d} \frac{1}{l_{k1} \cdots l_{kd}} \sum_{j_1=0}^{l_{k1}-1} \cdots \sum_{j_d=0}^{l_{kd}-1} \left| f\left(T_1^{n_{k1}+j_1} \cdots T_d^{n_{kd}+j_d} x\right)\right| \\ &\leq \sup_{(n_{k1}, l_{k1}) \in \Omega_1} \frac{1}{l_{k1}} \sum_{j_1=0}^{l_{k1}-1} \left| \cdots \sup_{(n_{kd}, l_{kd}) \in \Omega_d} \frac{1}{l_{kd}} \sum_{j_d=0}^{l_{kd}-1} \left| f\left(T_1^{n_{k1}+j_1} \cdots T_d^{n_{kd}+j_d} x\right)\right| \right| \\ &= M_{\Omega_1} \circ \cdots \circ M_{\Omega_d} f(x), \end{aligned}$$

where

$$M_{\Omega_i} f(x) := \sup_{(n_{ki}, l_{ki}) \in \Omega_i} \frac{1}{l_{ki}} \sum_{j=0}^{l_{ki}-1} \left| f\left(T^{n_{ki}+j} x\right)\right|$$

for $i = 1, \ldots, d$. Since the functions $M_{\mathcal{B}} f$ and $M_{\Omega_1} \circ \cdots \circ M_{\Omega_d} f$ are positive, we deduce that

$$\|M_{\mathcal{B}} f\|_p \leq \|M_{\Omega_1} \circ \cdots \circ M_{\Omega_d} f\|_p.$$

By [BJR90, Theorem 1], each operator M_{Ω_i} is of strong type (p, p). Then $M_{\mathcal{B}}$ is also of strong type (p, p).

2.2. **Proof of Theorem 1.7.** The proof in the case $d > 1$ relies once again on the case $d = 1$ and is analogous to that of Theorem 1.1. We therefore leave it to the reader.

In the case $d = 1$ we modify the proof of [BJR90, Theorem 1]. This requires some additional work. In what follows, for any $p \geq 1$ and any measure space Y we denote by $\mathscr{L}^p(Y)$ the set of p-integrable functions from Y to $\mathbb{R}$.

For any $f \in \mathscr{L}^1(X)$, $T > 0$, $x \in X$, and $t > 0$ put

$$f_{T,x}(t) := f(U_t x) \cdot \chi_{[-T,T]}(t).$$

Then $f_{T,x} \in \mathscr{L}^1(\mathbb{R})$. Let $f_{T,x}^*$ denote the one-sided Hardy-Littlewood maximal function associated to $f_{T,x}$, that is

$$f_{T,x}^*(\xi) := \sup_{r \neq 0} \frac{1}{|r|} \int_0^r |f(\xi + \tau)| \, \mathrm{d}\tau,$$

where the bounds of the integral are to be inverted if $r < 0$. By the Hardy-Littlewood maximal inequality [Tao11, Lemma 1.6.16] we have that for any $\lambda > 0$

$$\operatorname{Leb}\left(\xi : f_{T,x}^*(\xi) > \lambda\right) \leq \frac{\|f_{T,x}\|_1}{\lambda} = \frac{1}{\lambda} \int_{-T}^{T} |f(U_t x)| \, \mathrm{d}t. \tag{2.1}$$

Let us now fix $\lambda > 0$ and a function $f \in \mathscr{L}^1(X)$. Then

$$\{x : N_{\mathcal{B}} f(x) > \lambda\} = \bigcup_k \{x : R_k |f|(x) > \lambda\}.$$

Let $\varepsilon > 0$ and let $K_\varepsilon \geq 1$ such that

$$\mu\left(\bigcup_{k\leq K_\varepsilon}\{x : R_k|f|(x) > \lambda\}\right) \geq \mu\left(x : N_{\mathcal{B}}f(x) > \lambda\right) - \varepsilon.$$

Choose N_ε so large that

$$\frac{\max_{k\leq K_\varepsilon}\{w_k + s_k\}}{N_\varepsilon} \leq \varepsilon$$

and set

$$T_\varepsilon := N_\varepsilon + \max_{k\leq K_\varepsilon}\{w_k + s_k\}.$$

Further, define

$$Y_\varepsilon := \bigcup_{k\leq K_\varepsilon}\{(x,t) : R_k|f|(U_t x) > \lambda \text{ and } |t| < N_\varepsilon\}$$

and observe that, by the invariance of μ, we have that

$$\mu\otimes\mathrm{Leb}(Y_\varepsilon) = \int_{-N_\varepsilon}^{N_\varepsilon}\int_X \chi_{\{\sup_{k\leq K_\varepsilon} R_k|f|>\lambda\}}\circ U_t \,\mathrm{d}\mu\mathrm{d}t \tag{2.2}$$

$$2N_\varepsilon\cdot\mu\left(\bigcup_{k\leq K_\varepsilon}\{x : R_k|f|(x) > \lambda\}\right) \geq 2N_\varepsilon\cdot\mu\left(x : N_{\mathcal{B}}f(x) > \lambda\right) - 2N_\varepsilon\cdot\varepsilon.$$

Finally, for any $x \in X$ let

$$Y_\varepsilon(x) := \{\xi : (x,\xi) \in Y_\varepsilon\}.$$

Then we have the following.

Lemma 2.1. *For any $(x,t) \in Y_\varepsilon$ there exists a pair $(w_k, s_k) \in \Omega$ with $k \leq K_\varepsilon$ such that*

$$[t + w_k, t + w_k + s_k) \subset \left\{\xi : f^*_{T_\varepsilon,x}(\xi) > \lambda\right\}. \tag{2.3}$$

Proof. Fix $(x,t) \in Y_\varepsilon$. Then there exists $(w_k, s_k) \in \Omega$ with $k \leq K_\varepsilon$ such that

$$\frac{1}{s_k}\int_0^{s_k}|f(U_{t+w_k+\xi}x)|\,\mathrm{d}\xi = \frac{1}{s_k}\int_0^{s_k}|f_{T_\varepsilon,x}(t + w_k + \xi)|\,\mathrm{d}\xi > \lambda. \tag{2.4}$$

Assume by contradiction that (2.3) fails. Then for some $\xi_0 \in [t + w_k, t + w_k + s_k)$ we have that $f^*_{T_\varepsilon,x}(\xi_0) \leq \lambda$. Hence,

$$\int_{-(\xi_0-(t+w_k))}^{0}|f_{T_\varepsilon,x}(\xi_0 + \tau)|\,\mathrm{d}\tau \leq \lambda\cdot(\xi_0 - (t + w_k)) \quad \text{and}$$

$$\int_0^{(t+w_k+s_k)-\xi_0}|f_{T_\varepsilon,x}(\xi_0 + \tau)|\,\mathrm{d}\tau \leq \lambda\cdot((t + w_k + s_k) - \xi_0)$$

and

$$\int_{t+w_k}^{t+w_k+s_k}|f_{T_\varepsilon,x}(\tau)|\,\mathrm{d}\tau \leq \lambda\cdot s_k$$

– a contradiction to (2.4). □

Lemma 2.1 further implies the following.

Lemma 2.2. *For any $x \in X$ it holds that*

$$\mathrm{Leb}\left(Y_\varepsilon(x)\right) \leq AC_\alpha \cdot \mathrm{Leb}\left(\xi : f^*_{T_\varepsilon,x}(\xi) > \lambda\right),$$

where A and α are the constants in Condition $(\tilde{C}_i)$ (for $i = 1$) and C_α is an absolute constant only depending on α.

Proof. Fix $\delta > 0$ and $x \in X$ such that $Y_\varepsilon(x) \neq \emptyset$. Note that, by construction $Y_\varepsilon(x) \subset (-N_\varepsilon, N_\varepsilon)$. Let $O_x \subset (-N_\varepsilon, N_\varepsilon)$ be an open set with the following two properties:

$$O_x \supseteq Y_\varepsilon(x) \quad \text{and} \quad \mu\left(O_x \setminus Y_\varepsilon(x)\right) \leq \delta. \tag{2.5}$$

For each $\xi \in O_x$ put $r_\xi := \sup\{r > 0 : (\xi, \xi + r) \subset O_x\}$ and consider the covering of O_x given by

$$\bigcup_{\xi \in O_x} (\xi, \xi + r_\xi).$$

By the Vitali Covering Lemma (see [Coh13, Theorem 6.2.1]), there exists a countable sub-collection of *disjoint* open intervals $(\xi_i, \xi_i + r_i)$ (where $r_i = r_{\xi_i}$ for brevity) such that

$$O_x \subseteq \bigcup_i (\xi_i - 2r_i, \xi_i + 3r_i).$$

Fix $t \in Y_\varepsilon(x)$. By Lemma 2.1, we may find $k \leq K_\varepsilon$ such that

$$[t + w_k, t + w_k + s_k) \subset O_x.$$

Now, if there exists i such that

$$t + w_k + \frac{s_k}{3} < \xi_i < t + w_k + \frac{2s_k}{3},$$

by the definition of r_ξ it must be $r_i \geq s_k/3$. On the other hand, if for all i if holds that

$$\xi_i < t + w_k + \frac{s_k}{3} \quad \text{or} \quad \xi_i \geq t + w_k + \frac{2s_k}{3},$$

then, there must be i such that $t + w_k + s_k/2 \in (\xi_i - 2r_i, \xi_i + 3r_i)$. Hence, $3r_i \geq s_k/6$. In any case, we have that there exists i such that $r_i \geq s_k/18$ and $(\xi_i - 2r_i, \xi_i + 3r_i) \cap [t + w_k, t + w_k + s_k) \neq \emptyset$. For this i we therefore have that

$$(\xi_i - 21r_i, \xi_i + 21r_i) \supseteq [t + w_k, t + w_k + s_k).$$

Now, note that

$$|(\xi_i - 21r_i) - t - w_k| \leq |(\xi_i - 21r_i) - (t + w_k)| \leq 42r_i \leq \alpha(C_\alpha r_i - s_k)$$

for $C_\alpha = 42\alpha^{-1} + 18$. By definition of $\Omega^{(\alpha)}$, we conclude that

$$\xi_i - 21r_i - t \in \Omega^{(\alpha)}(C_\alpha r_i).$$

Hence, we have that

$$Y_\varepsilon(x) \subset \bigcup_i (\xi_i - 21r_i) - \Omega^{(\alpha)}(C_\alpha r_i).$$

This shows that

$$\begin{aligned}\mathrm{Leb}\left(Y_\varepsilon(x)\right) &\leq \sum_i \mathrm{Leb}\left(\Omega^{(\alpha)}(C_\alpha r_i)\right) \leq \sum_i AC_\alpha r_i \\ &\leq AC_\alpha \cdot \mathrm{Leb}(O_x) \leq AC_\alpha \cdot \left(\mathrm{Leb}\left(\xi : f^*_{T,x}(\xi) > \lambda\right) + \delta\right).\end{aligned}$$

By letting $\delta \to 0$, we conclude. □

On combining Lemma 2.2 and (2.1), we find that for any $x \in X$ it holds that

$$\mathrm{Leb}\left(Y_\varepsilon(x)\right) \leq AC_\alpha \cdot \mathrm{Leb}\left(\xi : f^*_{T_\varepsilon,x}(\xi) > \lambda\right) \leq \frac{AC_\alpha}{\lambda} \cdot \int_{-T_\varepsilon}^{T_\varepsilon} |f(U_t x)| \, \mathrm{d}t. \tag{2.6}$$

We now use transference. From (2.2) and (2.6), it follows that

$$\begin{aligned}&2N_\varepsilon \cdot \mu(x : N_{\mathcal{B}} f(x) > \lambda) - 2N_\varepsilon \cdot \varepsilon \\ &\quad \leq \mu \otimes \mathrm{Leb}(Y_\varepsilon) = \int_X \mathrm{Leb}(Y_\varepsilon(x)) \mathrm{d}\mu(x) \\ &\quad \leq \frac{AC_\alpha}{\lambda} \int_X \int_{-T_\varepsilon}^{T_\varepsilon} |f(U_t x)| \, \mathrm{d}t \mathrm{d}\mu \leq 2T_\varepsilon \cdot \frac{AC_\alpha}{\lambda} \|f\|_1.\end{aligned}$$

Dividing both sides by $2N_\varepsilon$ (recall that $1 \leq T_\varepsilon / N_\varepsilon \leq 1 + \varepsilon$) gives

$$\mu\left(x : N_{\mathcal{B}} f(x) > \lambda\right) \leq \frac{AC_\alpha}{\lambda} \|f\|_1 + \left(1 + \frac{AC_\alpha}{\lambda} \|f\|_1\right) \varepsilon.$$

On letting $\varepsilon \to 0$, we conclude that $N_{\mathcal{B}}$ is of weak type $(1,1)$. Since $N_{\mathcal{B}}$ is bounded from L^∞ to L^∞, by the Marcinkiewicz Interpolation Theorem [Gra14, Theorem 1.3.2], $N_{\mathcal{B}}$ is of strong type (p,p) for $p > 1$.

2.3. **Proof of Corollaries 1.3 and 1.8.** The strategy to prove an ergodic theorem given a maximal inequality is standard (see for example [EW11, Section 2.6.5]). We therefore only give a sketch of proof for the discrete case. Let us start with an observation.

Remark 2.3. If for some $1 \leq i \leq d$ Condition (C_i) is satisfied, then it must be $l_{ki} \to \infty$. In fact, if there is a sequence k_r such that $l_{k_r i} \leq C$ for a given constant $C \geq 1$, then for all integer $\lambda \geq C$ and all r we have that

$$(n_{k_r i}, \lambda) \in \Omega_i^{(\alpha)}(\lambda),$$

whence $\#\Omega_i^{(\alpha)}(\lambda) = \infty$.

The first step in the proof is the following lemma.

Lemma 2.4. *Let $f \in \mathrm{L}^2(X)$. Then there exists a function $f' \in \mathrm{L}^2(X)$ that is invariant under the transformations $T_1, \ldots, T_d$, such that*

$$\|A_k f - f'\|_2 \to 0.$$

Sketch of Proof. For $i = 1, \ldots, d$ define $U_{T_i} f := f \circ T_i$ and let

$$I := \{g \in \mathrm{L}^2(X) : U_{T_i} g = g \text{ for } i = 1, \ldots, d\}.$$

Let $B := I^{\perp}$. Then

$$B = \overline{B_1 \oplus \cdots \oplus B_d},$$

where

$$B_i := \{U_{T_i} g - g : g \in \mathrm{L}^2(X)\}.$$

Since $l_{ki} \to \infty$ for $i = 1, \ldots, d$ (see Remark 2.3) it is clear that for any $h \in B_i$ we have that $\|A_k h\|_2 \to 0$ as $k \to \infty$. Then the conclusion follows as in [EW11, Theorem 2.21]. □

Let $f \in \mathrm{L}^\infty(X)$. By Lemma 2.4, we know that $A_k f$ converges to an invariant function f' in $\mathrm{L}^2(X)$. Now, for every measurable $B \subset X$ we have that

$$\langle A_k f, \chi_B \rangle \leq \|f\|_\infty \cdot \mu(B).$$

Hence, the same must be true if $A_k f$ is replaced by f'. This shows that $f' \in \mathrm{L}^\infty(X)$.

We now need the following.

Lemma 2.5. *For any $k \geq h$ and any $f \in \mathrm{L}^\infty(X)$ we have that*

$$A_k \circ A_h f = A_k f + O_{\boldsymbol{n}_h, \boldsymbol{l}_h}\left((l_{k1} \cdots l_{kd})^{-1} \|f\|_\infty\right).$$

Proof. We have that

$$\begin{aligned} A_k \circ A_h f &= \frac{1}{l_{k1} \cdots l_{kd}} \sum_{j_1'=0}^{l_{k1}-1} \cdots \sum_{j_d'=0}^{l_{kd}-1} \\ &\qquad \frac{1}{l_{h1} \cdots l_{hd}} \sum_{j_1=0}^{l_{h1}-1} \cdots \sum_{j_d=0}^{l_{hd}-1} f\left(T_1^{n_{h1}+n_{k1}+j_1+j_1'} \cdots T_d^{n_{hd}+n_{kd}+j_d+j_d'} x\right) \\ &= \frac{1}{l_{h1} \cdots l_{hd}} \sum_{j_1=0}^{l_{h1}-1} \cdots \sum_{j_d=0}^{l_{hd}-1} \\ &\qquad \frac{1}{l_{k1} \cdots l_{kd}} \sum_{j_1'=n_{h1}+j_1}^{l_{k1}-1+n_{h1}+j_1} \cdots \sum_{j_d'=n_{hd}+j_d}^{l_{kd}-1+n_{hd}+j_d} f\left(T_1^{n_{k1}+j_1'} \cdots T_d^{n_{kd}+j_d'} x\right) \\ &= A_k f + O_{\boldsymbol{n}_h, \boldsymbol{l}_h}\left((l_{k1} \cdots l_{kd})^{-1} \|f\|_\infty\right), \end{aligned}$$

as desired. □

Let us show point-wise convergence in $\mathrm{L}^\infty(X)$. Let $f \in \mathrm{L}^\infty(X)$ and let f' be the function found in Lemma 2.4. Note that for any k it holds that $A_k f' = f'$, since f' is invariant. Fix $\varepsilon, \delta > 0$ and pick h so large that $\|A_h f - f'\|_2 \leq \delta$. Then, by Lemma 2.5, Remark 2.3, and Theorem 1.7, we have that

$$\mu\left(x : \limsup_k |A_k f - f'| > \varepsilon\right) = \mu\left(x : \limsup_k |A_k \circ A_h f - A_k f'| > \varepsilon\right)$$
$$\leq \mu\left(x : M_{\mathcal{B}}(A_h f - f') > \varepsilon\right) \ll \|A_h f - f'\|_2 \leq \delta.$$

This gives point-wise convergence. Finally, since $\mathrm{L}^\infty(X)$ is dense in $\mathrm{L}^p(X)$ for any $p > 1$, we may use Theorem 1.7 and a density argument to conclude.

3. Proof of Theorems 1.5 and 1.9

In this section we show that, provided Condition $(!C_i)$ (respectively $(!\tilde{C}_i)$) holds for some $1 \leq i \leq d$, the operators A_k (respectively R_k) enjoy the "strong sweeping out" property. This is a direct application of [BJR90, Theorem 3], which we recall below for the convenience of the reader.

Theorem 3.1. [BJR90, Theorem 3] *Let (X, Σ, μ) be a probability space and let $\{T_k\}$ be a sequence of linear operators on $\mathrm{L}^1(X)$ satisfying the following properties:*

- $T_k \geq 0$;
- $T_k 1 = 1$;
- *all T_k commute with a mixing family of measure-preserving transformations $\{S_h\}$ on X.*

For $n \in \mathbb{N}$ let $M_n f := \sup_{k \geqslant n} |T_k f|$ and assume that for each $\varepsilon > 0$ and $n \in \mathbb{N}$ there exists a sequence of sets $\{H_p\}$ in X such that

$$(\star) \qquad \sup_p \frac{\mu(M_n \chi_{H_p} > 1 - \varepsilon)}{\mu(H_p)} = \infty.$$

Then the sequence T_k has the "strong sweeping out" property (see Theorem 1.5).

When $T_k = A_k$ (or $T_k = R_k$), the first two assumptions in Theorem 3.1 are trivially true, while the third one follows by our hypothesis on the transformations $T_1, \ldots, T_d$ (or the flow $U_{\boldsymbol{t}}$). Hence, it is enough for us to verify that $(\star)$ holds. Note also that, when f is the characteristic function of a set in X, $M_n f$ coincides with $M_{\mathcal{B}'} f$ (or $N_{\mathcal{B}'} f$), where $\mathcal{B}' = \{B_k \in \mathcal{B} : k \geq n\}$.

3.1. **Proof of Theorem 1.5.** Take $\alpha = 1$ and assume, without loss of generality, that $(!C_i)$ holds for $i = 1$. Fix $p \in \mathbb{N}$ and choose an integer λ_p so that $\#\Omega_1^{(1)}(\lambda_p) \geq p \cdot (4\lambda_p + 1)$. Recall that, by definition of $\Omega_1^{(1)}(\lambda_p)$, for any $z \in \Omega_1^{(1)}(\lambda_p)$ there exists $(n_{k1}, l_{k1}) \in \Omega_1$ such that

$$(3.1) \qquad |z - n_{k1}| \leq \lambda_p - l_{k1} \leq \lambda_p.$$

In view of this, we may write

$$\Omega_1^{(1)}(\lambda_p) = \bigcup_k C_k^{(1)}(\lambda_p),$$

where

$$C_k^{(1)}(\lambda_p) := \{x \in \mathbb{Z} : |x - n_{k1}| \le (\lambda_p - l_{k1})\}.$$

Choose $K_p \ge 1$ large enough so that

$$\# \left(\bigcup_{k \le K_p} C_k^{(1)}(\lambda_p) \right) \ge p \cdot (4\lambda_p + 1)$$

and put

$$\Delta := \bigcup_{k \le K_p} C_k^{(1)}(\lambda_p).$$

Then for each $z \in \Delta$ there exists $k \le K_p$ such that (3.1) holds. Note that it must be $\lambda_p \ge l_{k1}$, so that for all $j = 0, \ldots, l_{k1} - 1$ we have that

$$|-z + n_{k1} + j| \le \lambda_p + \lambda_p = 2\lambda_p. \tag{3.2}$$

Now, let

$$N_1 := 2\lambda_p + \sup \Delta$$

and

$$N_j := \sup_{k \le K_p} |n_{kj} + l_{kj}|$$

for $j = 2, \ldots, d$. Form a Rokhlin tower as in [KW72, Theorem 1] with parameters $N_1, 3N_2, \ldots, 3N_d$. Then

$$\mu \left(\bigcup_{j_1=0}^{N_1-1} \bigcup_{j_2=0}^{3N_2-1} \cdots \bigcup_{j_d=0}^{3N_d-1} T_1^{j_1} T_2^{j_2} \ldots T_d^{j_d}(B) \right) > 1 - \delta$$

for some positive-measure set $B \subset X$ and some small $\delta > 0$, where the union is disjoint. Define

$$H_p := \bigcup_{j_1=N_1-4\lambda_p-1}^{N_1-1} \bigcup_{j_2=0}^{3N_2-1} \cdots \bigcup_{j_d=0}^{3N_d-1} T_1^{j_1} T_2^{j_2} \ldots T_d^{j_d}(B).$$

Fix $z \in \Delta$ and

$$x \in F_p := \bigcup_{j_2=N_2}^{2N_2-1} \cdots \bigcup_{j_d=N_d}^{2N_d-1} T_1^{N_1-2\lambda_p-1} T_2^{j_2} \ldots T_d^{j_d}(B).$$

Then, by (3.2), for some $k \le K_p$ it holds that

$$N_1 - 4\lambda_p - 1 \le (N_1 - 2\lambda_p - 1) - z + n_{k1} + j \le N_1 - 1$$

for all $j = 0, \ldots, l_{k1} - 1$. Since $|n_{kj} + l_{kj}| \le N_j$ for $j = 2, \ldots, d$, we conclude that

$$M_{\mathcal{B}} \chi_{H_p}(T_1^{-z} x) \ge \frac{1}{l_{k1} \cdots l_{kd}} \sum_{j_1=0}^{l_{k1}-1} \cdots \sum_{j_d=0}^{l_{kd}-1} \chi_{H_p} \left(T_1^{-z+n_{k1}+j_1} T_2^{n_{k2}+j_2} \cdots T_d^{n_{kd}+j_d} x \right) = 1.$$

This implies that

$$\{M_{\mathcal{B}} \chi_{H_p} > 1 - \varepsilon\} \supset \bigcup_{z \in \Delta} T_1^{-z} F_p,$$

whence

$$\mu \left(M_{\mathcal{B}} \chi_{H_p} > 1 - \varepsilon \right) \ge \#\Delta \cdot N_2 \ldots N_d \cdot \mu(B).$$

Moreover, it is clear that

$$\mu(H_p) = 3^{d-1}(4\lambda_p + 1) \cdot N_2 \ldots N_d \cdot \mu(B).$$

Combining these two observations, we find that

$$\frac{\mu\left(M_{\mathcal{B}}\chi_{H_p} > 1-\varepsilon\right)}{\mu(H_p)} \geq \frac{\#\Delta \cdot N_2 \ldots N_d \cdot \mu(B)}{3^{d-1}(4\lambda_p+1) \cdot N_2 \ldots N_d \cdot \mu(B)} \geq \frac{p}{3^{d-1}},$$

showing that $(\star)$ holds for the operator $M_{\mathcal{B}}$. To conclude, observe that if $\mathcal{B}' := \{B_k \in \mathcal{B} : k \geq n\}$, then $(!C_i)$ holds for the collection $\mathcal{B}'$ and the above argument also applies to the operator $M_{\mathcal{B}'}$.

3.2. **Proof of Theorem 1.9.** The continuous case follows from a similar argument, based on a Rokhlin tower construction for aperiodic flows proved in [Lin75], which we recall below.

Theorem 3.2. [Lin75, Theorem 1] *Let $U_{\boldsymbol{t}} = U_{t_1,\ldots,t_d}$ be a d-dimensional measure-preserving aperiodic flow on (X,μ). Let $L_1,\ldots,L_d,\delta > 0$ and let $Q = Q_L := [0,L_1) \times \ldots \times [0,L_d)$. Then there exists a set $B \subset X$ with the following properties:*

- *the sets $U_{\boldsymbol{t}}B$ for $\boldsymbol{t} \in Q$ are pairwise disjoint;*
- *the set $Y := \bigcup_{\boldsymbol{t}\in Q} U_{\boldsymbol{t}}B$ is measurable and $\mu(Y) > 1-\delta$;*
- *there exists a measure ν_B defined on B such that the map $\varphi : B \times Q \to X$ given by $\varphi(x,\boldsymbol{t}) := U_{\boldsymbol{t}}x$ is bijective and both φ and its inverse are measurable and measure-preserving with respect to the measures $\nu_B \otimes \mathrm{Leb}$ on $B \times Q$ and μ on X.*

In particular, the last part of Theorem 3.2 implies that for any $f \in \mathrm{L}^1(X)$ we have that

$$\int_Y f \,\mathrm{d}\mu = \int_B \int_{\boldsymbol{t}\in Q} f(U_{\boldsymbol{t}}x)\,\mathrm{d}\boldsymbol{t}\mathrm{d}\nu_B(x). \tag{3.3}$$

As in the proof of 1.5, take $\alpha = 1$ and suppose that $(!\tilde{C}_i)$ holds for $i = 1$. Fix $p \in \mathbb{N}$ and choose a real number $\lambda_p > 0$ so that $\mathrm{Leb}\left(\Omega_1^{(1)}(\lambda_p)\right) \geq p \cdot 4\lambda_p$. Then for each $z \in \Omega_1^{(1)}(\lambda_p)$ we have that

$$|z - w_{k1}| \leq \lambda_p - s_{k1} \tag{3.4}$$

for some $(w_{k1}, s_{k1}) \in \Omega_1$. Once again, this implies that

$$\Omega_1^{(1)}(\lambda_p) = \bigcup_k C_k^{(1)}(\lambda_p),$$

where

$$C_k^{(1)}(\lambda_p) := \{x \in \mathbb{R} : |x - w_{k1}| \leq (\lambda_p - s_{k1})\}.$$

Choose $K_p \geq 1$ large enough so that

$$\mathrm{Leb}\left(\bigcup_{k \leq K_p} C_k^{(1)}(\lambda_p)\right) \geq p \cdot 4\lambda_p$$

and put

$$\Delta := \bigcup_{k \le K_p} C_k^{(1)}(\lambda_p).$$

Then for each $z \in \Delta$ there exists $k \le K_p$ such that (3.4) holds. Note that it must be $\lambda_p \ge s_{k1}$ so that for all $0 \le t_1 < s_{k1}$ we have that

$$|-z + w_{k1} + t_1| \le 2\lambda_p. \tag{3.5}$$

Let

$$L_1 := 2\lambda_p + \sup \Delta$$

and

$$L_j := \sup_{k \le K_p} |w_{kj} + s_{kj}|$$

for $j = 2, \dots, d$. Form a tower as in Theorem 3.2 with parameters $L_1, 3L_2, \dots, 3L_d$. Define

$$Q_p := [L_1 - 4\lambda_p, L_1) \times [0, 3L_2) \times \cdots \times [0, 3L_d)$$

and $H_p := \{U_{\boldsymbol{t}} x : x \in B \text{ and } \boldsymbol{t} \in Q_p\}$, so that

$$\chi_{H_p}(y) = \begin{cases} \chi_{Q_p}(\boldsymbol{t}) & \text{if } y = U_{\boldsymbol{t}} x \text{ with } x \in B \text{ and } \boldsymbol{t} \in Q. \\ 0 & \text{otherwise.} \end{cases}$$

By (3.3), we deduce that

$$\begin{aligned} \mu(H_p) &= \int_Y \chi_{H_p}(y)\, \mathrm{d}\mu = \int_B \int_{\boldsymbol{t} \in Q} \chi_{H_p}(U_{\boldsymbol{t}} x)\, \mathrm{d}\boldsymbol{t} \mathrm{d}\nu_B(x) \\ &= \int_B \int_{\boldsymbol{t} \in Q} \chi_{Q_p}(\boldsymbol{t})\, \mathrm{d}\boldsymbol{t} \mathrm{d}\nu_B(x) = \frac{\mathrm{Leb}(Q_p)}{L_1 \cdot 3L_2 \cdots 3L_d} \cdot \mu(Y) = \frac{4\lambda_p}{L_1} \cdot \mu(Y). \end{aligned} \tag{3.6}$$

Now, let

$$F_p := \{U_{\boldsymbol{t}} x : x \in B \text{ and } \boldsymbol{t} \in \{L_1 - 2\lambda_p\} \times [L_2, 2L_2) \times \cdots \times [L_d, 2L_d)\}.$$

Fix $z \in \Delta$ and $y \in F_p$, so that $y = U_{\boldsymbol{a}} x$ with $x \in B$, $a_1 = L_1 - 2\lambda_p$, and $a_j \in [0, L_j)$ for $j = 2, \dots, d$. By (3.5), there exists $k \le K_p$ such that

$$|a_1 - z + w_{k1} + t_1| \in [L_1 - 4\lambda_p, L_1) \text{ for all } t_1 \in [0, s_{k1}).$$

Since $|w_{kj} + s_{kj}| \le L_j$ for $j = 2, \dots, d$, we conclude that

$$\begin{aligned} N_{\mathcal{B}} \chi_{H_p}(U_{-z,0,\dots,0} y) &= N_{\mathcal{B}} \chi_{H_p}(U_{-z,0,\dots,0} U_{\boldsymbol{a}} x) \\ &\ge \frac{1}{s_{k1} \dots s_{kd}} \int_{\boldsymbol{t} \in B_k} \chi_{H_p}(U_{-z+a_1+t_1, a_2+t_2, \dots, a_d+t_d} x)\, \mathrm{d}\boldsymbol{t} = 1. \end{aligned}$$

This implies that

$$\{N_{\mathcal{B}} \chi_{H_p} > 1 - \varepsilon\} \supset \bigcup_{z \in \Delta} U_{-z,0,\dots,0} F_p.$$

Thus, precisely as in (3.6), we have that

$$\mu\left(\bigcup_{z \in \Delta} U_{-z,0,\dots,0} F_p\right) = \frac{\mathrm{Leb}\,(\Delta) \cdot L_2 \cdots L_d}{L_1 \cdot 3L_2 \cdots 3L_d} \cdot \mu(Y) \ge \frac{p \cdot 4\lambda_p}{3^{d-1} \cdot L_1} \cdot \mu(Y),$$

whence

$$\frac{\mu\left(N_{\mathcal{B}}\chi_{H_p} > 1-\varepsilon\right)}{\mu(H_p)} \geq \frac{p\cdot 4\lambda_p}{3^{d-1}\cdot L_1}\cdot\frac{L_1}{4\lambda_p} = \frac{p}{3^{d-1}},$$

showing $(\star)$ for the operator $N_{\mathcal{B}}$. To conclude, note that if $\mathcal{B}' := \{B_k \in \mathcal{B} : k \geq n\}$, then $(!\tilde{C}_i)$ holds for the collection $\mathcal{B}'$, and the above argument also applies to the operator $N_{\mathcal{B}'}$.

4. Proof of Theorem 1.10

Choose vectors $\boldsymbol{u}, \boldsymbol{v}_1, \dots, \boldsymbol{v}_m \in \mathbb{R}^d$ such that

$$U := \{\boldsymbol{u} + \lambda_1\boldsymbol{v}_1 + \cdots + \lambda_m\boldsymbol{v}_m : \lambda_1, \dots, \lambda_m \in (0,1)\} \subset M \cap \pi$$

is an open set in M. Since π does not contain the origin, we may always assume that $\boldsymbol{u}$ and $\boldsymbol{v}_1, \dots, \boldsymbol{v}_m$ are linearly independent. Then for any measurable set $E \subset X$, any $x \in X$, and any $t > 0$ we have that

$$\int_{tM} \chi_E(\boldsymbol{a}.x)\,\mathrm{dvol}_m(\boldsymbol{a}) \geq \int_{tU} \chi_E(\boldsymbol{a}.x)\,\mathrm{dvol}_m(\boldsymbol{a}). \tag{4.1}$$

Let us study the integral at the right-hand side. Consider the parametrization of tU given by

$$\varphi_t(\boldsymbol{\lambda}) = t\boldsymbol{u} + t\lambda_1\boldsymbol{v}_1 + \cdots + t\lambda_m\boldsymbol{v}_m.$$

Then, if $V := (\boldsymbol{v}_1, \dots, \boldsymbol{v}_m)$, we have that

$$\int_{tU} \chi_E(\boldsymbol{a}.x)\,\mathrm{dvol}_m(\boldsymbol{a}) = \int_{(0,1)^m} \chi_E(\varphi_t(\boldsymbol{\lambda}).x)\cdot t^m\sqrt{\det\left(V^TV\right)}\,\mathrm{d}\boldsymbol{\lambda}. \tag{4.2}$$

Note that, since the vectors $\boldsymbol{v}_1, \dots, \boldsymbol{v}_m$ are linearly independent, the determinant is non-null. By combining (4.1) and (4.2), we deduce that

$$\frac{1}{t^m\cdot\mathrm{vol}_m(M)}\int_{tM} \chi_E(\boldsymbol{a}.x)\,\mathrm{dvol}_m(\boldsymbol{a}) \geq \frac{\sqrt{\det\left(V^TV\right)}}{\mathrm{vol}_m(M)}\int_{(0,1)^m} \chi_E(\varphi_t(\boldsymbol{\lambda}).x)\,\mathrm{d}\boldsymbol{\lambda}. \tag{4.3}$$

Now, let us consider a new action of $\mathbb{R}^{m+1}$ on X (which we denote by ".."), defined by the relations:

$$\boldsymbol{e}_0..x := \boldsymbol{u}.x \quad \text{and} \quad \boldsymbol{e}_i..x := \boldsymbol{v}_i.x \text{ for } i = 1, \dots, m.$$

Then we have that

$$\begin{aligned} \int_{(0,1)^m} \chi_E(\varphi(\boldsymbol{\lambda}).x)\,\mathrm{d}\boldsymbol{\lambda} &= \int_{(0,1)^m} \chi_E\Big((t\boldsymbol{e}_0 + t\lambda_1\boldsymbol{e}_1 + \cdots + t\lambda_m\boldsymbol{e}_m)..x\Big)\,\mathrm{d}\boldsymbol{\lambda} \\ &= \frac{1}{t^m}\int_{(0,t)^m} \chi_E\Big((t\boldsymbol{e}_0 + \lambda_1\boldsymbol{e}_1 + \cdots + \lambda_m\boldsymbol{e}_m)..x\Big)\,\mathrm{d}\boldsymbol{\lambda}. \end{aligned} \tag{4.4}$$

For $k \in \mathbb{N}$ let

$$B_k := [k-1, k) \times [0, k)^m.$$

Note that, in the notation of Theorem 1.9 we have that $s_{k1} = 1$ for all k. Thus, by Remark 2.3, Condition $(!\tilde{C}_i)$ holds for $i = 1$. Moreover, the action of $\mathbb{R}^{m+1}$ defined above is aperiodic, since $\boldsymbol{u}, \boldsymbol{v}_1, \dots, \boldsymbol{v}_m$ are linearly independent, and it commutes with a mixing family of transformations on X, by Remark 1.6. Thus, by Theorem 1.9, for any

$\varepsilon > 0$ there exist a measurable set $E_\varepsilon \subset X$ and a sequence of integers k_r such that for all r and μ-a.e. $x \in X$ it holds that

$$\frac{1}{k_r^m} \int_{[k_r-1,k_r)\times[0,k_r)^m} \chi_{E_\varepsilon}\Big((\lambda_0 \boldsymbol{e}_0 + \lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_m \boldsymbol{e}_m).x\Big)\, \mathrm{d}\lambda_0 \mathrm{d}\boldsymbol{\lambda} > 97/100.$$

By Fubini's Theorem, there must be a real number $t_r \in [k_r - 1, k_r)$ such that

$$\frac{1}{k_r^m} \int_{[0,k_r)^m} \chi_{E_\varepsilon}\Big((t_r \boldsymbol{e}_0 + \lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_m \boldsymbol{e}_m).x\Big) \mathrm{d}\boldsymbol{\lambda} > \frac{97}{100},$$

whence

$$\begin{aligned}&\frac{1}{k_r^m} \int_{(0,t_r)^m} \chi_{E_\varepsilon}\Big((t_r \boldsymbol{e}_0 + \lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_m \boldsymbol{e}_m).x\Big) \mathrm{d}\boldsymbol{\lambda}\\ &= \frac{1}{k_r^m} \int_{[0,k_r)^m} \chi_{E_\varepsilon}\Big((t_r \boldsymbol{e}_0 + \lambda_1 \boldsymbol{e}_1 + \cdots + \lambda_m \boldsymbol{e}_m).x\Big) \mathrm{d}\boldsymbol{\lambda} + O_m(k_r^{-1}) > \frac{97}{100} + O_m(k_r^{-1}).\end{aligned}$$

From this, we deduce that

$$\begin{aligned}\frac{1}{t_r^m} \int_{(0,t_r)^m} &\chi_{E_\varepsilon}\Big((t_r \boldsymbol{e}_0 + \lambda_1 \boldsymbol{e}_1 + \cdots \lambda_m \boldsymbol{e}_m).x\Big) \mathrm{d}\boldsymbol{\lambda}\\ &> \frac{97 k_r^m}{100 t_r^m} + O_m(t_r^{-1}) = \frac{97}{100} + O_m(t_r^{-1}).\end{aligned} \tag{4.5}$$

Combining (4.3), (4.4), and (4.5), we conclude that for all r it holds that

$$\frac{1}{t_r^m \cdot \mathrm{vol}_m(M)} \int_{t_r M} \chi_E(\boldsymbol{a}.x)\, \mathrm{dvol}_m(\boldsymbol{a}) \geq \frac{97\sqrt{\det(V^T V)}}{100 \mathrm{vol}_m(M)} + O_{m,M,V}(t_r^{-1}).$$

This contradicts the convergence to $\mu(\chi_{E_\varepsilon})$ if ε is sufficiently small.

References

[BF09] Michael Björklund and Alexander Fish. Equidistribution of dilations of polynomial curves in nilmanifolds. *Proceedings of the American Mathematical Society*, 137(6):2111–2123, 2009.

[BFK25] Prasuna Bandi, Reynold Fregoli, and Dmitry Kleinbock. Submanifold-genericity of $\mathbb{R}^d$-actions and uniform multiplicative Diophantine approximation. Preprint, arXiv:2504.02258 [math.NT] (2025), 2025.

[BJR90] Alexandra Bellow, Roger Jones, and Joseph Rosenblatt. Convergence for moving averages. *Ergodic Theory Dyn. Syst.*, 10(1):43–62, 1990.

[Coh13] Donald Cohn. *Measure Theory, Second Edition.* Birkhäuser Advanced Texts Basler Lehrbücher. Birkhäuser New York, NY, 2013.

[EW11] Manfred Einsiedler and Thomas Ward. *Ergodic theory. With a view towards number theory*, volume 259 of *Grad. Texts Math.* London: Springer, 2011.

[Gra14] Loukas Grafakos. *Classical Fourier Analysis.* Grad. Texts Math. Springer New York, NY, 2014.

[Jon93] Roger L. Jones. Ergodic averages on spheres. *J. Anal. Math.*, 61:29–45, 1993.

[KW72] Yitzhak Katznelson and Benjamin Weiss. Commuting measure-preserving transformations. *Israel Journal of Mathematics*, 12:161–173, 1972.

[Lac95] Michael T. Lacey. Ergodic averages on circles. *J. Anal. Math.*, 67:199–206, 1995.

[Lin75] D.A Lind. Locally compact measure preserving flows. *Advances in Mathematics*, 15(2):175–193, 1975.

[Lin01] Elon Lindenstrauss. Pointwise theorems for amenable groups. *Invent. Math.*, 146(2):259–295, 2001.

[Saw66] S. Sawyer. Maximal inequalities of weak type. *Annals of Mathematics*, 84(2):157–174, 1966.

[Tao11] Terence Tao. *An introduction to measure theory*, volume 126 of *Grad. Stud. Math.* Providence, RI: American Mathematical Society (AMS), 2011.

Department of Mathematics, University of Michigan, Ann Arbor MI, USA

Email address: djcheng@umich.edu

Department of Mathematics, University of Michigan, Ann Arbor MI, USA

Email address: reynldfregoli@gmail.com

Department of Mathematics, University of Michigan, Ann Arbor MI, USA

Email address: carlguo@umich.edu