Illinois Journal of Mathematics
Volume 59, Number 3, Fall 2015, Pages 663–674
S 0019-2082

A RANDOM POINTWISE ERGODIC THEOREM WITH HARDY FIELD WEIGHTS

BEN KRAUSE AND PAVEL ZORIN-KRANICH

ABSTRACT. Let a_n be the random increasing sequence of natural numbers which takes each value independently with probability n^{-a}, $0 < a < 1/2$, and let $p(n) = n^{1+\varepsilon}$, $0 < \varepsilon < 1$. We prove that, almost surely, for every measure-preserving system (X,T) and every $f \in L^1(X)$ the modulated, random averages

$$\frac{1}{N}\sum_{n=1}^{N} e\big(p(n)\big)T^{a_n(\omega)}f$$

converge to 0 pointwise almost everywhere.

1. Introduction

A sequence of integers $\{n_k\} \subset \mathbb{Z}$ is said to be *universally L^p-good* if for every measure-preserving system (X,μ,T) and every $f \in L^p(X)$ the subsequence averages

$$A_N^{\{n_k\}}f := \frac{1}{N}\sum_{k=1}^{N} T^{n_k}f$$

converge pointwise almost everywhere. In this language, Birkhoff's classical pointwise ergodic theorem [Bir31] states that the full sequence of integers is universally L^1-good.

Obtaining pointwise convergence results for rougher, sparser sequences is much more challenging. For instance, Bourgain's Polynomial Ergodic Theorem [Bou89] states that the sequence $\{P(n)\}$, P integer polynomial, is universally L^p-good for each $p > 1$. Note that $\{P(n)\}$ are zero-Banach-density subsequences of the integers; in fact, Bourgain used a probabilistic method to find *extremely* sparse universally good sequences. From now on, $\{X_n\}$ will

Received January 12, 2015; received in final form March 24, 2016.
2010 *Mathematics Subject Classification.* 37A30.

denote a sequence of independent $\{0,1\}$ valued random variables (on a probability space Ω) with expectations σ_n. The *counting function* $a_n(\omega)$ is the smallest integer subject to the constraint

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

THEOREM 1.1 ([Bou88, Proposition 8.2]). *Suppose*

$$\sigma_n=\frac{(\log\log n)^{B_p}}{n}, \quad B_p>\frac{1}{p-1}, 1<p\le 2.$$

Then, almost surely, $\{a_n\}$ is universally L^p-good.

In the years to follow random sequences became a widely used model for pointwise ergodic theorems. One indication at their amenability to analysis is LaVictoire's L^1 random ergodic theorem.

THEOREM 1.2 ([LaV09]). *Suppose $\sigma_n=n^{-a}$ with $0<a<1/2$. Then, almost surely, $\{a_n\}$ is universally L^1-good.*

Here, by the strong law of large numbers, almost surely

$$a_n(\omega)/n^{\frac{1}{1-a}}$$

converges to a non-zero number. For comparison, it is known that the sequences of dth powers, $d>1$ integer, are universally L^1 bad [BM07], [LaV11].

Random sequences have also been used as a model for multiple ergodic averages. Frantzikinakis, Lesigne, and Wierdl recently showed the following.

THEOREM 1.3 ([FLW12, Theorem 1.1]). *Suppose $\sigma_n=n^{-a}$, $0<a<1/14$. Then, almost surely, $(a_n)_n$ has the following property: for every pair of measure preserving transformations T,S on a probability space X and any functions $f,g\in L^\infty(X)$ the averages*

$$\sum_{n=1}^{N} g\big(S^n x\big) f\big(T^{a_n}x\big)$$

converge pointwise almost everywhere.

It is noted in their paper that the linear sequence of powers S^n can likely be replaced by other deterministic sequences, but their method of proof did not seem to allow this. In this article, we prove a related result in which we are able to replace the linear sequence of powers by a sequence drawn from a more general class at the cost of weakening the result in several other respects. More precisely, with $0<\varepsilon<1$ arbitrary but fixed, suppose $p:\mathbb{R}\to\mathbb{R}$ is a *logarithmico-exponential* function which satisfies

(1) the second-order difference relationship

$$p(x+y+z)-p(x+y)-p(x+z)+p(x)=O\big(x^{\varepsilon-1}yz\big)$$

for $x,y,z>0$ ("big-O" notation is recalled in the section on notation below); and

(2) for all $a(x) \in C \cdot \mathbb{Q}[x]$, the set of real constant multiples of rational polynomials, $\frac{|a(x)-p(x)|}{\log x} \to \infty$.

Good examples of such functions are $p(x) = x^{1+\varepsilon}$. We refer the reader to [B+05] for a more complete discussion of *logarithmico-exponential* functions; informally, these are all the functions one can get by combining real constants, the variable x, and the symbols exp, log, $\cdot$, and $+$ (e.g., $x^{1/2} = \exp(1/2 \cdot \log x)$ and $x^\pi / \log\log x$ are both logarithmico-exponential).

Our main result is the following theorem.

THEOREM 1.4. *Suppose $\sigma_n = n^{-a}$, $0 < a < 1/2$, and p is as above. Then, almost surely, the following holds*:

For each measure-preserving system (X, μ, T) and each $f \in L^1(X)$ the averages

$$\frac{1}{N}\sum_{n=1}^{N} e\big(p(n)\big) T^{a_n(\omega)} f$$

converge to zero pointwise almost everywhere (*here and later $e(t) := e^{2\pi i t}$*).

Pointwise ergodic theorems with exponential polynomial weights are collectively known as Wiener–Wintner type theorems, see, for example, [Ass03] for linear polynomials and [Les93] for general polynomials. If the random sequence $\{a_n\}$ is replaced by the linear sequence $\{n\}$ in Theorem 1.4, the result follows from the Wiener–Wintner theorem for Hardy field functions due to Eisner and the first author [EK15]. However, note that the full measure sets in our result depend on the choice of p. It would be interesting to remove this dependence. Also, the second order difference relation in the hypothesis of Theorem 1.4 can likely be replaced by a polynomial growth assumption; this would require an inductive application of van der Corput's inequality.

The structure of this paper is as follows: In Section 2, we introduce a few preliminary tools, discuss our proof strategy, and reduce our theorem to proving Proposition 2.11; and in Section 3, we prove Proposition 2.11, thereby completing the proof of Theorem 1.4.

2. Preliminaries

2.1. Notation and tools. With X_n, σ_n as above, we let $Y_n := X_n - \sigma_n$.

We will be dealing with sums of random variables, so we introduce the following compact notation:

$$S_N = \sum_{n=1}^{N} X_N \quad \text{and} \quad S_{M,N} = \sum_{n=M}^{N} X_n.$$

We also let

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

so that W_N grows as N^{1-a}.

We will make use of the modified Vinogradov notation. We use $X \lesssim Y$, or $Y \gtrsim X$ to denote the estimate $X \leq CY$ for an absolute constant C. If we need C to depend on a parameter, we shall indicate this by subscripts, thus for instance $X \lesssim_\omega Y$ denotes the estimate $X \leq C_\omega Y$ for some C_ω depending on ω.

We also make use of big-O notation: we let $O(Y)$ denote a quantity that is $\lesssim Y$, and similarly $O_\omega(Y)$ a quantity that is $\lesssim_\omega Y$.

The main probabilistic input in our argument is the following special case of Chernoff's inequality.

Lemma 2.1 (See, e.g., [TV10]). *Let $\{X_n\}$, $\{\sigma_n\}$ be as above. There exists an absolute constant $c > 0$ so that for each $A > 0$,*

$$\mathbb{P}\big(|S_N - W_N| \geq A\big) \lesssim \max\left\{\exp\left(-c\frac{A^2}{W_N}\right), \exp(-cA)\right\}.$$

Consequently,

$$\mathbb{P}\left(|S_N - W_N| \geq \frac{1}{2}W_N\right) \lesssim \exp(-cW_N) \lesssim \exp\big(-cN^{1-a}\big).$$

This also implies the following version of the law of large numbers:

$$S_N/W_N \to 1 \quad \text{almost surely.} \tag{2.2}$$

We will also need the Hilbert space van der Corput inequality.

Lemma 2.3 (See, e.g., [FLW12]). *Let $\{v_n\}$ be a sequence in a Hilbert space H and $1 \leq M \leq N$. Then*

$$\left\|\sum_{n=1}^N v_n\right\|^2 \leq 2\frac{N}{M}\sum_{n=1}^N \|v_n\|^2 + 4\frac{N}{M}\sum_{m=1}^M \left|\sum_{n=1}^{N-m}\langle v_{n+m}, v_n\rangle\right|. \tag{2.4}$$

2.2. Strategy. In proving his Random Ergodic theorem, LaVictoire showed that on a set of full probability, $\Omega' \subset \Omega$, the maximal function

$$f \mapsto \sup_N \frac{1}{N}\sum_{n=1}^N T^{a_n(\omega)}|f|,$$

is weakly bounded on $L^1(X)$ [LaV09]. In particular, for $\omega \in \Omega'$ the set of $f \in L^1(X)$ for which the averages

$$\frac{1}{N}\sum_{n=1}^N e\big(p(n)\big)T^{a_n(\omega)}f$$

tend to zero pointwise a.e. is closed in L^1. Hence it will be enough to prove pointwise convergence for $f \in L^\infty(X)$. Now, as observed in [RW95], for bounded functions it is enough to prove convergence along every lacunary sequence $\lfloor \rho^{\mathbb{N}} \rfloor = (\lfloor \rho^k \rfloor, k \in \mathbb{N})$, where $\rho > 1$ is taken from a countable sequence converging to 1.

We will fix some $\rho > 1$ throughout, and the averaging parameters N are assumed to belong to $\lfloor \rho^{\mathbb{N}} \rfloor$ unless mentioned otherwise.

We follow a similar plan to [FLW12]. We will prove Theorem 1.4 by showing that almost surely, for every measure-preserving system (X, μ, T) and every $f \in L^\infty(X)$, the following chain of asymptotic equivalences holds μ-almost everywhere:

$$\frac{1}{N}\sum_{n=1}^{N} e\big(p(n)\big)T^{a_n} f \approx \frac{1}{S_N}\sum_{n=1}^{N} X_n(\omega) e\big(p(S_n)\big)T^n f \tag{2.5}$$

$$\approx \frac{1}{W_N}\sum_{n=1}^{N} X_n(\omega) e\big(p(S_n)\big)T^n f \tag{2.6}$$

$$\approx \frac{1}{W_N}\sum_{n=1}^{N} \sigma_n e\big(p(S_n)\big)T^n f \tag{2.7}$$

$$\approx \bar{f} \cdot \frac{1}{W_N}\sum_{n=1}^{N} \sigma_n e\big(p(S_n)\big) \tag{2.8}$$

$$\approx \bar{f} \cdot \frac{1}{N}\sum_{n=1}^{N} e\big(p(n)\big) \tag{2.9}$$

$$\approx 0. \tag{2.10}$$

Here, the symbol $\approx$ means that the difference converges to 0 as $N \to \infty$ and $\bar{f} := \lim_N \frac{1}{N}\sum_{n=1}^{N} T^n f$ is the projection of f onto the invariant factor of T.

Let us now list the ingredients used to establish the above asymptotic equivalences.

(2.5) holds because the right-hand side equals the left-hand side with N replaced by S_N.

(2.6) holds by (2.2).

(2.7) is the key to our argument. We isolate this crucial step in the following

Proposition 2.11. *In the setting of Theorem 1.4, almost surely the following holds: for each measure-preserving system (X, μ, T), and each $f \in L^2(X)$, the sequence*

$$\left\| \frac{1}{W_N}\sum_{n=1}^{N} Y_n(\omega) e\big(p(S_n)\big)T^n f \right\|_{L^2(X)}^2$$

is summable over lacunary N, and in particular

$$\frac{1}{W_N}\sum_{n=1}^{N} Y_n(\omega) e\big(p(S_n)\big)T^n f \to 0 \quad \mu\text{-a.e.}$$

(2.8) for averages with weights σ_n follows by the partial summation formula

$$\frac{1}{W_N}\sum_{n=1}^{N}\sigma_n a_n = \frac{N\sigma_N}{W_N}A_N + \sum_{M=1}^{N-1}\frac{M(\sigma_M-\sigma_{M+1})}{W_N}A_M, \quad A_N = \frac{1}{N}\sum_{n=1}^{N}a_n,$$

from the following result on unweighted averages with $G = e\circ p$:

LEMMA 2.12. *Suppose $0 < a < 1$. Then, almost surely, for every measure-preserving system (X,μ,T) and every $f \in L^1(X,\mu)$ pointwise μ-a.e. we have*

$$\frac{1}{N}\sum_{n=1}^{N}G(S_n)T^n f \approx \bar{f}\cdot\frac{1}{N}\sum_{n=1}^{N}G(S_n)$$

for every bounded function $G:\mathbb{N}\to\mathbb{R}$ as $N\to\infty$.

This is a slight abstraction from [FLW12, Lemma 2.2], where a different function G was specified (but its special form not used in the proof). For completeness, the proof is reproduced below.

(2.9) follows by applying the above steps in reverse order, with $f = 1_X$; and

(2.10) reduces to a statement about trigonometric sums, namely

$$\frac{1}{N}\sum_{n=1}^{N}e\big(p(n)\big)\to 0,$$

which was proved in [Bos94, Theorem 1.3].

Proof of Lemma 2.12. By the usual maximal ergodic theorem, for each fixed ω the set of f for which asymptotic equivalence holds a.e. is closed in $L^1(X)$. Since the equivalence is clear in the case $f=\bar{f}$ and in view of the splitting $L^2(X)=\{f=\bar{f}\}\oplus\overline{\{Th-h, h\in L^\infty(X)\}}$, it suffices to consider the case when $f = Th-h$, $h\in L^\infty$, is a coboundary, so that in particular $\bar{f}=0$. Since $f\in L^\infty$ in this case, it suffices to obtain equivalence for $N\in\lfloor\rho^{\mathbb{N}}\rfloor$ with $\rho>1$ fixed but arbitrary.

Summation by parts gives

$$\frac{1}{N}\sum_{n=1}^{N}G(S_n)T^n(h-Th) = O\big(\|G\|_\infty/N\big)+\frac{1}{N}\sum_{n=1}^{N-1}\big(G(S_n)-G(S_{n+1})\big)T^n h.$$

The first summand is deterministic and converges to 0. The second summand is μ-a.e. bounded by

$$2\|G\|_\infty\|h\|_\infty\frac{1}{N}\sum_{n=1}^{N-1}X_{n+1}\le 2\|G\|_\infty\|h\|_\infty\frac{S_N}{N},$$

and this converges to 0 almost surely in view of (2.2). □

With this reduction complete, we now turn to the proof of Proposition 2.11.

3. Proof of Proposition 2.11

Throughout this section, we will view $0<\delta\ll 1$ as a (small) floating parameter, whose precise value will be fixed at the end of the proof; $0<\nu=\nu(\delta)=O(\delta)$ will be used to denote (possibly different) parameters (all of which grow linearly in δ); $0<\kappa=O(\delta)$ will be used similarly.

We begin with a criterion that guarantees that a bounded sequence $\{c_n\}$ is a good sequence of weights for a pointwise ergodic theorem along a lacunary sequence.

Lemma 3.1. *Let $0<a<b<1$ and fix $\rho>1$. Let $\{c_n\}$ be a bounded sequence such that the following holds*:

$$\sum_{n=1}^{N}|c_n|\lesssim N^{1-a},\quad N\in\lfloor\rho^{\mathbb{N}}\rfloor,\quad \textit{and} \tag{3.2}$$

$$\sum_{N\in\lfloor\rho^{\mathbb{N}}\rfloor}N^{2a-1-b}\sum_{m=1}^{N^b}\left|\sum_{n=N^{1-\delta}}^{N-m}c_{n+m}\bar{c}_n\right|<\infty. \tag{3.3}$$

Then for every measure-preserving system (X,μ,T) and $f\in L^2(X)$ we have

$$\sum_{N\in\lfloor\rho^{\mathbb{N}}\rfloor}\left\|\frac{1}{N^{1-a}}\sum_{n=1}^{N}c_nT^nf\right\|_{L^2(X)}^2<\infty.$$

Proof. Note that (3.2) with $N\in\lfloor\rho^{\mathbb{N}}\rfloor$ implies (3.2) with $N\in\mathbb{N}$, and we obtain

$$\begin{aligned}\left\|\frac{1}{N^{1-a}}\sum_{n=1}^{N^{1-\delta}}c_nT^nf\right\|_{L^2(X)}&\le\frac{1}{N^{1-a}}\sum_{n=1}^{N^{1-\delta}}|c_n|\|f\|_{L^2(X)}\\&\lesssim\frac{1}{N^{1-a}}N^{(1-\delta)(1-a)}\|f\|_{L^2(X)}\\&=N^{-\delta(1-a)}\|f\|_{L^2(X)},\end{aligned}$$

so we may replace the sum in the conclusion of the lemma by $\sum_{n=N^{1-\delta}}^{N}$.

Using van der Corput inequality (2.4) on the Hilbert space $H=L^2(X)$ with $M=N^b$, estimate

$$\begin{aligned}&\left\|\frac{1}{N^{1-a}}\sum_{n=N^{1-\delta}}^{N}c_nT^nf\right\|_{L^2(X)}^2 \\&\quad\lesssim N^{2a-2}\frac{N}{N^b}\sum_{n=N^{1-\delta}}^{N}\left\|c_nT^nf\right\|_{L^2(X)}^2\\&\qquad+N^{2a-2}\frac{N}{N^b}\sum_{m=1}^{N^b}\left|\sum_{n=N^{1-\delta}}^{N-m}\int_X c_{n+m}T^{n+m}f\bar{c}_nT^n\bar{f}\right|.\end{aligned}\tag{3.4}$$

The first term in (3.4) is bounded by

$$N^{2a-2}\frac{N}{N^b}\sum_{n=N^{1-\delta}}^{N}|c_n|^2\|f\|_{L^2(X)}^2,$$

and by the assumption (3.2) and boundedness of (c_n) this is $O(N^{a-b})$. By precomposing with T^{-n}, the second term in (3.4) is bounded by

$$N^{2a-1-b}\sum_{m=1}^{N^b}\left|\sum_{n=N^{1-\delta}}^{N-m}c_{n+m}\bar{c}_n\right|\left|\left\langle T^m f, f\right\rangle_{L^2(X)}\right|,$$

and this is summable by the assumption (3.3). □

PROPOSITION 3.5. *Let $p:\mathbb{R}\to\mathbb{R}$ be a function such that*

$$p(x+y+z)-p(x+y)=p(x+z)-p(x)+O\bigl(x^{\varepsilon-1}yz\bigr) \tag{3.6}$$

for $x,y,z>0$. Let also $0<a<1/2$ and fix $\rho>1$. Then there exists $b\in(a,1/2)$ such that, almost surely, the sequence $c_n=Y_n e(p(S_n))$ satisfies (3.3).

Proof. By Fubini's theorem, it suffices to show that the expectation

$$N^{2a-1-b}\sum_{m=1}^{N^b}\underbrace{\mathbb{E}\left|\sum_{n=N^{1-\delta}}^{N-m}Y_{n+m}e\bigl(p(S_{n+m})\bigr)Y_n e\bigl(-p(S_n)\bigr)\right|}_{=:I(m)}$$

is summable along the lacunary sequence $N\in\lfloor\rho^{\mathbb{N}}\rfloor$. By Cauchy–Schwarz, we have

$$I(m)^2\le\mathbb{E}\left|\sum_{n=N^{1-\delta}}^{N-m}Y_nY_{n+m}e\bigl(p(S_{n+m})-p(S_n)\bigr)\right|^2. \tag{3.7}$$

Using the van der Corput inequality (2.4) with values in the Hilbert space $H=L^2(\Omega)$ and $R=N^c$, $0<c<1$ to be chosen later, we obtain the estimate

$$\begin{aligned}
I(m)^2&\le I_1(m)^2+I_2(m)^2+I_3(m)^2\\
&:=\frac{N-m}{R}\sum_{n=N^{1-\delta}}^{N-m}\bigl\|Y_nY_{n+m}e\bigl(p(S_{n+m})-p(S_n)\bigr)\bigr\|_{L^2(\Omega)}^2\\
&\quad+\frac{N-m}{R}\left|\mathbb{E}\sum_{n=N^{1-\delta}}^{N-2m}Y_{n+2m}Y_{n+m}Y_{n+m}Y_n\right.\\
&\quad\left.\cdot e\bigl(p(S_{n+2m})-p(S_{n+m})-p(S_{n+m})+p(S_n)\bigr)\right|
\end{aligned}$$

$$+\frac{N-m}{R}\sum_{r=1,r\neq m}^{R}\left|\mathbb{E}\sum_{n=N^{1-\delta}}^{N-m-r}Y_{n+r}Y_{n+r+m}Y_nY_{n+m}\right.$$
$$\left.\cdot\, e\big(p(S_{n+r+m})-p(S_{n+r})-p(S_{n+m})+p(S_n)\big)\right|.$$

The task is now to show that, uniformly in $m\le N^b$, we have

$$I_j(m)^2\lesssim N^{2-4a-\kappa}\quad \text{for each } j=1,2,3,$$

for some $\kappa=\kappa(\delta,a,b,c)>0$.

To this end, we estimate the first term, $I_1(m)^2$, by

$$\frac{N}{R}\sum_{n=N^{1-\delta}}^{N-m}\|Y_n\|_{L^2(\Omega)}^2\|Y_{n+m}\|_{L^2(\Omega)}^2$$

by independence; this is bounded by

$$\frac{N}{R}\sum_{n=N^{1-\delta}}^{N-m}\sigma_n\sigma_{n+m}\lesssim N^{1-c}N^{1-2a}<N^{2-4a-\kappa}$$

provided we take $2a<c<1$.

We next turn to $I_2(m)^2$, which contributes at most

$$\frac{N}{R}\mathbb{E}\sum_{n=N^{1-\delta}}^{N-2m}|Y_{n+m}Y_{n+2m}Y_nY_{n+m}|,$$

and by independence this is bounded by

$$\begin{aligned}\frac{N}{R}\sum_{n=N^{1-\delta}}^{N-2m}\mathbb{E}|Y_{n+m}|^2\cdot\mathbb{E}|Y_{n+2m}|\cdot\mathbb{E}|Y_n| &\lesssim N^{1-c}\sum_{n=N^{1-\delta}}^{N-2m}\sigma_n\sigma_{n+m}\sigma_{n+2m}\\ &\lesssim N^{1-c}N^{1-3a+\nu}\\ &\lesssim N^{2-4a-\kappa},\end{aligned}$$

provided $c>2a$ (from above) and $\nu=\nu(\delta)>0$ is taken sufficiently small. (ν arises from the possibility that $3a>1$, in which case we may take e.g. $\nu=(3a-1)\delta$.)

The contribution of this term is also acceptable.

It remains to estimate $I_3(m)^2$, which we write in the form

$$I_3(m)^2=\frac{N-m}{R}\sum_{r=1,r\neq m}^{R}\left|\mathbb{E}\sum_{n=N^{1-\delta}}^{N-m-r}Y_{n+r}Y_{n+r+m}Y_nY_{n+m}\right.$$
$$\left.\cdot\, e\big(p(S_{n+s+t})-p(S_{n+t})-p(S_{n+s})+p(S_n)\big)\right|,$$

with $s = \min(r, m)$ and $t = \max(r, m)$. To recover independence, we apply (3.6) with

$$x = S_{n+t-1}, \qquad y = X_{n+t}, \qquad z = S_{n+t+1,n+t+s},$$

to the first two summands in the argument of e. This gives the estimate

$$\begin{aligned}
&\frac{N-m}{R} \sum_{r=1, r\neq m}^{R} \Big| \mathbb{E} \sum_{n=N^{1-\delta}}^{N-m-r} Y_{n+r} Y_{n+r+m} Y_n Y_{n+m} \\
&\qquad \cdot e\big(p(S_{n+t-1} + S_{n+t+1,n+t+s}) - p(S_{n+t-1}) - p(S_{n+s}) + p(S_n)\big) \Big| \\
&\qquad + \frac{N-m}{R} \sum_{r=1, r\neq m}^{R} \mathbb{E} \sum_{n=N^{1-\delta}}^{N-m-r} \big| Y_{n+r} Y_{n+r+m} Y_n Y_{n+m} \\
&\qquad \cdot \min\big(O\big(S_{n+t-1}^{\varepsilon-1} X_{n+t} S_{n+t+1,n+t+s}\big), 1\big) \big|.
\end{aligned}$$

The main feature of this splitting is that the exponential in the first term does not depend on X_{n+t}, so Y_{n+t} is independent from all other terms. Therefore, the first term vanishes identically. The second term is estimated by

$$\begin{aligned}
&\frac{N}{R} \sum_{r=1, r\neq m}^{R} \sum_{n=N^{1-\delta}}^{N-m-r} \mathbb{E}\big(|Y_{n+r} Y_{n+r+m} Y_n Y_{n+m}| \cdot \min\big(S_{n+t-1}^{\varepsilon-1} S_{n+t+1,n+t+s}, 1\big)\big) \\
&\quad \leq \frac{N}{R} \sum_{r=1, r\neq m}^{R} \sum_{n=N^{1-\delta}}^{N-m-r} \mathbb{E}\Big(|Y_{n+r} Y_{n+r+m} Y_n Y_{n+m}| \\
&\qquad \cdot \min\big(S_{n+t-1}^{\varepsilon-1}(S_{n+t+1,n+t+s-1} + 1), 1\big)\Big).
\end{aligned}$$

By Lemma 2.1, this is bounded by

$$\begin{aligned}
&\frac{N}{R} \sum_{r=1, r\neq m}^{R} \sum_{n=N^{1-\delta}}^{N-m-r} \mathbb{E}\Big(|Y_{n+r} Y_{n+r+m} Y_n Y_{n+m}| \\
&\qquad \cdot \min\big(W_{n+t-1}^{\varepsilon-1}(S_{n+t+1,n+t+s-1} + 1), 1_{E_{n+t-1}}\big)\Big),
\end{aligned}$$

where E_n is an exceptional set of measure $\lesssim \exp(-cn^{1-a})$. At this point, we estimate the minimum by a sum. We consider first the non-exceptional part. All remaining random variables are independent, so we get the estimate

$$\begin{aligned}
&\frac{N}{R} \sum_{r=1, r\neq m}^{R} \sum_{n=N^{1-\delta}}^{N-m-r} \sigma_{n+r}\sigma_{n+r+m}\sigma_n\sigma_{n+m}(n+t-1)^{(1-a)(\varepsilon-1)} \left(\sum_{j=n+t+1}^{n+t+s-1} \sigma_j + 1 \right) \\
&\quad \leq \frac{N}{R} \sum_{r=1, r\neq m}^{R} \sum_{n=N^{1-\delta}}^{N-m-r} n^{-4a} n^{(1-a)(\varepsilon-1)} \big(n^{-a} N^b + 1\big)
\end{aligned}$$

$$
\begin{aligned}
&\lesssim \frac{N}{R} \sum_{r=1, r\neq m}^{R} N^b \sum_{n=N^{1-\delta}}^{N-m-r} n^{-4a} n^{(1-a)(\varepsilon-1)} n^{-a} \\
&\lesssim N^{2-4a+(b-a+\nu)+(1-a)(\varepsilon-1)} \\
&\lesssim N^{2-4a-\kappa},
\end{aligned}
$$

provided that b is taken sufficiently close to a and δ is sufficiently small, since $(1-a)(\varepsilon-1) < 0$.

Finally, for the exceptional part we have superpolynomial decay in N. $\square$

Thus, we have verified that the assumption (3.3) of Lemma 3.1 holds almost surely in the setting of Proposition 2.11. The missing assumption (3.2) also holds almost surely because

$$\sum_{n=1}^{N} |Y_n| \leq S_N + W_N$$

and in view of (2.2). This completes the proof of Proposition 2.11 and hence of Theorem 1.4.

Acknowledgments. We are grateful to Nikos Frantzikinakis, Rowan Killip, and Alexander Volberg for helpful conversations, and to Tanja Eisner for suggesting the collaboration. We thank the Hausdorff Research Institute for Mathematics for hospitality during the Trimester Program "Harmonic Analysis and Partial Differential Equations".

References

[Ass03] I. Assani, *Wiener Wintner ergodic theorems*, World Scientific Publishing Co., River Edge, NJ, 2003. MR 1995517

[Bir31] G. D. Birkhoff, *Proof of the ergodic theorem*, Proc. Natl. Acad. Sci. USA **17** (1931), 656–660.

[B+05] M. Boshernitzan, G. Kolesnik, A. Quas and M. Wierdl, *Ergodic averaging sequences*, J. Anal. Math. **95** (2005), 63–103. MR 2145587

[Bos94] M. D. Boshernitzan, *Uniform distribution and Hardy fields*, J. Anal. Math. **62** (1994), 225–240. MR 1269206

[Bou88] J. Bourgain, *On the maximal ergodic theorem for certain subsets of the integers*, Israel J. Math. **61** (1988), no. 1, 39–72. MR 0937581

[Bou89] J. Bourgain, *Pointwise ergodic theorems for arithmetic sets*, Inst. Hautes Études Sci. Publ. Math. **69** (1989), 5–45. With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein. MR 1019960

[BM07] Z. Buczolich and R. D. Mauldin, *Concepts behind divergent ergodic averages along the squares*, Ergodic theory and related fields, Contemp. Math., vol. 430, Amer. Math. Soc., Providence, RI, 2007, pp. 41–56. MR 2331324

[EK15] T. Eisner and B. Krause, *(Uniform) convergence of twisted ergodic averages*, Ergodic Theory Dynam. Systems **36** (2016), 2172–2202.

[FLW12] N. Frantzikinakis, E. Lesigne and M. Wierdl, *Random sequences and pointwise convergence of multiple ergodic averages*, Indiana Univ. Math. J. **61** (2012), no. 2, 585–617. MR 3043589

[LaV09] P. LaVictoire, *An L^1 ergodic theorem for sparse random subsequences*, Math. Res. Lett. **16** (2009), no. 5, 849–859. MR 2576702

[LaV11] P. LaVictoire, *Universally L^1-bad arithmetic sequences*, J. Anal. Math. **113** (2011), 241–263. MR 2788358

[Les93] E. Lesigne, *Spectre quasi-discret et théorème ergodique de Wiener–Wintner pour les polynômes*, Ergodic Theory Dynam. Systems **13** (1993), no. 4, 767–784. MR 1257033

[RW95] J. M. Rosenblatt and M. Wierdl, *Pointwise ergodic theorems via harmonic analysis*, Ergodic theory and its connections with harmonic analysis (Alexandria, 1993), London Math. Soc. Lecture Note Ser., vol. 205, Cambridge University Press, Cambridge, 1995, pp. 3–151. MR 1325697

[TV10] T. Tao and V. H. Vu, *Additive combinatorics*, Cambridge Studies in Advanced Mathematics, vol. 105, Cambridge University Press, Cambridge, 2010. Paperback edition of MR 2289012. MR 2573797

Ben Krause, UCLA Math Sciences Building, Los Angeles, CA 90095-1555, USA

E-mail address: benkrause23@math.ucla.edu

Pavel Zorin-Kranich, Mathematisches Institut, Universität Bonn, Endenicher Allee 60, 53115 Bonn, Germany

E-mail address: pzorin@uni-bonn.de