An $L^1$ Ergodic Theorem for Sparse Random Subsequences

Patrick LaVictoire

December 16, 2008

Abstract

We prove an $L^1$ subsequence ergodic theorem for sequences chosen by independent random selector variables, thereby showing the existence of universally $L^1$-good sequences nearly as sparse as the set of squares. In the process, we prove that a certain deterministic condition implies a weak maximal inequality for a sequence of $\ell^1$ convolution operators (Prop. 3.1).

1 Introduction

Let $(X, \mathcal{F}, m)$ be a non-atomic probability space and $T$ a measure-preserving transformation on $X$; we call $(X, \mathcal{F}, m, T)$ a dynamical system. For a sequence of integers $n = \{n_k\}$ and any $f \in L^1(X)$, we may define the subsequence average

$$A^{(n)}_N f(x) := \frac{1}{N} \sum_{k=1}^{N} f(T^{n_k}x).$$

Given a sequence $n$, a major question is for which $1 \leq p \leq \infty$ and which $(X, \mathcal{F}, m, T)$ we have convergence of various sorts for all $f \in L^p(X)$. An important definition along these lines is as follows:

Definition A sequence of integers $n = \{n_k\}$ is universally $L^p$-good if for every dynamical system $(X, \mathcal{F}, m, T)$ and every $f \in L^p(X, m)$, $\lim_{N \to \infty} A^{(n)}_N f(x)$ exists for almost every $x \in X$.

Birkhoff’s Ergodic Theorem asserts, for instance, that the sequence $n_k = k$ is universally $L^1$-good. On the other extreme, the sequence $n_k = 2^k$ is not even universally $L^\infty$-good (lacunary sequences are bad for convergence of ergodic averages in various strong ways, see for example [10] or [1]). Between these extrema lie many results on the existence of universally $L^p$-good sequences of various sorts, beginning with Bourgain’s celebrated result [5] that $n_k = k^2$ is universally $L^2$-good; see [6] and [3] for extensions of this result to other sequences.

The most restrictive case $p = 1$ is more subtle than the others. A surprising illustration of the difference is the recent result of Buczolich and Mauldin that $n_k = k^2$ is not universally $L^1$-good [8]. Positive results in $L^1$ have been difficult to come by, particularly for sequences which are sparse in $N$.

Universally $L^1$-good sequences of density 0 had long been known to exist, but these were sparse block sequences, which consist of large ‘blocks’ of consecutive integers, separated by wide gaps. Bellow and Losert [2] showed that for any $F : \mathbb{N} \to \mathbb{R}^+$, there exists a universally $L^1$-good block sequence $\{n_k\}$ with $n_k \geq F(k)$.

To distinguish such block sequences from more uniformly distributed ones, we recall the notion of Banach density:

Definition A sequence of positive integers $\{n_k\}$ has Banach density $c$ if

$$\lim_{m \to \infty} \sup_{N} \frac{\left|\left\{n_k \in [N, N + m]\right\}\right|}{m} = c.$$
Note that block sequences with arbitrarily large block lengths have Banach density 1 (the sequences in [2] are all of this sort). The first example of a Banach density 0 universally \( L^1 \)-good sequence was constructed by Buczolich [7], and Urban and Zienkiewicz [13] subsequently proved that the sequence \([k^a]\) for \(1 < a < 1 + \frac{1}{1000}\) is universally \( L^1 \)-good.

Bourgain [5] noted that certain sparse random sequences were universally \( L^p \)-good with probability 1 for all \(p > 1\). These sequences are generated as follows: given a decreasing sequence of probabilities \( \{ \tau_j : j \in \mathbb{N} \} \), let \( \{ \xi_j : j \in \mathbb{N} \} \) be independent random variables on a probability space \( \Omega \) with \( P(\xi_j = 1) = \tau_j, P(\xi_j = 0) = 1 - \tau_j \). Then for each \( \omega \in \Omega \), define a random sequence by taking the set \( \{ n : \xi_n(\omega) = 1 \} \) in increasing order. (For \( \alpha > 0 \) and \( \tau_j = O(j^{-\alpha}) \), these sequences have Banach density 0 with probability 1; see Prop. 4.3 of this paper.)

In their treatment [11] of Bourgain’s method, Rosenblatt and Wierdl demonstrate by Fourier analysis that if \( \tau_j \to 0 \) slowly enough (e.g. \( \tau_j \geq \frac{c \log \log j}{j} \) suffices), then \( \{ n : \xi_n(\omega) = 1 \} \) is universally \( L^2 \)-good with probability 1 (see Example 4.7), thus proving the existence of superpolynomial universally \( L^2 \)-good sequences. However, their approach cannot be applied to the \( L^1 \) case.

In this paper, we apply a construction of [13] to these random sequences and achieve the following \( L^1 \) result:

**Theorem 1.1.** Let \( 0 < \alpha < 1/2 \), and let \( \xi_n \) be independent selector variables on \( \Omega \) with \( P(\xi_n = 1) = n^{-\alpha} \). Then there exists a set \( \Omega' \subset \Omega \) of probability 1 such that for every \( \omega \in \Omega' \), \( \{ n : \xi_n(\omega) = 1 \} \) is universally \( L^1 \)-good.

Thus we prove the existence of universally \( L^1 \)-good sequences which grow more rapidly than the ones obtained in [13] or [7], and which grow uniformly as compared to the sparse block sequences of [2]. In particular, with probability 1 these sequences have \( n_k = \Theta(k^{1/(1-\gamma)}) \) (that is, \( c_1 k^{1/(1-\gamma)} \leq n_k \leq c_2 k^{1/(1-\gamma)} \)), so Theorem 1.1 applies to random sequences nearly as sparse as the sequence of squares.

Our method is as follows: in Section 2 we define our notation and reduce the problem (by transference) to one of proving a weak maximal inequality on \( \mathbb{Z} \) for convolutions with a series of random \( \ell^1(\mathbb{Z}) \) functions \( \mu_n^{(\omega)} \). In Section 3, we use the framework of [13] to prove this inequality under an assumption about the convolutions of \( \mu_n^{(\omega)} \) with their reflections about the origin; and in Section 4, we establish that with probability 1, these random functions do indeed satisfy that assumption.

## 2 Definitions, and Reduction to a Weak Maximal Inequality

Let \( \{ \tau_n : n \in \mathbb{N} \} \) be a nonincreasing sequence of probabilities. Let \( \Omega \) be a probability space, and define independent mean \( \tau_n \) Bernoulli random variables \( \{ \xi_n(\omega) : n \in \mathbb{N} \} \) on \( \Omega \); that is, \( P(\xi_n = 1) = \tau_n \) and \( P(\xi_n = 0) = 1 - \tau_n \). Let

\[
\beta(N) := \sum_{n=1}^{N} \tau_n.
\]

**Definition** For a dynamical system \((X, \mathcal{F}, m, T)\) and \( f \in L^1(X) \), define the random average

\[
A_N^{(\omega)} f(x) := \beta(N)^{-1} \sum_{n=1}^{N} \xi_n(\omega) f(T^n x)
\]

and its \( L^1(X) \)-valued expectation

\[
E_\omega A_N^{(\omega)} f(x) := \beta(N)^{-1} \sum_{n=1}^{N} \tau_n f(T^n x).
\]
Remark $A_N^{(c)} f$ differs from the subsequence averages discussed before by the factor $\beta(N)^{-1} \sum_{n=1}^{N} \xi_n(\omega)$. However, if $\beta(N) \to \infty$, then with probability 1 in $\Omega$, $\beta(N)^{-1} \sum_{n=1}^{N} \xi_n(\omega) \to 1$; this follows quickly from an application of Chernoff’s Inequality, which we will use elsewhere in this paper:

**Theorem 2.1.** Let $\{X_n\}_{n=1}^{N}$ a sequence of independent random variables with $|X_n| \leq 1$ and $E X_n = 0$. Let $X = \sum_{n=1}^{N} X_n$, and $\sigma^2 = \text{Var} \ X = E X^2$. Then for any $\lambda > 0$, 

$$\mathbb{P}(|X| \geq \lambda \sigma) \leq 2 \max(e^{-\lambda^2/4}, e^{-\lambda \sigma/2}).$$

**Proof.** This is Theorem 1.8 in [12], for example. \hfill \Box

We restrict ourselves to the set $\Omega_1 \subset \Omega$ on which $\beta(N)^{-1} \sum_{n=1}^{N} \xi_n(\omega) \to 1$. The a.e. convergence of $A_N^{(\omega)} f(x)$ for every dynamical system $(X, \mathcal{F}, m, T)$ and every $f \in L^p(X)$ is then equivalent to the statement that $\{j \in N : \xi_j(\omega_0) = 1\}$ is universally $L^p$-good. We further remark that for a power law $\tau_n = n^{-\alpha}$, we have $N^{n-1} \beta(N) \to C \in (0, \infty)$ for $\alpha < 1$.

By Bourgain’s result in [5], there is a set $\Omega_2 \subset \Omega_1$ with $\mathbb{P}(\Omega_2) = 1$ such that for $\omega \in \Omega_2$ we have a.e. convergence of $A_N^{(\omega)} f$ for all $f \in L^2(X)$, which is dense in $L^1(X)$. Theorem 1.1 thus reduces to proving on a set of probability 1 the weak maximal inequality

$$\| \sup_{N} |A_N^{(\omega)} f| \|_{1, \infty} \leq C_\omega \| f \|_1 \forall f \in L^1(X). \tag{2.1}$$

As usual, it is enough to take this supremum over the dyadic subsequence $N \in \{2^j : j \in \mathbb{N}\}$, since $\frac{\beta(2^{j+1})}{\beta(2^j)} \leq 2$ and thus $0 \leq A_N^{(\omega)} f \leq 2 A_{2^j}^{(\omega)} f$ for $f \geq 0$ and $2^j \leq N < 2^{j+1}$. As in [4] and other papers, we can transfer this problem to the group algebra $\ell^1(\mathbb{Z})$. Namely, if we define the random $\ell^1(\mathbb{Z})$ functions

$$\mu_j^{(\omega)}(n) := \begin{cases} \beta(2^j)^{-1} \xi_n(\omega), & 1 \leq n \leq 2^j \\ 0 & \text{otherwise} \end{cases}$$

$$\mathbb{E} \mu_j(n) := \begin{cases} \beta(2^j)^{-1} \tau_n, & 1 \leq n \leq 2^j \\ 0 & \text{otherwise} \end{cases}$$

$$\nu_j^{(\omega)}(n) := \mu_j^{(\omega)}(n) - \mathbb{E} \mu_j^{(\omega)}(n),$$

then $\mu_j^{(\omega)}$ and $\mathbb{E} \mu_j$ correspond to the operators $A_{2^j}^{(\omega)}$ and $\mathbb{E} A_{2^j}^{(\omega)}$, respectively. It suffices to prove that with probability 1 in $\Omega$, 

$$\| \sup_{j} \varphi * \mu_j^{(\omega)} \|_{1, \infty} \leq C_\omega \| \varphi \|_1 \forall \varphi \in \ell^1(\mathbb{Z}). \tag{2.2}$$

We will use $\tilde{\mu}$ to denote the reflection of a function $\mu$ about the origin; as the adjoint of the operator given by convolution with $\mu$ is a convolution with $\tilde{\mu}$, this will be an important object. (It would be standard to use the notation $\mu^*$, but this becomes unwieldy when using other superscripts as above.)

### 3 Calderon-Zygmund Argument

The proof of (2.2) uses a generalization of a deterministic argument from the paper by Urban and Zienkiewicz [13], related to a construction of Christ in [9]:

---

3
Proposition 3.1. Let \( \mu_j \) and \( \nu_j \) be sequences of functions in \( \ell^1(\mathbb{Z}) \). Let \( r_j := \| \text{supp } \mu_j \| \) and suppose \( \text{supp } \nu_j \subset [-R_j, R_j] \). Assume there exist \( \epsilon > 0 \) and \( A, A_0, A_1 < \infty \) such that \( \sum_{j \leq k} r_j \leq A r_k \) \( \forall k \in \mathbb{N} \) and

\[
|\nu_j * \tilde{v}_j(x)| \leq A_0 r_j^{-1} \delta_0(x) + A_1 R_j^{-1(1+\epsilon)} , \quad \forall x \in \mathbb{Z}. \tag{3.1}
\]

If for all \( \varphi \in \ell^1 \), \( \| \varphi * \mu_j - \nu_j \|_{1, \infty} \leq C \| \varphi \|_1 \) and \( \| \sup_j \| \varphi * \mu_j \|_{p, \infty} \leq C_p \| \varphi \|_p \) for some \( 1 < p \leq \infty \), then

\[
\| \sup_j \| \varphi * \mu_j \|_{1, \infty} \leq C' \| \varphi \|_1 \quad \forall \varphi \in \ell^1(\mathbb{Z}). \tag{3.2}
\]

Proof. We will follow the argument in Section 3 of [13], which makes use of a Calderon-Zygmund type decomposition of \( \varphi \) depending on the index \( j \). We begin with the standard decomposition at height \( \lambda > 0 \): \( \varphi = g + b \), where

- \( b = \sum_{(s,k) \in B} b_{s,k} \) for some index set \( B \subset \mathbb{N} \times \mathbb{Z} \)
- \( b_{s,k} \) is supported on the dyadic cube \( Q_{s,k} = [k2^s, (k+1)2^s) \cap \mathbb{Z} \)
- \( \{Q_{s,k} : (s,k) \in B\} \) is a disjoint collection
- \( \| b_{s,k} \|_1 \leq \lambda |Q_{s,k}| = \lambda 2^s \)
- \( \sum_{(s,k) \in B} |Q_{s,k}| \leq \frac{C}{\lambda} \| \varphi \|_1 \) (\( C \) independent of \( \varphi \) and \( \lambda \)).

Let \( b_s = \sum_k b_{s,k} \). We will divide \( \sum_s b_s \) into two parts, splitting at the index \( s(j) := \min\{s : 2^s \geq R_j\} \).

We begin by noting \( \{ x : \sup_j |\varphi * \mu_j(x)| > 6\lambda \} \subset \)

\[
\{ \sup_j |g * \mu_j| > \lambda \} \cup \{ \sup_j |b * (\mu_j - \nu_j)| > \lambda \} \cup \{ \sup_j \sum_{s=s(j)}^{s(j)-1} b_s * \nu_j | > \lambda \} \cup \{ \sup_j \sum_{s=0}^{s(j)-1} b_s * \nu_j | > 3\lambda \}
\]

\[
= E_1 \cup E_2 \cup E_3 \cup E_4.
\]

By the weak \((p,p)\) inequality (if \( p < \infty \)), \( |E_1| \leq C \lambda^{-p} \| g \|_p \leq C \lambda^{-p} \| g \|_\infty \leq C \lambda^{-1} \| \varphi \|_1 \); if \( p = \infty \), consider instead that \( \{ x : \sup_j |g * \mu_j(x)| > C_\infty \lambda \} = \emptyset \) since \( \sup_j |g * \mu_j| \| \infty \leq C_\infty \| g \|_\infty = C_\infty \lambda \).

Next, \( |b * (\mu_j - \nu_j)(x)| \leq |b| \| \mu_j - \nu_j(\| x) \), so by the assumed weak \((1,1)\) inequality,

\[
|E_2| \leq | \{ \sup_j |b * (\mu_j - \nu_j)| > \lambda \} | \leq \frac{C}{\lambda} \| b \|_1 \leq \frac{C}{\lambda} \| \varphi \|_1.
\]

To bound \( |E_3| \), note that for \( s \geq s(j) \), \( \text{supp } (b_{s,k} * \nu_j) \subset Q_{s,k} + [-R_j, R_j] \subset Q^*_{s,k} \), an expansion of \( Q_{s,k} \) by a factor of 3. Thus

\[
E_3 \subset \bigcup_j \bigcup_{k \in \mathbb{Z}, s \geq s(j)} \text{supp } (b_{s,k} * \nu_j) \subset \bigcup_{k \in \mathbb{Z}, s \geq s(j)} Q^*_{s,k}
\]

and

\[
|E_3| \leq \sum_{(s,k) \in B} 3 |Q_{s,k}| \leq \frac{C}{\lambda} \| \varphi \|_1.
\]

4
We have thus reduced the problem to obtaining a bound on $|E_4|$. We will attempt this directly for heuristic purposes, and then modify our setup for the actual argument. By Chebyshev’s Inequality,

$$
|\{ x : \sup_j \left| \sum_{s=0}^{s(j)-1} b_s \ast \nu_j(x) \right| > \lambda \} | \leq \lambda^{-2} \sum_j \left| \sum_{s=0}^{s(j)-1} b_s \ast \nu_j(x) \right|^2 \leq \lambda^{-2} \sum_j \left\| \sum_{s=0}^{s(j)-1} b_s \ast \nu_j \right\|_{\ell^2}^2 \tag{3.3}
$$

and we will use our estimate on the convolution product $\nu_j \ast \tilde{\nu}_j$:

**Lemma 3.2.** Let $f, g \in \ell^1$ such that $\sum_{x \in Q_{s(j), k}} |g(x)| \leq \lambda^{2s(j)}$ for all $k$, and assume the $\nu_j$ satisfy (3.1). Then

$$
|\langle f \ast \nu_j, g \ast \nu_j \rangle| \leq A_0 r_j^{-1} |\langle f, g \rangle| + 10 A_1 \lambda R_j^{-\epsilon} \|f\|_1.
$$

**Proof.**

$$
|\langle f \ast \nu_j, g \ast \nu_j \rangle| = |\langle f \ast \nu_j, \tilde{\nu}_j, g \rangle| \leq A_0 r_j^{-1} |\langle f, g \rangle| + A_1 R_j^{-\epsilon} 1 \|f\|_1 g_1.
$$

We let $f_k = f \chi(Q_{s(j), k})$ and $g_l = g \chi(Q_{s(j), l})$; note that $\|g_l\|_1 \leq \lambda^{2s(j)} \leq 2 \lambda R_j$. If $|k - l| > 2$, then $\langle f_k \ast \nu_j, g_l \ast \nu_j \rangle = 0$ as the supports are disjoint; thus

$$
|\langle f \ast \nu_j, g \ast \nu_j \rangle| \leq \sum_k \sum_{i=-2}^2 |\langle f_k \ast \nu_j, g_{k+i} \ast \nu_j \rangle|
$$

$$
\leq \sum_k \sum_{i=-2}^2 A_0 r_j^{-1} |\langle f_k, g_{k+i} \rangle| + 2 A_1 \lambda R_j^{-\epsilon} 1 \|f_k\|_1
$$

$$
\leq A_0 r_j^{-1} |\langle f, g \rangle| + 10 A_1 \lambda R_j^{-\epsilon} 1 \|f\|_1.
$$

Therefore

$$
|\{ x : \sup_j \left| \sum_{s=0}^{s(j)-1} b_s \ast \nu_j(x) \right| > \lambda \} | \leq \lambda^{-2} \sum_j \sum_{0 \leq s_1, s_2 < s(j)} A_0 r_j^{-1} |\langle b_{s_1}, b_{s_2} \rangle| + 10 A_1 \lambda R_j^{-\epsilon} 1 \|b_{s_1}\|_1
$$

$$
\leq \lambda^{-2} \sum_j \sum_{0 \leq s < s(j)} A_0 r_j^{-1} \|b_s\|^2 + 10 A_1 \lambda s(j) R_j^{-\epsilon} 1 \|b_s\|_1
$$

$$
\leq A_0 \lambda^{-2} \sum_j r_j^{-1} \|b_j\|^2 + 10 A_1 \lambda^{-1} \sum_j \log_2(2 R_j) R_j^{-\epsilon} 1 \|b\|_1.
$$

Since $r_j$ (and thus $R_j$) grows faster than any polynomial by the assumption $\sum_{j \leq k} r_j \leq A_k \forall k \in \mathbb{N}$, the second term is $\leq \frac{C}{\lambda} \|b\|_1$ as desired. The first term does not, however, give us that bound. We will therefore decompose these functions further.
For each $j$, we decompose $b_{s,k} = b_{s,k}^{(j)} + B_{s,k}^{(j)}$, where $b_{s,k}^{(j)} = b_{s,k} \chi(|b_{s,k}| > \lambda r_j)$. Define $b_s^{(j)}, B_s^{(j)}, b^{(j)}, B^{(j)}$ by summing over one or both indices, respectively. Now we see that

$$E_4 \subset \{ \sup_j s_{(j)-1} \sum_{s=0}^{s_{(j)-1}} b_s^{(j)} * (\nu_j - \mu_j) | > \lambda \} \cup \{ \sup_j s_{(j)-1} \sum_{s=0}^{s_{(j)-1}} b_s^{(j)} * \mu_j | > \lambda \} \cup \{ \sup_j s_{(j)-1} \sum_{s=0}^{s_{(j)-1}} B_s^{(j)} * \nu_j | > \lambda \} = E_5 \cup E_6 \cup E_7.$$

We control $E_5$ just as we controlled $E_2$, since $|b^{(j)}| \leq |b|$; and

$$|E_6| \leq \sum_j \{ \{ x : |b^{(j)} * \mu_j(x) | > 0 \} \leq \sum_j \{ \text{supp } \mu_j \} \cdot \{ \{ x : |b(x) | > \lambda r_j \} \}
= \sum_j r_j \sum_{k \geq j} \{ \{ x : \lambda r_k < |b(x) | \leq \lambda r_{k+1} \} \}
= \sum_j \{ \{ x : \lambda r_k < |b(x) | \leq \lambda r_{k+1} \} \} \sum_j r_j
\leq \frac{A}{\lambda} \sum_k \lambda r_k \{ \{ x : \lambda r_k < |b(x) | \leq \lambda r_{k+1} \} \};$$

now since this sum is a lower sum for $|b|$, we have $|E_6| \leq \frac{A}{\lambda} \|b\|_1 \leq \frac{C}{\lambda} \| \varphi \|_1$.\hfill $\Box$

We proceed with $E_7$ just as we tried before, since Lemma 3.2 applies to the $B_s^{(j)}$ as well as to the $b_s$. We thus find

$$|E_7| \leq A_0 \lambda^2 \sum_j r_j^{-1} \|B^{(j)}(x)\|^2 + 10 A_1 \lambda^2 \sum_j \log_2(2R_j) R_j^{-e} \|B^{(j)}\|_1
\leq A_0 \lambda^2 \sum_j \sum_x r_j^{-1} |B^{(j)}(x)|^2 + \frac{C}{\lambda} \| \varphi \|_1.$$

Because $\sum_j r_j \leq A r_k \forall k \in \mathbb{N} \implies \exists N \text{ s.t. } r_{j+n} \geq 2r_j \forall j \in \mathbb{N}, n \geq N \implies \sum_{j=k}^{\infty} r_j^{-1} \leq A' r_k^{-1}$, for each $x$

$$\sum_j r_j^{-1} |B^{(j)}(x)|^2 \leq \sum_{j : \lambda r_j \geq |b(x)|} r_j^{-1} |b(x)|^2 \leq A' \lambda |b(x)|$$

so $|E_7| \leq \frac{C}{\lambda} \| \varphi \|_1$ and the proof of Proposition 3.1 is complete. \hfill $\Box$

## 4 Probabilistic Lemma, Conclusion of the Proof

Having established Proposition 3.1, it remains to show that the random measures $\mu_j^{(\omega)}$ and $\nu_j^{(\omega)}$ satisfy the assumptions with probability 1. Note first that $r_j = |\text{supp } \mu_j^{(\omega)}| = \sum_{1 \leq n \leq 2^j} \xi_n(\omega) = \Theta(\beta(2^j)) = \Theta(2^{1-\alpha j})$ on $\Omega_1$, and $R_j = 2^{j+1}$. We must prove the bound 3.1 on $\nu_j^{(\omega)} * \nu_j^{(\omega)}$.

**Lemma 4.1.** Let $E \subset \mathbb{Z}$, and let $\{X_n\}_{n \in E}$ be independent random variables with $|X_n| \leq 1$ and $\mathbb{E}X_n = 0$. Assume that $\sum_{n \in E} (\text{Var } X_n)^2 \geq 1$. Let $X$ be the random $\ell^1$ function $\sum_{n \in E} X_n \delta_n$, and let $Z^\times$ denote $\mathbb{Z} \setminus \{0\}$. Then for any $\theta > 0$,

$$\mathbb{P} \left( \|X \ast \hat{X}\|_{E^\times(Z^\times)} \geq \theta \left( \sum_{n=1}^{N} (\text{Var } X_n)^2 \right)^{1/2} \right) \leq 4 |E|^2 \max(e^{-\theta^2/16}, e^{-\theta/4}). \quad (4.1)$$

6
For $k \neq 0$, 
\[
X \ast \tilde{X}(k) = \sum_{n \in E \setminus E-k} X_n X_{n+k} = \sum_{n \in E} Y_n
\]
where $\mathbb{E}Y_n = 0$ and $|Y_n| \leq 1$ (of course $Y_n = 0$ if $n + k \notin E$). We want to apply Chernoff’s Inequality (Theorem 2.1), but the $Y_n$ are not independent.

However, we can easily partition $E$ into two subsets $E_1$ and $E_2$ such that $E_i \cap (E_i - k) = \emptyset$ for each $i$; then within each $E_i$, the $Y_n$ depend on distinct independent random variables, so they are independent.

Now $\sum_{n \in E_i} Y_n$ has variance $\sigma_i^2 = \sum_{n \in E_i} \text{Var} X_n \text{Var} X_{n+k} \leq \sum_{n \in E} (\text{Var} X_n)^2$ by Hölder’s Inequality. Chernoff’s Inequality states that for any $\lambda > 0$, $\mathbb{P}(\sum_{n \in E_i} Y_n \geq \lambda \sigma_i) \leq 2 \exp(-\lambda^2/4)$.

Take $\lambda_i = \theta \sigma_i^{-1} (\sum_{n \in E} (\text{Var} X_n)^2)^{1/2}$; then $\lambda_i \geq \theta$ and $\lambda_i \sigma_i = \theta (\sum_{n \in E} (\text{Var} X_n)^2)^{1/2} \geq \theta$.

\[
\mathbb{P}(\sum_{n \in E_i} Y_n \geq \lambda_i \sigma_i) \leq 4 \exp(-\theta^2/4).
\]

Since this holds for each $k \neq 0$ and $|\text{supp } X \ast \tilde{X}| \leq |E|^2$, the conclusion follows (replacing $2\theta$ with $\theta$).

Corollary 4.2. Let $\nu_j^{(\omega)}$ be the random measure defined as before, $0 < \alpha < 1/2$ and $\kappa > 0$. Then there is a set $\Omega_3 \subset \Omega_2$ with $\mathbb{P}(\Omega_3 = 1)$ such that for each $\omega \in \Omega_3$,

\[
|\nu_j^{(\omega)} \ast \tilde{\nu}_j^{(\omega)}(x)| \leq C_\omega \beta(2^j)^{-1} \delta_0(x) + C_\omega \beta(2^j)^{-2} 2^{\kappa j} \sqrt[4]{\sum_{n=1}^{2j} \tau_n^2}
\]

for $j$ sufficiently large, by Chernoff’s inequality. The Borel-Cantelli Lemma implies that $\nu_j^{(\omega)} \ast \tilde{\nu}_j^{(\omega)}(0) \leq 3 \beta(2^j)^{-1}$ for $j$ sufficiently large (depending on $\omega$), so there exists $C_\omega$ with $0 \leq \nu_j^{(\omega)} \ast \tilde{\nu}_j^{(\omega)}(0) \leq C_\omega \beta(2^j)^{-1}$ for all $j$.

For the other term, we note that $\text{Var} \xi_n \leq \tau_n$, so we set $\theta = 2^{\kappa j}$ and apply Lemma 4.1

\[
\mathbb{P} \left( \beta(2^j)^2 \|\nu_j^{(\omega)} \ast \tilde{\nu}_j^{(\omega)}\|_{\ell^\infty(\mathbb{Z}^d)} \geq 2^\kappa \sqrt[4]{\sum_{n=1}^{2j} \tau_n^2} \right) \leq 4 \cdot 2^{2j} \exp(-2^\kappa j/4)
\]

which sum over $j$. The Borel-Cantelli Lemma again proves the bound holds with probability 1.
Note that \( \sum_{n=1}^{2^j} \tau_n^2 = \Theta(2^{(1-2\alpha)j}) \); thus for \( \alpha < 1/2 \) and \( \kappa + \epsilon = 1/2 - \alpha \),

\[
\beta(2^j)^{-2} \left( \sum_{n=1}^{2^j} \tau_n^2 \right)^{1/2} = O(2^{(-1/2+\alpha+\kappa)j}) = O(R_j^{-(1+\epsilon)}).
\]

Therefore the measures \( \nu_j^{(\omega)} \) satisfy the bound \( (5.1) \), for all \( \omega \in \Omega_3 \). Since \( \mu_j^{(\omega)} - \nu_j^{(\omega)} = \mathbb{E} \mu_j \) is a weighted average of the regular ergodic averages, \( \sup_j |\varphi \ast \mathbb{E} \mu_j| \leq C \sup_N |\varphi \ast N^{-1} \chi[1, N]| \) so that Birkhoff’s Ergodic Theorem implies the needed weak \( \ell^1 \) bound; and the \( \ell^\infty \) maximal inequality for \( \mu_j^{(\omega)} \) is trivial. Thus Proposition 3.1 implies the weak maximal inequality \( (2.2) \), and we have proved Theorem 1.1.

**Remark** This argument does not require \( \tau_n \) to obey a power law. If \( \tau_n \) is decreasing and if \( \beta(2^j)^{-2} \sum_{n=1}^{2^j} \tau_n^2 \leq C2^{-(1+\epsilon)j} \) for some \( \epsilon > 0, C < \infty \) and all \( j \), the sequence \{\( n : \xi_n(\omega) = 1 \)\} will be universally \( L^1 \)-good with probability 1.

It remains, finally, to note that \( \{ n : \xi_n = 1 \} \) indeed has Banach density 0 (with probability 1) if the \( \tau_n \) decrease more rapidly than some power law. Conveniently enough, a converse result also holds:

**Proposition 4.3.** Let \( \{ \tau_n \} \) be a decreasing sequence of probabilities, and let \( \xi_n \) be independent Bernoulli random variables with \( \mathbb{P}(\xi_k = 1) = k^{-\alpha} \). Then if \( \tau_n = O(n^{-\alpha}) \) for some \( \alpha > 0 \), the sequence of integers \( \{ n : \xi_n = 1 \} \) has Banach density 0 with probability 1 in \( \Omega \); otherwise, it has Banach density 1 with probability 1 in \( \Omega \).

**Proof.** It is elementary to show that

\[
2^{-r\tau_r^m} r^{-1} \leq \mathbb{P} \left( \sum_{j=r}^{r(n+1)-1} \xi_j \geq m \right) \leq 2^{-r\tau_r^m}, \tag{4.3}
\]

(We majorize or minorize the \( \xi_j \) by i.i.d. Bernoulli variables and use the Binomial Theorem.) Then if \( \tau_n = O(n^{-\alpha}) \), let \( K > 0 \) and fix \( m, r \in \mathbb{N} \) such that \( ma > 1 \) and \( r > mK \); the probabilities above are then summable, so the first Borel-Cantelli Lemma implies that on a set \( \Omega_K \) of probability 1 in \( \Omega \), there exists an \( M_\omega \) such that for all \( n \geq M_\omega \), \( \sum_{j=r(n+1)}^{r(n+1)-1} \xi_j < m < \frac{r}{K} \); then it is clear that \( \{ n : \xi_n = 1 \} \) has Banach density less than \( 3K^{-1} \). Let \( \Omega' = \bigcap_K \Omega_K \); then \( \mathbb{P}(\Omega') = 1 \) and \( \{ n : \xi_n = 1 \} \) has Banach density 0 on \( \Omega' \).

For the other implication, note that if \( \tau_n \neq O(n^{-1/R}) \), there exists a sequence \( n_k \) with \( n_{k+1} \geq 2n_k \) such that \( \tau_{n_k} \geq n_k^{-1/R} \); then

\[
\sum_{n=1}^{\infty} \tau_{R_n} \geq R^{-1} \sum_{n=2}^{\infty} \tau_n^R \geq R^{-1} \sum_{k=2}^{\infty} (n_k - n_{k-1}) \tau_{n_k}^R \geq R^{-1} \sum_{k=2}^{\infty} \frac{1}{2^k} = \infty.
\]

Thus the probabilities in \( (4.3) \) are not summable in \( n \), for \( m = r = R \). Since the variables \( \xi_n \) are independent, the second Borel-Cantelli Lemma implies that there is a set \( \Omega_R \) of probability 1 on which \( \{ n : \xi_n(\omega) = 1 \} \) contains infinitely many blocks of \( R \) consecutive integers. Therefore if \( \tau(n) \neq O(n^{-\alpha}) \) for every \( \alpha > 0 \), let \( \Omega' = \bigcap_R \Omega_R \); on this set of probability 1, \( \{ n : \xi_n = 1 \} \) has Banach density 1.

The author thanks his dissertation advisor, M. Christ, for consultation and assistance throughout the composition of this paper, and M. Wierdl and J. Rosenblatt for many comments and suggestions.
References

[1] Akcoglu, Mustafa; Bellow, Alexandra; Jones, Roger L.; Losert, Viktor; Reinhold-Larsson, Karin; Wierdl, Máté The strong sweeping out property for lacunary sequences, Riemann sums, convolution powers, and related matters. Ergodic Theory Dynam. Systems 16 (1996), no. 2, 207–253.

[2] Bellow, A.; Losert, V. On sequences of density zero in ergodic theory. Conference in modern analysis and probability (New Haven, Conn., 1982), 49–60, Contemp. Math., 26, Amer. Math. Soc., Providence, RI, 1984.

[3] Boshernitzan, Michael; Kolesnik, Grigori; Quas, Anthony; Wierdl, Máté Ergodic averaging sequences. J. Anal. Math. 95 (2005), 63–103.

[4] Bourgain, Jean. An Approach to Pointwise Ergodic Theorems, Geometric aspects of functional analysis (1986/87), 204–223, Lecture Notes in Math., 1317, Springer, Berlin, 1988.

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

[6] Bourgain, Jean. Pointwise ergodic theorems for arithmetic sets. With an appendix by the author, Harry Furstenberg, Yitzhak Katznelson and Donald S. Ornstein. Inst. Hautes Études Sci. Publ. Math. No. 69 (1989), 5–45.

[7] Buczolich, Z. Universally $L^1$ good sequences with gaps tending to infinity. Acta Math. Hungar. 117 (2007), no. 1-2, 91–140.

[8] Buczolich, Z; Mauldin, D, Concepts behind divergent ergodic averages along the squares. Ergodic Theory and Related Fields, AMS, Contemporary Mathematics Vol. 430 (2007) 41-56.

[9] Christ, Michael. Weak type $(1,1)$ bounds for rough operators. Ann. of Math. (2) 128 (1988), no. 1, 19–42.

[10] Rosenblatt, Joseph. Universally bad sequences in ergodic theory. Almost everywhere convergence, II (Evanston, IL, 1989), 227–245, Academic Press, Boston, MA, 1991.

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

[12] Tao, Terence; Vu, Van, Additive Combinatorics. Cambridge Studies in Advanced Mathematics, 105. Cambridge University Press, Cambridge, 2006. xviii+512 pp. ISBN: 978-0-521-85386-6; 0-521-85386-9

[13] Urban, Roman; Zienkiewicz, Jacek, Weak Type $(1,1)$ Estimates for a Class of Discrete Rough Maximal Functions, Math. Res. Lett. 14 (2007), no. 2, 227–237.