Möbius disjointness for models of an ergodic system and beyond

H. El Abdalaoui[*], J. Kułaga-Przymus[*†], M. Lemańczyk[*†], T. de la Rue[*]

April 13, 2017

Abstract

We give a necessary and sufficient condition (called the strong MOMO property) for a uniquely ergodic model of an ergodic measure-preserving system $(Z, \mathcal{D}, \kappa, R)$ to have all uniquely ergodic models of the system Möbius disjoint. It follows that all uniquely ergodic models of: ergodic unipotent diffeomorphisms on nil-manifolds, discrete spectrum automorphisms, systems given by some substitutions of constant length (including the classical Thue-Morse and Rudin-Shapiro substitutions), systems determined by Kakutani sequences are Möbius disjoint. We also discuss the absence of the strong MOMO property in positive entropy systems.

Contents

[*]Research supported by the special program in the framework of the Jean Morlet semester "Ergodic Theory and Dynamical Systems in their Interactions with Arithmetic and Combinatorics".

[†]Research supported by Narodowe Centrum Nauki grant UMO-2014/15/B/ST1/03736.

1 Introduction

Assume that T is a continuous map of a compact metric space X. In 2010, P. Sarnak [38] stated the following conjecture: whenever the (topological) entropy of T is zero,

$$(1) \qquad \frac{1}{N}\sum_{n\leq N} f(T^n x)\boldsymbol{\mu}(n) \xrightarrow[N\to\infty]{} 0$$

for all $f \in C(X)$ and $x \in X$ (we recall that the Möbius function $\boldsymbol{\mu}$ is defined as $\boldsymbol{\mu}(p_1 \dots p_k) = (-1)^k$ for distinct primes $p_1, \dots, p_k$, $\boldsymbol{\mu}(1) = 1$ and $\boldsymbol{\mu}(n) = 0$ for the remaining natural numbers). When (1) holds (for all $f \in C(X)$ and $x \in X$), we also say that the system (X, T) is *Möbius disjoint.* Sarnak's conjecture is of purely topological dynamics nature. However, by looking at the limit points of empiric measures $\frac{1}{N}\sum_{n\leq N}\delta_{T^n x}$, $N \geq 1$, we see that the convergence in (1) is often determined by properties of measure-theoretic dynamical systems given by some invariant measures for (X, T), and dynamical properties of the subshift $X_{\boldsymbol{\mu}} \subset \{-1, 0, 1\}^{\mathbb{N}}$ given by $\boldsymbol{\mu}$ [1] – indeed, it is a playground, where certain relations between number theory and dynamics can be seen. One of the most motivating examples of such an interplay is that the famous Chowla conjecture for $\boldsymbol{\mu}$ [2] implies the validity of Sarnak's conjecture, see [4, 38, 40] for more details. Due to a recent result of Tao [39], we know that Sarnak's conjecture on its turn implies the logarithmic version of the Chowla conjecture. In particular, Sarnak's conjecture implies that all admissible blocks do appear on $\boldsymbol{\mu}$,[3] cf. comments on page 9 in [38]. While the Möbius disjointness has been proved for many particular classes of zero entropy dynamical systems (see the bibliography in [39]), to emphasize its measure-theoretic character, one can ask the following:

Question 1. Can measure-theoretic properties of a measurable system $(Z, \mathcal{D}, \kappa, R)$ imply the validity of (1) for any (X, T), $f \in C(X)$ provided that $x \in X$ is a generic point for a measure μ such that the measure-theoretic system $(X, \mathcal{B}(X), \mu, T)$ is measure-theoretically isomorphic to $(Z, \mathcal{D}, \kappa, R)$?

A particular instance of this approach would be the question whether some measure-theoretic properties of $(Z, \mathcal{D}, \kappa, R)$ can imply Möbius disjointness of all its uniquely ergodic models. This line of research was firstly taken up in

[1] $X_{\boldsymbol{\mu}}$ is given by the closure of the orbit of $\boldsymbol{\mu} \in \{-1, 0, 1\}^{\mathbb{N}}$ under the left shift.

[2] The Chowla conjecture says that $\frac{1}{N}\sum_{n\leq N}\boldsymbol{\mu}^{j_0}(n)\boldsymbol{\mu}^{j_1}(n+a_1)\dots\boldsymbol{\mu}^{j_r}(n+a_r) \xrightarrow[N\to\infty]{} 0$ for each choice $1 \leq a_1 < \dots < a_r$ and at least one j_k odd; equivalently, $\boldsymbol{\mu}$ is a generic point for the so-called Sarnak's measure on $\{-1, 0, 1\}^{\mathbb{N}}$ ([4],[38]).

[3] This observation has been communicated to us by W. Veech.

[2], where so-called AOP property (see (6) below) has been introduced, and it was shown that all uniquely ergodic models of AOP automorphisms are Möbius disjoint. It follows that all uniquely ergodic models of (zero entropy) affine automorphisms on compact connected Abelian groups are Möbius disjoint [2], or, more generally, all uniquely ergodic models of unipotent diffeomorphisms on nil-manifolds so are [16].[4] In particular, it follows from [2] that all uniquely ergodic models of totally ergodic discrete spectrum automorphisms are Möbius disjoint. In an unpublished (earlier) version of the present article it was shown how to use new remarkable results on the average behavior of non-pretensious multiplicative functions on, so-called, short intervals [33, 34], to see that all uniquely ergodic models of finite rotations are Möbius disjoint. Moreover, in [21], it has been proved recently that all uniquely ergodic models of an arbitrary discrete spectrum automorphism are Möbius disjoint. Following the above lines, the following natural question emerges:

Question 2. Does Möbius disjointness in a certain uniquely ergodic model of an ergodic system yield Möbius disjointness in all its uniquely ergodic models?

Clearly, in the zero entropy case, the potential positive answer to this question is supported by Sarnak's conjecture.

The main result of the paper, see Theorem 6, is an "almost" positive answer to Question 2. We will show that the existence of one "good" uniquely ergodic model of a system implies Möbius disjointness of all uniquely ergodic models of the system.

To formulate better our result, we need to introduce the notion of MOMO which was implicitly used in [2]. The definition is quite similar to (1), but we allow the orbit to be changed from time to time, less and less often. Let $\boldsymbol{u}\colon \mathbb{N} \to \mathbb{C}$ be an arbitrary arithmetic function. One might think of $\boldsymbol{u}$ as being e.g. the Möbius function, which justifies the terminology of the following definitions, but as our main results are valid independently of the nature of a sequence, we formulate them in this more abstract setting.

Definition 3 (MOMO property: Möbius Orthogonality on Moving Orbits)**.** We say that (X, T) satisfies the *MOMO property* [relatively to $\boldsymbol{u}$] if, for any increasing sequence of integers $0 = b_0 < b_1 < b_2 < \cdots$ with $b_{k+1} - b_k \to \infty$, for any sequence (x_k) of points in X, and any $f \in C(X)$,

$$\frac{1}{b_K} \sum_{k<K} \sum_{b_k \leq n < b_{k+1}} f(T^{n-b_k} x_k) \boldsymbol{u}(n) \xrightarrow[K\to\infty]{} 0. \tag{2}$$

The requirement in the next definition is still stronger.

Definition 4 (strong MOMO property)**.** We say that (X, T) satisfies the *strong MOMO property* [relatively to $\boldsymbol{u}$] if, for any increasing sequence of integers $0 = b_0 < b_1 < b_2 < \cdots$ with $b_{k+1} - b_k \to \infty$, for any sequence (x_k) of points in X, and any $f \in C(X)$,

$$\frac{1}{b_K} \sum_{k<K} \left| \sum_{b_k \leq n < b_{k+1}} f(T^{n-b_k} x_k) \boldsymbol{u}(n) \right| \xrightarrow[K\to\infty]{} 0. \tag{3}$$

[4]On the other hand, it is still unknown whether all uniquely ergodic models of horocycle flows are Möbius disjoint.

By taking $f = 1$ in Definition 4, we obtain that whenever strong MOMO holds, $\boldsymbol{u}$ has to satisfy:

$$\frac{1}{b_K}\sum_{k<K}\left|\sum_{b_k\leq n<b_{k+1}}\boldsymbol{u}(n)\right| \xrightarrow[K\to\infty]{} 0 \tag{4}$$

for every sequence $0 = b_0 < b_1 < b_2 < \cdots$ with $b_{k+1} - b_k \to \infty$.[5] In particular, $\frac{1}{N}\sum_{n<N}\boldsymbol{u}(n) \xrightarrow[N\to\infty]{} 0$.

Note that if $\boldsymbol{u}$ is bounded, by unique ergodicity, to verify the MOMO or the strong MOMO property we only need to check the relevant convergence for a linearly L^1-dense set of continuous functions. We will see later (see Corollary 17) that Sarnak's conjecture is equivalent to the strong MOMO property (relatively to $\boldsymbol{\mu}$) for each zero entropy system. However, directly from Definition 4, it follows that the strong MOMO property implies uniform convergence in (1).[6] In particular, we have the following:

Proposition 5 (Cf. Remark 19)**.** *If Sarnak's conjecture is true then for all zero entropy systems (X,T) and $f \in C(X)$, we have $\frac{1}{N}\sum_{n\leq N} f(T^n x)\boldsymbol{\mu}(n) \to 0$ uniformly in $x \in X$.*

By $M(X,T)$ we denote the set of T-invariant Borel probability measures, and $M^e(X,T)$ its subset of ergodic measures. If $\mu_1,\dots,\mu_t \in M(X,T)$ then by $\mathrm{conv}(\mu_1,\dots,\mu_t)$, we denote the corresponding convex envelope. Finally, given $x \in X$, by Q-gen(x) we denote the set of measures in $M(X,T)$ for which x is quasi-generic. We are now ready to state the main result of the paper (see also Theorem 14, for a slightly stronger version).

Theorem 6. *Let $(Z,\mathcal{D},\kappa,R)$ be an ergodic dynamical system. Let $\boldsymbol{u}\colon \mathbb{N}\to\mathbb{C}$ be an arithmetic function. The following conditions are equivalent:*

(P1) There exists a uniquely ergodic model (Y,S) of $(Z,\mathcal{D},\kappa,R)$ such that (Y,S) enjoys the strong MOMO property.

(P2) For any topological dynamical system (X,T) and any $x\in X$, if there exists a finite number of T-invariant measures μ_j, $1\leq j\leq t$, such that

- *for each j, $(X,\mathcal{B}(X),\mu_j,T)$ is measure-theoretically isomorphic to $(Z,\mathcal{D},\kappa,R)$,*
- *any measure for which x is quasi-generic is a convex combination of the measures μ_j, i.e. Q-gen$(x) \subset \mathrm{conv}(\mu_1,\dots,\mu_t)$,*

then $\frac{1}{N}\sum_{n\leq N} f(T^n x)\boldsymbol{u}(n) \xrightarrow[N\to\infty]{} 0$ for each $f\in C(X)$.

---

[5] We recall that by a result of Matomaki and Radziwiłł [33] it follows that $\boldsymbol{\mu}$ satisfies (4).

[6] Notice that if for a system (X,T) and $f \in C(X)$ we do not have uniform convergence of the sums $\frac{1}{N}\sum_{n\leq N} f(T^n\cdot)\boldsymbol{u}(n)$, then there exists $\varepsilon_0 > 0$ such that for each $k \geq 1$, we can find $b_k \geq k$ and $x_k \in X$ for which $\left|\frac{1}{b_k}\sum_{n\leq b_k} f(T^n x_k)\boldsymbol{u}(n)\right| \geq \varepsilon_0$. We can assume that the distances $b_k - b_{k-1}$ grow very rapidly so that we obtain $\frac{1}{b_K}\sum_{k<K}(b_k - b_{k-1})\left|\frac{1}{b_k-b_{k-1}}\sum_{b_{k-1}\leq n<b_k} f(T^n x_k)\boldsymbol{u}(n)\right| \geq \sum_{k<K}\frac{b_k-b_{k-1}}{b_K}\varepsilon_0/2 \geq \varepsilon_0/3$, whence the strong MOMO property (relative to $\boldsymbol{u}$) fails.

(P3) All uniquely ergodic models of $(Z, \mathcal{D}, \kappa, R)$ enjoy the strong MOMO property.

While the proof of implication (P2)⇒(P3) uses some ideas from [2], the proof of implication (P1)⇒(P2) heavily depends on the main ideas from [21].

We will now discuss some consequences of Theorem 6. Let $\mathbb{D}$ stand for the unit disc, $\mathbb{D}_L := L\mathbb{D}$ for $L > 0$ and S for the shift in $\mathbb{D}_L^{\mathbb{Z}}$. Given $\boldsymbol{u} \in (\mathbb{D}_L)^{\mathbb{Z}}$, by $X_{\boldsymbol{u}}$, we denote the subshift generated by $\boldsymbol{u}$.

Corollary 7. *Fix $\kappa \in M^e((\mathbb{D}_L)^{\mathbb{Z}}, S)$, $\kappa \neq \delta_{(\dots 0.00\dots)}$. Let (X, T) be any uniquely ergodic model of $((\mathbb{D}_L)^{\mathbb{Z}}, \kappa, S)$. Then for any $\boldsymbol{u} \in (\mathbb{D}_L)^{\mathbb{Z}}$ for which Q-gen$(\boldsymbol{u}) \subset$ $\mathrm{conv}(\kappa_1, \dots, \kappa_m)$, where $((\mathbb{D}_L)^{\mathbb{Z}}, \kappa_j, S)$ for $j = 1, \dots, m$ is measure-theoretically isomorphic to $((\mathbb{D}_L)^{\mathbb{Z}}, \kappa, S)$, the system (X, T) does not satisfy the strong MOMO property (relative to $\boldsymbol{u}$).*

In other words, see Section 2.5 for details, if (X, T) is fixed then all points $\boldsymbol{u}$ as above are "visible" in X in the following sense:

$$\begin{gathered}(\exists \varepsilon_0 > 0)\ (\exists f \in C(X))\ (\exists (x_k) \subset X) \\ \big(\exists A = \textstyle\bigcup_{k=1}^{\infty} [a_k, c_k) \subset \mathbb{N} \text{ of disjoint intervals}, c_k - a_k \to \infty,\ \overline{d}(A) > 0\big) \\ \text{such that} \\ \tfrac{1}{c_k - a_k} \left| \textstyle\sum_{n=0}^{c_k - a_k - 1} f(T^n x_k) \boldsymbol{u}(a_k + n) \right| \geq \varepsilon_0 \text{ for each } k \geq 1.\end{gathered} \tag{5}$$

Recently, Downarowicz and Serafin [14] constructed positive entropy homeomorphisms of arbitrarily large entropy which are Möbius disjoint. The following natural question arises:

Question 8. Does there exist an ergodic positive entropy measure-theoretic system all uniquely ergodic models of which are Möbius disjoint?

A partial answer is given by the following result:

Corollary 9. *Assume that $\boldsymbol{u} \in (\mathbb{D}_L)^{\mathbb{Z}}$ is generic for a Bernoulli measure κ. Then for each uniquely ergodic system (X, T) with $h(X, T) > h((\mathbb{D}_L)^{\mathbb{Z}}, \kappa, S)$, we do not have the strong MOMO property relative to $\boldsymbol{u}$.*

It follows that if we assume the validity of the Chowla conjecture for the Liouville function[7] $\boldsymbol{\lambda}$ then for no uniquely ergodic system (X, T) of entropy larger than $\log 2$, we have the strong MOMO property relative to $\boldsymbol{\lambda}$.[8] In fact, by the proof of implication (P2) ⇒ (P3) it follows that, whenever an ergodic measure-theoretic system $(Z, \mathcal{D}, \kappa, R)$ has the entropy larger than $\log 2$, there must exist a topological system which has at most three ergodic measures, all yielding systems isomorphic to R, and for which the Sarnak property (relative to $\boldsymbol{\lambda}$) does not hold.

Recall that an ergodic automorphism R is said to have AOP (asymptotically orthogonal powers) [2] if for each $f, g \in L_0^2(Z, \mathcal{D}, \kappa)$, we have

$$\lim_{\mathscr{P} \ni r, s \to \infty, r \neq s} \sup_{\kappa \in J^e(R^r, R^s)} \left| \int_{X \times X} f \otimes g \, d\kappa \right| = 0.\text{[9]} \tag{6}$$

[7]The Chowla conjecture is equivalent to saying that $\boldsymbol{\lambda}$ is a generic point for the Bernoulli measure $B(1/2, 1/2)$ for the shift on $\{-1, 1\}^{\mathbb{N}}$, cf. footnote 2.

[8]It seems to be an interesting challenge to construct a uniquely ergodic model of the Bernoulli system $B(1/2, 1/2)$ which has no strong MOMO property relative to $\boldsymbol{\lambda}$.

[9]$\mathscr{P}$ stands for the set of prime numbers. The set $J^e(R^r, R^s)$ consists of $R^r \times R^s$-invariant measures on $Z \times Z$ which are ergodic and whose both projections on Z are κ.

The AOP property implies zero entropy. It also implies total ergodicity. Indeed, as clearly AOP is closed under taking factors, we merely need to notice that $Rx = x + 1$ acting on $\mathbb{Z}/k\mathbb{Z}$ with $k \geq 2$ has no AOP property. The latter easily follows from the Dirichlet theorem on primes in arithmetic progressions.

Clearly, if the powers of R are pairwise disjoint[10] in the Furstenberg sense [19] then R enjoys the AOP property. The AOP property of $(Z, \mathcal{D}, \kappa, R)$ implies the MOMO property in every uniquely ergodic model of R [2] relatively to a multiplicative[11] $\boldsymbol{u}\colon \mathbb{N} \to \mathbb{C}$, $|\boldsymbol{u}| \leq 1$, satisfying $\frac{1}{N}\sum_{n\leq N} \boldsymbol{u}(n) \to 0$ when $N \to \infty$. In particular, the MOMO property relatively to $\boldsymbol{\mu}$ holds. We will show that AOP implies condition (P3) in Theorem 6 (see Section 2.4), which results in the following.

Theorem 10. *Let $\boldsymbol{u}\colon \mathbb{N} \to \mathbb{C}$ be multiplicative, $|\boldsymbol{u}| \leq 1$. Suppose that $(Z, \mathcal{D}, \kappa, R)$ satisfies AOP. Then the following are equivalent:*

- *$\boldsymbol{u}$ satisfies (4);*
- *The strong MOMO property relatively to $\boldsymbol{u}$ is satisfied in each uniquely ergodic model (X, T) of R.*

In particular, if the above holds, for each $f \in C(X)$, we have $\frac{1}{N}\sum_{n\leq N} f(T^n x)\boldsymbol{u}(n) \to 0$ uniformly in $x \in X$.

Corollary 11. *Assume that $(Z, \mathcal{D}, \kappa, R)$ enjoys the AOP property. Then in each uniquely ergodic model (X, T) of R, we have*

$$\frac{1}{M}\sum_{M\leq m<2M}\left|\frac{1}{H}\sum_{m\leq h<m+H} f(T^n x)\boldsymbol{\mu}(n)\right| \to 0 \text{ when } H \to \infty, H/M \to 0 \tag{7}$$

for all $f \in C(X)$, $x \in X$.

Note that even if we consider $(Z, \mathcal{D}, \kappa, R)$ with the property that all of its powers are disjoint, the assertion (7) true in each uniquely ergodic model (X, T) of R is a new result.[12]

In the second part of the paper, we will concentrate on examples of automorphisms for which (P2) in Theorem 6 holds. We will show a general criterion to lift the strong MOMO property to extensions and deal with coboundary extensions of homeomorphisms satisfying the strong MOMO property. The validity of (P2) in a large subclass of so-called generalized Morse sequences [26] then follows by exploiting the idea of lifting the strong MOMO property by some cocycle extensions (see Theorem 26 below). We will show in particular that (P2) in Theorem 6 holds (hence also the strong MOMO property holds) for all uniquely ergodic models of: all unipotent diffeomorphisms on nilmanifolds, all transformations with discrete spectrum, typical automorphism of a probability standard Borel space, systems coming from bijective substitutions and some other "close" to that; in particular, the classical Thue-Morse and Rudin-Shapiro systems and, finally, for systems determined by so-called Kakutani sequences. For the Möbius disjointness in the classes of systems listed above, see [6, 8, 10, 16, 20, 35, 41].

[10] This is a "typical" property of an automorphism of a probability standard Borel space [12].
[11] Multiplicativity means that $\boldsymbol{u}(1) = 1$ and $\boldsymbol{u}(mn) = \boldsymbol{u}(m)\boldsymbol{u}(n)$ whenever m and n are relatively prime.
[12] Möbius disjointness for this case is already noticed in [8].

We will now illustrate a meaning of the strong MOMO property for Kakutani sequences. Using the sequence $1, 2, 2^2, \ldots$, each natural number $n \geq 1$ can be written uniquely as $n = \sum_{j\geq 0} \varepsilon_j 2^j$ with $\varepsilon_j = 0$ or 1. Then, we can consider the sequence $s_2(n) := \sum_{j\geq 0} \varepsilon_j \bmod 2$, $n \geq 1$. Using the sequence $2 = p_1 < p_2 < p_3 < \ldots$ of consecutive prime numbers, each natural number $n \geq 2$ can be written uniquely as $n = \prod_{j\geq 1} p_j^{\alpha_j}$ with $\alpha_j \geq 0$, $j \geq 1$. Then, we can consider the sequence $b(n) := \sum_{j\geq 1} \alpha_j \bmod 2$, $n \geq 1$. Studying properties of the sequence $(s_2(n))$ concerns the additive structure of natural numbers, while studying properties of the sequence $(b(n))$ concerns the multiplicative structure of $\mathbb{N}$. We will now show that the sequences $(-1)^{s_2(n)}$ and $(-1)^{b(n)}$, $n \geq 1$, do not correlate in a strong way. Note that for the first sequence, we have $(-1)^{s_2(n)} = (-1)^{x(n)}$, where $x \in \{0,1\}^{\mathbb{N}}$ is the classical Thue-Morse sequence, while for the second, we have $\boldsymbol{\lambda}(n) := (-1)^{b(n)}$, where $\boldsymbol{\lambda}$ is the classical Liouville function.[13]

Before, we state a result, let us consider a generalization of the approach. Following B. Green [20], let $A \subset \mathbb{N}$. Consider $x = x_A \in \{0,1\}^{\mathbb{N}}$ such that

$$x(n) = \sum_{i\in A} \varepsilon_i \bmod 2, \text{ where } n = \sum_{i\geq 0} \varepsilon_i 2^i.$$

As explained in [3], each $x = x_A$ is a Kakutani sequence (and each Kakutani sequence determines an A).[14]

It follows from the strong MOMO property (with respect to $\boldsymbol{\lambda}$) of the system determined by x_A that

$$\frac{1}{b_K} \sum_{k<K} \left| \sum_{b_k \leq n < b_{k+1}} (-1)^{x_A(n)} \boldsymbol{\lambda}(n) \right| \xrightarrow[K\to\infty]{} 0.$$

Now, $\frac{1}{N}\sum_{n\leq N} (-1)^{x_A(n)} \to 0$ [15] and $\frac{1}{N}\sum_{n\leq N} \boldsymbol{\lambda}(n) \to 0$ (this is equivalent to the PNT). More than that, the same property holds on each short interval for the first sequence (by the unique ergodicity of the system determined by a Kakutani sequence) or on a typical short interval for the Liouville function by a result of [33]. Recall that if we have two random variables X, Y taking two values ± 1 with probability $1/2$ then they are independent if and only if $\int XY = 0$. By all this, we have proved the following form of independence of the sequences $(x_A(n))$ and $(b(n))$:

Proposition 12. *We have*

$$\frac{1}{M} \sum_{M\leq m<2M} \left| \frac{1}{H} \sum_{m\leq h<m+H} (-1)^{x_A(h)} \boldsymbol{\lambda}(h) \right| \to 0$$

when $H \to \infty$, $H/M \to 0$. *Moreover, for each* $e, f \in \{-1, 1\}$,

$$\frac{1}{M} \sum_{M\leq m<2M} \left| \frac{1}{H} \left| \{m \leq h < m+H : \ (-1)^{x_A(h)} = e, \boldsymbol{\lambda}(h) = f\} \right| - \frac{1}{4} \right| \to 0$$

when $H \to \infty$, $H/M \to 0$.

[13] As $n \mapsto b(n)$ is completely additive, the Liouville function is a completely multiplicative function.

[14] This observation is due to C. Mauduit.

[15] This follows from the fact that for the unique invariant measure μ_{x_A} for the subshift determined by x_A, we have $\int (-1)^{z(0)} \, d\mu_{x_A}(z) = 0$.

2 MOMO property in models of an ergodic system

Let $\boldsymbol{u}\colon \mathbb{N} \to \mathbb{C}$ be a fixed arithmetic function.

Definition 13 (Sarnak property)**.** A point $x \in X$ satisfies the *Sarnak property* [relatively to $\boldsymbol{u}$] if, for any $f \in C(X)$,

$$\frac{1}{N}\sum_{n<N} f(T^n x)\boldsymbol{u}(n) \xrightarrow[N\to\infty]{} 0. \tag{8}$$

We say that (X,T) satisfies the Sarnak property [relatively to $\boldsymbol{u}$] if any point $x \in X$ satisfies the Sarnak property.

Note that, if we restrict ourselves to *bounded* arithmetic functions $\boldsymbol{u}$, then the Sarnak property, the MOMO property (Definition 3) and the strong MOMO property (Definition 4) remain unchanged if we modify $\boldsymbol{u}$ on a subset of $\mathbb{N}$ of density 0.

2.1 Orthogonality properties of a measure-theoretic dynamical system

We consider now an ergodic measure-theoretic dynamical system $(Z,\mathcal{D},\kappa,R)$, and we introduce several properties of this system (still relatively to $\boldsymbol{u}$), cf. Theorem 6.

Property P1. There exists a topological dynamical system (Y,S), and an S-invariant probability measure ν on Y, such that

- (Y,S) satisfies the strong MOMO property,
- $(Y,\mathcal{B}(Y),\nu,S)$ is measure-theoretically isomorphic to $(Z,\mathcal{D},\kappa,R)$.

Property P2. For any topological dynamical system (X,T) and any $x \in X$, if there exists a finite number of T-invariant measures μ_j, $1 \le j \le t$, such that

- for each j, $(X,\mathcal{B}(X),\mu_j,T)$ is measure-theoretically isomorphic to $(Z,\mathcal{D},\kappa,R)$,
- any measure for which x is quasi-generic is a convex combination of the measures μ_j, i.e. $\text{Q-gen}(x) \subset \text{conv}(\mu_1,\dots,\mu_t)$,

then x satisfies the Sarnak property.

(Observe that P2 implies that any uniquely ergodic model of $(Z,\mathcal{D},\kappa,R)$ satisfies the Sarnak property.)

Property P3. Any uniquely ergodic model (Y,S) of $(Z,\mathcal{D},\kappa,R)$ satisfies the strong MOMO property.

Theorem 14. *Properties P1, P2 and P3 are equivalent.*

We observe that, since there always exists a uniquely ergodic model of $(Z,\mathcal{D},\kappa,R)$ by the Jewett-Krieger theorem, the implication P3 $\Longrightarrow$ P1 is obvious. The two other implications are treated in separate subsections.

2.2 Proof of P1 $\Longrightarrow$ P2

This proof strongly relies on ideas borrowed from [21]. Let (Y, S) and ν be given by P1, and let (X, T), $\mu_1, \dots, \mu_t$ and x be as in the assumptions of P2. In particular, for each $1 \le j \le t$, $(X, \mathcal{B}(X), \mu_j, T)$ is measure-theoretically isomorphic to $(Y, \mathcal{B}(Y), \nu, S)$. We also fix a continuous function f on X, and $0 < \varepsilon < \frac{1}{2}$.

Since the measures μ_j are ergodic for T, we can find T-invariant disjoint Borel subsets X_j, $1 \le j \le t$ with $\mu_j(X_j) = 1$, and measure-theoretic isomorphisms $\phi_j \colon (X_j, \mathcal{B}(X_j), \mu_j, T) \to (Y, \mathcal{B}(Y), \nu, S)$. For each $1 \le j \le t$, Lusin's theorem now provides a compact subset $W_j \subset X_j$, with $\mu_j(W_j) > 1 - \varepsilon^4$, such that the restriction $\phi_j|_{W_j}$ is continuous. Then this restriction is in fact a homeomorphism between W_j and its image $\phi_j(W_j)$, which is a compact subset of Y. The function $f \circ \phi_j^{-1}$ is continuous on this compact subset, so by the Tietze extension theorem it can be extended to a continuous function g_j on the entire space Y, with $\|g_j\|_\infty = \|f\|_\infty$. The following observation will be useful: for any $w \in X$ and any $s \ge 0$,

$$
[w \in W_j \text{ and } T^s w \in W_j] \Longrightarrow f(T^s w) = g_j\left(\phi_j(T^s w)\right) = g_j\left(S^s(\phi_j w)\right). \tag{9}
$$

Informally, the preceding observation means that these compact sets W_j can be used as "windows" through which we can see the behaviour of the dynamical system (Y, S). We would like to see this behaviour along long pieces of orbits, but these windows are not T-invariant. This is why we need to define, for each $1 \le j \le t$ and each integer $L \ge 1$, the following subset of W_j:

$$
B_j(L) := \left\{ w \in W_j : \frac{1}{L} \sum_{\ell < L} \mathbb{1}_{W_j}(T^\ell w) > 1 - \varepsilon^2 \right\}.
$$

Observe that $B_j(L)$ can be written as the finite union of all sets of the form $T^{-\ell_1} W_j \cap \cdots \cap T^{-\ell_r} W_j$, where $\{\ell_1, \dots, \ell_r\}$ ranges over all finite subsets of $\{0, \dots, L-1\}$ of cardinality larger than $1 - \varepsilon$. Each $T^{-\ell} W_j$ being compact, $B_j(L)$ is compact. On $W_j \setminus B_j(L)$, we have the inequality

$$
\frac{1}{L} \sum_{\ell < L} \mathbb{1}_{X \setminus W_j}(T^\ell w) \ge \varepsilon^2.
$$

Therefore, we have

$$
\varepsilon^2 \mu_j\big(W_j \setminus B_j(L)\big) \le \int_X \frac{1}{L} \sum_{\ell < L} \mathbb{1}_{X \setminus W_j} \circ T^\ell \, d\mu_j = \mu_j(X \setminus W_j) < \varepsilon^4,
$$

and this yields $\mu_j(B_j(L)) > 1 - \varepsilon$.

Let $d(\cdot, \cdot)$ be the distance on X. To each integer $L \ge 1$, we also associate a positive number $\eta(L)$, small enough so that for any $w, w' \in X$,

$$
d(w, w') < \eta(L) \Longrightarrow \forall 0 \le n < L, |f(T^n w) - f(T^n w')| < \varepsilon. \tag{10}
$$

The following lemma is the only place where we use the assumption on the quasi-genericity of the point x.

Lemma 15. *For each $L \geq 1$, let $B(L)$ be the disjoint union $B_1(L) \sqcup \cdots \sqcup B_t(L)$, which is also a compact subset of X. Then*

$$\limsup_{N\to\infty} \frac{1}{N} \# \Big\{ n \in \{0, \dots, N-1\} : d\big(T^n x, B(L)\big) \geq \eta(L) \Big\} < \varepsilon.$$

Proof. Let (N_i) be an increasing sequence of positive integers along which the convergence to the limit superior in the statement of the lemma holds. Extracting if necessary a subsequence, we can assume that x is quasi-generic, along this sequence (N_i), for a T-invariant measure μ which is of the form $\mu = \alpha_1\mu_1 + \cdots + \alpha_t\mu_t$, with $\alpha_j \geq 0$ and $\alpha_1 + \cdots + \alpha_t = 1$. Since $\mu_j(B(L)) \geq \mu_j(B_j(L)) > 1 - \varepsilon$ for each j, we also have $\mu(B(L)) > 1 - \varepsilon$.

Using again the Tietze extension theorem, we can construct a continuous function $h \in C(X)$, $0 \leq h \leq 1$, satisfying $h(w) = 0$ if $w \in B(L)$, and $h(w) = 1$ if $d\big(w, B(L)\big) \geq \eta(L)$. Then, we have

$$\begin{aligned}
&\limsup_{N\to\infty} \frac{1}{N} \# \Big\{ n \in \{0, \dots, N-1\} : d\big(T^n x, B(L)\big) \geq \eta(L) \Big\} \\
&= \lim_{i\to\infty} \frac{1}{N_i} \# \Big\{ n \in \{0, \dots, N_i-1\} : d\big(T^n x, B(L)\big) \geq \eta(L) \Big\} \\
&\leq \lim_{i\to\infty} \frac{1}{N_i} \sum_{n<N_i} h(T^n x) \\
&= \int_X h \, d\mu \leq \mu(X \setminus B(L)) < \varepsilon.
\end{aligned}$$

□

We now fix an increasing sequence of integers $1 \leq L_1 < L_2 < \cdots$. We also choose an increasing sequence of integers $(M_i)_{i\geq 1}$ such that

$$\lim_{i\to\infty} \frac{1}{M_i} \left| \sum_{n<M_i} f(T^n x)\boldsymbol{u}(n) \right| = \limsup_{N\to\infty} \frac{1}{N} \left| \sum_{n<N} f(T^n x)\boldsymbol{u}(n) \right|.$$

Set $M_0 := 0$. With the help of Lemma 15, we can assume, passing to a subsequence if necessary, that for each $i \geq 1$,

$$\frac{1}{M_i - M_{i-1}} \# \Big\{ n \in \{M_{i-1}, \dots, M_i - 1\} : d\big(T^n x, B(L_i)\big) \geq \eta(L_i) \Big\} < \varepsilon. \tag{11}$$

We can also assume that for each $i \geq 1$, M_i is large enough to have

$$L_i < \varepsilon M_i. \tag{12}$$

Then, for any integer $b \geq 0$, there exists a unique $i \geq 1$ such that $M_{i-1} \leq b < M_i$. We say that b is *good* if $d\big(T^b x, B(L_i)\big) < \eta(L_i)$. Observe that, by (11), the proportion of good integers between M_{i-1} and M_i is always at least $1 - \varepsilon$.

Finally, we inductively define a third increasing sequence of integers $0 = b_0 < b_1 < b_2 < \cdots$ in the following way. Let b_1 be the smallest good integer $b \geq 1$. Assume that the integer b_k has already been defined ($k \geq 1$), and that it is a good integer. Let $i \geq 1$ be such that $M_{i-1} \leq b_k < M_i$. Since b_k is good, there exist $1 \leq j_k \leq t$ and a point $x_k \in B_{j_k}(L_i)$ such that $d(T^{b_k} x, x_k) < \eta(L_i)$. By the definition of $\eta(L_i)$, this implies that for any $0 \leq s < L_i$, $|f(T^{b_k+s}x) - f(T^s x_k)| < \varepsilon$. But by the definition of $B_{j_k}(L_i)$, the number of integers s, $0 \leq s < L_i - 1$, such that $T^s x_k \in W_{j_k}$ is at least $(1-\varepsilon)L_i$. Moreover, since $x_k \in W_{j_k}$, by (9), for each such s, we have

$$f(T^s x_k) = g_{j_k}(S^s y_k), \text{ where } y_k := \phi_{j_k}(x_k),$$

which yields $|f(T^{b_k+s}x) - g_{j_k}(S^s y_k)| < \varepsilon$. We therefore get

$$\sum_{s<L_i} |f(T^{b_k+s}x) - g_{j_k}(S^s y_k)| < \varepsilon L_i + 2\varepsilon L_i \|f\|_\infty.$$

Now, we define b_{k+1} as the smallest integer $b \geq b_k + L_i$ which is good. The above inequality then gives

$$\text{(13)} \quad \sum_{b_k \leq n < b_{k+1}} |f(T^n x) - g_{j_k}(S^{n-b_k} y_k)| < \varepsilon(b_{k+1} - b_k)(1 + 2\|f\|_\infty) \\ + 2\|f\|_\infty \#\{b_k \leq n < b_{k+1} : n \text{ is not good}\}.$$

And since $L_i \to \infty$, we also have $b_{k+1} - b_k \to \infty$.

Now, fix $i \geq 1$. Let $K \geq 1$ be the largest integer such that $b_K \leq M_i$. We want to approximate the sum

$$S_{M_i} := \sum_{n<M_i} f(T^n x)\boldsymbol{u}(n)$$

by the following expression coming from the dynamical system (Y, S)

$$E_K := \sum_{k<K} \sum_{b_k \leq n < b_{k+1}} g_{j_k}(S^{n-b_k} y_k)\boldsymbol{u}(n).$$

Considering that (13) holds for $1 \leq k < K$, observing that all integers $n \in \{0 = b_0, \ldots, b_1 - 1\}$ are not good, and that the number of integers less than M_i which are not good is bounded by εM_i, we get

$$|S_{M_i} - E_K| < \varepsilon M_i(1 + 4\|f\|_\infty) + 2\|f\|_\infty (M_i - b_K).$$

Now, by the definition of K, $M_i - b_K$ can be bounded by $L_i + \#\{n < M_i : n \text{ is not good}\}$, which is itself bounded by $2\varepsilon M_i$ by (12). Therefore, there exists a constant $C > 0$ such that

$$\limsup_{N\to\infty} \frac{1}{N} \left| \sum_{n<N} f(T^n x)\boldsymbol{u}(n) \right| = \lim_{i\to\infty} \frac{1}{M_i} |S_{M_i}| \leq \limsup_{K\to\infty} \frac{1}{b_K} |E_K| + C\varepsilon.$$

It only remains to do the following estimation:

$$\begin{aligned} \frac{1}{b_K}|E_K| &\leq \frac{1}{b_K} \sum_{k<K} \left| \sum_{b_k \leq n < b_{k+1}} g_{j_k}(S^{n-b_k} y_k)\boldsymbol{u}(n) \right| \\ &\leq \sum_{1\leq j\leq t} \frac{1}{b_K} \sum_{k<K} \left| \sum_{b_k \leq n < b_{k+1}} g_j(S^{n-b_k} y_k)\boldsymbol{u}(n) \right|, \end{aligned}$$

which goes to 0 as $K \to \infty$ by the strong MOMO property of (Y, S). This concludes the proof of the implication P1 $\Longrightarrow$ P2.

2.3 Proof of P2 $\implies$ P3

Before we begin the proof, let us recall the following elementary result.

Lemma 16. *Assume that* $(c_n) \subset \mathbb{C}$ *and* $(m_n) \subset \mathbb{N}$*. Then:*

(A) If for each sequence $(\varepsilon_k) \subset \{-1,1\}^{\mathbb{N}}$*, we have*

$$\frac{1}{m_N}\sum_{n\le N}\varepsilon_n c_n \xrightarrow[N\to\infty]{} 0 \text{ then } \frac{1}{m_N}\sum_{n\le N}|c_n| \xrightarrow[N\to\infty]{} 0;$$

(B) If the sequence (c_n) *is contained in a convex cone* C *whose vertex is at* $0 \in \mathbb{C}$*, and which is not a half-plane, then* $\frac{1}{m_N}\sum_{n\le N} c_n \xrightarrow[N\to\infty]{} 0$ *if and only if* $\frac{1}{m_N}\sum_{n\le N}|c_n| \xrightarrow[N\to\infty]{} 0$.

Proof. To see (A) write ($c_n = a_n + ib_n$, $a_n, b_n \in \mathbb{R}$)

$$\frac{1}{m_N}\sum_{n\le N}\varepsilon_n c_n = \frac{1}{m_N}\sum_{n\le N}\varepsilon_n a_n + i\frac{1}{m_N}\sum_{n\le N}\varepsilon_n b_n \xrightarrow[N\to\infty]{} 0$$

to reduce the problem to $c_n \in \mathbb{R}$, and then set $\varepsilon_n := \mathrm{sign}(c_n)$.

To see (B), first multiply the whole sequence by some $e \in \mathbb{C}$, $|e| = 1$ to assume without loss of generality that C is the cone delimited by the half-lines $y = a_1 x$, $y = a_2 x$, $a_1 \neq 0 \neq a_2$, and $x \geq 0$. It follows that there is a constant $\gamma \geq 1$ such that for each $c \in C$, $|c| \leq \gamma \mathrm{Re}(c)$. If $\frac{1}{m_N}\sum_{n\le N} c_n \xrightarrow[N\to\infty]{} 0$, then $\frac{1}{m_N}\sum_{n\le N}\mathrm{Re}(c_n) \xrightarrow[N\to\infty]{} 0$, whence

$$\frac{1}{m_N}\sum_{n\le N}|c_n| \leq \gamma\frac{1}{m_N}\sum_{n\le N}\mathrm{Re}(c_n) \xrightarrow[N\to\infty]{} 0.$$

□

Let (Y,S) be a uniquely ergodic model of $(Z,\mathcal{D},\kappa,R)$, and let ν be the unique S-invariant measure. We fix a continuous function $f \in C(Y)$, an increasing sequence of integers $0 = b_0 < b_1 < b_2 < \cdots$ with $b_{k+1} - b_k \to \infty$, and a sequence of points $(y_k)_{k\geq 0}$ in Y. Let $\mathbb{A} := \{1, e^{i2\pi/3}, e^{i4\pi/3}\}$ be the set of third roots of unity. For each k, let $e_k \in \mathbb{A}$ be such that

$$e_k\left(\sum_{b_k\le n<b_{k+1}} f(S^{n-b_k}y_k)\boldsymbol{u}(n)\right)$$

belongs to the closed convex cone $\{0\}\cup\{z\in\mathbb{C} : \arg(z)\in[-\pi/3,\pi/3]\}$. Then, by Lemma 16, the convergence that we want to prove:

$$\frac{1}{b_K}\sum_{k<K}\left|\sum_{b_k\le n<b_{k+1}} f(S^{n-b_k}y_k)\boldsymbol{u}(n)\right| \xrightarrow[K\to\infty]{} 0$$

is equivalent to the convergence

$$\frac{1}{b_K}\sum_{k<K}\sum_{b_k\le n<b_{k+1}} e_k f(S^{n-b_k}y_k)\boldsymbol{u}(n) \xrightarrow[K\to\infty]{} 0. \tag{14}$$

Now, we introduce a new topological dynamical system: let $X := (Y \times \mathbb{A})^{\mathbb{N}}$, and let T be the shift on X. We define a particular point $x \in X$ by setting

$$x_n := (S^{n-b_k} y_k, e_k) \text{ whenever } b_k \le n < b_{k+1}. \tag{15}$$

Let μ be a measure for which x is quasi-generic, along a sequence (N_r). Since $b_{k+1} - b_k \to \infty$, we have

$$\left(\frac{1}{N_r} \sum_{n<N_r} \delta_{T^n x}\right) \Big(\{v \in X : (v_1, a_1) = (Sv_0, a_0)\}\Big) \xrightarrow[N_r \to \infty]{} 1.$$

Since the set $\{v \in X : (v_1, a_1) = (Sv_0, a_0)\}$ is closed, by Portmanteau theorem it must be of full measure μ in (X, T). Moreover, such a measure μ must be T-invariant, hence, it is concentrated on the set of sequences which are of the form $\Big((y, a), (Sy, a), (S^2 y, a), \dots\Big)$ for some $y \in Y$ and some $a \in \mathbb{A}$. Since μ is T-invariant, its marginal given by the first y-coordinate is S-invariant, so it is equal to ν by unique ergodicity. On the other hand, the marginal given by the a-coordinate must be of the form $\alpha_0 \delta_1 + \alpha_1 \delta_{e^{i2\pi/3}} + \alpha_2 \delta_{e^{i4\pi/3}}$. Using now the fact that any ergodic system is disjoint from the identity, μ must be the direct product of its marginals on $Y^{\mathbb{N}}$ and $\mathbb{A}^{\mathbb{N}}$. Hence, μ must be of the form $\alpha_0 \mu_0 + \alpha_1 \mu_1 + \alpha_2 \mu_2$, where μ_j is the pushforward of ν by the map $y \mapsto \Big((y, e^{i2\pi j/3}), (Sy, e^{i2\pi j/3}), (S^2 y, e^{i2\pi j/3}), \dots\Big)$. Then $(X, \mathcal{B}(X), \mu_j, S)$ is measure-theoretically isomorphic to $(Y, \mathcal{B}(Y), \nu, S)$, hence also to $(Z, \mathcal{D}, \kappa, R)$: the assumptions needed to apply P2 are fulfilled. It follows that x satisfies the Sarnak property, and if we write the corresponding convergence for the continuous function $g \in C(X)$ defined by $g\Big((w_0, a_0), (w_1, a_1), (w_2, a_2), \dots\Big) := a_0 f(w_0)$, we exactly get (14).[16]

Note that, if in the above proof $(Z, \mathcal{D}, \kappa, R)$ is an arbitrary system of zero entropy, and (Y, S) is any uniquely ergodic model of $(Z, \mathcal{D}, \kappa, R)$, then the topological entropy of S is zero. The same property holds for the topological system (X, T) constructed above. If Sarnak's conjecture holds, the system (X, T) is Möbius disjoint and hence the strong MOMO property holds for (Y, S). It follows that if Sarnak's conjecture holds, then the strong MOMO property is satisfied in each zero entropy uniquely ergodic system. In fact, we have even more than this:

Corollary 17. *Sarnak's conjecture holds if and only if the strong MOMO property (relatively to $\boldsymbol{\mu}$) is satisfied for each system of zero topological entropy.*

Proof. In view of the above proof of P2 $\Longrightarrow$ P3, what we need to show is that the orbit closure of $x \in X$ defined in (15) under T has zero topological entropy. Suppose first that x is quasi-generic for some measure μ. Denote by $\mu^{(1)}$ the marginal of μ given by the first (y, e)-coordinate. Arguing as above, we obtain that $\mu^{(1)}$ is a measure invariant under $S \times I$, in particular, it has zero entropy. Moreover, μ is the image of $\mu^{(1)}$ by the map $(I \times I) \times (S \times I) \times (S^2 \times I) \times \dots$, i.e. also has zero entropy.

Consider now z in the orbit closure of x and suppose that it is a quasi-generic point for some measure. If $n_j \to \infty$ and $z = \lim_{j\to\infty} T^{n_j} x$, then either

$$z = ((y, e), (Sy, e), (S^2 y, e), \dots) \text{ for some } (y, e) \in Y \times \mathbb{A}$$

[16] Note also that $\int_X g \, d\mu_j = a_0 \int_X f(w_0) \, d\mu_j = a_0 \int_Y f \, d\nu$.

or

$$z = ((y_1, e_1), (Sy_1, e_1), \dots, (S^\ell y_1, e_1), (y_2, e_2), (Sy_2, e_2), \dots) \\ \text{for some } (y_1, e_1), (y_2, e_2) \in Y \times \mathbb{A} \text{ and } \ell \geq 0.$$

Indeed, we can approximate any "window" $z[1, M]$ by $T^{n_j}[1, M] = x[n_j + 1, n_j + M]$ and when $n_j \to \infty$, such a window has at most one point of "discontinuity", that is, it contains at most once two consecutive coordinates which are not successive images by $S \times I$ of some $(y_k, e_k) \in Y \times \mathbb{A}$. Thus, to conclude, we can use the same argument as in the first part of the proof. □

Remark 18. Assume that $\boldsymbol{u}\colon \mathbb{N} \to \mathbb{C}$ is an arithmetic function relatively to which the Sarnak property holds for each zero entropy T. We have already noticed that $\boldsymbol{u}$ has to satisfy (4) but in fact, we obtain a stronger condition.

Fix (X, T) of zero entropy. We claim that we have the following arithmetic version of the strong MOMO property:

(16) for each $N \geq 1$, $h = 0, 1, \dots N-1$, $(b_k) \subset \mathbb{N}$ with $b_{k+1} - b_k \to \infty$, (x_k) and $f \in C(X)$, we have $\frac{1}{b_K} \sum_{k<K} \left| \sum_{b_k \leq n < b_{k+1}} f(T^n x_k) \boldsymbol{u}(Nn + h) \right| \xrightarrow[K\to\infty]{} 0$

Indeed, consider the N-discrete suspension $\widetilde{T}$ of T, i.e. $\widetilde{X} := X \times \{0, 1, \dots, N-1\}$ and let the homeomorphism $\widetilde{T}$ act by the formula $\widetilde{T}(x, j) := (x, j+1)$ when $0 \leq j < N-1$ and $\widetilde{T}(x, N-1) = (Tx, 0)$. Then $(\widetilde{X}, \widetilde{T})$ has still zero entropy, and, by Corollary 17, the strong MOMO property is satisfied for $(\widetilde{X}, \widetilde{T})$. Define $F \in C(\widetilde{X})$ by setting $F(x, h) = f(x)$ and 0 otherwise. Hence

$$\frac{1}{Nb_K} \sum_{k<K} \left| \sum_{Nb_k \leq n < Nb_{k+1}} F(\widetilde{T}^n(x_k, 0)) \boldsymbol{u}(n) \right| \xrightarrow[K\to\infty]{} 0.$$

Therefore

$$\frac{1}{Nb_K} \sum_{k<K} \left| \sum_{b_k \leq n < b_{k+1}} f(T^n(x_k)) \boldsymbol{u}(Nn + h) \right| \xrightarrow[K\to\infty]{} 0$$

and the claim (16) follows.

In particular, the function $\boldsymbol{u}$ has to satisfy

$$\frac{1}{b_K} \sum_{k<K} \left| \sum_{b_k \leq n < b_{k+1}} \boldsymbol{u}(Nn + h) \right| \xrightarrow[K\to\infty]{} 0. \tag{17}$$

Therefore, $\boldsymbol{u}$ is aperiodic[17] and in fact, it is aperiodic on "typical" short interval:

$$\frac{1}{M} \sum_{M \leq m < 2M} \left| \frac{1}{H} \sum_{m \leq g < m+H} \boldsymbol{u}(Ng + h) \right| \to 0 \tag{18}$$

[17] Note that any non-principal Dirichlet character of modulus q yields a (completely) multiplicative function for which we have (4) (since the sum of the values along the period equals zero and $b_{k+1} - b_k \to 0$) but which is not aperiodic.

when $H \to \infty$ and $H/M \to 0$.

Note that if above, for X we take the one-point space, then $(\widetilde{X}, \widetilde{T})$ stands for the rotation by 1 on $\mathbb{Z}/N\mathbb{Z}$ and the strong MOMO property follows from [34], see Subsection 4.1. It follows that (17) holds for $\boldsymbol{u} = \boldsymbol{\mu}$.

Remark 19. One can think of the assertion of Proposition 5 (with $\boldsymbol{\mu}$ replaced by a sequence $\boldsymbol{u}$) as of a statement on "unique ergodicity". For example, if (Y, S) is a topological system and $y \in Y$ is a generic point for a Bernoulli measure ν and if $g \in C(Y)$ has ν-mean zero then for every zero topological entropy system (X, T) and every $f \in C(X)$ the averages $\frac{1}{N}\sum_{n\leq N} f(T^n x)g(S^n y) \to 0$ uniformly in $x \in X$ (here $\boldsymbol{u}(n) = g(S^n y)$). However, it is easy to prove uniform convergence here by repeating any of classical proofs of the fact that in a uniquely ergodic system ergodic averages converge uniformly.[18]

2.4 AOP implies P1, P2 and P3 for Möbius

We return now to the case where $\boldsymbol{u}\colon \mathbb{N} \to \mathbb{C}$ is a multiplicative function, $|\boldsymbol{u}| \leq 1$, satisfying (4), in particular, we can take $\boldsymbol{u} = \boldsymbol{\mu}$.

The purpose of this section is to prove that, in this setting, the AOP property ensures the validity of P1, P2 and P3. For this, it will be useful to introduce the following weaker version of P2 for the measure-theoretic dynamical system $(Z, \mathcal{D}, \kappa, R)$, where we restrict the class of continuous functions for which we demand that convergence (8) holds.

Property P2*. For any topological dynamical system (X, T) and any $x \in X$, if there exists a finite number of T-invariant measures μ_j, $1 \leq j \leq t$, such that

- for each j, $(X, \mathcal{B}(X), \mu_j, T)$ is measure-theoretically isomorphic to $(Z, \mathcal{D}, \kappa, R)$,
- Q-gen$(x) \subset \mathrm{conv}(\mu_1, \dots, \mu_t)$,

then, for any $f \in C(X)$ satisfying $\forall 1 \leq j \leq t$, $\int_X f\, d\mu_j = 0$, convergence (8) holds.

Proposition 20. *If $\boldsymbol{u}$ is a multiplicative function, $|\boldsymbol{u}| \leq 1$, satisfying (4), then P2* $\Longrightarrow$P3.*

Proof. It follows by (4) that if we want to prove the strong MOMO property in a specific uniquely ergodic model (Y, S) with the unique invariant measure ν, it is enough to check the required convergence for a continuous function f with $\int_Y f\, d\nu = 0$. But then the continuous function g constructed at the end of the proof of P2 $\Longrightarrow$ P3 will satisfy $\int_X g\, d\mu_j = 0$ for $j = 0, 1, 2$ (see footnote 16), hence property P2* will be enough. ◻

We will need the following criterion:

[18] If $|\frac{1}{b_k}\sum_{n\leq b_k} f(T^n x_k)g(S^n y)| \geq \varepsilon_0$ (cf. footnote 6) then by considering $\frac{1}{b_k}\sum_{n\leq b_k}\delta_{(T^n x_k, S^n \boldsymbol{u})}$, we can assume that it converges to a joining of a measure with zero entropy and ν (which is Bernoulli). Hence it is the product measure, and we easily obtain a contradiction with the fact that $\int g\, d\nu = 0$).

Proposition 21 (KBSZ criterion, [25, 8], see also [2])**.** *Let $(a_n) \subset \mathbb{C}$ be bounded. If*

$$\limsup_{\mathscr{P}\ni r,s\to\infty, r\neq s} \limsup_{N\to\infty} \left| \frac{1}{N} \sum_{n<N} a_{rn}\overline{a}_{sn} \right| = 0 \tag{19}$$

then $\lim_{N\to\infty} \frac{1}{N}\sum_{n<N} a_n \boldsymbol{v}(n) = 0$ for each multiplicative function $\boldsymbol{v}\colon \mathbb{N}\to\mathbb{C}$, $|\boldsymbol{v}|\leq 1$.

Theorem 22. *If $\boldsymbol{u}$ is a multiplicative function, $|\boldsymbol{u}|\leq 1$, satisfying (4), then AOP $\Longrightarrow$ P2*. In particular, the implication holds for the Möbius function $\boldsymbol{\mu}$.*

Proof. Assume that $(Z,\mathcal{D},\kappa,R)$ enjoys the AOP property. Let (X,T), x, $\mu_1,\dots,\mu_t$ and f be as in the assumptions of Property P2*. In order to prove the convergence (8), we want to apply the KBSZ criterion to the sequence $\big(f(T^n x)\big)_{n\geq 0}$. Given $\varepsilon>0$, we have to show that, if $r\neq s$ are two different primes which are large enough, then

$$\limsup_{N\to\infty} \frac{1}{N}\left|\sum_{n<N} f(T^{rn}x)\overline{f}(T^{sn}x)\right| < \varepsilon.$$

Let (N_i) be an increasing sequence of integers along which the convergence to the above limit superior holds. Without loss of generality, we can also assume that the sequence of empirical measures

$$\sigma_{N_i} := \frac{1}{N_i}\sum_{n<N_i} \delta_{(T^{rn}x, T^{sn}x)}$$

converges weakly to some $T^r\times T^s$-invariant measure ρ on $X\times X$. We then have

$$\limsup_{N\to\infty} \frac{1}{N}\left|\sum_{n<N} f(T^{rn}x)\overline{f}(T^{sn}x)\right| = \left|\int_{X\times X} f\otimes\overline{f}\, d\rho\right|,$$

and it is enough to prove that, if r and s are large enough, then for each ergodic component γ of ρ, we have

$$\left|\int_{X\times X} f\otimes\overline{f}\, d\gamma\right| < \varepsilon.$$

Let ρ_1 (respectively ρ_2) be the marginal of ρ on the first (respectively the second) coordinate. Then ρ_1 is T^r-invariant, and

$$\frac{1}{r}\left(\rho_1 + T_*(\rho_1) + \cdots + T_*^{r-1}(\rho_1)\right) = \lim_{i\to\infty} \frac{1}{rN_i}\sum_{0\leq n<rN_i} \delta_{T^n x}$$

is a T-invariant probability measure on X for which x is quasi-generic. By assumption, this measure is a convex combination of $\mu_1,\dots,\mu_t$, and ρ_1 is absolutely continuous with respect to this convex combination. By the same argument, ρ_2 is also absolutely continuous with respect to some convex combination of $\mu_1,\dots,\mu_t$. Moreover, since $(Z,\mathcal{D},\kappa,R)$ has the AOP property, this system is totally ergodic, and by isomorphism, this also holds for each

$(X,\mathcal{B}(X),\mu_j,T)$. It follows that each ergodic component of ρ is an ergodic joining of $(X,\mathcal{B}(X),\mu_i,T^r)$ and $(X,\mathcal{B}(X),\mu_j,T^s)$ for some $i,j\in\{1,\dots,t\}$.

For each j, $1\le j\le t$, let $\varphi_j\colon (Z,\mathcal{D},\kappa,R)\to(X,\mathcal{B}(X),\mu_j,T)$ be an isomorphism of measure-theoretic dynamical systems. Set also $f_j:=f\circ\varphi_j\in L^2(\kappa)$. Let $\gamma\in J^e\big((X,\mathcal{B}(X),\mu_i,T^r),(X,\mathcal{B}(X),\mu_j,T^s)\big)$ be an ergodic component of ρ. Then the pushforward image $(\varphi_i\times\varphi_j)_*(\gamma)$ is an ergodic joining of $(Z,\mathcal{D},\kappa,R^r)$ and $(Z,\mathcal{D},\kappa,R^s)$, and we have

$$\left|\int_{X\times X} f\otimes\overline{f}\,d\gamma\right| = \left|\int_{Z\times Z} f_i\otimes\overline{f_j}\,d(\varphi_i\times\varphi_j)_*(\gamma)\right| \le \sup_{i,j\in\{1,\dots,t\}}\ \sup_{\eta\in J^e(R^r,R^s)}\left|\int_{Z\times Z} f_i\otimes\overline{f_j}\,d\eta\right|.$$

But by the AOP property, if r and s are large enough, the modulus of the RHS is bounded by ε. □

2.5 Strong MOMO property in positive entropy systems

In this section we prove Corollaries 7, 9 and equation (5).

Proof of Corollary 7. Suppose that the strong MOMO property (relative to $\boldsymbol{u}$) holds for (X,T). Then P1 from Theorem 14 holds. Equivalently, P2 holds. Notice that the assumptions of P2 are satisfied here (for $(\mathbb{D}_L^{\mathbb{Z}},S)$ and $\boldsymbol{u}$), whence the assertion of P2 also holds, i.e. $\boldsymbol{u}$ satisfies the Sarnak property. Take $g(y):=\overline{y(0)}$, and note that $\frac{1}{N}\sum_{n\le N}|\boldsymbol{u}(n)|^2=\frac{1}{N}\sum_{n\le N}g(S^n\boldsymbol{u})\boldsymbol{u}(n)\to 0$. Hence $\boldsymbol{u}$ is a generic point for $\delta_{(\dots 0.00\dots)}$, a contradiction. □

Notice that Corollary 7 puts some restrictions on dynamical properties of measures for which $\boldsymbol{u}$ is quasi-generic.

Corollary 23. *Let $\boldsymbol{u}$ be a bounded arithmetic function. Assume that Q-gen$(\boldsymbol{u})\subset$ $\mathrm{conv}(\kappa_1,\dots,\kappa_m)$, where all $(X_{\boldsymbol{u}},\kappa_j,S)$ are isomorphic and have the AOP property. Then $\kappa_1=\ldots=\kappa_m=\delta_{(\dots 00.00\dots)}$, i.e. $\boldsymbol{u}$ is equal to zero, up to a set of zero density.*

Proof. Let (X,T) be any uniquely ergodic model of $(X_{\boldsymbol{u}},\kappa_1,S)$. Then, since the AOP property is an isomorphism invariant, (X,T) also satisfies the AOP property. Since this system is additionally uniquely ergodic, we have the strong MOMO property for (X,T) and, by Corollary 7, the proof is complete. □

In particular, if we consider $\boldsymbol{\lambda}$ (or $\boldsymbol{\mu}$) then the set Q-gen$(\boldsymbol{\lambda})$ cannot be contained in $\mathrm{conv}(\kappa_1,\dots,\kappa_m)$, where the dynamical systems given by κ_j are all isomorphic and have the AOP property.[19] Studying properties of dynamical systems given by (potentially many) measures in Q-gen$(\boldsymbol{\lambda})$ is important in the light of a recent remarkable result of N. Frantzikinakis [18] which says that if Q-gen$(\boldsymbol{\lambda})$ consists solely of ergodic measures then the Chowla conjecture holds.

[19] The AOP property can be replaced by the existence of a uniquely ergodic model of the dynamical system associated to κ_1 for which we have the strong MOMO property.

Proof of Corollary 9. Assume that we have found a uniquely ergodic system (X,T), $h(X,T) > h((\mathbb{D}_L)^{\mathbb{Z}},\kappa,S)$ which has the strong MOMO property relative to $\boldsymbol{u}$. Let μ be the only invariant measure for T. Then, by Sinai's theorem, the Bernoulli automorphism $(\mathbb{D}_L^{\mathbb{Z}},\kappa,S)$ is a measure-theoretic factor of (X,μ,T). By a theorem of B. Weiss [43] there are uniquely (in fact, strictly) ergodic systems (X',T'), (Y',S') (with the unique invariant measures μ' and ν', respectively) and a continuous, equivariant map $\pi'\colon X'\to Y'$ such that (X',μ',T'), (Y',ν',S') are measure-theoretically isomorphic to (X,μ,T) and $(X_{\boldsymbol{u}},\kappa,S)$, respectively. Since (X,T) has the strong MOMO property relative to $\boldsymbol{u}$, so has (X',T') (by Theorem 6). Since (Y',S') is a topological factor of (X',T'), also (Y',S') enjoys the strong MOMO property relative to $\boldsymbol{u}$. But (Y',S') is a uniquely ergodic model of $(X_{\boldsymbol{u}},\kappa,S)$, and we obtain a contradiction with Corollary 7. □

Remark 24. *Even though in Corollary 9 we deal with systems of "sufficiently large" entropy, if a uniquely ergodic (X,T) has positive entropy then for some $k\geq 1$, the system (X,T^k) which we assume to remain uniquely ergodic, has sufficiently large entropy so that we can apply Corollary 7, hence the strong MOMO property (relative to $\boldsymbol{u}$) will be absent in it.*

The following technical lemma shows what happens if we contradict the condition from the definition of the strong MOMO property.

Lemma 25. *Let $f\in C(X)$, $(x_k)\subset X$, $(b_k)\subset\mathbb{N}$, $b_{k+1}-b_k\to\infty$. Assume that*

$$\limsup_{K\to\infty}\frac{1}{b_{K+1}}\sum_{k\leq K}\left|\sum_{b_k\leq n<b_{k+1}} f(T^nx_k)\boldsymbol{u}(n)\right|>0.$$

Then there exist $\varepsilon_0>0$ and a collection $\{[a_k,c_k)\subset\mathbb{N}:k\geq 1\}$ of disjoint intervals with $c_k-a_k\to\infty$ and $\overline{d}(\bigcup_{k\geq 1}[a_k,c_k))>0$ such that

$$\frac{1}{c_k-a_k}\left|\sum_{n=0}^{c_k-a_k-1} f(T^nx_k)\boldsymbol{u}(a_k+n)\right|\geq\varepsilon_0 \text{ for each } k\geq 1.$$

Proof. We begin the proof by the following simple observation: for $0\leq F\in L^\infty(X,\mathcal{B},\mu)$ and $\varepsilon_0\leq\int F\,d\mu$, we have

$$\varepsilon_0\leq\int_{[F\leq\varepsilon_0/2]}F\,d\mu+\int_{[F\geq\varepsilon_0/2]}F\,d\mu\leq\varepsilon_0/2+\|F\|_\infty\mu(\{x\in X: F(x)\geq\varepsilon_0/2\}),$$

whence

$$\mu(\{x\in X:\ F(x)\geq\varepsilon_0/2\})\geq\varepsilon_0/(2\|F\|_\infty).$$

Fix $(z_k)_{k\geq 1}$ and $(\alpha_k)_{k\geq 1}$ with $0\leq z_k\leq M$ and $\sum_{k\geq 1}\alpha_k=1$ and suppose that, for some $K\geq 1$, $\sum_{k\leq K}\alpha_kz_k\geq\varepsilon_0$. Set $y_k:=z_k$ for $k\leq K$ and $y_k:=0$ for $k>K$. Then clearly $\sum_{k\geq 1}\alpha_ky_k=\sum_{k\leq K}\geq\varepsilon_0$ and we can apply the above observation to obtain

$$\sum_{k\leq K,z_k\geq\varepsilon_0/2}\alpha_k=\sum_{k\leq K,y_k\geq\varepsilon_0/2}\alpha_k\geq\varepsilon_0/(2M). \tag{20}$$

Let now $\varepsilon_0 > 0$ and $(K_\ell)_{\ell\geq 1}$ be such that

$$\frac{1}{b_{K_\ell+1}} \sum_{k\leq K_\ell} \left| \sum_{b_k\leq n<b_{k+1}} f(T^n x_k)\boldsymbol{u}(n) \right| \geq \varepsilon_0. \tag{21}$$

Set $z_k := \left| \frac{1}{b_{k+1}-b_k} \sum_{b_k\leq n<b_{k+1}} f(T^n x_k)\boldsymbol{u}(n) \right|$ and $\alpha_k := \frac{b_{k+1}-b_k}{b_{K_\ell+1}}$. Then (21) takes the form

$$\sum_{k\leq K_\ell} \alpha_k z_k \geq \varepsilon_0.$$

Since

$$\overline{d}\Big(\bigcup_{k\geq 1, z_k\geq \varepsilon_0/2} [b_k, b_{k+1}) \Big) \geq \limsup_{\ell\to\infty} \frac{1}{K_\ell} \sum_{k\leq K_\ell, z_k\geq \varepsilon_0/2} (b_{k+1}-b_k),$$

it follows from (20) that

$$\overline{d}\Big(\bigcup_{k\geq 1, z_k\geq \varepsilon_0/2} [b_k, b_{k+1}) \Big) \geq \varepsilon_0/(2M).$$

Thus, we obtain a sequence $[a_k, c_k)$, $k \geq 1$, of disjoint intervals of the form $[b_{k_s}, b_{k_s+1})$ such that

$$\overline{d}\Big(\bigcup_{k\geq 1} [a_k, c_k) \Big) \geq \varepsilon_0/(2M) > 0$$

and

$$\left| \frac{1}{c_k - a_k} \sum_{a_k\leq n<c_k} f(T^n x_k)\boldsymbol{u}(n) \right| \geq \varepsilon_0/2.$$

This completes the proof. $\square$

Equation (5) immediately follows from the above lemma.

3 Lifting strong MOMO to extensions

In this section, we continue our considerations on orthogonality to bounded multiplicative functions $\boldsymbol{u}$ satisfying (4); in particular, we can take $\boldsymbol{u} = \boldsymbol{\mu}$.

3.1 Lifting strong MOMO to Rokhlin extensions

We start from a uniquely ergodic system (X, T) enjoying the strong MOMO property, and we denote by μ_X the unique T-invariant probability measure. We consider a continuous extension $\bar{T}$ of T to some product space $X \times Y$; Y is also a compact metric space, and $\bar{T}$ is a continuous transformation of $X \times Y$ which has the form $\bar{T}(x, y) = (Tx, S_x(y))$. We also assume that $\bar{T}$ is uniquely ergodic, its unique invariant measure having the form $\mu_X \otimes \mu_Y$ for some probability measure μ_Y on Y. Our purpose is to give a sufficient condition for the strong MOMO property to hold in the extension $(X \times Y, \bar{T})$. The condition we give can be seen as a form of relative disjointness of $\bar{T}^r$ and $\bar{T}^s$ over the base system (for large different prime integers r and s), and the proof relies on the same kind of arguments as in the proof of P2 $\Longrightarrow$ P3.

Theorem 26. *Suppose that, for all large enough prime numbers $r \neq s$, the following holds: each probability measure on $(X\times Y)\times(X\times Y)$ which is invariant and ergodic under the action of $\bar{T}^r \times \bar{T}^s$ is, up to a natural permutation of coordinates, of the form $\rho\otimes\mu_Y\otimes\mu_Y$, where ρ is some $T^r\times T^s$-invariant measure on $X\times X$. Then the strong MOMO property also holds in $(X\times Y, \bar{T})$.*

Proof. We fix an increasing sequence of integers $0 = b_0 < b_1 < b_2 < \cdots$ with $b_{k+1} - b_k \to \infty$, and a sequence of points $\big((x_k, y_k)\big)_{k\geq 0}$ in $X \times Y$. We also fix a continuous function f on $X\times Y$, and we assume that f is of the form $f = f_1\otimes f_2 : (x,y) \mapsto f_1(x)f_2(y)$ where $f_1 \in C(X)$ and $f_2 \in C(Y)$. Considering continuous functions of this type on $X\times Y$ is enough for our purposes, since they generate a dense subspace in $C(X\times Y)$ and $\bar{T}$ is uniquely ergodic. We thus have to prove the convergence

$$\frac{1}{b_K}\sum_{k<K}\left|\sum_{b_k\leq n<b_{k+1}} f_1\otimes f_2\left(\bar{T}^{n-b_k}(x_k,y_k)\right)\boldsymbol{u}(n)\right| \xrightarrow[K\to\infty]{} 0.$$

We observe that, subtracting $\int_Y f_2\,d\mu_Y$ from f_2 if necessary, which does not affect the above limit by the strong MOMO property of (X,T), we can always assume that

$$\int_Y f_2\,d\mu_Y = 0. \tag{22}$$

We again consider the set $\mathbb{A} := \{1, e^{i2\pi/3}, e^{i4\pi/3}\}$ of third roots of unity, and for each k, we choose $e_k\in\mathbb{A}$ such that

$$e_k\left(\sum_{b_k\leq n<b_{k+1}} f_1\otimes f_2\left(\bar{T}^{n-b_k}(x_k,y_k)\right)\boldsymbol{u}(n)\right)$$

belongs to the closed convex cone $\{0\}\cup\{z\in\mathbb{C} : \arg(z)\in[-\pi/3,\pi/3]\}$. Then again by Lemma 16, it is enough to prove that

$$\frac{1}{b_K}\sum_{k<K}\sum_{b_k\leq n<b_{k+1}} e_k f_1\otimes f_2\left(\bar{T}^{n-b_k}(x_k,y_k)\right)\boldsymbol{u}(n) \xrightarrow[K\to\infty]{} 0. \tag{23}$$

Now, we introduce the space $Z := (X\times Y\times\mathbb{A})^{\mathbb{N}}$, on which acts the shift map S. We define a particular point $z\in Z$ by setting $z_n := (\bar{T}^{n-b_k}(x_k,y_k), e_k)$ whenever $b_k \leq n < b_{k+1}$, and we consider the continuous function F on Z defined by

$$F\Big((u_0,v_0,a_0),(u_1,v_1,a_1),\dots\Big) := a_0 f_1(u_0) f_2(v_0).$$

Then, to get (23), it is enough to establish the orthogonality of $\boldsymbol{u}$ and $(F(S^n z))$, i.e.

$$\frac{1}{N}\sum_{n<N} F(S^n z)\boldsymbol{u}(n) \xrightarrow[N\to\infty]{} 0. \tag{24}$$

Using the KBSZ criterion, the above holds as soon as, for all large enough different prime numbers r and s, we have

$$\frac{1}{N}\sum_{n<N} F(S^{rn}z)\overline{F}(S^{sn}z) \xrightarrow[N\to\infty]{} 0. \tag{25}$$

Let (N_i) be an increasing sequence of positive integers along which the sequence of empirical measures

$$\frac{1}{N_i}\sum_{n<N_i}\delta_{(S^{rn}z,S^{sn}z)},\ i\geq 1,$$

converges to some $S^r\times S^s$-invariant probability measure $\mu_{Z\times Z}$ on $Z\times Z$. We therefore have

$$\frac{1}{N_i}\sum_{n<N_i}F(S^{rn}z)\overline{F}(S^{sn}z)\xrightarrow[i\to\infty]{}\int_{Z\times Z}F\otimes\overline{F}\,d\mu_{Z\times Z}.$$

Since $b_{k+1}-b_k\to\infty$, the measure $\mu_{Z\times Z}$ is concentrated on the set of pairs $(w,w')\in Z^2$ whose coordinates are of the form

$$w=\Big((u,v,a),(\bar{T}(u,v),a),(\bar{T}^2(u,v),a),\dots\Big)$$

and

$$w'=\Big((u',v',a'),(\bar{T}(u',v'),a'),(\bar{T}^2(u',v'),a'),\dots\Big)$$

for some u,u' in X, some v,v' in Y, and some a,a' in $\mathbb{A}$. By the invariance of $\mu_{Z\times Z}$ under $S^r\times S^s$, the marginal of $\mu_{Z\times Z}$ given by the coordinates $\big((u,v),(u',v')\big)$ is $\bar{T}^r\times\bar{T}^s$-invariant. Hence, if instead of $\mu_{Z\times Z}$ we consider an ergodic component γ of $\mu_{Z\times Z}$, the assumption of the theorem implies that the marginal given by the coordinates $\big((u,v),(u',v')\big)$ is of the form $\rho\otimes\mu_Y\otimes\mu_Y$ for some probability measure ρ on $X\times X$. Moreover, using again the disjointness of ergodicity and identity, we see that under γ the coordinates (a,a') are independent of $\big((u,v),(u',v')\big)$. If we denote by $\gamma_{\mathbb{A}\times\mathbb{A}}$ the marginal defined on $\mathbb{A}\times\mathbb{A}$ by the coordinates (a,a'), we thus have

$$\begin{aligned}\int_{Z\times Z}F\otimes\overline{F}\,d\gamma=\int_{\mathbb{A}\times\mathbb{A}}a\overline{a'}d\gamma_{\mathbb{A}\times A}(a,a')\int_{X\times X}f_1(u)\overline{f_1}(u')\,d\rho(u,u')\\ \times\int_Y f_2(v)\,d\mu_Y(v)\int_Y\overline{f_2}(v')\,d\mu_Y(v'),\end{aligned}$$

hence this integral vanishes by (22). Since this is true for all ergodic components of $\mu_{Z\times Z}$, we get

$$\lim_{i\to\infty}\frac{1}{N_i}\sum_{n<N_i}F(S^{rn}z)\overline{F}(S^{sn}z)=\int_{Z\times Z}F\otimes\bar{F}d\mu_{Z\times Z}=0,$$

so (25) follows and the proof is complete. □

3.2 Sarnak's conjecture for continuous extensions by coboundaries

In the theorem below we consider homeomorphisms for which the measure-theoretic systems determined by ergodic invariant measures are all isomorphic. However, the set of ergodic invariant measures is uncountable which makes a direct use of Theorem 6 questionable. On the other hand, this set is quite structured which allows one to repeat the main steps of the proof of the implication P1 $\Longrightarrow$ P2.

Theorem 27. *Suppose that (Y, S) is uniquely ergodic and satisfies the strong MOMO property [relatively to $\boldsymbol{u}$]. Let G be a compact Abelian group (with Haar measure λ_G) and let $\varphi\colon Y \to G$ be continuous with $\varphi = \psi \circ S - \psi$, where $\psi\colon Y \to G$ is measurable. Then $(Y \times G, S_\varphi)$, $S_\varphi(x, g) = (Sx, \varphi(x) + g)$, satisfies the Sarnak property [relatively to $\boldsymbol{u}$].*

Proof. First, we need to introduce some notation. For $g \in G$, let $\widetilde{A}_g$ be the graph of $\psi + g$, i.e. $\widetilde{A}_g := \{(y, \psi(y) + g) : y \in Y\}$ and let $\pi_g\colon \widetilde{A}_g \to Y$ stand for the projection onto the first coordinate. Let $\widetilde{\nu}_g := \left(\pi_g^{-1}\right)_* \nu$, where ν is the unique invariant measure for (Y, S). Then the ergodic decomposition of the product system $(Y \times G, \nu \otimes \lambda_G, S_\varphi)$ is given by

$$\nu \otimes \lambda_G = \int \widetilde{\nu}_g \, d\lambda_G(g).$$

But what is more important here is that $\{\widetilde{\nu}_g : g \in G\}$ is the set of S_φ-invariant ergodic measures, see e.g. [27]. Since $(Y, \mathcal{B}(Y), \nu, S) \simeq (Y \times G, \mathcal{B}(Y \times G), \widetilde{\nu}_0, S_\varphi)$ via the map $Id \times \psi\colon Y \to Y \times G$, it follows by Lusin's theorem that there exists a compact set $K \subset Y$ such that $\nu(K) > 1 - \varepsilon^4$ and such that the restriction of $Id \times \psi$ to K is continuous. Then $\widetilde{K}_0 := (Id \times \psi)(K) \subset \widetilde{A}_0$ is also compact. Define $\widetilde{K}_g := \widetilde{K}_0 + g$ and notice that $\bigcup_{g \in G} \widetilde{K}_g = K \times G$.

Now, fix $(\overline{y}, \overline{g}) \in Y \times G$. We will show that the Sarnak property holds in $(\overline{y}, \overline{g}) \in Y \times G$ by showing convergence (8) for functions of the form $F = f \otimes \chi$, where $f \in C(Y)$ and $\chi \in \widehat{G}$ is a character on G.[20] Let H_0 be a continuous extension of $F \circ \pi_0^{-1}|_K$ to the whole space Y, such that $\|H_0\|_\infty = \|F\|_\infty$ (such H_0 exists by the Tietze extension theorem). Let $H_h := \chi(h) H_0$ for $h \in G$. Then H_h is a continuous extension of $F \circ \pi_h^{-1}|_K$. Indeed, for $y \in K$, we have

$$\begin{aligned} H_h(y) &= \chi(h) H_0(y) = \chi(h) F \circ \pi_0^{-1}(y) = \chi(h) F(y, \psi(y)) \\ &= \chi(h) f(y) \chi(\psi(y)) = \chi(h + \psi(y)) f(y) = F(y, \psi(y) + h) = F \circ \pi_h^{-1}(y). \end{aligned}$$

Notice that

$$\text{if } (y, g) \in \widetilde{K}_h \text{ and } S_\varphi^s(y, g) \in \widetilde{K}_h \text{ then} \quad \begin{aligned} F(S_\varphi^s(y, g)) &= (F \circ \pi_h^{-1})(\pi_h(S_\varphi^s(y, g))) \\ &= H_h(\pi_h(S_\varphi^s(y, g))) = H_h(S^s(y)). \end{aligned} \tag{26}$$

For $L \geq 1$, define the following compact subset of $\widetilde{K}_0$:

$$B_0(L) := \left\{ (y, g) \in \widetilde{K}_0 : \frac{1}{L} \sum_{l \leq L} \mathbb{1}_{\widetilde{K}_0}(S_\varphi^s(y, g)) > 1 - \varepsilon^2 \right\}.$$

In the same way, we define $B_h(L)$ for $h \in G$. It follows by a straightforward calculation that $B_h(L) = B_0(L) + h$ for each $h \in G$. Finally, define $B(L) := \bigcup_{h \in G} B_h(L)$. Clearly, $B(L) = \pi_0(B_0(L)) \times G$, whence it is again a compact set. By repeating an argument from the proof of P1 $\Longrightarrow$ P2, we obtain $\widetilde{\nu}_h(B_h(L)) > 1 - \varepsilon$ for each $h \in G$.

[20] We recall that such functions form a set which is linearly dense in the uniform topology in $C(Y \times G)$.

Let $D(\cdot,\cdot)$ be the product distance on $Y\times G$. For each $L\geq 1$, define $\eta(L)>0$ such that for $(y,g),(y',g')\in Y\times G$, we have

$$\text{(27)}\quad D((y,g),(y',g'))<\eta(L)\Longrightarrow \\ |F(S_\varphi^n(y,g))-F(S_\varphi^n(y',g'))|<\varepsilon \text{ for all } 0\leq n<L.$$

By repeating the proof of Lemma 15, we obtain that

$$\text{(28)}\qquad \limsup_{N\to\infty}\frac{1}{N}\#\{0\leq n\leq N-1: D(S_\varphi^n(y,g),B(L))\geq\eta(L)\}<\varepsilon.^{21}$$

We now fix an increasing sequence of integers $1\leq L_1<L_2<\cdots$. Repeat the arguments from the proof of P1 $\Longrightarrow$ P2 to obtain sequences $(M_i)_{i\geq 0}$ and $(b_k)_{k\geq 0}$.

Now, fix $i\geq 1$. Let $K\geq 1$ be the largest integer such that $b_K\leq M_i$. We want to approximate the sum

$$S_{M_i}:=\sum_{n<M_i}F(S_\varphi^n(\overline{y},\overline{g}))\boldsymbol{u}(n)$$

by the following expression coming from the dynamical system (Y,S):

$$E_K:=\sum_{k<K}\sum_{b_k\leq n<b_{k+1}}H_{h_k}(S^{n-b_k}y_k)\boldsymbol{u}(n).$$

As in the proof of P1 $\Longrightarrow$ P2, we obtain that

$$\limsup_{N\to\infty}\frac{1}{N}\left|\sum_{n<N}F(S_\varphi^n(\overline{y},\overline{g}))\boldsymbol{u}(n)\right|=\lim_{i\to\infty}\frac{1}{M_i}|S_{M_i}|\leq\limsup_{K\to\infty}\frac{1}{b_K}|E_K|+C\varepsilon$$

for some constant $C>0$. However

$$\begin{aligned}\frac{1}{b_K}|E_K|&\leq\frac{1}{b_K}\sum_{k<K}\left|\sum_{b_k\leq n<b_{k+1}}H_{h_k}(S^{n-b_k}y_k)\boldsymbol{u}(n)\right|\\&=\frac{1}{b_K}\sum_{k<K}\left|\sum_{b_k\leq n<b_{k+1}}\chi(h_k)H_0(S^{n-b_k}y_k)\boldsymbol{u}(n)\right|\\&=\frac{1}{b_K}\sum_{k<K}\left|\sum_{b_k\leq n<b_{k+1}}H_0(S^{n-b_k}y_k)\boldsymbol{u}(n)\right|\end{aligned}$$

which goes to 0 as $K\to\infty$ by the strong MOMO property of (Y,S). This completes the proof. □

Remark 28. *See [32], where there is the first example of a continuous (in fact analytic) $f\colon\mathbb{T}\to\mathbb{R}$ considered over an irrational rotation T such that $T_{e^{2\pi if}}$ is minimal, f is a measurable coboundary and $T_{e^{2\pi if}}$ is Möbius disjoint. Cf. also [42] for Möbius disjointness of all analytic Anzai skew products.*

[21] Indeed, if $(\overline{y},\overline{g})$ is quasi-generic, along a sequence (N_i), for an S_φ-invariant measure then this measure has to be of the form $\widetilde{\mu}=\int_G\widetilde{\nu}_g\,dP(g)$, where P is a probability measure on G.

4 Examples

4.1 Ergodic systems with discrete spectrum

Recall that in [2] it has been proved that all uniquely ergodic models of totally ergodic systems with discrete spectrum are Möbius disjoint. The result has been extended to all uniquely ergodic models of all ergodic systems with discrete spectrum in [21]. By Halmos-von Neumann theorem it follows that if $(Z,\mathcal{D},\kappa,R)$ is an ergodic transformation with discrete spectrum then one of its uniquely ergodic models is a rotation $Tx=x+x_0$, where X is a compact, metric group and $\overline{\{nx_0 : n\in\mathbb{Z}\}}=X$. If $\chi\in\widehat{X}$, $(b_k)\subset\mathbb{N}$ with $b_{k+1}-b_k\to\infty$ and $(x_k)\subset X$ then for each $(y_k)\subset X$, we have

$$\frac{1}{b_K}\sum_{k<K}\sum_{b_k\leq n<b_{k+1}}\chi(T^n(x_k+y_k))\boldsymbol{u}(n)=\frac{1}{b_K}\sum_{k<K}\chi(y_k)\sum_{b_k\leq n<b_{k+1}}\chi(T^nx_k)\boldsymbol{u}(n).$$

It easily follows from Lemma 16 that in (X,T) we have the strong MOMO property whenever we have the MOMO property. But

$$\left|\frac{1}{b_K}\sum_{k<K}\sum_{b_k\leq n<b_{k+1}}\chi(T^n(x_k))\boldsymbol{u}(n)\right|\leq\frac{1}{b_K}\sum_{k<K}\left|\sum_{b_k\leq n<b_{k+1}}\left(\chi(x_0)\right)^n\boldsymbol{u}(n)\right|. \tag{29}$$

We have $\chi(x_0)=e^{2\pi i\alpha}$ for a unique $\alpha\in[0,1)$. Two cases now arise:

Case 1. If α is irrational then the RHS in (29) goes to zero by [2] for any $\boldsymbol{u}$ multiplicative, $|\boldsymbol{u}|\leq 1$.

Case 2. Assume that α is rational. Then it follows from Theorem 1.7 in [34] that the RHS in (29) goes to zero if $\boldsymbol{u}$ is multiplicative, $|\boldsymbol{u}|\leq 1$ and

$$\inf_{|t|\leq M,\xi \bmod q, q\leq Q} D(\boldsymbol{u}, n\mapsto \xi(n)n^{it}; M)^2\to\infty, \tag{30}$$

when $10\leq H\leq M$, $H\to\infty$ and $Q=\min(\log^{1/125}M,\log^5 H)$; here ξ runs over all Dirichlet characters of modulus $q\leq Q$ and

$$D(\boldsymbol{u},\boldsymbol{v};M):=\left(\sum_{p\leq M,p\in\mathscr{P}}\frac{1-\mathrm{Re}(\boldsymbol{u}(p)\overline{\boldsymbol{v}(p)})}{p}\right)^{1/2}$$

for each $\boldsymbol{u},\boldsymbol{v}:\mathbb{N}\to\mathbb{C}$ multiplicative satisfying $|\boldsymbol{u}|,|\boldsymbol{v}|\leq 1$. In particular, (30) implies (4). Moreover, classical multiplicative functions like $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ satisfy (30) [34].

Corollary 29. *Let $(Z,\mathcal{D},\kappa,R)$ be an ergodic system with discrete spectrum. If R is totally ergodic, then in each uniquely ergodic model of R we have the strong MOMO property relatively to any multiplicative function $\boldsymbol{u}$, $|\boldsymbol{u}|\leq 1$, satisfying (4). If the spectrum of R has a nontrivial rational eigenvalue then the strong MOMO property holds for any $\boldsymbol{u}$ satisfying additionally (30).*

4.2 Systems satisfying the AOP property

4.2.1 Systems whose powers are disjoint. Typical systems

Each totally ergodic transformation whose prime powers are disjoint satisfies the AOP property. This includes large classes of rank one transformations: [1, 7, 37], automorphisms with the minimal self-joining property [11], and recently it has been shown by Chaika and Eskin [9] that a.e. 3-interval transformation has sufficiently many disjoint prime powers. By Theorem 10, it follows that the first assertion of Corollary 29 holds for an arbitrary R belonging to any of those classes of transformations. In particular, by [12], it follows that all uniquely ergodic models of a typical system R of a standard Borel probability space satisfy this assertion.

4.2.2 Unipotent diffeomorphisms on nilmanifolds

As shown in [16], ergodic unipotent diffeomorphisms $T(x\Gamma) = uA(x)\Gamma$, where G is a connected, simply connected, nilpotent Lie group, $\Gamma \subset G$ a lattice, $A\colon G \to G$ is a unipotent automorphism with $A(\Gamma) = \Gamma$ and $u \in G$, enjoy the AOP property. It follows that all other uniquely ergodic models of such systems satisfy the strong MOMO property (relatively to $\boldsymbol{u}$ satisfying (4)), in particular, this holds for uniquely ergodic models of ergodic nil-rotations. A special case of unipotent diffeomorphisms on nilmanifolds are affine automorphisms on Abelian compact connected groups. In particular, we obtain that in all uniquely ergodic models of quasi-discrete spectrum systems, we have the strong MOMO property (with respect to $\boldsymbol{u}$), cf. [2].

A general unipotent case seems to be much less clear. We have been unable to answer the following:

Question 30. Does, for each horocycle flow on the unit tangent bundle of compact surfaces of constant negative curvature, the time-1 automorphism satisfy the (strong) MOMO property (relatively to $\boldsymbol{\mu}$)? (For Möbius disjointness of time automorphisms of horocycle flows, see [8]; such automorphisms do not possess the AOP property [2].)

4.2.3 Cocycle extensions of irrational rotations

As we have already noticed, all irrational rotations have the AOP property. The extensions of them considered in this subsection satisfy the assumptions of Theorem 26, but in fact we lift AOP.

Consider $f\colon \mathbb{T} \to \mathbb{R}$ which is $C^{1+\delta}$ and which is not a trigonometric polynomial. In [28], it is shown that for a G_δ and dense set of irrational α the corresponding Anzai skew product $T_{e^{2\pi i f}}$ on $\mathbb{T}\times\mathbb{S}^1$: $(x, z) \mapsto (x+\alpha, e^{2\pi i f(x)}\cdot z)$ enjoys the AOP property (cf. Corollary 2.5.6 in [28]). Moreover, it is proved (again for a "typical" α) in [29] that the Rokhlin skew product $T_{f,\mathcal{S}}$: $(x, y) \mapsto (x+\alpha, S_{f(x)}(y))$ enjoys the AOP property for each ergodic flow $\mathcal{S} = (S_t)_{t\in\mathbb{R}}$ acting on $(Y, \mathcal{C}, \nu)$ (on $X \times Y$ we consider product measure $\mu \otimes \nu$), see Proposition 5.1 and Corollary 5.2 in [29].

Consider an affine case: $f(x) = x - \frac{1}{2}$. Theorem 7.10 in [29] tells us that $T_{f,\mathcal{S}}$ has the AOP property whenever α has bounded partial quotients and the spectrum of the flow $\mathcal{S}$ on $L^2_0(Y, \mathcal{C}, \nu)$ does not contain any rational number.

Note that if we replace f by $f' := f + j - j\circ T$, where $j\colon \mathbb{T} \to \mathbb{R}$ is continuous, then the resulting skew products $T_{e^{2\pi i f'}}$ or $T_{f',\mathcal{S}}$[22] are uniquely ergodic models of $T_{e^{2\pi i f}}$ and $T_{f,\mathcal{S}}$ (whenever $\mathcal{S}$ is a uniquely ergodic flow), respectively. In

[22]Formally, we should slightly extend T using so-called Sturmian models [17, Chapter 6 by Arnoux], so that f becomes continuous, see [29] for details.

particular, $T_{e^{2\pi i f'}}$ and $T_{f',\mathcal{S}}$ enjoy the strong MOMO property (relatively to $\boldsymbol{u}$ satisfying (4)).

Question 31. Assume that $T_{e^{2\pi i f}}$ is an ergodic Anzai skew product with $f\colon \mathbb{T} \to \mathbb{R}$ analytic. Does $T_{e^{2\pi i f}}$ enjoy the strong MOMO property, the AOP property? (It has been shown by Wang [42] that all such skew products are Möbius disjoint.)

4.3 Cocycles extensions of odometers. Morse and Kakutani sequences

When T is an odometer then it is not totally ergodic, and the method of AOP fails. However, we have already shown that odometers satisfy the strong MOMO property. In this section, we will give examples of extensions of odometers illustrating Theorem 26. Because of the second assertion in Corollary 29, we will constantly assume that $\boldsymbol{u}$ is a bounded multiplicative arithmetic function which satisfies (30). (In particular, the following results are valid when $\boldsymbol{u}$ is the Möbius or the Liouville function.)

4.3.1 Odometers, Toeplitz systems and generalized Morse systems

Given an increasing sequence (n_t) of natural numbers with $n_t|n_{t+1}$, $t \geq 0$ ($n_0 = 1$) set $\lambda_t := n_{t+1}/n_t$ for $t \geq 0$ and let $X = \prod_{t\geq 0} \mathbb{Z}/\lambda_t\mathbb{Z}$. It is a compact, metric and monothetic group where the addition is coordinatewise with carrying the remainder to the right. If by μ_X we denote Haar measure on X then $(X, \mathcal{B}(X), \mu_X, T)$, where $Tx = x + \underline{1}$, $\underline{1} := (1, 0, 0, \ldots)$, is ergodic (in fact, (X, T) is uniquely ergodic). If

$$D^{(t)} = \{D_0^{(t)}, D_1^{(t)}, \ldots, D_{n_t-1}^{(t)}\}, \tag{31}$$

where $D_0^{(t)} = \{x \in X : x_0 = \ldots = x_{t-1} = 0\}$, $D_j^{(t)} = T^j D_0^{(t)}$, $j = 0, 1, \ldots, n_t - 1$ then $D^{(t)}$ is a partition of X and $\bigcup_{j=0}^{n_t-1} D_j^{(t)} = X$, that is, $D^{(t)}$ is a Rokhlin tower filling the whole space.

Let now $b^t \in \{0,1\}^{\lambda_t}$, $b^t(0) = 0$, for $t \geq 0$. The sequence

$$x := b^0 \times b^1 \times \ldots \tag{32}$$ [23]

is called a *generalized Morse sequence* [26].[24] Let $\widehat{x}$ be the sequence defined by

$$\widehat{x}(n) := x(n) + x(n+1) \bmod 2 \quad (n \geq 0). \tag{33}$$

Then $\widehat{x}$ is a Toeplitz sequence [22], i.e. for each j there is k_j such that $\widehat{x}(j) = \widehat{x}(j + mk_j)$ for each $m \geq 0$. Let X_x and $X_{\widehat{x}}$ stand for the (two-sided) subshifts determined by x and $\widehat{x}$, respectively. By [44], the odometer (X, T) is the maximal equicontinuous factor of $(X_{\widehat{x}}, S)$, where S stands for the left shift. We will constantly assume that $\widehat{x}$ is regular [22]. In this case $(X_{\widehat{x}}, S)$ is a

[23] The multiplication of blocks $b^0, b^1, \ldots$ is from the left to the right; $B \times C := (B + C(0))(B + C(1)) \ldots (B + C(\lambda - 1))$, where $\lambda := |C|$ stands for the length of C and $B + c := (B(0) + c)(B(1) + c) \ldots (B(|B| - 1) + c)$ (the addition mod 2 on each coordinate).

[24] As a matter of fact, there are some mild assumptions on the sequence (b^t) to obtain a non-trivial dynamical systems, see e.g. the concept of continuous Morse sequence in [26].

uniquely ergodic model of (X,T) [22]. Therefore, the strong MOMO property holds for $(X_{\widehat{x}},S)$. On the other hand, clearly, $(X_{\widehat{x}},S)$ is a topological factor of (X_x,S), and moreover, there is a topological isomorphism of (X_x,S) with $(X_{\widehat{x}}\times(\mathbb{Z}/2\mathbb{Z}),S_\varphi)$, where $\varphi(z)=z(0)$ for each $z\in X_{\widehat{x}}$ [31]. The cocycle φ has a special form (see the notion of Morse cocycle below), and we will show that in certain classical cases the assumptions of Theorem 26 are satisfied.

4.3.2 A little bit of algebra

Denote $\mathbb{Z}_N=\mathbb{Z}/N\mathbb{Z}$ and let $0\neq s\in\mathbb{Z}_N$, $(s,N)=1$. Then $s\in\mathbb{Z}_N^*$, i.e. s is in the group of invertible (under multiplication) elements in the ring $\mathbb{Z}_N$. Therefore, we can write $\frac{1}{s}$ for the inverse of s in $\mathbb{Z}_N$, and for any integer r, $\frac{r}{s}$ is well defined as an element of $\mathbb{Z}_N$.

In the ring $\mathbb{Z}_N$, consider the F-norm $\|i\|:=\min(i,N-i)$.

Lemma 32. *Assume that $r,s,k\geq 1$ are fixed, pairwise coprime, and let (n_t) be an increasing sequence of integers. We assume that $(s,n_t)=1$ while $k|n_t$ for each $t\geq 1$. Then there exists $\eta>0$ such that*

$$\liminf_{t\to\infty}\min_{0\leq j<k-1}\left\|\frac{r}{s}-j\frac{n_t}{k}\right\|\geq\eta n_t.$$

Proof. Denote $b_t:=\frac{r}{s}\in\mathbb{Z}_{n_t}$, i.e.

$$r=sb_t \bmod n_t, \tag{34}$$

and we also interpret b_t as an integer in $\{0,\ldots,n_t-1\}$. Fix $0<\varepsilon<\frac{1}{ks}$, and suppose that $b_t\in(j\frac{n_t}{k}-\varepsilon n_t,j\frac{n_t}{k}+\varepsilon n_t)$ for some $0\leq j\leq k$. Then

$$sb_t\in\left(sj\frac{n_t}{k}-s\varepsilon n_t,sj\frac{n_t}{k}+s\varepsilon n_t\right).$$

If $j=0$, then $sb_t\in(0,s\varepsilon n_t)$ and (34) is really $r=sb_t$ as soon as $n_t>r$, whence $s|r$, a contradiction.

If $0<j<k$, then the number $sj\frac{n_t}{k}$ is of the form $\ell n_t+\frac{j'}{k}n_t$ with $0<j'<k$ (remember that $(k,s)=1$), and by (34) we also have

$$r\in\left(sj'\frac{n_t}{k}-s\varepsilon n_t,sj'\frac{n_t}{k}+s\varepsilon n_t\right).$$

In particular, $r>\frac{n_t}{k}-\varepsilon sn_t=n_t\left(\frac{1}{k}-\varepsilon s\right)$, which is impossible when n_t is large.

If $j=k$, then $sb_t\in\big(sn_t(1-\varepsilon),sn_t\big)$. By (34), there exists an integer ℓ such that

$$r\in(\ell n_t-s\varepsilon n_t,\ell n_t).$$

As $r\geq 1$, we must have $\ell\geq 1$, but then $r\geq n_t(1-s\varepsilon)$, which is impossible when n_t is large. □

Proposition 33. *Assume that (X,T) is an odometer, $Tx=x+\underline{1}$ (with $\underline{1}=(1,0,0,\ldots)$) on X, where X is determined by the numbers $n_t|n_{t+1}$, $t\geq 0$. Assume that $r,s,k\geq 1$ are fixed, pairwise coprime, and that $(r,n_t)=1=(s,n_t)$ while $k|n_t$ for each $t\geq 1$. Then, there exists a unique automorphism W of*

$(X,\mathcal{B}(X),\mu_X)$ *such that* $W^r=T^s$. *More precisely, there exists a sequence* (b_t), $b_t\in\mathbb{Z}_{n_t}$, *depending only on* r,s *and* n_t, *such that*

$$\forall t\geq 0,\forall 0\leq i<n_t,\quad WD_i^{(t)}=D_{i+b_t}^{(t)}. \tag{35}$$

(Here the addition has to be understood in $\mathbb{Z}_{n_t}$*.) Moreover, there exists* $\eta>0$ *such that*

$$\liminf_{t\to\infty}\sup_{0\leq j\leq k-1}\frac{\left\|b_t-j\frac{n_t}{k}\right\|}{n_t}\geq\eta.$$

Proof. Let W satisfy $W^s=T^r$. Since $(r,n_t)=1$ for each t, T^r is ergodic, and the subsets $D_i^{(t)}$, $0\leq i<n_t$ are also the levels of a Rokhlin tower of height n_t for T^r (but in a different order). W commutes with T^r, hence the sets $WD_i^{(t)}$, $0\leq i<n_t$ also form a Rokhlin tower of height n_t for T^r. By ergodicity of T^r, there exists only one such Rokhlin tower (up to cyclic permutation of the levels). We thus conclude that W permutes the levels $D_i^{(t)}$ of the Rokhlin tower. In particular there exists $b_t\in\{0,\dots,n_t-1\}$ such that $WD_0^{(t)}=D_{b_t}^{(t)}$. Using again the fact that $(r,n_t)=1$, we see that there exists an integer m such that $T^{mr}=T$ on the finite σ-algebra generated by the sets $D_i^{(t)}$, $0\leq i<n_t$. As W commutes with T^r, W also commutes with T on this σ-algebra, thus (35) holds. Now the relation $W^s=T^r$ on this σ-algebra just says that $b_t=\frac{r}{s}$ in $\mathbb{Z}_{n_t}$.

Conversely, setting $b_t:=\frac{r}{s}$ in $\mathbb{Z}_{n_t}$ for each $t\geq 0$, (which exists because $(s,n_t)=1$), we see that, for each $t\geq 0$, $b_{t+1}=b_t \bmod n_t$ as $n_t|n_{t+1}$. Thus we can define a unique automorphism W by (35) (note that $\mathcal{B}(X)$) is the supremum of the increasing sequence of finite σ-algebras generated by the sets $D_i^{(t)}$). We then get an s-th root of T^r.

Finally, the existence of η follows from Lemma 32.
[25] □

Note that the conclusion of the preceding proposition also implies

$$\left\|(n_t-b_t)-i\frac{n_t}{k}\right\|\geq\eta n_t \tag{36}$$

for all $t\geq 1$.

We will say that an odometer (X,T) has *small rational spectrum* if the set

$$\mathrm{Spec}(T):=\{p\in\mathscr{P}:\ (\exists t\geq 1)\ p|n_t\}$$

is finite. (We may think of T as being "close" to an automorphism which is totally ergodic, in the sense, that most of its prime powers are ergodic.)

4.3.3 $\mathbb{Z}/2\mathbb{Z}$-extensions for which there are not too many roots. k-Morse cocycles

Let (X,T) be an odometer. Fix $k\geq 1$ and assume that $k|n_t$, $t\geq 1$.

Definition 34. A cocycle $\phi:X\to\mathbb{Z}/2\mathbb{Z}$ is said to be a *k-Morse cocycle* if, for each $t\geq 1$, ϕ is constant on each level $D_j^{(t)}$ except for $D_{\frac{i}{k}n_t-1}^{(t)}$ for $i=1,2,\dots,k$.

[25] The action of W on the tower $D^{(t)}$ is the "rotation" by b_t; sb_t corresponds to W^s which is r mod n_t (which corresponds to T^r).

Morse cocycles are, by definition, 1-Morse cocycles. The skew products determined by T_ϕ, where ϕ is a Morse cocycle, correspond to subshifts given by generalized Morse sequences (32). Moreover (see [31]), for each $t \geq 1$ and $i = 0, \ldots, n_t - 2$, we have

$$\phi|_{D_i^{(t)}} = \widehat{c}_t(i), \tag{37}$$

where $c_t := b^0 \times \ldots \times b^{t-1}$ and then inductively

$$\phi|_{D_{in_t}^{(t+1)}} = \widehat{C}_{t+1}(in_t) = b^{t+1}(i-1) + b^{t+1}(i) + c_t(n_t - 1),\ i = 1, \ldots, \lambda_{t+1} - 1. \tag{38}$$

The Morse sequence for which $b^t = 01$ for each $t \geq 0$ is the classical Thue-Morse sequence (e.g. [5]). More generally, the Morse sequences for which $b^t \in \{00, 01\}$ with infinitely many blocks equal to 01 are precisely Kakutani sequences, e.g. [30]. Note that in the case of Kakutani sequences, for the corresponding odometer (X, T), we have $\mathrm{Spec}(T) = \{2\}$, so we deal with small rational spectrum.

A similar situation arises when we consider the subshift given by the Rudin-Shapiro sequences (e.g. [5]). Indeed, in this case ϕ is a 2-Morse cocycle, see [31] for more details.

Definition 35. Assume that T has small rational spectrum and ϕ is a k-Morse cocycle. We say that ϕ is *probabilistic* if there exists $\eta > 0$ such that for infinitely many $t \geq 1$ the conditional distribution of 0 on each level $D^{(t)}_{\frac{i}{k}n_t - 1}$ $(i = 1, \ldots, k)$ is between η and $1 - \eta$. We will say that ϕ satisfies PC (the probabilistic condition).

Remark 36. *All Kakutani sequences satisfy PC in the following sense. First, notice that by introducing "parentheses"*

$$x = (b^0 \times \ldots \times b^{i_1 - 1}) \times (b^{i_1} \times \ldots \times b^{i_2 - 1}) \times \ldots = \overline{b}^0 \times \overline{b}^1 \times \ldots$$

we can obtain a new representation of a Morse sequence x, in which the corresponding odometer $(\overline{X}, \overline{T})$ is given by a subsequence of (n_t), and we look at the Morse cocycle ϕ only along this subsequence.

Now, take any $x = b^0 \times b^1 \times \ldots$, where $b^t = 00$ or 01 (with infinitely many b^t equal to 01). Then introduce "parentheses" putting together $01 \times 01 = 0110$ and $00 \times 01 = 0011$. Now, for the corresponding $\overline{b}^{t'}$s, we have $\widehat{\overline{b}^t}$ equal either to $101*$ or $010*$, so PC is satisfied (cf. (38)).

We can easily generalize this argument to obtain the following result.

Proposition 37. *All (continuous) Morse sequences $x = b^0 \times b^1 \times \ldots$ with bounded lengths of blocks yield ϕ satisfying PC.*

Proof. By introducing parentheses, if necessary, we can assume that $3 \leq |b^t| \leq C$, $t \geq 0$. Then if we see infinitely many blocks different from: $0\ldots0$, $01\ldots01$ or $01\ldots010$, then we are done. If not then if we have infinitely many blocks $0\ldots0$ then

$$0\ldots0 \times 01\ldots01, \text{ or } 0\ldots0 \times 01\ldots010$$

yield also "good" blocks, so by introducing more parentheses, we obtain a new representation which is good, by looking at the last positions of appearances of

$0\ldots0$. If not, then assume that we have infinitely many blocks $01\ldots01$. Then both blocks

$$01\ldots01\times01\ldots01 \text{ or } 01\ldots01\times01\ldots010$$

are "good", and we are done since otherwise, starting from some place, we must have $b^t=01\ldots010$ which means that x is periodic. $\square$

Remark 38. *For Rudin-Shapiro sequences, it follows from [31] that the corresponding 2-Morse cocycles satisfy PC.*

Theorem 39. *Assume that an odometer (X,T) has small rational spectrum, $\phi: X\to\mathbb{Z}/2\mathbb{Z}$ is a k-Morse cocycle which satisfies PC. Then for all prime numbers $r\neq s$ sufficiently large, $(T_\phi)^r$ has no sth root. In particular, $(T_\phi)^r$ and $(T_\phi)^s$ are not isomorphic. In fact, the only ergodic joinings between $(T_\phi)^r$ and $(T_\phi)^s$ are the relatively independent extensions of isomorphisms between T^r and T^s.*[26]

Proof. Assume that $(T_\phi)^r$ and $(T_\phi)^s$ are isomorphic (we assume that r,s are coprime with k). Then $(T_\phi)^r$ has an s-th root which is in the centralizer of $(T_\phi)^r$. But, by Lemma 4.3 in [15] it follows that $C(T_\phi)=C((T_\phi)^r)$, hence this s-th root is in $C(T_\phi)$. It has to be of the form $\widetilde{W}=W_\xi$, where $(\widetilde{W})^s=(T_\phi)^r$. It follows that $W=T^{r/s}$. In other words, we can lift the rotation by r/s to $C(T_\phi)$. Hence

$$\phi\circ W-\phi=\xi\circ T-\xi. \tag{39}$$

Recall that at the stage t, W is represented by b_t (i.e. on the tower D^t, W acts as T^{b_t}). In view of (36),

$$\left\|(n_t-b_t)-i\frac{n_t}{k}\right\|\geq\eta n_t \tag{40}$$

uniformly in t and $i\in\{0,\ldots,k-1\}$.

Fix $\varepsilon>0$ small. Then for t large enough the function ξ will be well approximated by the levels of $D^{(t)}$. Hence for $(1-\varepsilon)n_t$ levels of $D^{(t)}$ we will have that ξ is up to a set of conditional measure ε constant. In what follows we will speak about ξ being ε-constant on a level.

Consider $D^{(t)}_{b_t}=T^{b_t}D^{(t)}_0$. In view of (40) there exists $0\leq i\leq k-1$ such that

$$\frac{i}{k}n_t<b_t<\frac{i+1}{k}n_t$$

with $b_t-\frac{i}{k}n_t\geq\eta n_t$ and $\frac{i+1}{k}n_t-b_t|\geq\eta n_t$. It follows that we can find $0\leq\ell<\frac{1}{k}n_t$ [27], $\ell<\frac{i+1}{k}n_t-b_t$ such that $\xi|_{D^{(t)}_\ell}$ is ε-constant. Moreover $\phi|_{WD^{(t)}_\ell}$ and $\phi|_{D^{(t)}_\ell}$ are also constant, and (39) makes $\xi|_{D^{(t)}_{\ell+1}}$ ε-constant either with the same distribution as $\xi|_{D^{(t)}_\ell}$ or by replacing 0 by 1 and vice versa. We now repeat the same argument with ℓ replaced by $\ell+1$. We keep going in the same manner and we obtain consecutive levels on which ξ is ε-constant until ℓ reaches the value $\frac{i+1}{k}n_t-b_t$. Now, $\phi\circ W$ will have the same distribution as that of ϕ on $D^{(t)}_{\frac{i+1}{k}n_t-1}$,

[26] We recall that T^r is isomorphic to T whenever T^r is ergodic. The last assertion in the theorem means that the assumptions of Theorem 26 are satisfied.

[27] Note that $0<\frac{i+1}{k}n_t-b_t<\frac{1}{k}n_t$.

while ϕ on $D^{(t)}_{\frac{i+1}{k}n_t-b_t}$ is still constant and ξ is ε-constant. It follows that up to ε the conditional distribution of ξ on $TD^{(t)}_{\frac{i+1}{k}n_t-b_t}$ is that of ϕ on $D^{(t)}_{\frac{i+1}{k}n_t}$ or its "mirror". When we consider the next step $\phi\circ W$ and ϕ will be again constant, so ξ on the next level will have the same distribution as on the previous level (or its "mirror"). This will be continued for ηn_t levels and because ϕ satisfies PC, ξ cannot be measurable, a contradiction.

The last assertion follows from non-isomorphism of the powers and Corollary 4.7 in [15]. □

Remark 40. *For Morse cocycles, the above result can also be deduced from a theorem proved by Kwiatkowski and Rojek [23] about the centralizer of Morse subshifts.*

Now, we obtain that whenever the (uniquely ergodic) subshift (X_x,S) determined by a k-Morse sequence $x\in\{0,1\}^{\mathbb{N}}$ is given by a k-Morse cocycle satisfying PC, then the uniquely ergodic system (X_x,S) is an extension of $(X_{\widehat{x}},S)$ for which the assumptions of Theorem 26 hold and therefore, it satisfies the strong MOMO property. Using Theorem 6 (and Remark 38), we hence obtain the following.

Corollary 41. *In each uniquely ergodic model of the system determined by a Kakutani sequence*[28] *(in particular, by the Thue-Morse sequence) we have the strong MOMO property (relatively to $\boldsymbol{u}$ satisfying* (30)*). The same result holds for any Rudin-Shapiro sequence.*

4.4 Substitutions of constant length

Let A be a finite alphabet, $\#A\geq 2$. Let $q\geq 2$ be a fixed integer and $\theta: A\to A^q$ be a primitive aperiodic substitution of constant length q [36]. Recall that θ is extended to a morphism of the monoid A^* by the formula

$$\theta(a_0\cdots a_{\ell-1}) := \theta(a_0)\cdots\theta(a_{\ell-1}).$$

Similarly, we can extend θ to a map defined on $A^{\mathbb{N}}$. We denote by X_θ the two-sided associated subshift:

$$X_\theta := \Big\{x=(x(n),\ n\in\mathbb{Z})\in A^{\mathbb{Z}}: \\ \forall m\leq n, \exists t\geq 0, \exists a\in A, x(m,n) \text{ is a subblock of } \theta^t(a)\Big\}.$$

Let S be the shift map on X_θ. We recall that (X_θ,S) is uniquely ergodic, and we denote by μ_θ the unique S-invariant probability measure on X_θ. To each θ we can associate the *column number*.

Definition 42 (Kamae [24])**.** The *column number* of the substitution θ is the number

$$c(\theta) := \min_{t\geq 1}\min_{0\leq \ell\leq q^t-1}\#\{\theta^t(a)(\ell): a\in A\}.$$

If by X_q we denote the odometer determined by $n_t:=q^t$, $t\geq 0$, then (X_q,T) is the maximal equicontinuous factor of (X_θ,S) and moreover (see [24, 36])

(41) $\quad (X_\theta,\mathcal{B}(X_\theta),\mu_\theta,S)$ is an a.e. c-extension of $(X_q,\mathcal{B}(X),\mu_{X_q},T)$.

[28] It has been already known that the subshift determined by any Kakutani sequence is Möbius disjoint [6, 20, 15, 41].

4.4.1 Bijective substitutions

A substitution $\theta\colon A\to A^q$ (as above) is called *bijective* if the map

$$\tau_i(a):=\theta(a)(i),\ a\in A$$

is a bijection of A for each $i=0,1,\ldots,q-1$. Using the notion of group substitution, it is implicitly proved in [15][29] that the assumptions of Theorem 26 are satisfied. It follows that:

Corollary 43. *If θ is bijective then the strong MOMO property (relatively to a bounded multiplicative $\boldsymbol{u}$ satisfying (30)) is satisfied in each uniquely ergodic model of $(X_\theta,\mathcal{B}(X_\theta),\mu_\theta,S)$.*

4.4.2 The synchronized case

For one more case of substitutions the assertion of Proposition 43 easily holds. Namely, this is the case when the column number $c(\theta)=1$. Indeed, in this case, by (41), the factor map

$$(X_\theta,\mathcal{B}(X_\theta),\mu_\theta,S)\to(X_q,\mathcal{B}(X_q),\mu_{X_q},T)$$

is a.e. 1-1. It easily follows that (X_θ,S) is a uniquely ergodic model of the odometer (X_q,T), hence the result follows from Corollary 29.

It is well-known (Cobham's theorem) that fixed points of substitutions of constant length are in one-to-one correspondence with automatic sequences, i.e. sequences generated by deterministic complete automata [36]. Those automatic sequences which correspond to synchronized automata are called *synchronized*, and the substitutions with trivial column number are in 1-1 correspondence with synchronized automatic sequences [35]. An independent proof of Möbius disjointness in the synchronized case has been done in [13].

As all subshifts given by substitutions of constant length are Möbius disjoint by a recent result of Müllner [35], it is natural to ask:

Question 44. Is it true that all subshifts generated by substitutions of constant length satisfy the (strong) MOMO property (relatively to $\boldsymbol{\mu}$)?

Remark 45. It seems that the main problem to obtain Corollary 43 without any restriction on θ is a full description of the cocycle $\varphi\colon X_q\to\mathcal{S}(\{0,1,\ldots,c(\theta)-1)\})$ which is behind the statement (41).

References

[1] E. H. El Abdalaoui, M. Lemańczyk, and T. de la Rue, *On spectral disjointness of powers for rank-one transformations and Möbius orthogonality*, http://arxiv.org/abs/1301.0134.

[2] ——, *Asymptotic orthogonality of powers for quasi-discrete spectrum automorphisms*, International Math. Res. Notices (2016), http://arxiv.org/abs/1507.04132.

[29] This follows from the proofs of Proposition 4.5 and Theorem 5.4 (see also Proposition 4.2) in [15].

[3] H. El Abdalaoui, S. Kasjan, and M. Lemańczyk, *0-1 sequences of the Thue-Morse type and Sarnak's conjecture*, Proc. Amer. Math. Soc. **144** (2016), no. 1, 161–176. MR 3415586

[4] H. El Abdalaoui, J. Kułaga-Przymus, M. Lemańczyk, and T. de la Rue, *The Chowla and the Sarnak conjectures from ergodic theory point of view (extended version)*, http://arxiv.org/abs/1410.1673.

[5] J.-P. Allouche and J. Shallit, *Automatic sequences*, Cambridge University Press, Cambridge, 2003, Theory, applications, generalizations. MR 1997038 (2004k:11028)

[6] J. Bourgain, *Möbius-Walsh correlation bounds and an estimate of Mauduit and Rivat*, J. Anal. Math. **119** (2013), 147–163. MR 3043150

[7] ———, *On the correlation of the Moebius function with rank-one systems*, J. Anal. Math. **120** (2013), 105–130. MR 3095150

[8] J. Bourgain, P. Sarnak, and T. Ziegler, *Disjointness of Möbius from horocycle flows*, From Fourier analysis and number theory to Radon transforms and geometry, Dev. Math., vol. 28, Springer, New York, 2013, pp. 67–83. MR 2986954

[9] J. Chaika and A. Eskin, *Möbius disjointness for interval exchange transformations on three intervals*, https://arxiv.org/abs/1606.02357.

[10] H. Davenport, *On some infinite series involving arithmetical functions. II*, Quart. J. Math. Oxford **8** (1937), 313–320.

[11] A. del Junco and D. Rudolph, *On ergodic actions whose self-joinings are graphs*, Ergodic Theory Dynam. Systems **7** (1987), no. 4, 531–557. MR 922364 (89e:28029)

[12] A. del Junco, *Disjointness of measure-preserving transformations, minimal self-joinings and category*, Ergodic theory and dynamical systems, I (College Park, Md., 1979–80), Progr. Math., vol. 10, Birkhäuser, Boston, Mass., 1981, pp. 81–89. MR 633762 (82m:28035)

[13] J.-M. Deshouillers, M. Drmota, and C. Müllner, *Automatic sequences generated by synchronizing automata fulfill the Sarnak conjecture*, Studia Math. **231** (2015), no. 1, 83–95. MR 3460628

[14] T. Downarowicz and J. Serafin, *Almost full entropy subshifts uncorrelated to the möbius function*, https://arxiv.org/abs/1611.02084.

[15] S. Ferenczi, J. Kułaga-Przymus, M. Lemańczyk, and C. Mauduit, *Substitutions and Möbius disjointness*, Ergodic Theory, Dynamical Systems, and the Continuing Influence of John C. Oxtoby, Contemp. Math., vol. 678, Amer. Math. Soc., Providence, RI, 2016, pp. 151–173.

[16] L. Flaminio, K. Frączek, J. Kułaga-Przymus, and M. Lemańczyk, *Approximate orthogonality of powers for ergodic affine unipotent diffeomorphisms on nilmanifolds*, https://arxiv.org/abs/1609.00699.

[17] N. P. Fogg, *Substitutions in Dynamics, Arithmetics and Combinatorics*, Lecture Notes in Mathematics, vol. 1794, Springer-Verlag, Berlin, 2002, Edited by V. Berthé, S. Ferenczi, C. Mauduit and A. Siegel. MR 1970385 (2004c:37005)

[18] N. Frantzikinakis, *Ergodicity of the Liouville system implies the Chowla conjecture*, https://arxiv.org/abs/1611.09338.

[19] H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Math. Systems Theory **1** (1967), 1–49. MR 0213508 (35 #4369)

[20] B. Green, *On (not) computing the Möbius function using bounded depth circuits*, Combin. Probab. Comput. **21** (2012), no. 6, 942–951. MR 2981162

[21] W. Huang, Z. Wang, and G. Zhang, *Möbius disjointness for topological models of ergodic systems with discrete spectrum*, https://arxiv.org/abs/1608.08289.

[22] K. Jacobs and M. Keane, $0-1$*-sequences of Toeplitz type*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **13** (1969), 123–131. MR 0255766 (41 #426)

[23] ———, $0-1$*-sequences of Toeplitz type*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **13** (1969), 123–131. MR 0255766 (41 #426)

[24] T. Kamae, *A topological invariant of substitution minimal sets*, J. Math. Soc. Japan **24** (1972), 285–306. MR 0293611 (45 #2688)

[25] I. Kátai, *A remark on a theorem of H. Daboussi*, Acta Math. Hungar. **47** (1986), no. 1-2, 223–225. MR 836415 (87j:11072)

[26] M. Keane, *Generalized Morse sequences*, Z. Wahrscheinlichkeitstheorie und Verw. Gebiete **10** (1968), 335–353. MR 0239047 (39 #406)

[27] H. B. Keynes and D. Newton, *The structure of ergodic measures for compact group extensions*, Israel J. Math. **18** (1974), 363–389. MR 0369660

[28] J. Kułaga-Przymus and M. Lemańczyk, *The Möbius function and continuous extensions of rotations*, Monatsh. Math. **178** (2015), no. 4, 553–582. MR 3422903

[29] J. Kułaga-Przymus and M. Lemańczyk, *Möbius disjointness along ergodic sequences for uniquely ergodic actions*, https://arxiv.org/abs/1703.02347.

[30] J. Kwiatkowski, *Spectral isomorphism of Morse dynamical systems*, Bull. Acad. Polon. Sci. Sér. Sci. Math. **29** (1981), no. 3-4, 105–114. MR 638749 (83g:28033)

[31] M. Lemańczyk, *Toeplitz* Z_2*-extensions*, Ann. Inst. H. Poincaré Probab. Statist. **24** (1988), no. 1, 1–43. MR 937955 (90b:28020)

[32] J. Liu and P. Sarnak, *The Möbius function and distal flows*, Duke Math. J. **164** (2015), no. 7, 1353–1399. MR 3347317

[33] K. Matomäki and M. Radziwił ł, *Multiplicative functions in short intervals*, Ann. of Math. (2) **183** (2016), no. 3, 1015–1056. MR 3488742

[34] K. Matomäki, M. Radziwiłł, and T. Tao, *An averaged form of chowla's conjecture*, Algebra & Number Theory **9** (2015), 2167–2196.

[35] C. Müllner, *Automatic sequences fulfill the Sarnak conjecture*, https://arxiv.org/abs/1602.03042.

[36] M. Queffélec, *Substitution dynamical systems—spectral analysis*, second ed., Lecture Notes in Mathematics, vol. 1294, Springer-Verlag, Berlin, 2010. MR 2590264 (2011b:37018)

[37] V. V. Ryzhikov, *Bounded ergodic constructions, disjointness, and weak limits of powers*, Trans. Moscow Math. Soc. (2013), 165–171. MR 3235793

[38] P. Sarnak, *Three lectures on the Möbius function, randomness and dynamics*, http://publications.ias.edu/sarnak/.

[39] T. Tao, *Equivalence of the logarithmically averaged chowla and sarnak conjectures*, https://arxiv.org/abs/1605.04628.

[40] T. Tao, *What's new. Blog. the chowla conjecture and the sarnak conjecture*, https://terrytao.wordpress.com/2012/10/14/the-chowla-conjecture-and-the-sarnak-conjecture/.

[41] W. A. Veech, *Möbius orthogonality for generalized Morse-Kakutani flows*, to appear in Amer. J. Math.

[42] Z. Wang, *Möbius disjointness for analytic skew products*, to appear in Inventiones Math., https://arxiv.org/abs/1509.03183.

[43] B. Weiss, *Strictly ergodic models for dynamical systems*, Bull. Amer. Math. Soc. (N.S.) **13** (1985), no. 2, 143–146. MR 799798

[44] S. Williams, *Toeplitz minimal flows which are not uniquely ergodic*, Z. Wahrsch. Verw. Gebiete **67** (1984), no. 1, 95–107. MR 756807 (86k:54062)