RANDOM SEQUENCES AND POINTWISE CONVERGENCE OF MULTIPLE ERGODIC AVERAGES

N. FRANTZIKINAKIS, E. LESIGNE, AND M. WIERDL

Abstract. We prove pointwise convergence, as $N \to \infty$, for the multiple ergodic averages $\frac{1}{N}\sum_{n=1}^{N} f(T^n x)\cdot g(S^{a_n}x)$, where T and S are commuting measure preserving transformations, and a_n is a random version of the sequence $[n^c]$ for some appropriate $c > 1$. We also prove similar mean convergence results for averages of the form $\frac{1}{N}\sum_{n=1}^{N} f(T^{a_n} x)\cdot g(S^{a_n}x)$, as well as pointwise results when T and S are powers of the same transformation. The deterministic versions of these results, where one replaces a_n with $[n^c]$, remain open, and we hope that our method will indicate a fruitful way to approach these problems as well.

1. Introduction

1.1. **Background and new results.** Recent advances in ergodic theory have sparked an outburst of activity in the study of the limiting behavior of multiple ergodic averages. Despite the various successes in proving mean convergence results, progress towards the corresponding pointwise convergence problems has been very scarce. For instance, we still do not know whether the averages

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x)\cdot g(S^n x) \tag{1}$$

converge pointwise when T and S are two commuting measure preserving transformations acting on the same probability space and f and g are bounded measurable functions. Mean convergence for such averages was shown in [10] and was recently generalized to an arbitrary number of commuting transformations in [26]. On the other hand, the situation with pointwise convergence is much less satisfactory. Partial results that deal with special classes of transformations can be found in [2], [3], [21], [22], [1]. Without imposing any strictures on the possible classes of transformations considered, pointwise convergence is only known when T and S are powers of the same transformation [7] (see also [11] for an alternate proof), a result that has not been improved for twenty years.

More generally, for fixed $\alpha, \beta \in [1, +\infty)$, one would like to know whether the averages

$$\frac{1}{N}\sum_{n=1}^{N} f(T^{[n^\alpha]} x)\cdot g(S^{[n^\beta]} x) \tag{2}$$

2000 *Mathematics Subject Classification.* Primary: 37A30; Secondary: 28D05, 05D10, 11B25.

Key words and phrases. Ergodic averages, mean convergence, pointwise convergence, multiple recurrence, random sequences, commuting transformations.

The first author was partially supported by Marie Curie IRG 248008 and the third author by NSF grant DMS-0801316.

converge pointwise. Mean convergence for these and related averages has been extensively studied, partly because of various links to questions in combinatorics, and convergence is known in various cases ([17], [24], [14], [9]). Regarding pointwise convergence, again, the situation is much less satisfactory. When α, β are integers, some partial results for special classes of transformations can be found in [12] and [23]. Furthermore, pointwise convergence is known for averages of the form $\frac{1}{N}\sum_{n=1}^{N} f(T^{[n^{\alpha}]}x)$ with no restrictions on the transformation T ([6] for integers α, and [27] or [5] for non-integers α). But for general commuting transformations T and S, no pointwise convergence result is known, not even when $T = S$ and $\alpha \neq \beta$.

The main goal of this article is to make some progress related to the problem of pointwise convergence of the averages (2) by considering randomized versions of fractional powers of n, in place of the deterministic ones, for various suitably chosen exponents α and β. In our first result, we study a variation of the averages (2) where the iterates of T are deterministic and the iterates of S are random. More precisely, we let a_n be a random version of the sequence $[n^{\beta}]$ where $\beta \in (1, 14/13)$ is arbitrary. We prove that almost surely (the set of probability 1 is universal) the averages

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x) \cdot g(S^{a_n} x) \tag{3}$$

converge pointwise, and we determine the limit explicitly. This is the first pointwise convergence result for multiple ergodic averages of the form $\frac{1}{N}\sum_{n=1}^{N} f(T^{a_n}x) \cdot g(S^{b_n}x)$, where a_n, b_n are strictly increasing sequences and T, S are general commuting measure preserving transformations. In fact, even for mean convergence the result is new, and this is an instance where convergence of multiple ergodic averages involving sparse iterates is obtained without the use of rather deep ergodic structure theorems and equidistribution results on nilmanifolds.

Let us also remark that although convergence of the averages (1) for not necessarily commuting transformations is known to fail in general,[1] it is unclear to us if this is the case for the averages (2) when $\alpha \neq \beta$, or the averages (3).

In our second result, we study a randomized version of the averages (2) when $\alpha = \beta$. In this case, we let a_n be a random version of the sequence $[n^{\alpha}]$ where $\alpha \in (1, 2)$ is arbitrary, and prove that almost surely (the set of probability 1 is universal) the averages

$$\frac{1}{N}\sum_{n=1}^{N} f(T^{a_n} x) \cdot g(S^{a_n} x) \tag{4}$$

converge in the mean, and conditionally to the pointwise convergence of the averages (1), they also converge pointwise. Even for mean convergence, this gives the first examples of sparse sequences of integers a_n for which the averages (4) converge for general commuting measure preserving transformations T and S.

Because our convergence results come with explicit limit formulas, we can easily deduce some related multiple recurrence results. Using the correspondence principle of Furstenberg, these results translate to statements in combinatorics about configurations that can be found in every subset of the integers, or the integer lattice, with positive upper density.

[1] See example 7.1 in [2], or let $T, S\colon \mathbb{T} \to \mathbb{T}$, given by $Tx = 2x$, $Sx = 2x + \alpha$, and $f(x) = e^{-2\pi i x}, g(x) = e^{2\pi i x}$, where $\alpha \in [0, 1]$ is chosen so that the averages $\frac{1}{N}\sum_{n=1}^{N} e^{2\pi i \cdot 2^n \alpha}$ diverge.

We state the exact results in the next section, where we also give precise definitions of the concepts used throughout the paper.

1.2. Precise statements of new results.

1.2.1. *Our setup.* We work with random sequences of integers that are constructed by selecting a positive integer n to be a member of our sequence with probability $\sigma_n \in [0,1]$. More precisely, let $(\Omega, \mathcal{F}, \mathbb{P})$ be a probability space, and let $(X_n)_{n\in\mathbb{N}}$ be a sequence of independent random variables with

$$\mathbb{P}(X_n = 1) = \sigma_n \quad \text{and} \quad \mathbb{P}(X_n = 0) = 1 - \sigma_n.$$

In the present article we always assume that $\sigma_n = n^{-a}$ for some $a \in (0,1)$. The random sequence $(a_n(\omega))_{n\in\mathbb{N}}$ is constructed by taking the positive integers n for which $X_n(\omega) = 1$ in increasing order. Equivalently, $a_n(\omega)$ is the smallest $k \in \mathbb{N}$ such that $X_1(\omega) + \cdots + X_k(\omega) = n$. We record the identity

$$X_1(\omega) + \cdots + X_{a_n(\omega)}(\omega) = n \tag{5}$$

for future use.

The sequence $(a_n(\omega))_{n\in\mathbb{N}}$ is what we called random version of the sequence $n^{1/(1-a)}$ in the previous subsection. Indeed, using a variation of the strong law of large numbers (see Lemma 4.6), we have that almost surely $\frac{1}{\sum_{k=1}^N \sigma_k} \sum_{k=1}^N X_k(\omega)$ converges to 1. Using the implied estimate for $a_n(\omega)$ in place of N where n is suitably large, and using (5), we deduce that almost surely $a_n(\omega)/n^{1/(1-a)}$ converges to a non-zero constant.

We say that a certain property holds almost surely for the sequences $(a_n(\omega))_{n\in\mathbb{N}}$, if there exists a universal set $\Omega_0 \in \mathcal{F}$, such that $\mathbb{P}(\Omega_0) = 1$, and for every $\omega \in \Omega_0$ the sequence $(a_n(\omega))_{n\in\mathbb{N}}$ satisfies the given property.

1.2.2. *Different iterates.* In our first result we study a randomized version of the averages (2) when $\alpha = 1$.

Theorem 1.1. *With the notation of Section 1.2.1, let $\sigma_n = n^{-a}$ for some $a \in (0, 1/14)$. Then almost surely the following holds: For every probability space $(X, \mathcal{X}, \mu)$, commuting measure preserving transformations $T, S\colon X \to X$, and functions $f, g \in L^\infty(\mu)$, for almost every $x \in X$ we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^n x) \cdot g(S^{a_n(\omega)} x) = \tilde{f}(x) \cdot \tilde{g}(x) \tag{6}$$

where $\tilde{f} := \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^N T^n f = \mathbb{E}(f|\mathcal{I}(T))$, $\tilde{g} := \lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^N S^n g = \mathbb{E}(g|\mathcal{I}(S))$.[2]

We remark that the conclusion of Theorem 1.1 can be easily extended to all functions $f \in L^p$, $g \in L^q$, where $p \in [1, +\infty]$ and $q \in (1, +\infty]$ satisfy $1/p + 1/q \leq 1$.[3]

[2]If $(X, \mathcal{X}, \mu)$ is a probability space, $f \in L^\infty(\mu)$, and $\mathcal{Y}$ a sub-σ-algebra of $\mathcal{X}$, we denote by $\mathbb{E}(f|\mathcal{Y})$ the conditional expectation of f given $\mathcal{Y}$. If $T\colon X \to X$ is a measure preserving transformation, by $\mathcal{I}(T)$ we denote the sub-σ-algebra of sets that are left invariant by T.

[3]To see this, one uses a standard approximation argument and the fact that the averages $\frac{1}{N}\sum_{n=1}^N T^n f$ converge pointwise for $f \in L^p$ when $p \in [1, +\infty]$, and the same holds for the averages $\frac{1}{N}\sum_{n=1}^N S^{a_n(\omega)} f$ for $f \in L^q$ when $q \in (1, +\infty]$ (see for example exercise 3 on page 78 of [25]).

Combining the limit formula of Theorem 1.1 with the estimate (see Lemma 1.6 in [8])

$$\int f \cdot \mathbb{E}(f|\mathcal{X}_1) \cdot \mathbb{E}(f|\mathcal{X}_2)\ d\mu \geq \Big(\int f\ d\mu \Big)^3,$$

that holds for every non-negative function $f \in L^\infty(\mu)$ and sub-σ-algebras $\mathcal{X}_1$ and $\mathcal{X}_2$ of $\mathcal{X}$, we deduce the following:

Corollary 1.2. *With the assumptions of Theorem 1.1, we get almost surely, that for every $A \in \mathcal{X}$ we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \mu(A \cap T^{-n}A \cap S^{-a_n(\omega)}A) \geq \mu(A)^3.$$

The *upper density* $\bar{d}(E)$ of a set $E \subset \mathbb{Z}^2$ is defined by $\bar{d}(E) = \limsup_{N\to\infty} \frac{|E\cap[-N,N]^2|}{|[-N,N]^2|}$. Combining the previous multiple recurrence result with a multidimensional version of Furstenberg's correspondence principle [16], we deduce the following:

Corollary 1.3. *With the notation of Section 1.2.1, let $\sigma_n = n^{-a}$ for some $a \in (0, 1/14)$. Then almost surely, for every $v_1, v_2 \in \mathbb{Z}^2$ and $E \subset \mathbb{Z}^2$ we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \bar{d}\big(E \cap (E - nv_1) \cap (E - a_n(\omega)v_2)\big) \geq (\bar{d}(E))^3.$$

We remark that throughout the previous statement we could have used the upper Banach density d^* in place of the upper density $\bar{d}$. This is defined by $d^*(E) = \limsup_{|I|\to\infty} \frac{|E\cap I|}{|I|}$, where $|I|$ denotes area of a rectangle I and the lim sup is taken over all rectangles of $\mathbb{Z}^2$ with side lengths that increase to infinity. The same holds for the statement of Corollary 1.6 below.

1.2.3. *Same iterates.* In our next result we study a randomized version of the averages (1).

Theorem 1.4. *With the notation of Section 1.2.1, let $\sigma_n = n^{-a}$ for some $a \in (0, 1/2)$. Then almost surely the following holds: For every probability space $(X, \mathcal{X}, \mu)$, commuting measure preserving transformations $T, S \colon X \to X$, and functions $f, g \in L^\infty(\mu)$, the averages*

$$\frac{1}{N} \sum_{n=1}^{N} T^{a_n(\omega)} f \cdot S^{a_n(\omega)} g \tag{7}$$

converge in $L^2(\mu)$ and their limit equals the L^2-limit of the averages $\frac{1}{N}\sum_{n=1}^{N} T^n f \cdot S^n g$ (this exists by [10]*). Furthermore, if T and S are powers of the same transformation, then the averages* (7) *converge pointwise.*

We remark that our argument also gives pointwise convergence of the averages (7), conditionally to the pointwise convergence of the averages (1). Furthermore, using our method, one can also get similar convergence results for other random multiple ergodic averages. For instance, our method can be modified and combined with the results from [26] and [9] to show that if $\sigma_n = n^{-a}$ and a is small enough, then almost surely the averages $\frac{1}{N}\sum_{n=1}^{N} T_1^{a_n(\omega)} f_1 \cdots T_\ell^{a_n(\omega)} f_\ell$ and $\frac{1}{N}\sum_{n=1}^{N} T_1^{a_n(\omega)} f_1 \cdot T_2^{(a_n(\omega))^2} f_2 \cdots T_\ell^{(a_n(\omega))^\ell} f_\ell$ converge in the mean.

Combining Theorem 1.4 with the multiple recurrence result of Furstenberg and Katznelson [16], we deduce the following:

Corollary 1.5. *With the assumptions of Theorem 1.4, we get almost surely, that if $A \in \mathcal{X}$ has positive measure, then there exists $n \in \mathbb{N}$ such that*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \mu(A \cap T^{-a_n(\omega)}A \cap S^{-a_n(\omega)}A) > 0.$$

Combining the previous multiple recurrence result with the correspondence principle of Furstenberg [15], we deduce the following:

Corollary 1.6. *With the notation of Section 1.2.1, let $\sigma_n = n^{-a}$ for some $a \in (0, 1/2)$. Then almost surely, for every $v_1, v_2 \in \mathbb{Z}^2$, and every $E \subset \mathbb{Z}^2$ with $\bar{d}(E) > 0$, we have*

$$\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} \bar{d}\big(E \cap (E + a_n(\omega)v_1) \cap (E + a_n(\omega)v_2)\big) > 0.$$

1.3. **Further directions.** The restrictions on the range of the eligible parameter a in Theorem 1.1, Theorem 1.4, and the related corollaries, is far from best possible.[4] In fact, any $a < 1$ is expected to work, but it seems that new techniques are needed to prove this. This larger range of parameters is known to work for pointwise convergence of the averages $\frac{1}{N}\sum_{n=1}^{N} f(T^{a_n(\omega)}x)$ (see [4] for mean convergence, [6] for pointwise, and [25] for a survey of related results). Furthermore, when $\sigma_n = \sigma \in (0, 1)$ for every $n \in \mathbb{N}$, it is not clear whether the conclusion of Theorem 1.1 holds (see Theorem 4 in [20] for a related negative pointwise convergence result).

Regarding Theorem 1.1, it seems very likely that similar convergence results hold when the iterates of the transformation T are given by other "good" deterministic sequences, like polynomial sequences. Our argument does not give such an extension because it relies crucially on the linearity of the iterates of T. Furthermore, it seems likely that similar convergence results hold when the iterates of T and S are both given by random versions of different fractional powers, chosen independently. Again our present argument does not seem to apply to this case.

Lastly, as we mentioned in the introduction, it may be the case that the conclusion of Theorem 1.1 holds for general measure preserving transformations T and S (without imposing any commutativity assumption). Showing this for a single instance of a positive parameter α, or showing that no such parameter exists, would be very interesting.

1.4. **General conventions and notation.** We use the symbol $\ll$ when some expression is majorized by a constant multiple of some other expression. If this constant depends on the variables $k_1, \ldots, k_\ell$, we write $\ll_{k_1,\ldots,k_\ell}$. We say that $a_n \sim b_n$ if a_n/b_n converges to a non-zero constant. We denote by $o_N(1)$ a quantity that converges to zero when $N \to \infty$ and all other parameters are fixed. We say that two sequences are *asymptotically equal* whenever convergence of one implies convergence of the other and both limits coincide. If $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space, and X is a random variable, we set $\mathbb{E}_\omega(X) = \int X \, d\mathbb{P}$. We say that a property holds almost surely if it holds outside of a set with probability zero. We often suppress writing the variable x when we refer to functions and the variable ω when we refer to random variables or random sequences. Lastly, the following notation will be used throughout the article: $\mathbb{N} = \{1, 2, \ldots\}$, $Tf = f \circ T$, $e(t) = e^{2\pi i t}$.

[4]Any improvement in the range of the eligible parameter a in the statement of Proposition 2.1 or Proposition 3.1, would give corresponding improvements in the statement of Theorem 1.1 and Theorem 1.4 and the related corollaries.

2. Convergence for independent random iterates

In this section we prove Theorem 1.1. Throughout, we use the notation introduced in Section 1.2.1.

2.1. **Strategy of the proof.** Roughly speaking, in order to prove Theorem 1.1 we go through the following successive comparisons:

$$\begin{aligned}\frac{1}{N}\sum_{n=1}^{N} f(T^n x)\cdot g(S^{a_n(\omega)}x) &\approx \frac{1}{W_N}\sum_{n=1}^{N} X_n(\omega)\cdot f(T^{X_1(\omega)+\cdots+X_n(\omega)}x)\cdot g(S^n x)\\ &\approx \frac{1}{W_N}\sum_{n=1}^{N} \sigma_n\cdot f(T^{X_1(\omega)+\cdots+X_n(\omega)}x)\cdot g(S^n x)\\ &\approx \frac{1}{N}\sum_{n=1}^{N} f(T^{X_1(\omega)+\cdots+X_n(\omega)}x)\cdot g(S^n x)\\ &\approx \tilde{g}(x)\cdot\frac{1}{N}\sum_{n=1}^{N} f(T^{X_1(\omega)+\cdots+X_n(\omega)}x)\\ &\approx \tilde{g}(x)\cdot\frac{1}{N}\sum_{n=1}^{N} f(T^n x)\\ &\approx \tilde{f}(x)\cdot\tilde{g}(x),\end{aligned}$$

where $A_N(\omega,x)\approx B_N(\omega,x)$ means that almost surely (the set of probability 1 is universal), the expression $A_N(\omega,x)$ is asymptotically equal to $B_N(\omega,x)$ for almost every $x\in X$. The second comparison is the most crucial one; essentially one has to get good estimates for the L^2 norm of the averages $\frac{1}{W_N}\sum_{n=1}^{N}(X_n(\omega)-\sigma_n)\cdot T^{X_1(\omega)+\cdots+X_n(\omega)}f\cdot S^n g$. We do this in two steps. First we use an elementary estimate of van der Corput twice to get a bound that depends only on the random variables Y_n, and then estimate the resulting expressions using the independence of the variables Y_n. Let us also mention that the fifth comparison follows immediately by applying the first three for $g=1$.

2.2. **A reduction.** Let

$$Y_n := X_n-\sigma_n, \quad W_N := \sum_{n=1}^{N}\sigma_n.$$

We remark that if $\sigma_n=n^{-a}$ for some $a\in(0,1)$, then $W_N\sim N^{1-a}$.

Our first goal it to reduce Theorem 1.1 to proving the following result:

Proposition 2.1. *Suppose that $\sigma_n=n^{-a}$ for some $a\in(0,1/14)$ and let $\gamma>1$ be a real number. Then almost surely the following holds: For every probability space $(X,\mathcal{X},\mu)$, commuting measure preserving transformations $T,S\colon X\to X$, and functions $f,g\in L^\infty(\mu)$, we have*

$$\sum_{k=1}^{\infty}\left\|\frac{1}{W_{[\gamma^k]}}\sum_{n=1}^{[\gamma^k]} Y_n(\omega)\cdot T^{X_1(\omega)+\cdots+X_n(\omega)}f\cdot S^n g\right\|_{L^2(\mu)} < +\infty. \tag{8}$$

We are going to establish this reduction in the next subsections.

2.2.1. *First step.* We assume, as we may, that both functions $|f|$ and $|g|$ are pointwise bounded by 1 for all points in X. By (5) for every $\omega \in \Omega$ and $x \in X$ we have

$$\frac{1}{N}\sum_{n=1}^{N} f(T^n x)\cdot g(S^{a_n(\omega)}x) = \frac{1}{N}\sum_{n=1}^{N} f(T^{X_1(\omega)+\cdots+X_{a_n(\omega)}(\omega)}x)\cdot g(S^{a_n(\omega)}x).$$

A moment of reflection shows that for every bounded sequence $(b_n)_{n\in\mathbb{N}}$, for every $\omega \in \Omega$, the averages

$$\frac{1}{N}\sum_{n=1}^{N} b_{a_n(\omega)}$$

and the averages

$$\frac{1}{W_N(\omega)}\sum_{n=1}^{N} X_n(\omega)\cdot b_n,$$

where $W_N(\omega) := X_1(\omega)+\cdots+X_N(\omega)$, are asymptotically equal as $N \to \infty$. Moreover, Lemma 4.6 in the Appendix gives that almost surely $\lim_{N\to\infty} W_N(\omega)/W_N = 1$. Therefore, the last averages are asymptotically equal to the averages

$$\frac{1}{W_N}\sum_{n=1}^{N} X_n(\omega)\cdot b_n.$$

Putting these observations together, we see that for almost every $\omega \in \Omega$ the averages in (6) and the averages

$$\frac{1}{W_N}\sum_{n=1}^{N} X_n(\omega)\cdot f(T^{X_1(\omega)+\cdots+X_n(\omega)}x)\cdot g(S^n x) \tag{9}$$

are asymptotically equal for every $x \in X$.

2.2.2. *Second step.* Next, we study the limiting behavior of the averages (9) when the random variables X_n are replaced by their mean. Namely, we study the averages

$$\frac{1}{W_N}\sum_{n=1}^{N} \sigma_n\cdot f(T^{X_1(\omega)+\cdots+X_n(\omega)}x)\cdot g(S^n x). \tag{10}$$

By Lemma 4.3 in the Appendix, for every $\omega \in \Omega$ and $x \in X$ they are asymptotically equal to the averages

$$\frac{1}{N}\sum_{n=1}^{N} f(T^{X_1(\omega)+\cdots+X_n(\omega)}x)\cdot g(S^n x). \tag{11}$$

Lemma 2.2. *Suppose that $\sigma_n = n^{-a}$ for some $a \in (0,1)$. Then almost surely the following holds: For every probability space $(X, \mathcal{X}, \mu)$, measure preserving transformations $T, S\colon X \to X$, and functions $f, g \in L^\infty(\mu)$, we have*

$$\lim_{N\to\infty}\Big(\frac{1}{N}\sum_{n=1}^{N} f(T^{X_1(\omega)+\cdots+X_n(\omega)}x)\cdot g(S^n x) - \frac{1}{N}\sum_{n=1}^{N} f(T^{X_1(\omega)+\cdots+X_n(\omega)}x)\cdot \mathbb{E}(g|\mathcal{I}(S))(x)\Big) = 0$$

for almost every $x \in X$.

Proof. It suffices to show that almost surely, if $\mathbb{E}(g|\mathcal{I}(S)) = 0$, then $\lim_{N\to\infty} A_N(f, g, \omega, x) = 0$ for almost every $x \in X$, where

$$A_N(f, g, \omega, x) := \frac{1}{N}\sum_{n=1}^{N} f(T^{X_1(\omega)+\cdots+X_n(\omega)}x) \cdot g(S^n x).$$

First we consider functions g of the form $h - Sh$ where $h \in L^\infty(\mu)$. Assuming, as we may, that both $|f|$ and $|h|$ are pointwise bounded by 1 for all points in X, partial summation gives that

$$A_N(f, h-Sh, \omega, x) = \frac{1}{N}\sum_{n=1}^{N} \big(f(T^{X_1(\omega)+\cdots+X_n(\omega)}x) - f(T^{X_1(\omega)+\cdots+X_{n-1}(\omega)}x)\big) \cdot h(S^n x) + o_N(1).$$

The complex norm of the last expression is bounded by a constant times the average

$$\frac{1}{N}\sum_{n=1}^{N} \mathbf{1}_{E_n}(\omega)$$

where $E_n := \{\omega\colon X_n(\omega) = 1\}$. Since $\mathbb{P}(E_n) = n^{-a}$, combining our assumption with Lemma 4.5 in the Appendix, we get that the last average converges almost surely to 0 as $N \to \infty$. Therefore, on a set Ω_0 of probability 1, that depends only on the random variables X_n, we have

$$\lim_{N\to\infty} A_N(f, h-Sh, \omega, x) = 0 \tag{12}$$

for almost every $x \in X$.

Furthermore, using the trivial estimate

$$|A_N(f, g, \omega, x)| \leq \frac{1}{N}\sum_{n=1}^{N} |g|(S^n x),$$

and then applying the pointwise ergodic theorem for the transformation S and the Cauchy-Schwarz inequality, we get for every $\omega \in \Omega$ that

$$\int \limsup_{N\to\infty} |A_N(f, g, \omega, \cdot)| \; d\mu \leq \|g\|_{L^2(\mu)}\,. \tag{13}$$

Since every function $g \in L^\infty(\mu)$ that satisfies $\mathbb{E}(g|\mathcal{I}(S)) = 0$ can be approximated in $L^2(\mu)$ arbitrarily well by functions of the form $h - Sh$ with $h \in L^\infty(\mu)$, combining (12) and (13), we get for every $\omega \in \Omega_0$, that if $\mathbb{E}(g|\mathcal{I}(S)) = 0$, then $\lim_{N\to\infty} A_N(f, g, \omega, x) = 0$ for almost every $x \in X$. This completes the proof. □

2.2.3. *Third step.* We next turn our attention to the study of the limiting behavior of the averages

$$\frac{1}{N}\sum_{n=1}^{N} f(T^{X_1(\omega)+\cdots+X_n(\omega)}x). \tag{14}$$

Lemma 2.3. *Let $\sigma_n = n^{-a}$ for some $a \in (0, 1/14)$. Then almost surely the following holds: For every probability space $(X, \mathcal{X}, \mu)$, measure preserving transformation $T\colon X \to X$, and function $f \in L^\infty(\mu)$, the averages in (14) converge to $\mathbb{E}(f|\mathcal{I}(T))(x)$ for almost every $x \in X$.*

Remark. Improving the range of the parameter a would not lead to corresponding improvements in our main results. On the other hand, the restricted range we used enables us to give a succinct proof using Proposition 2.1.

Proof. We assume, as we may, that the function $|f|$ is pointwise bounded by 1 for all points in X. First notice that by Lemma 4.3 in the Appendix, for every $\omega \in \Omega$ and $x \in X$, the averages in (14) are asymptotically equal to the averages

$$\frac{1}{W_N} \sum_{n=1}^{N} \sigma_n \cdot f(T^{X_1(\omega)+\cdots+X_n(\omega)}x)$$

where $W_N := \sum_{n=1}^N n^{-a} \sim N^{1-a}$. Combining this observation with Corollary 4.2 on the Appendix, we deduce that it suffices to show that almost surely the following holds: For every probability space $(X, \mathcal{X}, \mu)$, measure preserving transformation $T\colon X \to X$, function $f \in L^\infty(\mu)$, and $\gamma \in \{1 + 1/k, k \in \mathbb{N}\}$, we have

$$\lim_{N\to\infty} \frac{1}{W_{[\gamma^N]}} \sum_{n=1}^{[\gamma^N]} \sigma_n \cdot f(T^{X_1(\omega)+\cdots+X_n(\omega)}x) = \mathbb{E}(f|\mathcal{I}(T))(x) \tag{15}$$

for almost every $x \in X$.

Using Proposition 2.1 for $g = 1$, we get that almost surely (the set of probability 1 depends only on the random variables X_n), for every $\gamma \in \{1 + 1/k, k \in \mathbb{N}\}$, the averages in (15) are asymptotically equal to the averages

$$\frac{1}{W_{[\gamma^N]}} \sum_{n=1}^{[\gamma^N]} X_n(\omega) \cdot f(T^{X_1(\omega)+\cdots+X_n(\omega)}x)$$

for almost every $x \in X$. Hence, it suffices to study the limiting behavior of the averages

$$\frac{1}{W_N} \sum_{n=1}^{N} X_n(\omega) \cdot f(T^{X_1(\omega)+\cdots+X_n(\omega)}x).$$

Repeating the argument used in Section 2.2.1 (with $g = 1$), we deduce that for every $\omega \in \Omega$ and $x \in X$, they are asymptotically equal to the averages

$$\frac{1}{N} \sum_{n=1}^{N} f(T^{X_1(\omega)+\cdots+X_{a_n(\omega)}(\omega)}x) = \frac{1}{N} \sum_{n=1}^{N} f(T^n x)$$

where the last equality follows from (5). Finally, using the pointwise ergodic theorem we get that the last averages converge to $\mathbb{E}(f|\mathcal{I}(T))(x)$ for almost every $x \in X$. This completes the proof. □

2.2.4. *Last step.* We prove Theorem 1.1 by combining Proposition 2.1 with the arguments in the previous three steps. We start with Proposition 2.1. It gives that there exists a set $\Omega_0 \in \mathcal{F}$ of probability 1 such that for every $\omega \in \Omega_0$ the following holds: For every probability space $(X, \mathcal{X}, \mu)$, commuting measure preserving transformations $T, S\colon X \to X$, functions $f, g \in L^\infty(\mu)$, and $\gamma \in \{1 + 1/k, k \in \mathbb{N}\}$, we have

$$\sum_{N=1}^{\infty} \left\| S_{[\gamma^N]}(\omega, \cdot) \right\|_{L^2(\mu)} < +\infty \tag{16}$$

where

$$S_N(\omega, x) := \frac{1}{W_N} \sum_{n=1}^{N} Y_n(\omega) \cdot f(T^{X_1(\omega)+\cdots+X_n(\omega)}x) \cdot g(S^n x).$$

In the remaining argument ω is assumed to belong to the aforementioned set Ω_0. Notice that (16) implies that

$$\lim_{N\to\infty} S_{[\gamma^N]}(\omega, x) = 0 \qquad \text{for almost every } x \in X.$$

We conclude that for almost every $x \in X$, for every $\gamma \in \{1 + 1/k, k \in \mathbb{N}\}$, the difference

$$\frac{1}{W_{[\gamma^N]}} \sum_{n=1}^{[\gamma^N]} X_n(\omega) \cdot f(T^{X_1(\omega)+\cdots+X_n(\omega)}x) \cdot g(S^n x) - \frac{1}{W_{[\gamma^N]}} \sum_{n=1}^{[\gamma^N]} \sigma_n \cdot f(T^{X_1(\omega)+\cdots+X_n(\omega)}x) \cdot g(S^n x)$$

converges to 0 as $N \to \infty$. In Sections 2.2.2 and 2.2.3 we proved that for almost every $x \in X$ we have

$$\lim_{N\to\infty} \frac{1}{W_N} \sum_{n=1}^{N} \sigma_n \cdot f(T^{X_1(\omega)+\cdots+X_n(\omega)}x) \cdot g(S^n x) = \tilde{f}(x) \cdot \tilde{g}(x),$$

where $\tilde{f} := \mathbb{E}(f|\mathcal{I}(T))$, and $\tilde{g} := \mathbb{E}(g|\mathcal{I}(S))$. We deduce from the above that for almost every $x \in X$, for every $\gamma \in \{1 + 1/k, k \in \mathbb{N}\}$, we have that

$$\lim_{N\to\infty} \frac{1}{W_{[\gamma^N]}} \sum_{n=1}^{[\gamma^N]} X_n(\omega) \cdot f(T^{X_1(\omega)+\cdots+X_n(\omega)}x) \cdot g(S^n x) = \tilde{f}(x) \cdot \tilde{g}(x).$$

Since the sequence W_N satisfies the assumptions of Corollary 4.2 in the Appendix, we conclude that for non-negative functions $f, g \in L^\infty(\mu)$, for almost every $x \in X$, we have

$$\lim_{N\to\infty} \frac{1}{W_N} \sum_{n=1}^{N} X_n(\omega) \cdot f(T^{X_1(\omega)+\cdots+X_n(\omega)}x) \cdot g(S^n x) = \tilde{f}(x) \cdot \tilde{g}(x). \tag{17}$$

Splitting the real and imaginary part of the function f as a difference of two non-negative functions, doing the same for the function g, and using the linearity of the operator $f \to \tilde{f}$, we deduce that (17) holds for arbitrary $f, g \in L^\infty(\mu)$.

Lastly, combining the previous identity and the argument used in Section 2.2.1, we deduce that for almost every $x \in X$ we have

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} f(T^n x) \cdot g(S^{a_n(\omega)}x) = \tilde{f}(x) \cdot \tilde{g}(x).$$

We have therefore established:

Proposition 2.4. *If Proposition 2.1 holds, then Theorem 1.1 holds.*

In the next subsection we prove Proposition 2.1.

2.3. **Proof of Proposition 2.1.** The proof of Proposition 2.1 splits in two parts. First we estimate the L^2 norm of the averages $\frac{1}{W_N}\sum_{n=1}^{N} Y_n \cdot T^{X_1+\cdots+X_n} f \cdot S^n g$ by an expression that is independent of the transformations T, S and the functions f, g. The main idea is to use van der Corput's Lemma (see Lemma 4.4 in the Appendix) enough times to get the desired cancelation, allowing enough flexibility on the parameters involved to ensure that certain terms become negligible. Subsequently, using moment estimates, we show that the resulting expression is almost surely summable along exponentially growing sequences of integers.

Before delving into the details we make some preparatory remarks that will help us ease our notation. We assume that both functions f, g are bounded by 1. We remind the reader that

$$\sigma_n = n^{-a}, \quad W_N \sim N^{1-a}$$

for some $a \in (0,1)$. We are going to use parameters M and R that satisfy

$$M := N^b, \quad R := N^c$$

for some $b, c \in (0,1)$ at our disposal. We impose more restrictions on a, b, c as we move on.

2.3.1. *Eliminating the dependence on the transformations and the functions.* To simplify our notation, in this subsection, when we write $\sum_{n=1}^{N^\alpha}$ we mean $\sum_{n=1}^{[N^\alpha]}$.

Using Lemma 4.4 in the Appendix with $M = [N^b]$ and $v_n = Y_n \cdot T^{X_1+\cdots+X_n} f \cdot S^n g$, we get that

$$A_N := \Big\| N^{-1+a} \sum_{n=1}^{N} Y_n \cdot T^{X_1+\cdots+X_n} f \cdot S^n g \Big\|_{L^2(\mu)}^2 \ll A_{1,N} + A_{2,N}, \tag{18}$$

where

$$A_{1,N} := N^{-1+2a-b} \cdot \sum_{n=1}^{N} \big\| Y_n \cdot T^{X_1+\cdots+X_n} f \cdot S^n g \big\|_{L^2(\mu)}^2$$

and

$$A_{2,N} := N^{-1+2a-b} \cdot \sum_{m=1}^{N^b} \Big| \sum_{n=1}^{N-m} \int Y_{n+m} \cdot Y_n \cdot T^{X_1+\cdots+X_{n+m}} f \cdot S^{n+m} g \cdot T^{X_1+\cdots+X_n} \bar{f} \cdot S^n \bar{g} \; d\mu \Big|.$$

We estimate $A_{1,N}$. Since $\mathbb{E}(Y_n^2) = \sigma_n - \sigma_n^2 \sim n^{-a}$, Lemma 4.6 in the Appendix gives for every $a \in (0,1)$ that $\sum_{n=1}^{N} Y_n^2 \sim \sum_{n=1}^{N} \mathbb{E}(Y_N^2) \sim N^{1-a}$. Therefore, almost surely we have

$$A_{1,N} \ll N^{-1+2a-b} \sum_{n=1}^{N} Y_n^2 \ll_\omega N^{-1+2a-b} \cdot N^{1-a} = N^{a-b}.$$

It follows that $A_{1,N}$ is bounded by a negative power of N as long as

$$b > a.$$

We estimate $A_{2,N}$. We compose with S^{-n} and use the Cauchy-Schwarz inequality. We get

$$A_{2,N} \ll N^{-1+2a-b} \cdot \sum_{m=1}^{N^b} \Big\| \sum_{n=1}^{N-m} Y_{n+m} \cdot Y_n \cdot S^{-n} T^{X_1+\cdots+X_{n+m}} f \cdot S^{-n} T^{X_1+\cdots+X_n} \bar{f} \Big\|_{L^2(\mu)}.$$

Furthermore, since

$$N^{-1+2a-b}\sum_{m=1}^{N^b} m \ll N^{-1+2a+b},$$

we get the estimate

$$A_{2,N} \ll N^{-d_1} + N^{-1+2a-b}\cdot \sum_{m=1}^{N^b}\Big\|\sum_{n=1}^{N} Y_{n+m}\cdot Y_n\cdot S^{-n}T^{X_1+\cdots+X_{n+m}}f\cdot S^{-n}T^{X_1+\cdots+X_n}\bar{f}\Big\|_{L^2(\mu)}$$

where $d_1 := 1-2a-b$ is positive as long as

$$2a+b<1.$$

Using the Cauchy-Schwarz inequality we get

$$A_{2,N}^2 \ll N^{-2d_1} + N^{-2+4a-b}\cdot \sum_{m=1}^{N^b}\Big\|\sum_{n=1}^{N} Y_{n+m}\cdot Y_n\cdot S^{-n}T^{X_1+\cdots+X_{n+m}}f\cdot S^{-n}T^{X_1+\cdots+X_n}\bar{f}\Big\|_{L^2(\mu)}^2.$$

Next we use Lemma 4.4 in the Appendix with $R=[N^c]$ and the obvious choice of functions v_n, in order to estimate the square of the L^2 norm above. We get the estimate

$$A_{2,N}^2 \ll N^{-2d_1} + A_{3,N} + A_{4,N},$$

where $A_{3,N}$, $A_{4,N}$, can be computed as before. Using Lemma 4.7 in the Appendix, and the estimate $\mathbb{E}(Y_n^2)\sim n^{-a}$, we deduce that almost surely, for every $a\in(0,1/6)$ we have

$$A_{3,N} \ll N^{-1+4a-b-c}\sum_{m=1}^{N^b}\sum_{n=1}^{N} Y_{n+m}^2 Y_n^2 \ll_\omega N^{2a-c} = N^{-d_2}$$

where $d_2>0$ as long as

$$2a<c.$$

Composing with $T^{-(X_1+\cdots+X_n)}S^n$, using that T and S commute, and the Cauchy-Schwarz inequality, we see that

$$A_{4,N} \ll N^{-1+4a-b-c}\cdot \sum_{m=1}^{N^b}\sum_{r=1}^{N^c}\Big\|\sum_{n=1}^{N-r} Y_{n+m+r}\cdot Y_{n+r}\cdot Y_{n+m}\cdot Y_n\cdot$$
$$T^{X_{n+1}+\cdots+X_{n+m+r}}S^{-r}f\cdot T^{X_{n+1}+\cdots+X_{n+r}}S^{-r}\bar{f}\cdot T^{X_{n+1}+\cdots+X_{n+m}}\bar{f}\Big\|_{L^2(\mu)}.$$

Since for every $k\in\mathbb{N}$ we have $X_{n+1}+\cdots+X_{n+k}\in\{0,\ldots,k\}$, it follows that

$$A_{4,N} \ll A_{5,N} := N^{-1+4a-b-c}\cdot \sum_{m=1}^{N^b}\sum_{r=1}^{N^c}\sum_{k_1=0}^{m+r}\sum_{k_2=0}^{r}\sum_{k_3=0}^{m}\Big|\sum_{n=1}^{N-r} Y_{n+m+r}\cdot Y_{n+r}\cdot Y_{n+m}\cdot Y_n\cdot$$
$$\mathbf{1}_{\sum_{i=1}^{m+r}X_{n+i}=k_1}(n)\cdot \mathbf{1}_{\sum_{i=1}^{r}X_{n+i}=k_2}(n)\cdot \mathbf{1}_{\sum_{i=1}^{m}X_{n+i}=k_3}(n)\Big|. \tag{19}$$

Summarizing, we have just shown that as long as

$$a<b,\qquad 2a+b<1,\qquad 2a<c,\qquad a\in(0,1/6),\qquad b,c\in(0,1), \tag{20}$$

almost surely the following holds: For every probability space $(X,\mathcal{X},\mu)$, commuting measure preserving transformations $T,S\colon X\to X$, and functions $f,g\in L^\infty(\mu)$ with $\|f\|_{L^\infty(\mu)}\leq 1$ and $\|g\|_{L^\infty(\mu)}\leq 1$, we have

$$A_N \ll_\omega N^{-d_3} + A_{5,N} \tag{21}$$

for some $d_3>0$, where $A_{5,N}$ is defined in (19). Notice that the expression $A_{5,N}$ depends only on the random variables X_n. Therefore, in order to complete the proof of Proposition 2.1, it suffices to show that almost surely $A_{5,N}$ is summable along exponentially growing sequences of integers.

2.3.2. *Estimating $A_{5,N}$ (End of proof of Proposition 2.1).* Assuming that

$$b<c, \tag{22}$$

equation (19) gives that

$$\mathbb{E}_\omega(A_{5,N}) \leq N^{-1+4a-b-c}\sum_{m=1}^{N^b}\sum_{r=1}^{N^c}\sum_{k_1=0}^{2N^c}\sum_{k_2=0}^{N^c}\sum_{k_3=0}^{N^b}\mathbb{E}_\omega\left|\sum_{n=1}^{N-r} Y_n\cdot Z_{n,m,r,k_1,k_2,k_3}\right| \tag{23}$$

where

$$Z_{n,m,r,k_1,k_2,k_3} := Y_{n+m+r}\cdot Y_{n+r}\cdot Y_{n+m}\cdot \mathbf{1}_{\sum_{k=1}^{m+r}X_{n+k}=k_1}(n)\cdot \mathbf{1}_{\sum_{k=1}^{r}X_{n+k}=k_2}(n)\cdot \mathbf{1}_{\sum_{k=1}^{m}X_{n+k}=k_3}(n).$$

Using the Cauchy-Schwarz inequality we get

$$\mathbb{E}_\omega\left|\sum_{n=1}^{N-r} Y_n\cdot Z_{n,m,r,k_1,k_2,k_3}\right| \leq \left(\mathbb{E}_\omega\left|\sum_{n=1}^{N-r} Y_n\cdot Z_{n,m,r,k_1,k_2,k_3}\right|^2\right)^{1/2}. \tag{24}$$

We expand the square in order to compute its expectation. It is equal to

$$\sum_{1\leq n_1,n_2\leq N-r}\mathbb{E}_\omega(Y_{n_1}\cdot Z_{n_1,m,r,k_1,k_2,k_3}\cdot Y_{n_2}\cdot Z_{n_2,m,r,k_1,k_2,k_3}).$$

Notice that if $n_1<n_2$, then for every $m,r\in\mathbb{N}$, and non-negative integers k_1,k_2,k_3, the random variable Y_{n_1} is independent of the variables Y_{n_2}, Z_{n_1,m,r,k_1,k_2,k_3}, and Z_{n_2,m,r,k_1,k_2,k_3}. Since Y_n has zero mean, it follows that if $n_1\neq n_2$, then

$$\mathbb{E}_\omega(Y_{n_1}\cdot Z_{n_1,m,r,k_1,k_2,k_3}\cdot Y_{n_2}\cdot Z_{n_2,m,r,k_1,k_2,k_3})=0.$$

Therefore, the right hand side of equation (24) is equal to

$$\left(\sum_{n=1}^{N-r}\mathbb{E}_\omega(Y_n^2)\cdot\mathbb{E}_\omega(Z^2_{n,m,r,k_1,k_2,k_3})\right)^{1/2} \leq \left(\sum_{n=1}^{N-r}\mathbb{E}_\omega(Y_n^2)\cdot\mathbb{E}_\omega(Y^2_{n+m}\cdot Y^2_{n+r}\cdot Y^2_{n+m+r})\right)^{1/2}.$$

If r,m,n are fixed and $r\neq m$, then the variables $Y^2_{n+m},Y^2_{n+r},Y^2_{n+m+r}$ are independent, and as a consequence the right hand side is almost surely bounded by

$$\left(\sum_{n=1}^{N}\sigma_n^4\right)^{1/2} \ll N^{1/2-2a}.$$

On the other hand, if r, m, n are fixed and $r = m$, then the random variables Y_{n+r}^4, Y_{n+2r}^2 are independent, and as a consequence the right hand side is almost surely bounded by

$$\left(\sum_{n=1}^{N} \sigma_n^3\right)^{1/2} \ll N^{1/2-3a/2}.$$

Combining these two estimates with (23), we deduce that

$$\mathbb{E}_\omega(A_{5,N}) \ll N^{-1+4a-b-c}(N^{1/2-2a+2b+3c}+N^{1/2-3a/2+2b+2c}) = N^{-1/2+2a+b+2c}+N^{-1/2+5a/2+b+c}.$$

For fixed $\varepsilon > 0$, letting $a \in (0, 1/6)$, b be greater and very close to a, and c be greater and very close to $2a$, we get that the conditions (20) and (22) are satisfied, and

$$\mathbb{E}_\omega(A_{5,N}) \ll N^{(-1+14a)/2+\varepsilon} + N^{(-1+11a)/2+\varepsilon} = N^{-d_4} \tag{25}$$

for some d_4 that satisfies

$$d_4 > (1 - 14a)/2 - \varepsilon. \tag{26}$$

Therefore, for every $a \in (0, 1/14)$, if ε is small enough, then the estimates (21) and (25) hold for some $d_3, d_4 > 0$.

Equation (25) gives that for every $\gamma > 1$ we have

$$\sum_{N=1}^{\infty} \mathbb{E}_\omega(A_{5,[\gamma^N]}) < +\infty.$$

As a consequence, for every $\gamma > 1$ we have almost surely that

$$\sum_{N=1}^{\infty} A_{5,[\gamma^N]}(\omega) < +\infty. \tag{27}$$

Recalling the definition of A_N in (18), and combining (21) and (27), we get that for every $a \in (0, 1/14)$ and $\gamma > 1$, almost surely the following holds: For every probability space $(X, \mathcal{X}, \mu)$, commuting measure preserving transformations $T, S\colon X \to X$, and $f, g \in L^\infty(\mu)$, we have

$$\sum_{N=1}^{\infty} \left\| S_{[\gamma^N]}(\omega, \cdot)\right\|_{L^2(\mu)} < +\infty$$

where

$$S_N(\omega, \cdot) := N^{-1+a} \sum_{n=1}^{N} Y_n(\omega) \cdot T^{X_1+\cdots+X_n} f \cdot S^n g.$$

This finishes the proof of Proposition 2.1.

3. Convergence for the same random iterates

In this section we prove Theorem 1.4. Throughout, we use the notation introduced in Section 1.2.1 and the beginning of Section 2.2.

3.1. **Strategy of the proof.** In order to prove Theorem 1.4 we go through the following successive comparisons:

$$\begin{aligned}\frac{1}{N}\sum_{n=1}^{N} f(T^{a_n(\omega)}x)\cdot g(S^{a_n(\omega)}x) &\approx \frac{1}{W_N}\sum_{n=1}^{N} X_n(\omega)\cdot f(T^n x)\cdot g(S^n x)\\ &\approx \frac{1}{W_N}\sum_{n=1}^{N} \sigma_n\cdot f(T^n x)\cdot g(S^n x)\\ &\approx \frac{1}{N}\sum_{n=1}^{N} f(T^n x)\cdot g(S^n x),\end{aligned}$$

where our notation was explained in Section 2.1. The key comparison is the second. One needs to get good estimates for the L^2 norm of the averages $\frac{1}{W_N}\sum_{n=1}^{N} Y_n(\omega)\cdot T^n f\cdot S^n g$, where $Y_n := X_n - \sigma_n$. We do this in two steps. First we use van der Corput's estimate and Herglotz's theorem to get a bound that depends only on the random variables Y_n. The resulting expressions turn out to be random trigonometric polynomials that can be estimated using classical techniques.[5]

3.2. **A reduction.** Arguing as in Section 2.2 (in fact the argument is much simpler in the current case) we reduce Theorem 1.4 to proving the following result:

Proposition 3.1. *Suppose that $\sigma_n = n^{-a}$ for some $a \in (0,1/2)$ and let $\gamma > 1$ be a real number. Then almost surely the following holds: For every probability space $(X,\mathcal{X},\mu)$, commuting measure preserving transformations $T,S\colon X\to X$, and functions $f,g\in L^\infty(\mu)$, we have*

$$(28)\qquad \sum_{k=1}^{\infty}\left\|\frac{1}{W_{[\gamma^k]}}\sum_{n=1}^{[\gamma^k]} Y_n(\omega)\cdot T^n f\cdot S^n g\right\|_{L^2(\mu)} < +\infty$$

where $W_N := \sum_{n=1}^{N}\sigma_n$.

We prove this result in the next subsection.

3.3. **Proof of Proposition 3.1.** As was the case with the proof of Proposition 2.1 the proof of Proposition 3.1 splits in two parts.

3.3.1. *Eliminating the dependence on the transformations and the functions.* We assume that both functions f,g are bounded by 1. We start by using Lemma 4.4 for $M = N$ and $v_n := Y_n\cdot T^n f\cdot S^n g$. We get that

$$(29)\qquad A_N := \left\|N^{-1+a}\sum_{n=1}^{N} Y_n\cdot T^n f\cdot S^n g\right\|_{L^2(\mu)}^2 \ll A_{1,N} + A_{2,N}$$

where

$$A_{1,N} := N^{-2+2a}\cdot\sum_{n=1}^{N}\|Y_n\cdot T^n f\cdot S^n g\|_{L^2(\mu)}^2$$

[5] A faster way to get such an estimate is to apply van der Corput's Lemma twice. The drawback of this method is that the resulting expression converges to zero only when $\sigma_n = n^{-a}$ for some $a\in(0,1/4)$.

and

$$A_{2,N} := N^{-2+2a} \cdot \sum_{m=1}^{N} \left| \sum_{n=1}^{N-m} \int Y_{n+m} \cdot Y_n \cdot T^{n+m} f \cdot S^{n+m} g \cdot T^n \bar{f} \cdot S^n \bar{g} \, d\mu \right|.$$

We estimate $A_{1,N}$. Since $\mathbb{E}_\omega(Y_n^2) \sim n^{-a}$, Lemma 4.6 gives $\sum_{n=1}^{N} Y_n^2 \ll_\omega N^{1-a}$. It follows that almost surely we have

$$A_{1,N} \ll N^{-2+2a} \sum_{n=1}^{N} Y_n^2 \ll_\omega N^{-2+2a} \cdot N^{1-a} = N^{a-1}. \tag{30}$$

Therefore, $A_{1,N}$ is bounded by a negative power of N for every $a \in (0,1)$.

We estimate $A_{2,N}$. Composing with S^{-n} and using the Cauchy-Schwarz inequality we get

$$A_{2,N} \ll N^{-2+2a} \cdot \sum_{m=1}^{N} \left\| \sum_{n=1}^{N-m} Y_{n+m} \cdot Y_n \cdot S^{-n} T^{n+m} f \cdot S^{-n} T^n \bar{f} \right\|_{L^2(\mu)}.$$

Using that T and S commute and letting $R = TS^{-1}$ and $f_m = T^m f \cdot \bar{f}$, we rewrite the previous estimate as

$$A_{2,N} \ll N^{-2+2a} \cdot \sum_{m=1}^{N} \left\| \sum_{n=1}^{N-m} Y_{n+m} \cdot Y_n \cdot R^n f_m \right\|_{L^2(\mu)}.$$

Using Herglotz theorem on positive definite sequences, and the fact that the functions f_m are uniformly bounded, we get that the right hand side is bounded by a constant multiple of

$$A_{3,N} := N^{-1+2a} \cdot \max_{1 \le m \le N} \max_{t \in [0,1]} \left| \sum_{n=1}^{N-m} Y_{n+m} \cdot Y_n \cdot e(nt) \right|.$$

Summarizing, we have shown that

$$A_N \ll N^{a-1} + A_{3,N}. \tag{31}$$

Therefore, in order to prove Proposition 3.1 it remains to show that almost surely $A_{3,N} \ll_\omega N^{-d}$ for some $d > 0$. We do this in the next subsection.

3.3.2. *Estimating $A_{3,N}$ (End of proof of Proposition 3.1).* The goal of this section is to prove the following result:

Proposition 3.2. *Suppose that $\sigma_n \sim n^{-a}$ for some $a \in (0, 1/2)$. Then almost surely we have*

$$\max_{1 \le m \le N} \max_{t \in [0,1]} \left| \sum_{n=1}^{N-m} Y_{n+m} \cdot Y_n \cdot e(nt) \right| \ll_\omega N^{1/2-a} \sqrt{\log N},$$

Notice that by combining this estimate with (31) we get a proof of Proposition 3.1, and as a consequence a proof of Theorem 1.4.

The key ingredient in the proof of Proposition 3.2 is a strengthening of an estimate of Bourgain [6] regarding random trigonometric polynomials. To prove it we are going to use a variant of a classical argument of Salem and Zygmund (see Section 6.2 of [18]). We were motivated to use this argument, over the one given in [6], after reading a paper of Fan and Schneider (in particular, the proof of Theorem 6.4 in [13]).

Lemma 3.3. *For $m, N \in \mathbb{N}$ let $\Lambda_{m,N}$ be (deterministic) subsets of the intervals $[1, N]$. Furthermore, let $(X_{m,n})_{m,n\in\mathbb{N}}$ be a family of random variables with values 0 or 1, and let $\rho_n := \sup_{m\in\mathbb{N}} \mathbb{P}(X_{m,n} = 1)$ and $W_N := \sum_{n=1}^{N} \rho_n$. Suppose that*

- *for every fixed $m \in \mathbb{N}$ the random variables $X_{m,1}, X_{m,2}, \dots$ are independent;*
- $\lim_{N\to\infty} \frac{W_N}{\log N} = +\infty$.

Then almost surely we have

$$\max_{1\le m\le N} \max_{t\in[0,1]} \left| \sum_{n\in\Lambda_{m,N}} (X_{m,n} - \mathbb{E}_\omega(X_{m,n})) \cdot e(nt) \right| \ll_\omega \sqrt{\log N \cdot W_N}.$$

Proof. It suffices to get the announced estimate for

$$M_N := \max_{1\le m\le N} \max_{t\in[0,1]} |P_{m,N}(t)|$$

where

$$P_{m,N}(t) := \sum_{n\in\Lambda_{m,N}} Y_{m,n} \cdot \cos(2\pi nt)$$

and

$$Y_{m,n} := X_{m,n} - \mathbb{E}_\omega(X_{m,n}).$$

In a similar way we get an estimate with $\sin(2\pi nt)$ in place of $\cos(2\pi nt)$.

Let

$$\rho_{m,n} := \mathbb{E}_\omega(X_{m,n}).$$

Since $1 + x \le e^x$ for every $x \in \mathbb{R}$, and $e^x \le 1 + x + (e/2)x^2$ for every $x \in (-\infty, 1]$, for fixed $m \in [1, N]$ and $\lambda \in [-1, 1]$ we have

$$\mathbb{E}_\omega(e^{\lambda Y_{m,n}}) = e^{-\lambda\rho_{m,n}}(1 + \rho_{m,n}(e^\lambda - 1)) \le e^{\rho_{m,n}(e^\lambda - 1 - \lambda)} \le e^{A\rho_{m,n}\lambda^2} \le e^{A\rho_n\lambda^2}$$

where $A := e/2$. Therefore, for every $\lambda \in [-1, 1]$ and $t \in [0, 1]$ we get that

$$\mathbb{E}_\omega(e^{\lambda P_{m,N}(t)}) = \prod_{n\in\Lambda_{m,N}} \mathbb{E}_\omega(e^{\lambda Y_{m,n}\cos(2\pi nt)}) \le \prod_{n\in\Lambda_{m,N}} e^{A\rho_n(\lambda\cos(2\pi nt))^2} \le e^{AW_N\lambda^2}. \tag{32}$$

Next notice that for $\lambda \in [0, 1]$ we have

(33)

$$\mathbb{E}_\omega(e^{\lambda M_N}) = \mathbb{E}_\omega\big(\max_{1\le m\le N} e^{\lambda \max_t |P_{m,N}(t)|}\big) \le \mathbb{E}_\omega\Big(\sum_{m=1}^{N} e^{\lambda \max_t |P_{m,N}(t)|}\Big) \le N \max_{1\le m\le N} \mathbb{E}_\omega(e^{\lambda M_{m,N}})$$

where

$$M_{m,N} := \max_{t\in[0,1]} |P_{m,N}(t)|.$$

As is well known, $\max_{t\in[0,1]} |P'_{m,N}(t)| \le N^2 M_{m,N}$. Therefore, there exist random intervals $I_{m,N}$ of length $|I_{m,N}| \ge N^{-2}$ such that $|P_{m,N}(t)| \ge M_{m,N}/2$ for every $t \in I_{m,N}$. Using this, we get that

$$\mathbb{E}_\omega(e^{\lambda_N M_{m,N}/2}) \ll N^2 \cdot \mathbb{E}_\omega\Big(\int_{I_{m,N}} (e^{\lambda_N P_{m,N}(t)} + e^{-\lambda_N P_{m,N}(t)})\ dt\Big) \leq$$
$$N^2 \cdot \mathbb{E}_\omega\Big(\int_{[0,1]} (e^{\lambda_N P_{m,N}(t)} + e^{-\lambda_N P_{m,N}(t)})\ dt\Big)$$

where $\lambda_N \in [0,1]$ are numbers at our disposal. Using (32) we get that

$$\mathbb{E}_\omega\Big(\int_{[0,1]} (e^{\lambda_N P_{m,N}(t)} + e^{-\lambda_N P_{m,N}(t)})\ dt\Big) = \int_{[0,1]} \mathbb{E}_\omega\big(e^{\lambda_N P_{m,N}(t)} + e^{-\lambda_N P_{m,N}(t)}\big)\ dt \leq 2e^{AW_N\lambda_N^2}.$$

Therefore,

$$\mathbb{E}_\omega(e^{\lambda_N M_{m,N}/2}) \ll N^2 \cdot e^{AW_N\lambda_N^2}.$$

Combining this estimate with (33), we get

$$\mathbb{E}_\omega(e^{\lambda_N M_N/2}) \ll N^3 \cdot e^{AW_N\lambda_N^2}.$$

Therefore, there exists a universal constant C such that

$$\mathbb{E}_\omega\Big(\exp\big(\lambda_N/2(M_N - 2AW_N\lambda_N - 2\log(CN^5)\lambda_N^{-1})\big)\Big) \leq \frac{1}{N^2}.$$

As a consequence,

$$\mathbb{P}\big(M_N \geq 2AW_N\lambda_N + 2\log(CN^5)\lambda_N^{-1}\big) \leq \frac{1}{N^2}. \tag{34}$$

For α, β positive, the function $f(\lambda) = \alpha\lambda + \beta\lambda^{-1}$ achieves a minimum $\sqrt{\alpha\beta}$ for $\lambda = \sqrt{\beta/\alpha}$. So letting $\lambda_N = \sqrt{\log(CN^5)/(AW_N)}$ (by assumption this converges to 0, so for large N it is less than 1) in (34) gives

$$\mathbb{P}\Big(M_N \geq \sqrt{4AW_N\log(CN^5)}\Big) \leq \frac{1}{N^2}.$$

By the Borel-Cantelli Lemma, we get almost surely that

$$M_N \ll_\omega \sqrt{W_N \log N}.$$

This completes the proof. □

Finally we use Lemma 3.3 to prove Proposition 3.2.

Proof of Proposition 3.2. Our goal is to apply Lemma 3.3 for the random variables $Y_{n+m} \cdot Y_n$ where $Y_n = X_n - \sigma_n$. In order to do this we have to take care of some technical issues first.

To get random variables with values 0 or 1, notice that the identity

$$Y_{n+m} \cdot Y_n = (X_{n+m}X_n - \sigma_{n+m}\sigma_n) - \sigma_n(X_{n+m} - \sigma_{n+m}) - \sigma_{n+m}(X_n - \sigma_n)$$

expresses each $Y_{n+m} \cdot Y_n$ as a linear combination of random variables of the form $Z - \mathbb{E}_\omega(Z)$ where Z is $0-1$ valued.

To get independence, we divide the positive integers into two classes:

$$\Lambda_{1,m} := \{n\colon 2km < n \leq (2k+1)m \ \text{ for some non-negative integer } k\}$$

and

$$\Lambda_{2,m} := \{n\colon (2k+1)m < n \leq (2k+2)m \ \text{ for some non-negative integer } k\}.$$

Then for fixed $m \in \mathbb{N}$, the random variables $X_{n+m}X_n$, $n \in \Lambda_{1,m}$, are independent, and the same holds for the random variables $X_{n+m}X_n$, $n \in \Lambda_{2,m}$.

Summarizing, the sum

$$\sum_{n=1}^{N-m} Y_{n+m} \cdot Y_n \cdot e(nt)$$

splits into four pieces each of which can be estimated by using Lemma 3.3. We only explain how to estimate

$$M_N := \max_{1\leq m\leq N} \max_{t\in[0,1]} \Big| \sum_{n\in\Lambda_{1,m}\cap[1,N-m]} (X_{n+m}X_n - \sigma_{n+m}\sigma_n) \cdot e(nt)\Big|,$$

the other three pieces can be estimated in a similar fashion. We set $X_{m,n} := X_{n+m}X_n$ and $\Lambda_{m,N} := \Lambda_{1,m} \cap [1, N-m]$. Notice that $\mathbb{E}_\omega(X_{m,n}) = \sigma_{n+m}\sigma_n \leq \sigma_n^2 \sim n^{-2a}$ and $\sum_{n=1}^N n^{-2a} \ll N^{1-2a}$. Lemma 3.3 gives that almost surely we have

$$M_N \ll_\omega N^{1/2-a}\sqrt{\log N}.$$

This completes the proof. □

4. Appendix

We prove some results that were used in the main part of the article.

4.1. **Lacunary subsequence trick.** We are going to give a variation of a trick that is often used to prove convergence results for averages (see [25] for several such instances).

Lemma 4.1. *Let $(a_n)_{n\in\mathbb{N}}$ be a sequence of non-negative real numbers, and $(W_n)_{n\in\mathbb{N}}$ be an increasing sequence of positive real numbers that satisfies*

$$\lim_{\gamma\to1^+} \limsup_{n\to\infty} \frac{W_{[\gamma n]}}{W_n} = 1.$$

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

$$A_N := \frac{1}{W_N}\sum_{n=1}^N a_n.$$

Suppose that there exists $L \in \mathbb{R} \cup \{+\infty\}$, and a sequence of real numbers $\gamma_k \in (1, +\infty)$, with $\gamma_k \to 1$, and such that for every $k \in \mathbb{N}$ we have

$$\lim_{N\to\infty} A_{[\gamma_k^N]} = L.$$

Then

$$\lim_{N\to\infty} A_N = L.$$

Proof. Fix $k \in \mathbb{N}$ and for $N \in \mathbb{N}$ let $M = M(k, N) \in \mathbb{N}$ be such that

$$\gamma_k^M \leq N \leq \gamma_k^{M+1}.$$

Since $a_n \geq 0$ for every $n \in \mathbb{N}$ and W_n is increasing, we have

$$A_N = \frac{1}{W_N}\sum_{n=1}^N a_n \leq \frac{1}{W_{\gamma_k^M}} \sum_{n=1}^{\gamma_k^{M+1}} a_n \leq c_{k,M} A_{[\gamma_k^{M+1}]} \quad \text{where} \quad c_{k,M} := W_{[\gamma_k^{M+1}]}/W_{\gamma_k^M}.$$

Similarly we have

$$A_N \geq c_{k,M}^{-1} A_{[\gamma_k^M]}.$$

Putting the previous estimates together we get

$$c_{k,M}^{-1} A_{[\gamma_k^M]} \leq A_N \leq c_{k,M} A_{[\gamma_k^{M+1}]}. \tag{35}$$

Notice that our assumptions give that

$$\lim_{k\to\infty} \limsup_{M\to\infty}\ c_{k,M} = 1. \tag{36}$$

Since $M = M(k,N) \to \infty$ as $N \to \infty$ and k is fixed, letting $N \to \infty$ and then $k \to \infty$ in (35), and combining equation (36) with our assumption $\lim_{N\to\infty} A_{[\gamma_k^N]} = L$, we deduce that

$$\liminf_{N\to\infty} A_N = \limsup_{N\to\infty} A_N = L.$$

This completes the proof. □

Corollary 4.2. *Let $(X, \mathcal{X}, \mu)$ be a probability space, $f_n \colon X \to \mathbb{R}$, $n \in \mathbb{N}$, be non-negative measurable functions, $(W_n)_{n\in\mathbb{N}}$ be as in the previous lemma, and for $N \in \mathbb{N}$ let*

$$A_N(x) := \frac{1}{W_N} \sum_{n=1}^{N} f_n(x).$$

Suppose that there exists a function $f \colon X \to \mathbb{R}$ and a sequence of real numbers $\gamma_k \in (1,\infty)$, with $\gamma_k \to 1$, and such that for every $k \in \mathbb{N}$ we have for almost every $x \in X$ that

$$\lim_{N\to\infty} A_{[\gamma_k^N]}(x) = f(x). \tag{37}$$

Then

$$\lim_{N\to\infty} A_N(x) = f(x) \quad \text{for almost every } x \in X.$$

Proof. It suffices to notice that for almost every $x \in X$ equation (37) is satisfied for every $k \in \mathbb{N}$, and then apply Lemma 4.1. □

4.2. **Weighted averages.** The following lemma is classical and can be proved using summation by parts (also the assumptions on the weights w_n can be weakened).

Lemma 4.3. *Let $(v_n)_{n\in\mathbb{N}}$ be a bounded sequence of vectors in a normed space, $(w_n)_{n\in\mathbb{N}}$ be a decreasing sequence of positive real numbers that satisfies $w_n \sim n^{-a}$ for some $a \in (0,1)$, and for $N \in \mathbb{N}$ let $W_N := w_1 + \cdots + w_N$. Then the averages $\frac{1}{N}\sum_{n=1}^{N} v_n$ and the averages $\frac{1}{W_N}\sum_{n=1}^{N} w_n v_n$ are asymptotically equal.*

4.3. **Van der Corput's lemma.** We state a variation of a classical elementary estimate of van der Corput.

Lemma 4.4. *Let V be an inner product space, $N \in \mathbb{N}$, and $v_1, \ldots, v_N \in V$. Then for every integer M between 1 and N we have*

$$\Big\| \sum_{n=1}^{N} v_n \Big\|^2 \leq 2M^{-1}N \cdot \sum_{n=1}^{N} \|v_n\|^2 + 4M^{-1}N \sum_{m=1}^{M} \Big| \sum_{n=1}^{N-m} < v_{n+m}, v_n > \Big|.$$

In the case where $V = \mathbb{R}$ and $\|\cdot\| = |\cdot|$, the proof can be found, for example, in [19]. The proof in the general case is essentially identical.

4.4. **Borel-Cantelli in density.** We are going to use the following Borel-Cantelli type lemma:

Lemma 4.5. *Let E_n, $n \in \mathbb{N}$, be events on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ that satisfy $\mathbb{P}(E_n) \ll (\log n)^{-1-\varepsilon}$ for some $\varepsilon > 0$. Then almost surely the set $\{n \in \mathbb{N} \colon \omega \in E_n\}$ has zero density.*[6]

Proof. Let

$$A_N(\omega) := \frac{1}{N} \sum_{n=1}^{N} \mathbf{1}_{E_n}(\omega).$$

Our assumption gives

$$\mathbb{E}_\omega(A_N(\omega)) \ll (\log N)^{-1-\varepsilon}.$$

Therefore, for every $\gamma > 1$

$$\sum_{N=1}^{\infty} A_{[\gamma^N]}(\omega) < +\infty$$

almost surely. This implies that for every $\gamma > 1$

$$\lim_{N\to\infty} A_{[\gamma^N]}(\omega) = 0 \quad \text{almost surely.}$$

Since $\gamma > 1$ is arbitrary we conclude by Corollary 4.2 that

$$\lim_{N\to\infty} A_N(\omega) = 0 \quad \text{almost surely.}$$

This proves the advertised claim. □

4.5. **Estimates for sums of random variables.** We use some straightforward moment estimates to get two bounds for sums of independent random variables that were used in the proofs.

Lemma 4.6. *Let X_n be non-negative, uniformly bounded, independent random variables, with $\mathbb{E}_\omega(X_n) \sim n^{-a}$ for some $a \in (0, 1)$. Suppose that the random variables $X_n - \mathbb{E}_\omega(X_n)$ are orthogonal. Then almost surely we have*

$$\lim_{N\to\infty} \frac{1}{W_N} \sum_{n=1}^{N} X_n = 1,$$

where, as usual, $W_N := \sum_{n=1}^{N} \mathbb{E}_\omega(X_n)$.

Proof. We can assume that $X_n(\omega) \leq 1$ for every $\omega \in \Omega$ and $n \in \mathbb{N}$. We let

$$A_N := \frac{1}{W_N} \sum_{n=1}^{N} Y_n$$

where

$$Y_n := X_n - \mathbb{E}_\omega(X_n).$$

Since Y_n are zero mean orthogonal random variables and $\mathbb{E}_\omega(Y_n^2) \leq \mathbb{E}_\omega(X_n)$, we have

$$\mathbb{E}_\omega(A_N^2) = \frac{1}{W_N^2} \sum_{n=1}^{N} \mathbb{E}_\omega(Y_n^2) \ll \frac{1}{W_N}.$$

[6]On the other hand, it is not hard to construct a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and events E_n, $n \in \mathbb{N}$, such that $\mathbb{P}(E_n) \leq (\log n)^{-1}$, and almost surely the set $\{n \in \mathbb{N} \colon \omega \in E_n\}$ has positive upper density.

Combining this estimate with the fact $W_N \sim N^{1-a}$, we conclude that for every $\gamma > 1$ we have

$$\sum_{N=1}^{\infty} \mathbb{E}_\omega(A^2_{[\gamma^N]}) < +\infty.$$

Therefore, for every $\gamma > 1$ we have

$$\lim_{N\to\infty} A_{[\gamma^N]} = 0 \quad \text{almost surely},$$

or equivalently, that

$$\lim_{N\to\infty} \frac{1}{W_{[\gamma^N]}} \sum_{n=1}^{[\gamma^N]} X_n = 1 \quad \text{almost surely}.$$

Since the sequence $(W_n)_{n\in\mathbb{N}}$ satisfies the assumptions of Corollary 4.2, and X_n is non-negative, we conclude that

$$\lim_{N\to\infty} \frac{1}{W_N} \sum_{n=1}^{N} X_n = 1 \quad \text{almost surely}.$$

This completes the proof. □

Lemma 4.7. *Let X_n be uniformly bounded random variables with $\mathbb{E}_\omega(X_n) \sim n^{-a}$ for some $a \in (0, 1/6)$, and let b be any positive real number. Then almost surely we have*

$$\Big| \sum_{m=1}^{N^b} \sum_{n=1}^{N} X_{n+m} X_n \Big| \ll_\omega N^{b+1-2a}.$$

We remark that using a subsequence trick, similar to the one used in the proof of Lemma 4.6, one shows that the conclusion actually holds for every $a \in (0, 1/4)$.

Proof. Let

$$S_N := \sum_{m=1}^{N^b} \sum_{n=1}^{N} (X_{n+m} X_n - \mathbb{E}_\omega(X_{n+m}) \cdot \mathbb{E}_\omega(X_n))$$

and

$$A_N := N^{-c} S_N \quad \text{ where } \quad c := b + 1 - 2a,$$

Since

$$N^{-c} \sum_{m=1}^{N^b} \sum_{n=1}^{N} \mathbb{E}_\omega(X_{n+m}) \cdot \mathbb{E}_\omega(X_n) \ll 1,$$

it suffices to show that almost surely we have $\lim_{N\to\infty} A_N = 0$.

Expanding S_N^2 and using the independence of the random variables X_n, we see that

$$\mathbb{E}_\omega(S_N^2) \ll |\{(m, m', n, n') \in [1, N^b]^2 \times [1, N]^2 \colon n, n', n+m, n'+m' \text{ are not distinct } \}| \ll N^{1+2b}.$$

Therefore,

$$\mathbb{E}_\omega(A_N^2) \ll N^{-(1-4a)}.$$

It follows that if $k \in \mathbb{N}$ satisfies $k(1 - 4a) > 1$, then

$$\sum_{N=1}^{\infty} \mathbb{E}_\omega(A^2_{N^k}) < +\infty.$$

As a consequence,

$$\lim_{N\to\infty} A_{N^k} = 0 \quad \text{for every } k \in \mathbb{N} \text{ satisfying } \ k(1-4a) > 1.$$

For any fixed $k \in \mathbb{N}$ that satisfies $k(1-4a) > 1$, and for $N \in \mathbb{N}$, let $M \in \mathbb{N}$ be an integer such that $M^k \leq N \leq (M+1)^k$. Then

$$\begin{aligned} |A_N - A_{M^k}| &\leq |(N^{-c}M^{kc} - 1)A_{M^k}| + N^{-c} \sum_{M^k < n \leq (M+1)^k} |Y_n| \\ &\ll |(N^{-c}M^{kc} - 1)A_{M^k}| + N^{-c}M^{k-1}. \end{aligned}$$

The first term converges almost surely to zero as $N \to \infty$, since this is the case for A_{M^k} and $N^{-1}M^k \leq 1$. The second term converges to zero if $kc > k-1$, or equivalently, if $k(2a-b) < 1$.

Combining the above estimates, we get almost surely that $\lim_{N\to\infty} A_N = 0$, provided that there exists $k \in \mathbb{N}$ such that $k(2a-b) < 1 < k(1-4a)$. If $a < 1/6$, then $k = 3$ is such a value. This completes the proof. □

References

[1] I. Assani. Pointwise convergence of nonconventional averages. *Colloq. Math.* **102** (2005), no. 2, 245-262.

[2] B. Berend. Joint ergodicity and mixing. *J. Analyse Math.* **45** (1985), 255–284.

[3] B. Berend. Multiple ergodic theorems. *J. Analyse Math.* **50** (1988), 123–142.

[4] M. Boshernitzan. Homogeneously distributed sequences and Poincaré sequences of integers of sublacunary growth. *Monatsh. Math.* **96** (1983), no. 3, 173–181.

[5] M. Boshernitzan, G. Kolesnik, A. Quas, M. Wierdl. Ergodic averaging sequences. *J. Analyse Math.* **95** (2005), 63–103.

[6] J. Bourgain. On the maximal ergodic theorem for certain subsets of the positive integers. *Israel J. Math.* **61** (1988), 39–72.

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

[8] Q. Chu. Multiple recurrence for two commuting transformations. To appear in *Ergodic Theory Dynam. Systems*. Available at arXiv:0912.3381

[9] Q. Chu, N. Franzikinakis, B. Host. Ergodic averages of commuting transformations with distinct degree polynomial iterates. To appear in *Proc. Lond. Math. Soc.* Available at arXiv:0912.2641

[10] J-P. Conze, E. Lesigne. Théorèmes ergodiques pour des mesures diagonales. *Bull. Soc. Math. France* **112** (1984), no. 2, 143–175.

[11] C. Demeter. Pointwise convergence of the ergodic bilinear Hilbert transform. *Illinois J. Math.* **51** (2007), no. 4, 1123-1158.

[12] J-M. Derrien, E. Lesigne. Un théorème ergodique polynomial ponctuel pour les endomorphismes exacts et les K-systèmes. *Ann. Inst. H. Poincaré Probab. Statist.* **32** (1996), no. 6, 765-778.

[13] A. Fan, D. Schneider. Recurrence properties of sequences of integers. *Sci. China Math.* 53 (2010), no. 3, 641–656.

[14] N. Frantzikinakis. Multiple recurrence and convergence for Hardy sequences of polynomial growth. To appear in *J. Analyse Math.* Available at arXiv:0903.0042

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

[16] H. Furstenberg, Y. Katznelson. An ergodic Szemerédi theorem for commuting transformations. *J. Analyse Math.* **34** (1979), 275–291.

[17] B. Host, B. Kra. Convergence of polynomial ergodic averages. *Isr. J. Math.* **149** (2005), 1–19.

[18] J-P. Kahane. Some random series of functions. Second edition. *Cambridge Studies in Advanced Mathematics, 5.* Cambridge University Press, Cambridge, 1985.

[19] L. Kuipers, H. Niederreiter. Uniform distribution of sequences. Pure and Applied Mathematics. *Wiley-Interscience*, New York-London-Sydney, (1974).

[20] M. Lacey, K. Petersen, M. Wierdl, D. Rudolph. Random ergodic theorems with universally representative sequences. *Ann. Inst. H. Poincare Probab. Statist.* **30** (1994), no. 3, 353-395.

[21] E. Lesigne. Équations fonctionnelles, couplages de produits gauches et théorèmes ergodiques pour mesures diagonales. *Bull. Soc. Math. France* **121** (1993), no. 3, 315–351.

[22] E. Lesigne, B. Rittaud, T. de la Rue. Weak disjointness of measure preserving dynamical systems. *Ergodic Theory Dynam. Systems* **23** (2003), 1173–1198.

[23] A. Leibman. Pointwise convergence of ergodic averages for polynomial sequences of rotations of a nilmanifold. *Ergodic Theory Dynam. Systems* **25** (2005), no. 1, 201–213.

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

[25] J. Rosenblatt, M. Wierdl. Pointwise ergodic theorems via harmonic analysis. Ergodic theory and its connections with harmonic analysis (Alexandria, 1993). *London Math. Soc. Lecture Note Ser.*, **205**, Cambridge Univ. Press, Cambridge, (1995), 3–151.

[26] T. Tao. Norm convergence of multiple ergodic averages for commuting transformations. *Ergodic Theory Dynam. Systems* **28** (2008), no. 2, 657–688.

[27] M. Wierdl. Almost everywhere convergence and recurrence along subsequences in ergodic theory. PhD Thesis, Ohio State University, 1989.

(Nikos Frantzikinakis) University of Crete, Department of mathematics, Knossos Avenue, Heraklion 71409, Greece
E-mail address: `frantzikinakis@gmail.com`

(Emmanuel Lesigne) Laboratoire de Mathématiques et Physique Théorique (UMR CNRS 6083), Université François Rabelais Tours, Fédération de Recherche Denis Poisson, Parc de Grandmont, 37200 Tours, France
E-mail address: `emmanuel.lesigne@lmpt.univ-tours.fr`

(Máté Wierdl) Department of Mathematics, University of Memphis, Memphis, TN, 38152, USA
E-mail address: `wierdlmate@gmail.com`