

Arbitrarily slow decay in the Möbius disjointness conjecture

AMIR ALGOM^{†,‡} and ZHIREN WANG[‡]

[†] *Department of Mathematics, The University of Haifa at Oranim,
Tivon 3600600, Israel*

(e-mail: amir.algom@math.haifa.ac.il)

[‡] *Department of Mathematics, The Pennsylvania State University,
University Park, PA 16802, USA*

(e-mail: zhirenw@psu.edu, aka5983@psu.edu)

(Received 20 June 2022 and accepted in revised form 2 August 2022)

Abstract. Sarnak's Möbius disjointness conjecture asserts that for any zero entropy dynamical system (X, T) , $(1/N)\sum_{n=1}^N f(T^n x)\mu(n) = o(1)$ for every $f \in \mathcal{C}(X)$ and every $x \in X$. We construct examples showing that this $o(1)$ can go to zero arbitrarily slowly. In fact, our methods yield a more general result, where in lieu of $\mu(n)$, one can put any bounded sequence a_n such that the Cesàro mean of the corresponding sequence of absolute values does not tend to zero. Moreover, in our construction, the choice of x depends on the sequence a_n but (X, T) does not.

Key words: Möbius disjointness, zero topological entropy, topological dynamics
2020 Mathematics Subject Classification: 37B05 (Primary); 11N37 (Secondary)

1. Introduction

A topological dynamical system is a pair (X, T) where X is compact metric space and $T \in \mathcal{C}(X)$. If the system (X, T) has zero topological entropy, then Sarnak's Möbius disjointness conjecture [16, Main Conjecture] predicts that

$$\frac{1}{N} \sum_{n=1}^N \mu(n) f(T^n x) = o(1) \quad \text{for every } f \in \mathcal{C}(X) \text{ and every } x \in X. \quad (1)$$

Many special cases of Sarnak's conjecture have been established. A very partial list of examples consists of [2, 5, 8, 9]. We refer to the surveys of Ferenczi, Kułaga-Przymus, and Lemańczyk [6] and of Kułaga-Przymus and Lemańczyk [11] for excellent expositions on the subject, and many more references.

The goal of this paper is to study the rate of decay in Sarnak's conjecture. That is, to study the nature of the $o(1)$ as in equation (1). We will show that there are systems

for which this $o(1)$ decays to zero arbitrarily slowly. Nevertheless, all the examples we construct to this end satisfy Sarnak’s conjecture. Here is our main result.

THEOREM 1.1. *For every decreasing and strictly positive sequence $\tau(n) \rightarrow 0$, there is a dynamical system (X, T) with zero topological entropy that satisfies the following.*

(1) *There exist $x \in X$ and $f \in \mathcal{C}(X)$ such that:*

$$\limsup_{N \rightarrow \infty} \frac{(1/N) \sum_{n=1}^N f(T^n x) \mu(n)}{\tau(n)} > 0.$$

(2) *The system (X, T) satisfies Sarnak’s conjecture in equation (1).*

Several remarks are in order. First, Sarnak [15, the remark following Main Conjecture] remarks that rates are not required in the conjecture, and this is formally justified by Theorem 1.1. Second, it is natural to ask if Theorem 1.1 may be upgraded by finding a zero entropy dynamical system (X, T) and $f \in \mathcal{C}(X)$ such that for every rate function τ , we can find $x \in X$ that satisfies Theorem 1.1(1). Doing so is as hard as solving the full Möbius disjointness conjecture. Indeed, by [4, Corollary 10], if the conjecture is true, then for every zero entropy system (X, T) and $f \in \mathcal{C}(X)$, equation (1) holds uniformly in $x \in X$. This cannot hold concurrently with the aforementioned upgraded version of Theorem 1.1. In other words, Theorem 1.1 is conjecturally optimal. Next, we remark that in many cases (possibly in all cases), it is known [17] that a sufficiently fast rate in Sarnak’s conjecture implies that the system (X, T) satisfies a prime number theorem (PNT) in the sense discussed in [6, Section 11.2]. Thus, recent examples [7, 10] of zero entropy systems failing to satisfy a PNT can be viewed as evidence toward Theorem 1.1. We also mention the recent interesting counterexamples to polynomial Sarnak’s conjecture constructed by Lian and Shi [12] and Kanigowski, Lemańczyk, and Radziwiłł [10] that, while not directly related to Theorem 1.1 as they focus on a sparse sequence of observations instead of the entire trajectory, are similar in spirit to our work.

Finally, we remark that our construction was partially inspired by the idea of building a sufficiently complex zero entropy system as a skew product from the recent work of Dolgopyat *et al* [3], where they exhibit some new classes of zero entropy smooth systems that satisfy the central limit theorem. In this paper, we construct a symbolic skew product instead of a smooth one to code more precise information carried by $\{a_n\}$.

We will derive Theorem 1.1 from a more general statement. This is the following theorem, which forms the main technical result of this paper.

THEOREM 1.2. *For every decreasing and strictly positive sequence $\tau(n) \rightarrow 0$, there is a zero entropy dynamical system (X, T) and some $f \in \mathcal{C}(X)$ that satisfy the following.*

(1) *Every sequence $|a_n| \leq 1$ with $\limsup_{N \rightarrow \infty} (1/N) \sum_{n=1}^N |a_n| > 0$ admits some $x \in X$ such that*

$$\limsup_{N \rightarrow \infty} \frac{(1/N) \sum_{n=1}^N f(T^n x) a_n}{\tau(n)} > 0.$$

(2) *The system (X, T) satisfies Sarnak’s conjecture in equation (1).*

In fact, we will show that any subsequence N_j such that

$$\lim_{j \rightarrow \infty} \frac{1}{N_j} \sum_{n=1}^{N_j} |a_n| = \theta > 0 \tag{2}$$

admits a further subsequence N_{j_k} such that for all k large enough,

$$\frac{1}{N_{j_k}} \sum_{n=1}^{N_{j_k}} f(T^n x) a_n \geq \theta \cdot \tau(N_{j_k}).$$

We emphasize that in Theorem 1.2, the system (X, T) and the function $f \in \mathcal{C}(X)$ only depend on the rate function τ , while the point $x \in X$ depends also on the sequence a_n . (Indeed, (X, T) is always a subsystem of the same ambient system, which is the product of four skew product systems with Bernoulli fiber and Bernoulli base and an addition finite system. When regarded as a function on this ambient system, f is also independent of τ .)

The derivation of Theorem 1.1 from Theorem 1.2 is straightforward. It is well known that the Möbius function μ satisfies

$$\lim_{N \rightarrow \infty} \frac{\sum_{n=1}^N |\mu(n)|}{N} = \frac{6}{\pi^2} > 0,$$

see e.g. [1, Corollary 1.6]. Thus, Theorem 1.2 applied with $a_n = \mu(n)$ gives Theorem 1.1.

We end this introduction with a brief explanation of our construction. We consider subshifts of $(\{-1, 0, 1\}^{\mathbb{N}} \times \{-1, 0, 1\}^{\mathbb{Z}}, T)$, where $T(y, z) = (\sigma y, \sigma^{y_1} z)$ and σ is the left shift. Given a rate function τ , we first construct a certain rapidly growing sequence $q_k \rightarrow \infty$. We then construct a subshift such that its base comes from words of length $q_{k+1} - q_k$ that have non-zero entries at distance at least q_k from each other. Our space X is a product of four spaces constructed this way, together with a finite set $\{0, 1, 2, 3\}$. The function f is taken to be

$$f((y^{(0)}, z^{(0)}), (y^{(1)}, z^{(1)}), (y^{(2)}, z^{(2)}), (y^{(3)}, z^{(3)}), i) = z_0^{(i)}.$$

We need four spaces in this construction for the reasons below. To retrieve positive correlation between observation $f(T^n x)$ and a_n from positive correlation between $a(n)$ and the sequence $\gamma_n = \text{sign } a(n)$, we make $\{f(T^{q_k j + b} x)\}$ mimic $\{\gamma_{q_k j + c}\}$ for $n = q_k j + b \in [q_k, q_{k+1}/3]$. For this analysis to be applied to most steps in the trajectory, we shall use two different sequences $\{q_k^{(0)}\}, q_k^{(1)}$ such that the intervals $[q_k^{(1)}, q_{k+1}^{(1)}/3]$ and $[q_k^{(0)}, q_{k+1}^{(0)}/3]$ together cover \mathbb{N} . In addition, to express the average of $f(T^n x) a_n$ as an approximate linear combination of that of $\gamma_{q_k j + c} a_n = \gamma_{q_k j + c} a_{q_k j + b}$, one has to explore different pairs of congruences (b, c) and use two different values q and $q - 1$ for q , as explained in the paragraph below. Thus, two more different sequences $\{q_k^{(2)} = q_k^{(0)} - 1\}$ and $\{q_k^{(3)} = q_k^{(1)} - 1\}$ are needed. Each of the four sequences $\{q_k^{(i)}\}$ corresponds to a different space $X^{(i)}$. We remark that if $\lim_{N \rightarrow \infty} (1/N) \sum_{n=1}^N |a_n| > 0$ is assumed instead of $\limsup_{N \rightarrow \infty} (1/N) \sum_{n=1}^N |a_n| > 0$, then only two spaces will be needed as the first concern above is no longer an issue in this case.

Given a_n as in Theorem 1.2(1), our construction of the point $x \in X$ relies on the following observation. Assuming $a_n \in \mathbb{R}$ (otherwise one can pass to either $\text{Re}(a_n)$ or $\text{Im}(a_n)$), let $\gamma_n := \text{sign}(a_n)$, and let N_j, θ be as in equation (2). For every $q, M \gg 1$, one may show that

$$\max_{c,d \in [0,q] \cap \mathbb{Z}} \left\{ \frac{1}{qM} \sum_{b=c}^{q-1+c} \sum_{n=1}^M \gamma_{qn+c} \cdot a_{qn+b}, \frac{-1}{qM} \sum_{b=d}^{q-1+d} \sum_{n=1}^M \gamma_{qn+d+1} \cdot a_{qn+b} \right\} \geq \frac{\theta}{4}.$$

Here we pick $k = k(j)$ in some convenient way, $q = q_k$, and $M \approx N_j/q_k$. We then construct our point x via working in one of the subshifts in our space—the exact choice depends on certain technical issues coming from the relation between N_j and q_k . To set up x , we carefully concatenate pieces of arithmetic progressions in γ or $-\gamma$ in the fiber (using the equation above), with the base living in the corresponding shift space and behaving nicely along the observable f . This will allow us to find a subsequence of N_j where the linear correlations as in Theorem 1.2(1) are well approximated by the average giving the max in the equation above. Thus, with some more work, we bound these correlations from below by $\tau(N_j) \cdot \theta$.

Finally, to derive Theorem 1.2(2), we apply the Matomäki–Radziwiłł bound [13] on averages of multiplicative functions along short intervals. To do this, we exploit some strong periodic behavior that exists in the systems we construct.

2. Proof of Theorem 1.2(1).

2.1. Preliminaries. Let (X, T) be a dynamical system, where we recall that X is a compact metric space and $T \in \mathcal{C}(X)$. We denote the metric on X by d_X . Let us recall the Bowen–Dinaburg definition of topological entropy (as in e.g. [18]). For every $n \in \mathbb{N}$, we define a metric on X via

$$d_n(x, y) = \max\{d_X(T^i(x), T^i(y)) : 0 \leq i < n\}.$$

A Bowen ball $B_n(x, \epsilon)$ of depth n centered at $x \in X$ of radius $\epsilon > 0$ is the corresponding (open) ball in the metric d_n ,

$$B_n(x, \epsilon) = \{y \in X : d_n(x, y) < \epsilon\}.$$

For any set $E \subseteq X$, let $N(E, n, \epsilon)$ denote the minimal number of Bowen balls of depth n and radius ϵ needed to cover E . The topological entropy of (X, T) is then defined as

$$h(T) := \lim_{\epsilon \rightarrow 0} \left(\limsup_{n \rightarrow \infty} \frac{\log N(X, n, \epsilon)}{n} \right).$$

Next, let $\sigma : \{-1, 0, 1\}^{\mathbb{Z}} \rightarrow \{-1, 0, 1\}^{\mathbb{Z}}$ denote the left shift. On $\{-1, 0, 1\}^{\mathbb{Z}}$ and $\{-1, 0, 1\}^{\mathbb{N}}$, we define the metric

$$d(x, y) = 3^{-\min\{|n|:x_n \neq y_n\}}.$$

Also, for every $x \in \{-1, 0, 1\}^{\mathbb{N}}$ and $k > l \in \mathbb{N}$, let $x|_l^k \in \{-1, 0, 1\}^{k-l}$ be the word

$$x|_l^k := (x_l, x_{l+1}, \dots, x_k),$$

and we use similar notation in the space $\{-1, 0, 1\}^{\mathbb{Z}}$ as well. Next, let

$$Z := \{-1, 0, 1\}^{\mathbb{N}} \times \{-1, 0, 1\}^{\mathbb{Z}}$$

and endow Z with the sup-metric on both its coordinates. Note that open balls in this metric are also closed, and thus for every $n \in \mathbb{N}$, $x \in X$, and $\epsilon > 0$, the Bowen ball $B_n(x, \epsilon)$ is closed. Also, we denote by Π_i , $i = 1, 2$, the coordinate projections in Z . Finally, we define the skew-product $T : Z \rightarrow Z$ via

$$T(y, z) = (\sigma(y), \sigma^{y_1}(z)).$$

We say that $X \subseteq Z$ is a subshift if it is closed and T -invariant.

We will require the following lemma.

LEMMA 2.1. *The system (Z, T) satisfies that for every $n \in \mathbb{N}$, $\epsilon > 0$, and $x = (y, z) \in Z$, the following hold.*

(1) *We have*

$$T^n(y, z) = (\sigma^n y, \sigma^{\sum_{i=1}^n y_i} z).$$

(2) *Let $m = m(n, y) = \min\{\min_{1 \leq k \leq n} \sum_{i=1}^k y_i, 0\}$ and $M := M(n, y) = \max\{\max_{1 \leq k \leq n} \sum_{i=1}^k y_i, 0\}$. Then for any $l \in \mathbb{N}$, the Bowen ball $d_n(x, 3^{-l})$ equals*

$$\{(a, b) \in Z : a|_1^{n+u} = y|_1^{n+u}, b|_{m-u}^{M+u} = z|_{m-u}^{M+u}\}.$$

(3) *For any set $E \subseteq Z$,*

$$N(E, n, \epsilon) = N(\text{cl}(E), n, \epsilon),$$

where $\text{cl}(E)$ is the closure of the set E .

Proof. Part (1) follows immediately from the definition of the map T . Part (2) follows from part (1). Finally, part (3) is an immediate consequence of the fact that in (Z, T) , Bowen balls are closed. □

2.2. *Construction of some zero entropy systems.* Fix a sequence $\tau(n) \rightarrow 0$ as in Theorem 1.2. We begin by constructing a rapidly growing sequence $q_k \rightarrow \infty$ (that depends on τ) such that for every $k \in \mathbb{N}$, we have the following.

- (1) $q_{k+1} > q_k^4 + 3q_k$.
- (2) $\tau(q_{k+1}/3) < 1/16q_k$.

We now use q_k to define four sequences:

$$q_k^{(0)} := q_{2k}, \quad q_k^{(1)} = q_{2k+1}, \quad q_k^{(2)} := q_k^{(0)} - 1, \quad q_k^{(3)} := q_k^{(1)} - 1.$$

Notice that property (1) above also holds for $q_k^{(i)}$ for every $i \in \{0, 1, 2, 3\}$. In particular,

$$\lim_{k \rightarrow \infty} \frac{q_{k+1}^{(i)}}{q_k^{(i)}} = \infty \quad \text{for every } i \in \{0, 1, 2, 3\}.$$

Next, for every $i \in \{0, 1, 2, 3\}$ and every k , let

$$A_k^{(i)} := \{j \cdot q_k^{(i)} : j \in \mathbb{Z}, q_k^{(i)} \leq j \cdot q_k^{(i)} \leq q_{k+1}^{(i)}\}.$$

For every $i \in \{0, 1, 2, 3\}$ and every $k \in \mathbb{N}$, we construct elements $s_k^{(i)} \in \{-1, 0, 1\}^{\mathbb{N}}$ such that the following hold.

- (1) $s_k^{(i)}(n) = 0$ for every integer $n \notin A_k^{(i)}$.
- (2) For every $j \cdot q_k^{(i)} \in A_k^{(i)}$,

$$s_k^{(i)}(j \cdot q_k^{(i)}) = 1 \quad \text{if } j \leq \left\lceil \frac{q_{k+1}^{(i)}}{3q_k^{(i)}} \right\rceil,$$

and

$$s_k^{(i)}(j \cdot q_k^{(i)}) = -1 \quad \text{if } \left\lceil \frac{q_{k+1}^{(i)}}{3q_k^{(i)}} \right\rceil < j \leq 2 \left\lceil \frac{q_{k+1}^{(i)}}{3q_k^{(i)}} \right\rceil.$$

Next, for every element $x \in \{-1, 0, 1\}^{\mathbb{N}}$ and $p \in \mathbb{N}_0$, we define $\sigma^{-p}x \in \{-1, 0, 1\}^{\mathbb{N}}$ as $\sigma^{-p}x = x$ if $p = 0$, and otherwise

$$(\sigma^{-p}x)|_1^p = (0, \dots, 0) \quad \text{and for all } n > p, \sigma^{-p}x(n) = x(n - p).$$

The following lemma is an immediate consequence of our construction.

LEMMA 2.2. For every $i \in \{0, 1, 2, 3\}$, $k \in \mathbb{N}$, and $p = 0, \dots, q_k^{(i)}$, we have

$$\sum_{n \in [q_k^{(i)}, q_{k+1}^{(i)}) \cap \mathbb{Z}} (\sigma^{-p} s_k^{(i)})(n) = 0.$$

Proof. This follows since by our construction,

$$|\{j \cdot q_k^{(i)} \in A_k^{(i)} : s_k^{(i)}(j \cdot q_k^{(i)}) = 1\}| = |\{j \cdot q_k^{(i)} \in A_k^{(i)} : s_k^{(i)}(j \cdot q_k^{(i)}) = -1\}|,$$

as well as the fact that for all $n \in A_k^{(i)}$ and $p = 0, \dots, q_k^{(i)}$, $n + p$ is still in the interval $[q_k^{(i)}, q_{k+1}^{(i)}) \cap \mathbb{Z}$. □

Next, for every $i \in \{0, 1, 2, 3\}$ and $k \in \mathbb{N}$, define the truncations

$$R_k^{(i)} = \{(\sigma^{-p} s_k^{(i)})|_{q_k^{(i)}}^{q_{k+1}^{(i)}-1} : p = 0, \dots, q_k^{(i)}\} \subseteq \{-1, 0, 1\}^{q_{k+1}^{(i)}-q_k^{(i)}}.$$

We now define the space $P^{(i)}$ of all infinite sequences that have, for every k , some word from $R_k^{(i)}$ between their $q_k^{(i)}$ and $q_{k+1}^{(i)} - 1$ digits. Formally,

$$P^{(i)} = \{x \in \{-1, 0, 1\}^{\mathbb{N}} : x|_{q_k^{(i)}}^{q_{k+1}^{(i)}-1} \in R_k^{(i)} \text{ and } x|_1^{q_1^{(i)}-1} = (0, \dots, 0)\}.$$

The following lemma is an immediate consequence of Lemma 2.2.

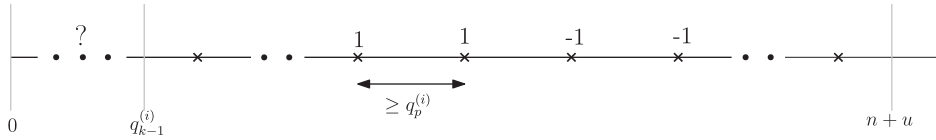


FIGURE 1. Illustration for Lemma 2.4.

LEMMA 2.3. For every $i \in \{0, 1, 2, 3\}$, $k \in \mathbb{N}$, and $y \in P^{(i)}$,

$$\sum_{j=1}^{q_k^{(i)}-1} y(j) = 0.$$

Finally, for every $i \in \{0, 1, 2, 3\}$, we define the subshift of (Z, T) ,

$$X_i = \text{cl}\left(\bigcup_{n \in \mathbb{N}_0} T^n(P^{(i)} \times \{-1, 0, 1\}^{\mathbb{Z}})\right).$$

LEMMA 2.4. For every $i \in \{0, 1, 2, 3\}$, we have $h(X_i, T) = 0$.

Proof. Fix $n, u \in \mathbb{N}$. We count how many Bowen balls of radius $1/3^u$ and depth n are needed to cover X_i . Recall that we denote this quantity by $N(X_i, n, 1/3^u)$. By Lemma 2.1(3), this is the same number as

$$N\left(\bigcup_{l \in \mathbb{N}_0} T^l(P^{(i)} \times \{-1, 0, 1\}^{\mathbb{Z}}), n, \frac{1}{3^u}\right).$$

So, we work with the latter space (that is, without taking the closure).

Let $k = k(n + u, i)$ be such that

$$q_k^{(i)} \leq n + u < q_{k+1}^{(i)}. \tag{3}$$

Our first observation is that we can write

$$\bigcup_{l \in \mathbb{N}_0} T^l(P^{(i)} \times \{-1, 0, 1\}^{\mathbb{Z}}) = A_1 \cup A_2 \cup A_3.$$

To define the sets A_i , we first note that every $x \in \bigcup_{l \in \mathbb{N}_0} T^l(P^{(i)} \times \{-1, 0, 1\}^{\mathbb{Z}})$ admits some $l \in \mathbb{N}_0$ and $\tilde{x} \in P^{(i)} \times \{-1, 0, 1\}^{\mathbb{Z}}$ such that $x = T^l \tilde{x}$. We denote by $p = p(x) \in \mathbb{N}$ the unique integer such that $q_{k-1}^{(i)} + l \in [q_p^{(i)}, q_{p+1}^{(i)})$. Note that $p \geq k - 1$. Then,

$$A_1 = \{x : p(x) \geq k + 1\}, \quad A_2 = \{x : p(x) = k\}, \quad A_3 = \{x : p(x) = k - 1\}.$$

Thus, we bound the covering numbers for A_1, A_2, A_3 separately. Before doing so, we notice that for any $x \in A_j$ for $j = 1, 2, 3$, there are at most $3^{q_{k-1}^{(i)}}$ possibilities for the first $q_{k-1}^{(i)}$ digits of $\Pi_1(x)$ (see Figure 1).

(i) Covering A_1 . For any $x \in A_1$, the word $(\Pi_1 x) \Big|_{q_{k-1}^{(i)}}^{n+u}$ always consists of zeros separated by 1 or -1 , and in this case, the non-zero entries appear at a distance at least $q_{k+1}^{(i)} > n + u$ from each other. Since there can be only one non-zero entry, there are at most $2(n + u)$

options for the configuration of this word. So, with the notation of Lemma 2.1(2), we see that

$$|m|, M \leq q_{k-1}^{(i)} + 1.$$

Thus, taking into account also the first $q_{k-1}^{(i)}$ digits, and via Lemma 2.1(2), the number of Bowen balls we need here is at most

$$(3^{q_{k-1}^{(i)}} \times 2(n + u)) \times (3^{u+q_{k-1}^{(i)}+1})^2.$$

(ii) Covering A_2 . The word $(\Pi_1 x) \Big|_{q_{k-1}^{(i)}}^{n+u}$ consists of zeros separated by 1 or -1 , and in this case, the first non-zero entries appear at a distance at least $q_k^{(i)} \leq n + u$ from each other. We also know that the first non-zero digit needs to appear within the first $q_k^{(i)}$ digits. Another factor that needs to be taken into consideration is the possibility that $[q_{k-1}^{(i)} + l, n + u + l]$ intersects $[q_{k+1}^{(i)}, \infty)$. So, with the notation of Lemma 2.1(2), we see that

$$|m|, M \leq q_{k-1}^{(i)} + \frac{n + u}{q_k^{(i)}} + 1.$$

Taking all these factor into account, the number of Bowen balls we need here is at most

$$(3^{q_{k-1}^{(i)}} \times q_k^{(i)} \times 2(n + u)) \times (3^{u+q_{k-1}^{(i)}+(n+u)/q_k^{(i)}+1})^2.$$

(iii) Covering A_3 . The word $(\Pi_1 x) \Big|_{q_{k-1}^{(i)}}^{n+u}$ consists of zeros separated by 1 or -1 , and in this case, the first non-zero entries appear at a distance at least $q_{k-1}^{(i)}$ from each other. We also know that the first non-zero digit needs to appear within the first $q_{k-1}^{(i)}$ digits. Another factor that needs to be taken into consideration is the possibility that $[q_{k-1}^{(i)} + l, n + u + l]$ intersects $[q_k^{(i)}, \infty)$. So, with the notation of Lemma 2.1(2), we see that

$$|m|, M \leq q_{k-1}^{(i)} + \frac{q_k^{(i)}}{q_{k-1}^{(i)}} + \frac{n + u}{q_k^{(i)}} + 1.$$

Taking all these factors into account, the number of Bowen balls we need here is at most

$$(3^{q_{k-1}^{(i)}} \times q_{k-1}^{(i)} \times q_k^{(i)} \times 2(n + u)) \times (3^{u+q_{k-1}^{(i)}+q_k^{(i)}/q_{k-1}^{(i)}+(n+u)/q_k^{(i)}+1})^2.$$

Thus, we see that

$$N\left(X_i, n, \frac{1}{3^u}\right) \leq 3 \cdot \max_{i=1,2,3} N\left(A_i, n, \frac{1}{3^u}\right) = 3 \cdot N\left(A_3, n, \frac{1}{3^u}\right),$$

which has been computed in point (iii) above. So, making use of equation (3),

$$\begin{aligned} & \frac{\log N(X_i, n, 1/3^u)}{n} \\ & \leq \frac{\log 3 + \log(3^{q_{k-1}^{(i)}} \cdot 2(n + u) \cdot q_{k-1}^{(i)} \cdot q_k^{(i)}) \cdot (3^{u+q_{k-1}^{(i)}+q_k^{(i)}/q_{k-1}^{(i)}+(n+u)/q_k^{(i)}+1)^2}{n} \\ & \leq \frac{\log 6}{n} + \frac{q_{k-1}^{(i)} \cdot \log 3}{n} + \frac{\log(n + u)}{n} + \frac{2 \log q_k^{(i)}}{n} \end{aligned}$$

$$\begin{aligned}
 & + \frac{\left(u + q_{k-1}^{(i)} + q_k^{(i)}/q_{k-1}^{(i)} + n + u/q_k^{(i)} + 1\right) \log 9}{n} \\
 \leq & C_1 \cdot \left(\frac{\log q_k^{(i)}}{n} + \frac{q_{k-1}^{(i)}}{n} + \frac{\log(n+u)}{n} + \frac{q_k^{(i)}}{q_{k-1}^{(i)} \cdot (n+u)} \cdot \frac{n+u}{n} + \frac{n+u}{q_k^{(i)} \cdot n}\right) \\
 \leq & C_1 \cdot \frac{n+u}{n} \cdot \left(\frac{2 \log(n+u)}{n} + \frac{q_{k-1}^{(i)}}{q_k^{(i)}} + \frac{1}{q_{k-1}^{(i)}} + \frac{1}{q_k^{(i)}}\right).
 \end{aligned}$$

Here, C_1 is a large constant that depends variously on u and the other constants appearing in the second equation. We conclude that, fixing u ,

$$\lim_{n \rightarrow \infty} \frac{\log N(X, n, 1/3^u)}{n} = 0,$$

and the claim is proved. □

2.3. Finding correlations along arithmetic progressions. Let a_n be a sequence as in Theorem 1.2(1), that is, such that $\limsup_{N \rightarrow \infty} (1/N) \sum_{n=1}^N |a_n| > 0$. By moving to either $\text{Re}(a_n)$ or $\text{Im}(a_n)$, we may assume a_n is a real valued sequence. We define a new sequence $\gamma_n \in \{-1, 0, 1\}$ via

$$\gamma_n := \text{sign}(a_n).$$

In particular,

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N \gamma_n \cdot a_n = \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^N |a_n| > 0.$$

Let $\theta := \limsup (1/N) \sum_{n=1}^N |a_n| > 0$, and let N_j be a subsequence such that

$$\lim_{j \rightarrow \infty} \frac{1}{N_j} \sum_{n=1}^{N_j} |a_n| = \theta.$$

Definition 2.5. For every $j \in \mathbb{N}$ large enough, we define $k' = k(j) \in \mathbb{N}$ and $i' = i(j) \in \{0, 1\}$ as the unique integers such that:

$$\text{if } N_j \in \left[\frac{q_{k'}^{(1)}}{3}, \frac{q_{k'+1}^{(0)}}{3}\right) \text{ then } i' = 0; \text{ and}$$

$$\text{if } N_j \in \left[\frac{q_{k'+1}^{(0)}}{3}, \frac{q_{k'+1}^{(1)}}{3}\right) \text{ then } i' = 1.$$

We also define an integer

$$M_{k'}^{(i')} := \left\lfloor \frac{N_j}{q_{k'}^{(i')}} \right\rfloor.$$

Note that by definition and the construction of the sequence q_k ,

$$\frac{(q_{k'}^{(i')})^3}{3} = \frac{(q_{k'}^{(i')})^4}{3q_{k'}^{(i')}} < M_{k'}^{(i')} \leq \frac{q_{k'+1}^{(i')}}{3q_{k'}^{(i')}}. \tag{4}$$

Next, recall the definition of Z from §2.1 and let $g : Z \rightarrow \{-1, 0, 1\}$ be the function

$$g(y, z) = z_0.$$

For every $q, M \gg 1$ and r, c such that $r, c \in [0, q]$, let

$$A_{r,c}^{q,M} := \frac{1}{qM} \sum_{b=r}^{q-1+r} \sum_{n=1}^M \gamma(qn + c) \cdot a(qn + b).$$

Finally, we also define

$$M_{k'}^{(i'+2)} := \left[\frac{q_{k'}^{(i')} M_{k'}^{(i')}}{q_{k'}^{(i')} - 1} \right] = \left[\frac{q_{k'}^{(i')} M_{k'}^{(i')}}{q_{k'}^{(i'+2)}} \right]$$

and note that $M_{k'}^{(i'+2)} \approx M_{k'}^{(i')}$. In the following lemma, we use the construction from §2.2.

LEMMA 2.6. *For every j and $u \in \{0, 1\}$ writing $\ell = i' + 2u$, for every two integers $c, r \in [0, q_{k'}^{(\ell)}]$, let $x \in P^{(\ell)} \times \{-1, 0, 1\}^{\mathbb{Z}} \subseteq X_\ell$ be any element such that for every $q_{k'}^{(\ell)} \leq n < q_{k'+1}^{(\ell)}$,*

$$x(n) = (s_{k'}^{(\ell)}(n - r), \gamma(q_{k'}^{(\ell)} \cdot n + c)).$$

Then

$$\frac{1}{q_{k'}^{(\ell)} M_{k'}^{(\ell)}} \sum_{n=1}^{q_{k'}^{(\ell)} M_{k'}^{(\ell)}} g(T^n x) a(n) = A_{r,c}^{q_{k'}^{(\ell)}, M_{k'}^{(\ell)}} + O\left(\frac{q_{k'}^{(\ell)}}{M_{k'}^{(\ell)}}\right).$$

Note that by the construction of $P^{(\ell)} \times \{-1, 0, 1\}^{\mathbb{Z}}$ in §2.2, there exists an element x as in the statement of the lemma in that space.

Proof. In this proof, we suppress the ℓ, k' in our notation and simply write q, M . First, for every two integers $j \in [1, M]$ and $b \in [r, q + r - 1]$,

$$\begin{aligned} \sum_{d=1}^{qj+b} (\Pi_1 x)(d) &= \sum_{d=1}^{q-1} (\Pi_1 x)(d) + \sum_{d=q}^{qj+b-1} (\Pi_1 x)(d) \\ &= \sum_{d=q}^{qj+b-1} s_k^{(\ell)}(d - r) \\ &= \sum_{d=q-r}^{qj+b-r-1} s_k^{(\ell)}(d) = j. \end{aligned}$$

Note the use of Lemma 2.3 in the second equality. Moreover, in the last equality, we use the fact that $M \leq q_{k'+1}^{(\ell)}/3q_{k'}^{(\ell)}$ and the definition of $s_k^{(\ell)}$ to guarantee that all summands are either 0 or 1. Therefore,

$$\begin{aligned} \frac{1}{qM} \sum_{n=1}^{qM} g(T^n x)a(n) &= \frac{1}{qM} \sum_{n=q}^{qM} g(T^n x)a(n) + O\left(\frac{1}{M}\right) \\ &= \frac{1}{qM} \sum_{j=1}^M \sum_{b=r}^{q+r-1} g(T^{q \cdot j+b} x)a(q \cdot j + b) + O\left(\frac{1}{M}\right) \\ &= \frac{1}{qM} \sum_{j=1}^M \sum_{b=r}^{q+r-1} g(\sigma^{qj+b} \Pi_1 x, \sigma^{\sum_{d=1}^{qj+b} (\Pi_1 x)(d)} \Pi_2 x)a(q \cdot j + b) \\ &\quad + O\left(\frac{1}{M}\right) \\ &= \frac{1}{qM} \sum_{j=1}^M \sum_{b=r}^{q+r-1} g(\sigma^{qj+b-r} s_{k'}^{(\ell)}, \sigma^j \Pi_2 x)a(q \cdot j + b) \\ &\quad + O\left(\frac{1}{M}\right) \\ &= \frac{1}{qM} \sum_{j=1}^M \sum_{b=r}^{q+r-1} \gamma(q \cdot j + c) \cdot a(q \cdot j + b) + O\left(\frac{q}{M}\right) \\ &= A_{r,c}^{q,M} + O\left(\frac{q}{M}\right). \end{aligned}$$

Indeed, the first equality follows since $g(T^n x)$ and a_n are both bounded sequences; in the third equality, we use Lemma 2.1(1); and in the fourth equality, we are using the previous equation array and the definition of x . This definition along with the definition of $s_k^{(\ell)}$ justify the fifth equality. The last equality is simply the definition of $A_{r,c}^{q,M}$. \square

Remark 2.7. In the setup of Lemma 2.6, we may similarly find another $x \in P^{(\ell)} \times \{-1, 0, 1\}^{\mathbb{Z}}$ that satisfies the conclusion of Lemma 2.6, but for $-A_{r,c}^{q_{k'}, M_{k'}^{(\ell)}}$. Indeed, this follows from the very same proof by picking $x \in P^{(\ell)} \times \{-1, 0, 1\}^{\mathbb{Z}}$ to be any element such that for every $q_{k'}^{(\ell)} \leq n < q_{k'+1}^{(\ell)}$,

$$x(n) = (s_{k'}^{(\ell)}(n - r), -\gamma(q_{k'}^{(\ell)} \cdot n + c)).$$

We will also require the following lemma.

LEMMA 2.8. *For every j large enough, there is either some $c \in [0, q_{k'}^{(i)})$ such that*

$$A_{c,c}^{q_{k'}^{(i)}, M_{k'}^{(i)}} \geq \frac{\theta}{8q_{k'}^{(i)}}, \tag{5}$$

or some $d \in [0, q_{k'}^{(i'+2)})$ with

$$-A_{d+1,d}^{q_{k'}^{(i'+2)}, M_{k'}^{(i'+2)}} \geq \frac{\theta}{8q_{k'}^{(i')}}.$$

Proof. In this proof, we again suppress the i', k', u in our notation, and write instead q, M , for $q_{k'}^{(i')}$ and $M_{k'}^{(i')}$, respectively (the terms corresponding to $i' + 2$ will come up in the proof later). Now, for every $c, r \in [0, q]$,

$$\sum_{c=0}^{q-1} A_{c+r,c}^{q,M} = \frac{1}{qM} \sum_{m=1}^{qM} \gamma(m) \cdot (a(m+r) + \dots + a(m+r+q-1)) + O\left(\frac{1}{M}\right).$$

So,

$$qM \cdot \sum_{c=0}^{q-1} A_{c,c}^{q,M} = \sum_{m=1}^{qM} \gamma(m) \cdot (a(m) + \dots + a(m+q-1)) + O(q)$$

and

$$\begin{aligned} (q-1) \left[\frac{qM}{q-1} \right] \sum_{c=1}^{q-1} A_{c+1,c}^{q-1, [qM/(q-1)]} \\ = \sum_{m=1}^{qM} \gamma(m) \cdot (a(m+1) + \dots + a(m+q-1)) + O(q^2). \end{aligned}$$

Combining the last two displayed equations,

$$\begin{aligned} qM \cdot \sum_{c=0}^{q-1} A_{c,c}^{q,M} - (q-1) \left[\frac{qM}{q-1} \right] \sum_{c=1}^{q-1} A_{c+1,c}^{q-1, [qM/(q-1)]} \\ = \sum_{m=1}^{qM} \gamma(m)a(m) + O(q^2) \geq \theta/2 \cdot qM + O(q^2). \end{aligned}$$

It follows that, assuming q is large enough and via equation (4),

$$\sum_{c=0}^{q-1} A_{c,c}^{q,M} - \sum_{d=1}^{q-1} A_{d+1,d}^{q-1, [qM/(q-1)]} \geq \theta/2 - O\left(\frac{q}{M}\right) \geq \theta/2 - O\left(\frac{1}{q^2}\right) \geq \theta/4.$$

Recalling our definition of $q_{k'}^{(i'+2)}$ and $M_{k'}^{(i'+2)}$, this implies the lemma. □

2.4. *Construction of the point and system as in Theorem 1.2.* Recalling Lemma 2.8, by perhaps moving to a further subsequence, we may assume that the inequality from Lemma 2.8 is always given by the term corresponding to $q_{k'}^{(i'+2u)}$, where $u = u(j)$ is either 0 or 1, and both the quantities $i' = i(j)$ (defined in Definition 2.5) and u are assumed to be constant in j . Let us denote this constant value $i' + 2u \in \{0, 1, 2, 3\}$ by ℓ . Recalling Definition 2.5, and passing to a subsequence if needed, we assume that the map $j \mapsto k(j) = k'$ is injective.

We now construct a point $x^{(\ell)} \in P^{(\ell)} \times \{-1, 0, 1\}^{\mathbb{Z}} \subseteq X_\ell$ as follows. For every $j \in \mathbb{N}$ and $q_{k(j)}^{(\ell)} \leq n < q_{k(j)+1}^{(\ell)}$, $x^{(\ell)}(n) = x(n)$, where x is the element as in Lemma 2.6 (if $u = 0$) or Remark 2.7 (if $u = 1$), corresponding to j, ℓ as in the paragraph above, and either $r = c$ and c (if $u = 0$) or $r = d + 1$ and $c = d$ (if $u = 1$) yielding the inequality from Lemma 2.8. Note that here we need the map $j \mapsto k(j)$ to be injective so this is well defined (that is, the intervals $[q_{k(j)}^{(\ell)}, q_{k(j)+1}^{(\ell)})$ do not overlap). Note that so far we have only specified the digits $n \in \bigcup_{j \in \mathbb{N}} [q_{k(j)}^{(\ell)}, q_{k(j)+1}^{(\ell)})$, and (since we have passed to a subsequence) it is possible that this union does not cover all of \mathbb{N} . So, for all digits not covered, we make some choice that ensures $x^{(\ell)} \in P^{(\ell)} \times \{-1, 0, 1\}^{\mathbb{Z}}$. Note that by Lemma 2.6 and the construction of $P^{(\ell)}$, such a choice is readily available.

We now take our space to be

$$X := X_0 \times X_1 \times X_2 \times X_3 \times \{0, 1, 2, 3\}, \tag{6}$$

with the self-mapping $\hat{T} \in \mathcal{C}(X)$ being

$$\hat{T}(p^{(0)}, p^{(1)}, p^{(2)}, p^{(3)}, i) = (Tp^{(0)}, Tp^{(1)}, Tp^{(2)}, Tp^{(3)}, i).$$

The function $f \in \mathcal{C}(X)$ is taken to be

$$f((y^{(0)}, z^{(0)}), (y^{(1)}, z^{(1)}), (y^{(2)}, z^{(2)}), (y^{(3)}, z^{(3)}), i) = z_0^{(i)}.$$

We next choose our point x to be any $x \in X$ such that its projection to X_ℓ is $x^{(\ell)}$, and its projection to $\{0, 1, 2, 3\}$ is ℓ .

We now prove Theorem 1.2(1) via the following two claims.

LEMMA 2.9. *We have $h(X, \hat{T}) = 0$.*

Proof. By Claim 2.4, each factor in the product space X has zero entropy, which implies the assertion via standard arguments. □

LEMMA 2.10. *For all j large enough,*

$$\frac{1}{N_j} \sum_{n=1}^{N_j} f(\hat{T}^n x) a(n) \geq \theta \cdot \tau(N_j).$$

In particular,

$$\limsup_{N \rightarrow \infty} \frac{(1/N) \sum_{n=1}^N f(\hat{T}^n x) a(n)}{\tau(N)} > 0.$$

Proof. Fix j large, and let us write N, q, M, x , suppressing the dependence on k', ℓ, j (except in parts of the proof where we wish to emphasize this dependence). Note that

$$qM \in [N - q, N].$$

Now,

$$\begin{aligned}
 \frac{1}{N} \sum_{n=1}^N f(\hat{T}^n x)a(n) &= \frac{1}{qM} \sum_{n=1}^{qM} f(\hat{T}^n x)a(n) + O\left(\frac{1}{M}\right) \\
 &= \frac{1}{qM} \left(\sum_{n=1}^{q-1} f(\hat{T}^n x)a(n) + \sum_{n=q}^{qM} f(\hat{T}^n x)a(n) \right) + O\left(\frac{1}{M}\right) \\
 &= \frac{1}{qM} \left(\sum_{n=1}^{q-1} f(\hat{T}^n x)a(n) + \sum_{n=1}^{qM} g(T^n x^{(\ell)})a(n) - \sum_{n=1}^{q-1} g(T^n x^{(\ell)})a(n) \right) \\
 &\quad + O\left(\frac{1}{M}\right) \\
 &= \frac{1}{q_{k'}^{(\ell)} M_{k'}^{(\ell)}} \sum_{n=1}^{q_{k'}^{(\ell)} M_{k'}^{(\ell)}} g(T^n x^{(\ell)})a(n) + O\left(\frac{1}{M_{k'}^{(\ell)}}\right) \\
 &\geq \frac{\theta}{8q_{k'}^{(i')}} + O\left(\frac{q_{k'}^{(\ell)}}{M_{k'}^{(\ell)}}\right).
 \end{aligned}$$

Note that in the third equality, we are again using Lemma 2.3 in a similar fashion to the proof of Lemma 2.6, which is allowed since $x^{(\ell)} \in P^{(\ell)} \times \{-1, 0, 1\}^{\mathbb{Z}}$. For the last inequality, we are using Lemmas 2.8 and 2.6 along with the definition of x .

We conclude that

$$\frac{1}{N_j} \sum_{n=1}^{N_j} f(\hat{T}^n x)a(n) \geq \frac{\theta}{8q_{k'}^{(i')}} + O\left(\frac{q_{k'}^{(\ell)}}{M_{k'}^{(\ell)}}\right).$$

By equation (4),

$$O\left(\frac{q_{k'}^{(\ell)}}{M_{k'}^{(\ell)}}\right) \leq O\left(\left(\frac{1}{q_{k'}^{(i')}}\right)^2\right),$$

and so, as long as j is large enough,

$$\frac{1}{N_j} \sum_{n=0}^{N_j-1} f(\hat{T}^n x)a(n) \geq \frac{\theta}{16q_{k'}^{(i')}}.$$

Finally, it follows from our choice of N_j that N_j is larger than the element of the sequence $q_k/3$ that comes after $q_{k'}^{(i')}/3$. So, by the choice of the sequence q_k ,

$$\frac{1}{16q_{k'}^{(i')}} \geq \tau(N_j).$$

Combining the last two displayed equations implies the claim. □

3. Proof of Theorem 1.2(2).

In this section, we prove Theorem 1.2(2). That is, we show that the system (X, \hat{T}) given in equation (6) satisfies the Möbius disjointness conjecture in equation (1). The proof

will be an application of Matomäki–Radziwiłł’s bound [13] on averages of multiplicative functions along short intervals. The Matomäki–Radziwiłł bound as well as its extension by Matomäki, Radziwiłł and Tao [14] have recently become a powerful tool to establish Möbius disjointness for systems with strong periodic behavior.

Denote a point $x \in X$ as

$$(x^{(0)}, x^{(1)}, x^{(2)}, x^{(3)}, i) \text{ where } x^{(\ell)} = (y^{(\ell)}, z^{(\ell)}).$$

For each $p = (y, z) \in \{-1, 0, 1\}^{\mathbb{N}} \times \{-1, 0, 1\}^{\mathbb{Z}}$ and $M \in \mathbb{N}$, denote by $[p]_M$ the truncation

$$[p]_M := ((y_1, \dots, y_M), (z_{-M}, \dots, z_M)).$$

Write $\mathcal{C}_M(X)$ for the space of cylinder functions $f(x)$ that only depends on $([x^{(\ell)}]_M)_{0 \leq \ell \leq 3}$ and the fifth coordinate $i \in \{0, 1, 2, 3\}$. Then $\bigcup_{M=1}^{\infty} \mathcal{C}_M(X)$ is dense in $\mathcal{C}(X)$ with respect to C^0 norm. In consequence, it suffices to verify equation (1) for all cylinder functions $f \in \mathcal{C}_M(X)$ for every M .

The main technical lemma that we need is the following.

LEMMA 3.1. *For all $0 \leq \ell \leq 3$ and $M, H \in \mathbb{N}$, and $x \in X$, there exists a set $\Lambda^{(\ell)}(M, H, x) \subseteq \mathbb{N}$ that satisfies:*

- (1) $\lim_{N \rightarrow \infty} (1/N) \#(\{1, \dots, N\} \cap \Lambda^{(\ell)}(M, H, x)) = 1;$
- (2) *for all $n \in \Lambda^{(\ell)}(M, H, x)$, $[T^{n+h}x^{(\ell)}]_M$ is constant for $0 \leq h \leq H - 1$.*

Proof. Since

$$x^{(\ell)} \in X_{\ell} = \text{cl} \left(\bigcup_{b \in \mathbb{N}_0} T^b(P^{(\ell)} \times \{-1, 0, 1\}^{\mathbb{Z}}) \right),$$

for each ℓ and all $N \in \mathbb{N}_0$, there exists $x^{(N, \ell)} \in \bigcup_{b \in \mathbb{N}_0} T^b(P^{(\ell)} \times \{-1, 0, 1\}^{\mathbb{Z}})$ such that

$$[x^{(N, \ell)}]_N = [x^{(\ell)}]_N \text{ for all } n \leq N.$$

We also choose $b^{(N, \ell)} \in \mathbb{N}_0$ and $\tilde{x}^{(N, \ell)} \in P^{(\ell)} \times \{-1, 0, 1\}^{\mathbb{Z}}$ such that $x^{(N, \ell)} = T^{b^{(N, \ell)}} \tilde{x}^{(N, \ell)}$.

Then, for $1 \leq n \leq N$ and $0 \leq h \leq H - 1$,

$$[T^{n+h}x^{(\ell)}]_M = [T^{n+h}x^{(N+H+M, \ell)}]_M = [T^{n+b^{(N+H+M, \ell)}+h}\tilde{x}^{(N+H+M, \ell)}]_M.$$

Therefore, by Lemma 2.1(1), $[T^{n+h}x^{(\ell)}]_M$ is constant for $0 \leq h \leq H - 1$ if

$$\Pi_1 \tilde{x}^{(N+H+M, \ell)}(n + b^{(N, \ell)} + h') = 0 \text{ for all } 0 \leq h' \leq H + M - 1. \tag{7}$$

Since $\tilde{x}^{(N+H+M, \ell)} \in P^{(\ell)} \times \{-1, 0, 1\}^{\mathbb{Z}}$, for every $k \in \mathbb{N}$, there is some $0 \leq r_k^{(\ell)} \leq q_k^{(\ell)} - 1$ such that

$$\Pi_1 \tilde{x}^{(N+H+M, \ell)}(n') = s_k^{(\ell)}(n' - r_k^{(\ell)}) \text{ for } q_k^{(\ell)} \leq n' < q_{k+1}^{(\ell)}.$$

In particular, $\Pi_1 \tilde{x}^{(N+H+M, \ell)}(n') = 0$ for all $q_k^{(\ell)} \leq n' < q_{k+1}^{(\ell)}$ with $n' \not\equiv r_k^{(\ell)} \pmod{q_k^{(\ell)}}$.

It follows that for each k , equation (7) holds on the set

$$\Lambda_{N,k}^{(\ell)}(M, H, x) := \{1 \leq n \leq N : q_k^{(\ell)} \leq n + b^{(N+H+M,\ell)} \leq q_{k+1}^{(\ell)} - H - M; \\ n + b^{(N+H+M,\ell)} \not\equiv r_k^{(\ell)} - H - M + 1, \dots, r_k^{(\ell)} - 1, r_k^{(\ell)} \pmod{q_k^{(\ell)}}\}.$$

Set $\Lambda_N^{(\ell)}(M, H, x) = \bigcup_{k=1}^{\infty} \Lambda_{N,k}^{(\ell)}(M, H, x) \subseteq \{1, \dots, N\}$. Then $[T^{n+h}x^{(\ell)}]_M$ is constant for $0 \leq h \leq H - 1$ if $n \in \Lambda_N^{(\ell)}(M, H, x)$.

Finally,

$$\lim_{N \rightarrow \infty} \frac{1}{N} \#\{1, \dots, N\} \cap \Lambda_N^{(\ell)}(M, H, x) = 1$$

because of the following facts: H and M are fixed, $b^{(N+H+M,\ell)} \geq 0$, $\lim_{k \rightarrow \infty} q_k^{(\ell)} = \infty$, and $\lim_{k \rightarrow \infty} (q_{k+1}^{(\ell)} / q_k^{(\ell)}) = \infty$. We conclude the proof by defining

$$\Lambda^{(\ell)}(M, H, x) := \bigcup_{N=1}^{\infty} \Lambda_N^{(\ell)}(M, H, x). \quad \square$$

COROLLARY 3.2. *For all $M, H \in \mathbb{N}$, and $x \in X$, there exists a set $\Lambda(M, H, x) \subseteq \mathbb{N}$ that satisfies the following:*

- (1) $\lim_{N \rightarrow \infty} (1/N) \#\{1, \dots, N\} \cap \Lambda(M, H, x) = 1$;
- (2) for all $f \in \mathcal{C}_M(X)$ and any given $n \in \Lambda(M, H, x)$, $f(\hat{T}^{n+h}x)$ is constant for $0 \leq h \leq H - 1$.

Proof. Let $\Lambda^{(\ell)}(M, H, x)$ be as in Lemma 3.1, and set

$$\Lambda(M, H, x) := \bigcap_{0 \leq \ell \leq 3} \Lambda^{(\ell)}(M, H, x) \subset \mathbb{N}.$$

Then clearly, we still have

$$\lim_{N \rightarrow \infty} \frac{1}{N} \#\{1, \dots, N\} \cap \Lambda(M, H, x) = 1.$$

Next, let $f \in \mathcal{C}_M(X)$. Since $f(\hat{T}^{n+h}x)$ only depends on $([T^{n+h}x^{(\ell)}]_M)_{0 \leq \ell \leq 3}$ and the i coordinate (that does not change when we apply \hat{T}), given $n \in \Lambda^{(\ell)}(i, M, H, x)$, it is constant for $0 \leq h \leq H - 1$ by Lemma 3.1. □

We are now ready to establish Möbius disjointness.

Proof of Theorem 1.2(2). As remarked in the beginning of this section, we may assume $f \in \mathcal{C}_M(X)$ for some M and $|f| \leq 1$. Let $x \in X$. Then for a fixed H , as $N \rightarrow \infty$,

$$\begin{aligned} \left| \frac{1}{N} \sum_{n=1}^N f(\hat{T}^n x) \mu(n) \right| &= \left| \frac{1}{N} \sum_{n=1}^N \frac{1}{H} \sum_{h=0}^{H-1} f(\hat{T}^{n+h} x) \mu(n) \right| + O\left(\frac{H}{N}\right) \\ &= \left| \frac{1}{N} \sum_{\substack{1 \leq n \leq N \\ n \in \Lambda(M, H, x)}} \frac{1}{H} \sum_{h=0}^{H-1} f(\hat{T}^{n+h} x) \mu(n) \right| + o_H(1) + O\left(\frac{H}{N}\right) \\ &\leq \frac{1}{N} \sum_{\substack{1 \leq n \leq N \\ n \in \Lambda(M, H, x)}} \left| \frac{1}{H} \sum_{h=0}^{H-1} f(\hat{T}^{n+h} x) \mu(n+h) \right| + o_H(1) + O\left(\frac{H}{N}\right). \end{aligned}$$

Here, $o_H(1)$ stands for a quantity that tends to 0 as $N \rightarrow \infty$ for a fixed H .

By Corollary 3.2, $f(\hat{T}^{n+h} x) = f(\hat{T}^n x)$ for every $n \in \Lambda(M, H, x)$ and $0 \leq h \leq H - 1$. So,

$$\begin{aligned} \left| \frac{1}{N} \sum_{n=1}^N f(\hat{T}^n x) \mu(n) \right| &\leq \frac{1}{N} \sum_{\substack{1 \leq n \leq N \\ n \in \Lambda(M, H, x)}} \left| \frac{1}{H} \sum_{h=0}^{H-1} f(\hat{T}^n x) \mu(n+h) \right| + o_H(1) + O\left(\frac{H}{N}\right) \\ &\leq \frac{1}{N} \sum_{\substack{1 \leq n \leq N \\ n \in \Lambda(M, H, x)}} \left| \frac{1}{H} \sum_{h=0}^{H-1} \mu(n+h) \right| + o_H(1) + O\left(\frac{H}{N}\right) \\ &\leq \frac{1}{N} \sum_{n=1}^N \left| \frac{1}{H} \sum_{h=0}^{H-1} \mu(n+h) \right| + o_H(1) + O\left(\frac{H}{N}\right) \\ &= O\left(\left(\frac{1}{\log H}\right)^{0.01} + \left(\frac{\log H}{\log N}\right)^{0.01} \right) + o_H(1) + O\left(\frac{H}{N}\right). \end{aligned}$$

The last step is given by [13, Theorem 1].

By letting $H \rightarrow \infty$ first, and then $N \rightarrow \infty$ for each fixed H , we see that

$$\frac{1}{N} \sum_{n=1}^N f(\hat{T}^n x) \mu(n) = o(1) \quad \text{as } N \rightarrow \infty. \quad \square$$

Acknowledgments. We thank the anonymous referee for helpful comments. Z.W. was supported by NSF grant DMS-1753042. A. A. acknowledges support from the Hebrew University of Jerusalem, where some of this research was done.

REFERENCES

[1] P. T. Bateman and H. G. Diamond. *Analytic Number Theory: An Introductory Course (Monographs in Number Theory, 1)*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2004.
 [2] J. Bourgain, P. Sarnak and T. Ziegler. Disjointness of Moebius from horocycle flows. *From Fourier Analysis and Number Theory to Radon Transforms and Geometry (Developments in Mathematics, 28)*. In memory of Leon Ehrenpreis. Eds. H. M. Farkas, R. C. Gunning, M. I. Knopp and B. A. Taylor. Springer, New York, 2013, pp. 67–83.

- [3] D. Dolgopyat, C. Dong, A. Kanigowski and P. Nándori. Flexibility of statistical properties for smooth systems satisfying the central limit theorem. *Invent. Math.* doi: [10.1007/s00222-022-01121-0](https://doi.org/10.1007/s00222-022-01121-0). Published online 9 June 2022.
- [4] E. H. el Abdalaoui, J. Kułaga-Przymus, M. Lemańczyk and T. de la Rue. Möbius disjointness for models of an ergodic system and beyond. *Israel J. Math.* **228**(2) (2018), 707–751.
- [5] E. H. el Abdalaoui, M. Lemańczyk and T. de la Rue. On spectral disjointness of powers for rank-one transformations and Möbius orthogonality. *J. Funct. Anal.* **266**(1) (2014), 284–317.
- [6] S. Ferenczi, J. Kułaga-Przymus and M. Lemańczyk. Sarnak’s conjecture: what’s new. *Ergodic Theory and Dynamical Systems in Their Interactions with Arithmetics and Combinatorics (Lecture Notes in Mathematics, 2213)*. CIRM Jean-Morlet Chair, Fall 2016. With a forward by P. Sarnak. Eds. S. Ferenczi, J. Kułaga-Przymus and M. Lemańczyk. Springer, Cham, 2018, pp. 163–235.
- [7] K. Frączek, A. Kanigowski and M. Lemańczyk. Prime number theorem for regular Toeplitz subshifts. *Ergod. Th. & Dynam. Sys.* **42** (2022), 1446–1473.
- [8] N. Frantzikinakis and B. Host. The logarithmic Sarnak conjecture for ergodic weights. *Ann. of Math. (2)* **187**(3) (2018), 869–931.
- [9] B. Green and T. Tao. The Möbius function is strongly orthogonal to nilsequences. *Ann. of Math. (2)* **175**(2) (2012), 541–566.
- [10] A. Kanigowski, M. Lemańczyk and M. Radziwiłł. Prime number theorem for analytic skew products. *Preprint*, 2020, [arXiv:2004.01125](https://arxiv.org/abs/2004.01125).
- [11] J. Kułaga-Przymus and M. Lemańczyk. Sarnak’s conjecture from the ergodic theory point of view. *Encycl. Complex. Syst. Sci.*, to appear.
- [12] Z. Lian and R. Shi. A counter-example for polynomial version of Sarnak’s conjecture. *Adv. Math.* **384** (2021), Paper no. 107765, 14 pp.
- [13] K. Matomäki and M. Radziwiłł. Multiplicative functions in short intervals. *Ann. of Math. (2)* **183**(3) (2016), 1015–1056.
- [14] K. Matomäki, M. Radziwiłł and T. Tao. An averaged form of Chowla’s conjecture. *Algebra Number Theory* **9**(9) (2015), 2167–2196.
- [15] P. Sarnak. Möbius randomness and dynamics. Lecture slides summer, 2010. <https://publications.ias.edu/sarnak/paper/546>.
- [16] P. Sarnak. Möbius randomness and dynamics. *Notices S. Afr. Math. Soc.* **43**(2) (2012), 89–97.
- [17] T. Tao. The Chowla conjecture and the Sarnak conjecture. <https://terrytao.wordpress.com/2012/10/14/the-chowla-conjecture-and-the-sarnak-conjecture/>.
- [18] P. Walters. *An Introduction to Ergodic Theory (Graduate Texts in Mathematics, 79)*. Springer, New York–Berlin, 1982.