



**HAL**  
open science

## On arithmetic functions orthogonal to deterministic sequences

Adam Kanigowski, Joanna Kulaga-Przymus, Mariusz Lemańczyk, Thierry de La Rue

► **To cite this version:**

Adam Kanigowski, Joanna Kulaga-Przymus, Mariusz Lemańczyk, Thierry de La Rue. On arithmetic functions orthogonal to deterministic sequences. 2021. hal-03233409v2

**HAL Id: hal-03233409**

**<https://hal.science/hal-03233409v2>**

Preprint submitted on 1 Jun 2021 (v2), last revised 10 Sep 2021 (v3)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# On arithmetic functions orthogonal to deterministic sequences

Adam Kanigowski      Joanna Kułaga-Przymus  
Mariusz Lemańczyk      Thierry de la Rue

June 1, 2021

## Abstract

We prove Veech’s conjecture on the equivalence of Sarnak’s conjecture on Möbius orthogonality with a Kolmogorov type property of Furstenberg systems of the Möbius function. This yields a combinatorial condition on the Möbius function itself which is equivalent to Sarnak’s conjecture. As a matter of fact, our arguments remain valid in a larger context: we characterize all bounded arithmetic functions orthogonal to all topological systems whose all ergodic measures yield systems from a fixed characteristic class (zero entropy class is an example of such a characteristic class) with the characterization persisting in the logarithmic setup. As a corollary, we obtain that the logarithmic Sarnak’s conjecture holds if and only if the logarithmic Möbius orthogonality is satisfied for all dynamical systems whose ergodic measures yield nilsystems.

## Contents

<b>1</b>	<b>Introduction</b>	<b>2</b>
<b>2</b>	<b>Characteristic classes</b>	<b>15</b>
2.1	Definition, examples, basic properties . . . . .	15
2.2	The smallest nontrivial and the largest proper characteristic class . . . . .	17
2.3	Characteristic classes given by ergodic components . . . . .	19
2.3.1	Largest characteristic factor . . . . .	20
2.3.2	A class vs. its ec-class . . . . .	24
2.3.3	Strong $u$ -MOMO property of systems whose visible measures yield systems in an ec-class . . . . .	27
<b>3</b>	<b>Lifting lemma</b>	<b>30</b>
3.1	Good sequences of partitions . . . . .	31
3.2	Proof of Proposition 3.1 . . . . .	32
3.3	Logarithmic case . . . . .	38

<b>4</b>	<b>Proofs of main theorems</b>	<b>40</b>
4.1	Proof of Theorem B . . . . .	40
4.2	Proof of Theorem A . . . . .	40
<b>5</b>	<b>Combinatorics</b>	<b>42</b>
5.1	Orthogonality to zero entropy systems . . . . .	42
5.1.1	Cancellations. Proof of Corollaries C and C' . . . . .	42
5.1.2	Conditional cancellations. Remark 1.6 . . . . .	43
5.2	Orthogonality to $\mathcal{C}_{\text{ID}}$ . Proof of Corollary 1.8 . . . . .	45
5.3	Orthogonality to $\mathcal{C}_{\text{DISP}(G)}$ with $G$ countable . . . . .	46
5.4	Furstenberg systems and the strong $\mathbf{u}$ -MOMO property . . . . .	47
5.5	Orthogonality to $\mathcal{C}_{\text{DISP}_{\text{ec}}}$ . Proof of Corollary 1.10 . . . . .	49
5.6	Orthogonality to $\mathcal{C}_{\text{DISP}}$ . Averaged Chowla property for multiplicative functions . . . . .	50
<b>6</b>	<b>No strong <math>\mathbf{u}</math>-MOMO in positive entropy</b>	<b>52</b>
6.1	Proof of Theorem D . . . . .	52
6.2	Proof of Lemma 6.4 . . . . .	54
<b>A</b>	<b>From averaged double to averaged multiple correlations</b>	<b>60</b>

## 1 Introduction

All transformations in this paper are assumed to be invertible. A topological dynamical system is a pair  $(X, T)$ , where  $T$  is a homeomorphism of a compact metric space  $X$ . A measure-theoretic dynamical system is a system of the form  $(Z, \mathcal{B}(Z), \kappa, R)$  where  $R$  is an automorphism (invertible measure-preserving transformation) of a standard Borel probability space  $(Z, \mathcal{B}(Z), \kappa)$ .

**Sarnak's conjecture** Given a topological dynamical system  $(X, T)$  and a bounded arithmetic function  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$ , we consider the corresponding problem of orthogonality:

$$(1) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} f(T^n x) \mathbf{u}(n) = 0 \quad \text{for all } f \in C(X) \text{ and } x \in X.$$

Once (1) holds for  $(X, T)$ , we say that the system  $(X, T)$  *satisfies the Sarnak property with respect to  $\mathbf{u}$*  and write  $\mathbf{u} \perp (X, T)$ . When a class  $\mathcal{C}$  of topological dynamical systems is given, and  $\mathbf{u} \perp (X, T)$  for each  $(X, T) \in \mathcal{C}$  then we write  $\mathbf{u} \perp \mathcal{C}$ .

The main motivation to study this orthogonality problem is Sarnak's conjecture [46] on Möbius orthogonality in which  $\mathbf{u}$  is the Möbius function  $\boldsymbol{\mu}$  (or, equivalently [15], the Liouville function  $\boldsymbol{\lambda}$ ) and  $\mathcal{C}$  is the class  $\mathcal{C}_{\text{ZE}}$  of zero

topological entropy dynamical systems. Sequences of the form  $(f(T^n x))$  with  $(X, T)$  running over  $\mathcal{C}_{ZE}$ ,  $f \in C(X)$  and  $x \in X$  are often called *deterministic* sequences. Focusing on a special  $\mathbf{u}$  in (1), say, being multiplicative, is important if we count on applications in number theory – the main motivation of Sarnak himself for the Möbius orthogonality conjecture was to “attack” the celebrated Chowla conjecture on auto-correlations of the Möbius function dynamically (indeed the Chowla conjecture implies Sarnak’s conjecture [46], see also [2, 49]).<sup>1</sup> However, if we aim at providing an **internal** characterization of  $\mathbf{u}$  being orthogonal to  $\mathcal{C}_{ZE}$ , then dropping the assumption of multiplicativity of  $\mathbf{u}$  seems to be reasonable. Let us give an argument for that. First, note that the class of deterministic sequences equipped with the coordinatewise multiplication and addition is a ring (this is an easy consequence of the fact that the zero entropy class is closed under taking joinings and factors). Moreover, the class of bounded sequences  $\mathbf{u}$  orthogonal to  $\mathcal{C}_{ZE}$  is a module over this ring. In other words, even if our “starting”  $\mathbf{u}$  displays some additional arithmetic property (like multiplicativity), the characterization which we aim at must still hold for  $\mathbf{u} \cdot \mathbf{v}$ , where  $\mathbf{v}$  is an **arbitrary** deterministic sequence.<sup>2</sup> Of course, properties like the multiplicativity of  $\mathbf{u} \cdot \mathbf{v}$  changes dramatically if  $\mathbf{v}$  is arbitrary, while the characterization we are looking for has to be stable under such a change of  $\mathbf{u}$ .

**Visible measures and Furstenberg systems** Let us now briefly discuss the matter of the topological and measure-theoretic aspects of choosing a class  $\mathcal{C}$ . Recall that  $M(X)$  stands for the set of probability (Borel) measures on  $X$  and  $M(X, T)$  is the (always non-empty) subset of  $T$ -invariant measures. Both spaces are compact in the weak- $*$ -topology. There is a third natural subspace  $V(X, T) \subset M(X, T)$  which is the set of ( $T$ -invariant) *visible* measures, i.e. measures possessing a quasi-generic point. Formally,  $\nu \in V(X, T)$  if, for some point  $x \in X$  and some increasing sequence  $(N_\ell)$ , we have

$$\lim_{\ell \rightarrow \infty} \frac{1}{N_\ell} \sum_{n \leq N_\ell} \delta_{T^n x} = \nu.^3$$

We also write  $\nu \in V_T(x)$  and say that  $x$  is *generic for  $\nu$  along  $(N_\ell)$* . Note that the set  $V(X, T)$  of visible measures contains the set  $M^e(X, T)$  of all ergodic

<sup>1</sup>On a potential equivalence of the Chowla and Sarnak’s conjectures see [26, 46, 49, 50, 51] and a resumé of that in the survey articles [15, 37].

<sup>2</sup>In what follows, we replace  $\mathcal{C}_{ZE}$  by  $\mathcal{C}_{\mathcal{F}}$ , where  $\mathcal{F}$  is a general **characteristic** class. Note that the argument  $\mathbf{u} \cdot \mathbf{v} \perp \mathcal{C}_{\mathcal{F}}$  whenever  $\mathbf{u} \perp \mathcal{C}_{\mathcal{F}}$  and  $\mathbf{v}$  is an arbitrary  $\mathcal{F}$ -sequence (cf. Def. 1.1) persists.

<sup>3</sup>Not all invariant measures, even in a transitive subshift, have to be visible: let  $X_y \subset \{0, 1, 2\}^{\mathbb{N}}$  be given by  $y = 0^{k_1} 1^{k_1} 2^{k_1} 1^{k_1} 0^{k_2} 1^{k_2} 2^{k_2} 1^{k_2} \dots$  with  $k_1 < k_2 < \dots$ . Let  $\nu = \frac{1}{2} \delta_0 + \frac{1}{2} \delta_2$  (two fixed points). If  $A := \{x \in X_y : x_0 = 0 \text{ or } x_0 = 2\}$ , then  $\nu(A) = 1$  but no window in  $y$  has the property that the frequency of 0 jointly with 2 on it is close to the value 1.

measures, but it may be strictly smaller than  $M(X, T)$ . To illustrate this, consider the automorphism  $(x, y) \mapsto (x, x + y)$  of  $\mathbb{T}^2$  in which each point is generic (i.e. generic along the whole sequence of natural numbers) and the corresponding measure-theoretic system is either a rational or an irrational rotation on the circle. Hence, each visible measure gives rise to a system with discrete spectrum, while there are many other invariant measures which yield partly continuous spectrum.

Given a bounded  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$ ,  $|\mathbf{u}| \leq L$ , we first extend it to  $\mathbb{Z}$  by setting  $\mathbf{u}(-n) = \mathbf{u}(n)$  for  $n \in \mathbb{N}$ ,  $\mathbf{u}(0) \in \mathbb{D}_L := \{z \in \mathbb{C} : |z| \leq L\}$ , and then take the closure  $X_{\mathbf{u}} := \overline{\{S^k \mathbf{u} : k \in \mathbb{Z}\}}$  of the orbit of  $\mathbf{u} \in \mathbb{D}_L^{\mathbb{Z}}$  under the left shift  $S: \mathbb{D}_L^{\mathbb{Z}} \rightarrow \mathbb{D}_L^{\mathbb{Z}}$ ,  $S((v_k)_{k \in \mathbb{Z}}) = (v_{k+1})_{k \in \mathbb{Z}}$  (recall that we only consider invertible systems). Hence,  $(X_{\mathbf{u}}, S)$  is the *subshift* determined by  $\mathbf{u}$ . By the very definition,  $\mathbf{u} \in X_{\mathbf{u}}$ . Each measure  $\kappa \in V_S(\mathbf{u})$  yields a measure-theoretic system  $(X_{\mathbf{u}}, \mathcal{B}(X_{\mathbf{u}}), \kappa, S)$  called a *Furstenberg system* of  $\mathbf{u}$  (by some abuse of vocabulary, we may call  $\kappa$  itself a Furstenberg system of  $\mathbf{u}$ ). Fix  $x \in X$  and consider the sequence  $(\frac{1}{N} \sum_{n \leq N} \delta_{(T^n x, S^n \mathbf{u})})$ . By compactness of  $M(X \times X_{\mathbf{u}})$ , we can select an increasing sequence  $(N_\ell)$  so that

$$\frac{1}{N_\ell} \sum_{n \leq N_\ell} \delta_{(T^n x, S^n \mathbf{u})} \rightarrow \rho,$$

where  $\rho \in V_{T \times S}((x, \mathbf{u})) \subset V(X \times X_{\mathbf{u}}, T \times S)$ . The projections of  $\rho$  on  $X$  and  $X_{\mathbf{u}}$  are denoted by  $\nu$  and  $\kappa$ , respectively, so that  $\rho$  is a joining of  $\nu$  and  $\kappa$ . Since  $\rho$  is a visible measure, so are  $\nu$  and  $\kappa$ :

$$(2) \quad \nu \in V_T(x),$$

and  $\kappa$  yields a Furstenberg system of  $\mathbf{u}$ . Now, setting

$$(3) \quad \pi_0 : \begin{cases} X_{\mathbf{u}} \rightarrow \mathbb{C} \\ z = (z_n)_{n \in \mathbb{Z}} \mapsto z_0, \end{cases}$$

we have

$$(4) \quad \begin{aligned} \lim_{\ell \rightarrow \infty} \frac{1}{N_\ell} \sum_{n \leq N_\ell} f(T^n x) \mathbf{u}(n) &= \lim_{\ell \rightarrow \infty} \int f \otimes \pi_0 d \left( \frac{1}{N_\ell} \sum_{n \leq N_\ell} \delta_{(T^n x, S^n \mathbf{u})} \right) \\ &= \int f \otimes \pi_0 d\rho = \int \mathbb{E}^\rho(f | X_{\mathbf{u}}) \cdot \pi_0 d\kappa. \end{aligned}$$

So, the fact that the limit is 0 in (1), i.e.  $\mathbf{u} \perp (X, T)$ , depends on the joinings of measures from  $V(X, T)$  with those from  $V_S(\mathbf{u})$  – the invariant measures from  $M(X, T) \setminus V(X, T)$  are irrelevant in this context. More precisely, what will matter is the “geometric position” of the **single** continuous function  $\pi_0$  in the  $L^2$ -space of all Furstenberg systems  $\kappa \in V_S(\mathbf{u})$  for such joinings. Namely, inside the  $L^2(\rho)$ -space, we want  $\pi_0$  to be orthogonal to  $L^2(\nu)$  (in

condition (4) we use **all** continuous functions  $f$  on  $X$  and we take into account all  $\nu \in V(X, T)$ ). Of course, whether this orthogonality can be established without referring to the (visible) joinings  $\rho$  (and remaining at the level of  $\kappa \in V_S(\mathbf{u})$ ) is another question. A kind of surprise is that the answer to this question will turn out to be positive for  $\mathcal{C}_{ZE}$  and some other characteristic classes, see Theorem A, below.

**Characteristic classes and Veech's conjecture** With the above in mind, a knowledge of some fundamental facts on joinings, especially on disjointness in the sense of Furstenberg in ergodic theory, strongly suggests to consider only the situation when the measures  $\nu$  appearing above yield measure-theoretic systems  $(X, \mathcal{B}(X), \nu, T)$  belonging to one of so called *characteristic classes of measure-theoretic dynamical systems*.

**Definition 1.1.** A class  $\mathcal{F}$  of measure-theoretic dynamical systems is called *characteristic* if it is closed under taking isomorphisms, factors and (countable) joinings. Given a characteristic class  $\mathcal{F}$ , by an  $\mathcal{F}$ -factor of a measure-theoretic dynamical system  $(Z, \mathcal{D}, \kappa, R)$  we call any factor sub- $\sigma$ -algebra of  $\mathcal{D}$  on which the action of  $R$  yields a system in the class  $\mathcal{F}$ . We denote by  $\mathcal{C}_{\mathcal{F}}$  the class of topological systems  $(X, T)$  for which we have  $(X, \mathcal{B}(X), \nu, T) \in \mathcal{F}$  for each  $\nu \in V(X, T)$ . The sequences of the form  $(f(T^n x))_{n \in \mathbb{Z}}$  with  $f \in C(X)$ ,  $x \in X$  and  $(X, T) \in \mathcal{C}_{\mathcal{F}}$  are called  $\mathcal{F}$ -sequences.

It follows from the above definition that:

- (a) every measure-theoretic dynamical system  $(Z, \mathcal{D}, \kappa, R)$  has a *largest  $\mathcal{F}$ -factor* (in the sense of inclusion of sub- $\sigma$ -algebras), which we denote by  $\mathcal{D}_{\mathcal{F}}$ ,
- (b) any joining of  $(Z, \mathcal{D}, \kappa, R)$  with a system from  $\mathcal{F}$  is uniquely determined by its restriction to the joining of the largest  $\mathcal{F}$ -factor  $\mathcal{D}_{\mathcal{F}}$  of  $R$  with the given system from  $\mathcal{F}$

(see Section 2 for details). The class ZE of zero (Kolmogorov-Sinai) entropy is of course characteristic. The largest zero entropy factor  $\mathcal{D}_{ZE}$  of  $(Z, \mathcal{D}, \kappa, R)$  is called the *Pinsker factor* and is denoted by  $\Pi(R)$  or  $\Pi(\kappa)$ . Note that (via the variational principle) the family  $\mathcal{C}_{ZE}$  is precisely the family of all topological systems whose all invariant measures yield systems in ZE. Returning to Sarnak's conjecture, in [52], Veech proves the following result:

**Theorem 1.1** ([52]). *If*

$$(5) \quad \pi_0 \perp L^2(\Pi(\kappa)) \text{ for each Furstenberg system } \kappa \in V_S(\boldsymbol{\mu})$$

*then Sarnak's conjecture holds, i.e.  $\boldsymbol{\mu} \perp \mathcal{C}_{ZE}$ .*<sup>4</sup>

---

<sup>4</sup>An ergodic proof (which goes back to a suggestion of Sarnak in [46]) of this result is

Then he formulates the following conjecture (Conjecture 24.3 page 88 in [52]):

**Conjecture 1** (Veech’s conjecture). *Condition (5) is equivalent to Sarnak’s conjecture.*

One of motivations for the present work was to prove the above conjecture. Let us first formulate (5) in full generality.

**Definition 1.2.** Given a characteristic class  $\mathcal{F}$  we say that a (bounded) arithmetic function  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  satisfies the *Veech condition* with respect to  $\mathcal{F}$  if

$$(6) \quad \pi_0 \perp L^2((\mathcal{B}(X_{\mathbf{u}}), \kappa)_{\mathcal{F}}) \text{ for each Furstenberg system } \kappa \in V_S(\mathbf{u}).$$

Given a characteristic class  $\mathcal{F}$ , we denote by  $\mathcal{F}_{ec}$  the class of those automorphisms  $R$  of  $(Z, \mathcal{D}, \kappa)$  such that a.e. ergodic component of  $\kappa$  yields a system in  $\mathcal{F}$ . Then (see Section 2.3),  $\mathcal{F}_{ec}$  is also a characteristic class,

$$(7) \quad \mathcal{C}_{\mathcal{F}} \subset \mathcal{C}_{\mathcal{F}_{ec}}$$

and

$$(8) \quad \mathcal{C}_{\mathcal{F}_{ec}} = \{(X, T) : (X, \mathcal{B}(X), \nu, T) \in \mathcal{F} \text{ for each ergodic } \nu \in M(X, T)\}.$$

Our main result is the following:

**Theorem A.** *Assume that  $\mathcal{F}$  is a characteristic class. Let  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  be a bounded arithmetic function. Then  $\mathbf{u} \perp \mathcal{C}_{\mathcal{F}_{ec}}$  if and only if  $\mathbf{u}$  satisfies the Veech condition (6) (with respect to  $\mathcal{F}_{ec}$ ).*

**Remark 1.2.** A subsequence version of this result also holds: If  $(N_\ell)$  is an increasing sequence of integers, then  $\mathbf{u}$  is  $(N_\ell)$ -orthogonal to  $\mathcal{C}_{\mathcal{F}_{ec}}$  (i.e.  $\frac{1}{N_\ell} \sum_{n \leq N_\ell} f(T^n x) \mathbf{u}(n) \rightarrow 0$  for each  $(X, T) \in \mathcal{C}_{\mathcal{F}_{ec}}$  and all  $f \in C(X)$ ,  $x \in X$ ) if and only if the Veech condition (6) holds for each measure  $\kappa$  for which  $\mathbf{u}$  is generic along some subsequence of  $(N_\ell)$ . The reason for the validity of this “local” version is that all tools used in the proof of Theorem A work well on subsequences, see also Remarks 1.4 and 2.20 to cope with the strong  $\mathbf{u}$ -MOMO property along subsequences.

See [8] for the validity of the alternative: *either there are no Siegel zeros or there exists a (universal) subsequence along which Sarnak’s conjecture holds* and [27] for a density version of Sarnak’s conjecture.

---

implicit in [2], where the implication “Chowla conjecture  $\Rightarrow$  Sarnak’s conjecture” has been proved using joinings. Veech cites [2] (which was on arXiv two years before Veech’s lecture notes [52] appeared) but instead he gives his own (slightly complicated) proof using the concept of quasi-factors of Glasner and Weiss.

Theorem A for  $\mathbf{u} = \boldsymbol{\mu}$  and  $\mathcal{F} = \text{ZE}$  (note that  $\text{ZE}_{\text{ec}} = \text{ZE}$ ) has the following consequence:

**Corollary 1.3.** *Veech's conjecture holds. Moreover, if Sarnak's conjecture holds then all Furstenberg systems  $\kappa \in V_S(\boldsymbol{\mu})$  have positive entropy.*

We also generalize Veech's theorem (Theorem 1.1) to the setting of characteristic classes:

**Theorem B.** *Assume that  $\mathcal{F}$  is a characteristic class. Let  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  be a bounded arithmetic function. If  $\mathbf{u}$  satisfies the Veech condition (6) with respect to  $\mathcal{F}$  then  $\mathbf{u} \perp \mathcal{C}_{\mathcal{F}}$ .*

Let us now briefly discuss the methods involved in the proofs of the above theorems. The proof of Theorem B is provided in Section 4.1. It is a straightforward application of the above Property (b) of the largest  $\mathcal{F}$ -factor (see Proposition 2.2 for more details). The proof of Theorem A which we provide in Section 4.2 relies on:

- so called Hansel models [29] of possibly highly non-ergodic measure-preserving systems,
- our version of a lifting lemma by Conze, Downarowicz, Serafin [9] on the existence of generic points for joinings, valid in a general context and using the concept of quasi-generic sequences (Proposition 3.1),
- general joining techniques.

Our first proof of Theorem A was based on a different lifting lemma by Bergelson, Downarowicz and Vandehey (Theorem 5.16 in [6]) and a recent result by Downarowicz and Weiss showing the existence, for each zero-entropy measure-theoretic system, of a special Hansel model which is symbolic [13]. The proof we finally chose to present here does not require the use of this symbolic model, as our lifting lemma works for all topological systems. Its additional advantage is that it is also valid in the context of logarithmic averages (see Section 3.3). This makes all our results also true in the logarithmic set up: for example, the logarithmic Sarnak's conjecture (denoted as  $\boldsymbol{\mu} \perp_{\log} \mathcal{C}_{\text{ZE}}$ ) is equivalent to the Veech condition for  $\boldsymbol{\mu}$  for all logarithmic Furstenberg systems  $\kappa \in V_S^{\log}(\boldsymbol{\mu})$ , see Corollary 1.12 for more.

In Theorem A, it is crucial that we deal with a characteristic class of the form  $\mathcal{F}_{\text{ec}}$  since, by Proposition 2.17,  $\mathbf{u} \perp \mathcal{C}_{\mathcal{F}_{\text{ec}}}$  is equivalent to the so called *strong  $\mathbf{u}$ -MOMO property* for systems in  $\mathcal{C}_{\mathcal{F}_{\text{ec}}}$ . This property, whose definition is recalled below, has been introduced in [3, 27]. We leave as an open problem whether Theorem A holds for an arbitrary characteristic class.



**Definition 1.3.** A topological system  $(X, T)$  satisfies the *strong  $\mathbf{u}$ -MOMO property*, if

$$(9) \quad \lim_{K \rightarrow \infty} \frac{1}{b_K} \sum_{k < K} \left\| \sum_{b_k \leq n < b_{k+1}} \mathbf{u}(n) f \circ T^n \right\|_{C(X)} = 0$$

for each  $f \in C(X)$  and each increasing sequence  $(b_k) \subset \mathbb{N}$  such that  $b_{k+1} - b_k \rightarrow \infty$ .

Clearly, the strong  $\mathbf{u}$ -MOMO property implies (1) uniformly in  $x \in X$ . The concept of strong  $\mathbf{u}$ -MOMO is formally stronger than the usual orthogonality. To see the difference between the usual orthogonality and strong MOMO, consider the system  $(x, y) \mapsto (x, x + y)$  on  $\mathbb{T}^2$  whose Möbius orthogonality follows easily from the DDKBSZ criterion,<sup>5</sup> see e.g. [15]) (in fact, the orthogonality holds even uniformly due to the Davenport estimate [10]), while the strong  $\boldsymbol{\mu}$ -MOMO property (apply the definition to  $f(x, y) = e^{2\pi i y}$ ) yields

$$\frac{1}{b_K} \sum_{k < K} \sup_{x \in \mathbb{T}} \left| \sum_{b_k \leq n < b_{k+1}} \boldsymbol{\mu}(n) e^{2\pi i n x} \right| \rightarrow 0$$

when  $K \rightarrow \infty$  which is an open problem,<sup>6</sup> see Sections 5.4 and 5.5 for more details.

**Remark 1.4.** Note that in the definition of the strong  $\mathbf{u}$ -MOMO property, convergence (9) can be replaced by

$$(10) \quad \lim_{N \rightarrow \infty} \frac{1}{N} \left( \sum_{k < K_N} \left\| \sum_{b_k \leq n < b_{k+1}} \mathbf{u}(n) f \circ T^n \right\|_{C(X)} + \left\| \sum_{b_{K_N} \leq n < N} \mathbf{u}(n) f \circ T^n \right\|_{C(X)} \right) = 0,$$

where  $K_N := \max\{k : b_k < N\}$ . As a matter of fact, the definition is unchanged if we only restrict to sequences  $(b_k)$  which further satisfy the condition

$$\lim_{k \rightarrow \infty} \frac{b_{k+1} - b_k}{b_k} = 0.$$

(If we have to consider a sequence  $(b_k)$  which does not satisfy the above condition, we can always add more integers to the set  $\{b_k : k \geq 1\}$  so that

<sup>5</sup>DDKBSZ stands for Daboussi-Delange-Kátai-Bourgain-Sarnak-Ziegler [7], [35].

<sup>6</sup>In fact, it is open whether a non-periodic, zero entropy, continuous, algebraic automorphism of  $\mathbb{T}^2$  satisfies the strong  $\boldsymbol{\mu}$ -MOMO property.

this convergence holds for the new sequence. And the validity of (9) or (10) for the new sequence is stronger than the same for the former sequence  $(b_k)$ .

Then it is easy to define also the strong  $\mathbf{u}$ -MOMO property along an increasing sequence  $(N_\ell)$ , by restricting convergence (10) to the subsequence  $(N_\ell)$ .

**Veech condition and combinatorics** Given a characteristic class  $\mathcal{F}$ , Theorems B and A determine a natural strategy to describe the arithmetic functions  $\mathbf{u}$  orthogonal to all  $\mathcal{F}$ -sequences. Namely, we need to describe the  $\mathcal{F}$ -factors and understand the orthogonality to their  $L^2$ -space (i.e. the Veech condition), hoping that this description can be expressed (for  $\pi_0$  in  $X_{\mathbf{u}}$  and a Furstenberg system  $\kappa$  of  $\mathbf{u}$ ) by some asymptotics of the integrals of continuous functions on  $X_{\mathbf{u}}$ . The final step would be to use the definition of a Furstenberg system to obtain a combinatorial condition on  $\mathbf{u}$  itself. The first part of the strategy should be seen as an extension of the theory of characteristic factors  $\mathcal{Z}_s(T)$  (given an automorphism  $T$ ) and the Gowers-Host-Kra (GHK in what follows) seminorms  $\|\cdot\|_{u^s}$  for  $s \geq 1$  [31]. In this perspective the Veech condition on  $\pi_0$  is the counterpart of  $\|\pi_0\|_{u^s} = 0$  for each Furstenberg system  $\kappa \in V_S(\mathbf{u})$  (and has its combinatorial translation in terms of the GHK seminorm of  $\mathbf{u}$ ). We will give more details on this shortly.

Let us discuss this strategy for the class ZE, see Section 5.1 for details. Here, the characteristic factor of a measure-preserving system is the Pinsker factor. The reader has certainly noticed that even though we study dynamical properties of Furstenberg systems, as a matter of fact, at the end we deal with a process  $(\pi_0 \circ S^n)_{n \in \mathbb{Z}}$ , stationary with respect to  $\kappa \in V_S(\mathbf{u})$  (each such measure is invariant in the subshift  $(X_{\mathbf{u}}, S)$ ). Now, the Veech condition leads to the following concept.

**Definition 1.4.** A centered stationary process  $\underline{X} = (X_n)$  taking finitely many values is called a *Sarnak process* if  $\mathbb{E}(X_0 \mid \sigma(X_N, X_{N+1}, \dots)) \rightarrow 0$  in  $L^2$  (or a.e.); equivalently  $\mathbb{E}(X_0 \mid \Pi(\underline{X})) = 0$ , where  $\Pi(\underline{X})$  stands for the tail  $\sigma$ -algebra.

Understanding the structure of Sarnak processes seems to be a problem of an independent interest and it will be studied elsewhere.

Now, when  $\mathbf{u}$  takes finitely many values, our main result (Theorem A applied to ZE) can be reformulated in the following manner:

**Corollary 1.5.** *Let  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  be an arithmetic function taking finitely values. Then  $\mathbf{u} \perp \mathcal{C}_{ZE}$  if and only if all stationary processes  $(\pi_0 \circ S^n)_{n \in \mathbb{Z}}$  determined by  $\kappa \in V_S(\mathbf{u})$  are Sarnak.*

From the ergodic theory point of view we are close to the concept of the **relative** Kolmogorov property (K-property) which is however perturbed by the fact that we need this property for a single function. But even though

only  $\pi_0$  is involved, the dynamical idea of the equivalence between the K-property and K-mixing (uniform mixing) works, and we can apply K-mixing of  $\pi_0$  against the family of functions depending on finitely many non-negative coordinates. This leads to studying the asymptotics of integrals of some continuous functions and finally gives the following combinatorial characterization of the orthogonality of  $\mathbf{u}$  to all deterministic sequences. In the following corollary, we use the fact that when  $\mathbf{u}$  takes its values in a finite set, a subset  $A \subset X_{\mathbf{u}}$  depends on finitely many non-negative coordinates if and only if there exists  $\ell \geq 1$  and a set  $C$  of blocks of length  $\ell$  appearing in  $\mathbf{u}$  such that  $\mathbb{1}_A(y) = \mathbb{1}_{(y(0), y(1), \dots, y(\ell-1)) \in C}$ .

**Corollary C.** *Let  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  be an arithmetic function taking finitely many values. Then  $\mathbf{u} \perp \mathcal{C}_{ZE}$  if and only if, for each subsequence  $(N_k)$  defining a Furstenberg system of  $\mathbf{u}$  and each subset  $A \subset X_{\mathbf{u}}$  depending on finitely many non-negative coordinates, we have the cancellation phenomenon of the values of  $\mathbf{u}$  uniformly along sufficiently large shifts of the set of visits of  $\mathbf{u}$  in  $A$ : for each  $\varepsilon > 0$ , there exists  $M \geq 1$  such that for each  $\ell \geq 1$  and each set  $C$  of blocks of length  $\ell$ , we have for each  $m \geq M$*

$$(11) \quad \lim_{k \rightarrow \infty} \left| \frac{1}{N_k} \sum_{n \leq N_k} \mathbf{u}(n) \underbrace{\mathbb{1}_{(\mathbf{u}(m+n), \mathbf{u}(m+n+1), \dots, \mathbf{u}(m+n+\ell-1)) \in C}}_{\mathbb{1}_A(S^{m+n}\mathbf{u})} \right| \leq \varepsilon.$$

Note that  $M$  above depends on  $(N_k)$  and  $\varepsilon$ . This combinatorial condition looks more attractive if we assume that  $\mathbf{u}$  is generic:

**Corollary C'.** *Let  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  be an arithmetic function taking finitely many values. If  $\mathbf{u}$  is generic then  $\mathbf{u} \perp \mathcal{C}_{ZE}$  if and only if*

$$\lim_{m \rightarrow \infty} \lim_{N \rightarrow \infty} \left| \frac{1}{N} \sum_{n \leq N} \mathbf{u}(n) \mathbb{1}_{(\mathbf{u}(m+n), \mathbf{u}(m+n+1), \dots, \mathbf{u}(m+n+\ell-1)) \in C} \right| = 0$$

*uniformly in  $\ell \geq 1$  and in  $C$ , a set of blocks of length  $\ell$ .*

**Remark 1.6.** Note however that the above two corollaries do not say much if the (clopen) sets  $C$  are already of small measures (e.g., in the most interesting case of blocks of large length). In fact, condition (11) in Corollary C is equivalent to the following: for each subsequence  $(N_k)$  defining a Furstenberg system of  $\mathbf{u}$  we have the **conditional** cancellation phenomenon of the values of  $\mathbf{u}$  uniformly along sufficiently large shifts of the set of visits of  $\mathbf{u}$  to “typical” blocks. More precisely, for each  $\varepsilon > 0$  there exists  $M \geq 1$  such that

$$\lim_{k \rightarrow \infty} \frac{\left| \sum_{n \leq N_k} \mathbf{u}(n) \mathbb{1}_{(\mathbf{u}(m+n), \mathbf{u}(m+n+1), \dots, \mathbf{u}(m+n+\ell-1)) \in C} \right|}{\sum_{n \leq N_k} \mathbb{1}_{(\mathbf{u}(n), \dots, \mathbf{u}(n+\ell-1)) \in C}} \leq \varepsilon$$

for all  $m \geq M$ , all  $\ell$  sufficiently large and blocks  $C$  of length  $\ell$  forming a family of measure  $> 1 - \varepsilon$ .

If, additionally,  $\mathbf{u}$  is generic, then the above condition reduces to

$$\lim_{m \rightarrow \infty} \lim_{N \rightarrow \infty} \frac{|\sum_{n \leq N} \mathbf{u}(n) \mathbb{1}_{(\mathbf{u}(m+n), \mathbf{u}(m+n+1), \dots, \mathbf{u}(m+n+\ell-1))=C}|}{\sum_{n \leq N} \mathbb{1}_{\mathbf{u}(m+n), \dots, \mathbf{u}(m+n+\ell-1)=C}} = 0$$

uniformly in  $m$ , for “good” blocks of length  $\ell$  sufficiently large.

**Remark 1.7.** Note the basic difference between the sums

$$\sum_{n \leq N_k} \mathbf{u}(n) \mathbb{1}_{(\mathbf{u}(n), \mathbf{u}(n+1), \dots, \mathbf{u}(n+\ell-1))=C} = C[0] \cdot \sum_{n \leq N_k} \mathbb{1}_{(\mathbf{u}(n), \mathbf{u}(n+1), \dots, \mathbf{u}(n+\ell-1))=C}$$

and

$$\sum_{n \leq N_k} \mathbf{u}(n) \mathbb{1}_{(\mathbf{u}(m+n), \mathbf{u}(m+n+1), \dots, \mathbf{u}(m+n+\ell-1))=C},$$

namely, the first one does not display **any** cancellation (that is, along the return times to a fixed block, we have no cancellation) and in the second one cancellations are possible and the fact that along further and further shifts of the set of return times we observe more and more cancellations, characterizes the Sarnak property.

### Veech’s and Sarnak’s conjectures for other characteristic classes

The family of all characteristic classes is enormous, see Section 2 for natural examples. Here, let us just notice that the discrete spectrum automorphisms form a characteristic class and it contains uncountably many characteristic subclasses whose pairwise intersections are equal to the class of all identities (indeed, discrete spectrum automorphisms whose group of eigenvalues is contained in  $\mathbb{Z}\alpha$ , with  $\alpha$  irrational, is a characteristic class). Moreover, there are the largest proper and the smallest non-trivial characteristic classes. Indeed, although our study of the zero entropy class was originally motivated by Sarnak’s conjecture, yet, ZE plays a special role since it is the largest (proper) characteristic class. In fact, we have

$$(12) \quad \{*\} \subset \text{ID} \subset \mathcal{F} \subset \text{ZE} \subset \text{ALL}$$

for each characteristic class  $\mathcal{F}$ , where ID stands for the (characteristic) class of identities (of all standard Borel probability spaces) and ALL stands for the (characteristic) class of all systems. Note that

$$\text{ID} = \{*\}_{\text{ec}} \text{ and } \text{ZE} = \text{ZE}_{\text{ec}}.$$

Clearly, the topological class  $\mathcal{C}_{\text{ALL}}$  consists of all topological systems and the only  $\mathbf{u}$  orthogonal to all of them is  $\mathbf{u}(n) = 0$  on a subset of  $n$  of full

density which is compatible with the Veech condition (which in this setting means that  $\pi_0$  equals (a.e.) zero for each Furstenberg system).

The topological class  $\mathcal{C}_{\{*\}}$  consists of topological systems whose all visible measures are given by fixed points. The reader can check that the Veech condition here is just  $\int_{X_{\mathbf{u}}} \pi_0 d\kappa = 0$  for each  $\kappa \in V_S(\mathbf{u})$ . The combinatorial condition (11) from Corollary C (equivalent to the Veech condition) in this setting reduces to  $\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \mathbf{u}(n) = 0$ . Clearly, in this setting, the latter is nothing but the Sarnak condition.

It is not hard to see that the topological class  $\mathcal{C}_{\text{ID}}$  consists of topological systems whose only ergodic measures are Dirac measures at fixed points. The Veech condition here is  $\pi_0 \perp L^2(\mathcal{I}_{\kappa})$  for each  $\kappa \in V_S(\mathbf{u})$ , where  $\mathcal{I}_{\kappa}$  stands for the  $\sigma$ -algebra of invariant sets. Finally, the counterparts of Corollaries C and C' are the following.

**Corollary 1.8.** *Let  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  be a bounded arithmetic function. Then  $\mathbf{u} \perp \mathcal{C}_{\text{ID}}$  if and only if for each  $\varepsilon > 0$  and each subsequence  $(N_k)$  defining a Furstenberg system of  $\mathbf{u}$ , there exists  $H_0 \geq 1$  such that for each  $H \geq H_0$ ,*

$$(13) \quad \lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} \left| \frac{1}{H} \sum_{h \leq H} \mathbf{u}(n+h) \right|^2 \leq \varepsilon.$$

Note that if, additionally,  $\mathbf{u}$  is generic and it satisfies a non-quantitative version of the Matomäki-Radziwiłł [41] convergence on a typical short interval:

$$\lim_{M, H \rightarrow \infty, H=o(M)} \frac{1}{M} \sum_{n \leq M} \left| \frac{1}{H} \sum_{h \leq H} \mathbf{u}(n+h) \right|^2 = 0$$

then  $\mathbf{u} \perp \mathcal{C}_{\text{ID}}$ .

**Remark 1.9.** It is also worth mentioning that the ID-sequences are precisely the *mean slowly varying functions* (see Proposition 5.1 in [28]), i.e. (bounded) arithmetic functions  $\mathbf{v}: \mathbb{N} \rightarrow \mathbb{C}$  for which

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n < N} |\mathbf{v}(n+1) - \mathbf{v}(n)| = 0.$$

Therefore, sequences satisfying (13) are precisely those orthogonal to all mean slowly varying functions.

Notice that it follows from (12) that, for a non-trivial class  $\mathcal{F}$ , the “zero mean condition on a typical short interval” (13) is a necessary condition for  $\mathbf{u} \perp \mathcal{C}_{\mathcal{F}}$ , whereas the condition given by Corollary C is sufficient for  $\mathbf{u} \perp \mathcal{C}_{\mathcal{F}}$ .

In Section 5.5, we discuss the case  $\mathcal{F} = \text{DISP}_{\text{ec}}$ , where DISP stands for the (characteristic) class of discrete spectrum automorphisms. In view of (8),  $\mathcal{C}_{\text{DISP}_{\text{ec}}}$  consists of homeomorphisms whose all **ergodic** measures

yield systems with discrete spectrum. Let  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  be bounded. For an increasing sequence of integers  $(N_k)$ , we set

$$(14) \quad \|\mathbf{u}\|_{u^1((N_k))}^2 := \lim_{H \rightarrow \infty} \frac{1}{H} \sum_{h \leq H} \left( \lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} \mathbf{u}(n) \overline{\mathbf{u}(n+h)} \right)$$

and, for  $s > 1$ ,

$$(15) \quad \|\mathbf{u}\|_{u^s((N_k))}^{2s+1} := \lim_{H \rightarrow \infty} \frac{1}{H} \sum_{h \leq H} \|\mathbf{u}(\cdot + h) \cdot \mathbf{u}(\cdot)\|_{u^s((N_k))}^{2s},$$

whenever all the above limits exist. If  $N_k = k$ , we set  $\|\mathbf{u}\|_{u^s} := \|\mathbf{u}\|_{u^s((N_k))}$ .

**Corollary 1.10.** *Let  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  be bounded. Then  $\mathbf{u} \perp \mathcal{C}_{\text{DISP}_{ec}}$  if and only if  $\|\mathbf{u}\|_{u^2((N_k))} = 0$  for each subsequence  $(N_k)$  along which  $\mathbf{u}$  is generic. In particular, if  $\mathbf{u}$  is generic then  $\|\mathbf{u}\|_{u^2} = 0$ .*

The main assertion of Corollary 1.10 is equivalent to  $\|\pi_0\|_{u^2(\kappa)} = 0$  for each Furstenberg system  $\kappa \in V_S(\mathbf{u})$ . The reason for the validity of this result is that given an automorphism  $(Z, \mathcal{D}, \rho, R)$ , we have the equality

$$\mathcal{D}_{\text{DISP}_{ec}} = \mathcal{Z}_1(R, \rho)$$

which is a consequence of  $\text{DISP}_{ec} = \text{NIL}_1$ , where  $\text{NIL}_s$  stands for the class of automorphisms whose a.a. ergodic components are inverse limits of  $s$ -step nil-automorphisms (see Section 2 for more details).

When we turn to the classes  $\text{NIL}_s (= (\text{NIL}_s)_{ec})$  and return to the original Sarnak's conjecture (for  $\mathbf{u}$ ) then clearly  $\pi_0 \perp L^2(\Pi(\kappa))$  implies  $\pi_0 \perp L^2(\mathcal{Z}_s(\kappa))$  for each  $s \geq 1$ . We hence obtain one more necessary condition for  $\mathbf{u}$  to be orthogonal to all deterministic sequences:

**Corollary 1.11.** *Let  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  be bounded. If  $\mathbf{u} \perp \mathcal{C}_{\text{ZE}}$  then  $\|\mathbf{u}\|_{u^s((N_k))} = 0$  for each  $s \in \mathbb{N}$  (for each subsequence  $(N_k)$  along which  $\mathbf{u}$  is generic).*

In Section 2.3.2, we prove that the Sarnak property of  $\mathbf{u}$  for the fundamental (in ergodic theory) class of distal automorphisms is equivalent to the Veech property of  $\mathbf{u}$ . We leave as an open problem whether the Veech property can be expressed combinatorially in the distal case.

**The logarithmic Sarnak's conjecture** As we have already noticed, our results also hold in the logarithmic case. In the corollary below we put together conditions which are equivalent to the logarithmic Sarnak's conjecture.<sup>7</sup>

<sup>7</sup>Our thanks go to N. Frantzikinakis who pointed out to us one of crucial equivalences: (iv)  $\Leftrightarrow$  (v).

**Corollary 1.12.** *Let  $\mathbf{u} = \boldsymbol{\mu}$  or  $\boldsymbol{\lambda}$ . The following conditions are equivalent:*

- (i)  $\mathbf{u} \perp_{\log} \mathcal{C}_{ZE}$  (i.e. zero entropy systems satisfy the logarithmic Sarnak property with respect to  $\mathbf{u}$ ),
- (ii)  $\mathbf{u} \perp_{\log} \mathcal{C}_{NIL_s}$  for each  $s \geq 1$ ,
- (iii)  $\pi_0 \perp L^2(\Pi(\kappa))$  for each  $\kappa \in V_S^{\log}(\mathbf{u})$  (i.e.  $\mathbf{u}$  satisfies the Veech condition for each logarithmic Furstenberg system of  $\mathbf{u}$ ),
- (iv)  $\pi_0 \perp L^2(\mathcal{Z}_s(\kappa))$  for each  $\kappa \in V_S^{\log}(\mathbf{u})$  and  $s \geq 1$ ,
- (v)  $\mathbf{u}$  satisfies the logarithmic Chowla conjecture.

The equivalence of (ii) and (iv) is due to Theorem A (in its logarithmic form), the equivalence of (iv) and (v) (based on the facts proved by Tao [50] for the equivalence of the logarithmic Sarnak’s and Chowla’s conjectures for the Liouville function) is formally proved in [16] (implicit in Section 2.7 therein). Other (needed) implications are standard. Recall also that, by [27], (i) is equivalent to the logarithmic strong  $\mathbf{u}$ -MOMO property of all zero entropy systems.

Note that the equivalence of (i) and (ii) in Corollary 1.12 yields immediately the following.

**Corollary 1.13.** *The logarithmic Sarnak’s conjecture holds if and only if for each  $s \geq 1$ ,  $\boldsymbol{\mu}$  is orthogonal to all systems whose all ergodic measures yield  $NIL_s$ -systems.*

**Remark 1.14.** As Theorem A is true in a larger context, also Corollary 1.12 can be formulated for more general **multiplicative** functions bounded by 1, cf. Theorem 1.8 in [16].

**Averaged Chowla property** The “iff” assertion of Theorem A cannot be applied to the class  $\mathcal{C}_{DISP}$ . However, in Section 5.6, we will show that the Sarnak and Veech conditions are equivalent in this setting for each bounded  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  such that

$$(16) \quad \text{all circle rotations satisfy the strong } \mathbf{u}\text{-MOMO property}$$

(strong  $\mathbf{u}$ -MOMO property has been defined in Definition 1.3).

**Corollary 1.15.** *Assume that  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  is bounded by 1 and satisfies (16). Then  $\mathbf{u}$  satisfies the Veech condition for the  $\mathcal{F} = DISP$ . In particular, the Sarnak and Veech conditions are equivalent for  $\mathcal{F} = DISP$ . Moreover, for every sequence  $(N_k)$  along which  $\mathbf{u}$  is generic,  $\mathbf{u}$  satisfies the averaged 2-Chowla property:*

$$(17) \quad \lim_{H \rightarrow \infty} \frac{1}{H} \sum_{h \leq H} \lim_{k \rightarrow \infty} \frac{1}{N_k} \left| \sum_{n \leq N_k} \mathbf{u}(n) \overline{\mathbf{u}(n+h)} \right| = 0$$

and, for all sequences  $\mathbf{v}_1, \dots, \mathbf{v}_k$  (bounded by 1), we have

$$(18) \quad \lim_{H \rightarrow \infty} \frac{1}{H^k} \sum_{h_1, \dots, h_k \leq H} \lim_{k \rightarrow \infty} \frac{1}{N_k} \left| \sum_{n \leq N_k} \mathbf{u}(n) \prod_{i=1}^k \mathbf{v}_i(n + h_i) \right| = 0.$$

Property (18), called the *averaged Chowla property* (cf. [42]), follows from (17) – this will be shown in Appendix A. For an alternative approach to obtain the assertions of the above corollary, see the method in [19]: for (17) cf. Thm. 4.1 and Prop. 4.3 [19], and for (18) cf. Thm. 2.1 and Prop. 5.1 therein.

**Corollary 1.16.** *Let  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  be a multiplicative function bounded by 1. If, for each Dirichlet character  $\chi$ ,  $\mathbf{u} \cdot \chi$  satisfies the short interval behaviour (13)<sup>8</sup> then  $\mathbf{u}$  satisfies the averaged Chowla property (18) (along each sequence  $(N_k)$  for which  $\mathbf{u}$  is generic).*

The above result with  $(N_k)$  being all positive integers has been proved by Matomäki, Radziwiłł and Tao in [42]. Note that to obtain Corollary 1.16 it is enough to show that irrational rotations satisfy the strong  $\mathbf{u}$ -MOMO property. Via the DDKBSZ criterion, this follows from the ergodic property AOP introduced and proved to hold for totally ergodic rotations in [1]. Thus, we obtain an ergodic proof of the averaged Chowla property for each sequence  $(N_k)$  as above for all multiplicative  $\mathbf{u}$  enjoying the special short interval behavior. In particular, assuming that  $\mathbf{u}$  is generic, we get a non-quantitative version of [42].

In Section 6, we prove (see Theorem D) that for each  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  taking finitely many values, if it satisfies the Sarnak property for the class  $\mathcal{C}_{ZE}$ , then no positive entropy system satisfies the strong  $\mathbf{u}$ -MOMO property (this was previously known for the Liouville function assuming the Chowla conjecture [3]).

## 2 Characteristic classes

### 2.1 Definition, examples, basic properties

Recall that a class  $\mathcal{F}$  of measure-theoretic dynamical systems is characteristic if it is closed under taking isomorphisms, factors and (countable) joinings. Recall also the following classical result on such classes (see e.g. [45]).

**Proposition 2.1.** *Given a characteristic class  $\mathcal{F}$ , each automorphism  $R$  on  $(Z, \mathcal{D}, \kappa)$  has a largest  $\mathcal{F}$ -factor, denoted by  $\mathcal{D}_{\mathcal{F}}$ .*

The following result whose proof is based on a fundamental non-disjointness lemma from [39] will be crucial for us:

---

<sup>8</sup>This is equivalent to saying that along arithmetic sequences the averages on a typical short interval vanish.



**Proposition 2.2** ([45]). *Let  $(X, \mathcal{B}, \nu, T)$  be a measure-theoretic dynamical system in the characteristic class  $\mathcal{F}$ , and let  $(Z, \mathcal{D}, \kappa, R)$  be any measure-theoretic dynamical system. Then any joining of  $R$  and  $T$  is relatively independent over the largest  $\mathcal{F}$ -factor  $\mathcal{D}_{\mathcal{F}}$  of  $R$ . That is: if  $g \in L^2(Z, \kappa)$  is such that  $\mathbb{E}^{\kappa}[g|\mathcal{D}_{\mathcal{F}}] = 0$ , and if  $\rho$  is a joining of  $T$  and  $R$ , then for any  $f \in L^2(X, \nu)$  we have*

$$(19) \quad \mathbb{E}^{\rho}(f \otimes g) = 0.$$

**Examples of characteristic classes** (some acronyms are used for those which will be used in the sequel):

- ALL: all automorphisms of standard Borel probability spaces;
- $\{*\}$ : the identity on the one-point space;
- ID: identity automorphisms (of all standard Borel probability spaces);
- DISP: discrete spectrum automorphisms;
- RDISP: rational discrete spectrum automorphisms;
- $\text{DISP}(G)$ : discrete spectrum automorphisms whose group of eigenvalues is contained in fixed countable subgroup  $G$  of the circle;
- $\text{NIL}_s$ : automorphisms whose a.a. ergodic components are inverse limits of  $s$ -step nilautomorphisms. The fact that ergodic joinings of nilsystems remain nil, see Proposition 15, page 186 in the book [31], and the same holds for inverse limits (this is actually Lemma A.4 in [20]). Regarding factors of ergodic nilsystems, see Theorem 11 in page 230 [31]. Here  $s \in \mathbb{N}$ .
- DIST: distal automorphisms are those which are given as a transfinite (indexed by ordinals smaller than a fixed countable ordinal) sequence of consecutive extensions each of which either has relative discrete spectrum or (in case of a limit ordinal) is the corresponding inverse limit. The structural theorem Theorem 6.17 together with the concluding remark (for  $\mathbb{Z}$ -actions) on page 139 [22] tell us that each system has a largest distal factor, hence DIST is closed under countable joinings. In Lemma 2.15, we note that an automorphism is distal if and only if a.a. its ergodic components are distal. To see that this class is closed under taking factors, let us first recall that this fact holds for ergodic automorphisms (see Theorem 10.18 [25]). If now  $\mathcal{A} \subset \mathcal{D}$  is a factor of a distal automorphism  $(Z, \mathcal{D}, \kappa, R)$  then  $\mathcal{A}$  (as an  $R$ -invariant  $\sigma$ -algebra) is also a factor of a.e. of its ergodic components. So a double use of Lemma 2.15 together with Theorem 10.18 from [25] gives that  $\mathcal{A}$  is also distal.

- ZE: zero entropy automorphisms;
- $\text{RIG}_{(q_n)}$ : automorphisms with a fixed sequence  $(q_n)$  of rigidity;
- multipliers  $\mathcal{M}(\mathcal{D}^\perp)$  of a class  $\mathcal{D}^\perp$  ( $\mathcal{D}$  is **any** class of automorphisms and by  $\mathcal{D}^\perp$  we mean the set of automorphisms disjoint from all systems from  $\mathcal{D}$ , and by  $\mathcal{M}(\mathcal{D}^\perp)$  we mean the set of systems whose all joinings with any element of  $\mathcal{D}^\perp$  remain in  $\mathcal{D}^\perp$ ); interesting classes of multipliers arise e.g. for  $\mathcal{D}$ =all weakly mixing (cf. Proposition 5.1 in [38]) or all mixing automorphisms; see [24], [38].
- the class of factors of all infinite self-joinings of a fixed automorphism  $R$  (the smallest characteristic class containing  $R$ ); especially in case of MSJ and simple automorphisms (cf. [25], Chapter 12). Characteristic classes of such type were used in [40].

Note also that the intersection of any family of characteristic classes yields again a characteristic class. In Section 2.3, we will show that each characteristic class  $\mathcal{F}$  determines another characteristic class  $\mathcal{F}_{\text{ec}}$  consisting of those automorphisms whose ergodic components are in  $\mathcal{F}$ .

## 2.2 The smallest nontrivial and the largest proper characteristic class

An obvious observation has been made already in the introduction that the family ALL of all automorphisms is the largest characteristic class, while the one-element  $\{*\}$  family (which is the one-point space automorphism) is the smallest characteristic class. It is more interesting however that the smallest non-trivial and the largest proper characteristic classes exist.

**Proposition 2.3.** *ID is the smallest non-trivial characteristic class.*

*Proof.* Let us first notice that the system  $([0, 1], \text{Leb}, \text{Id})$  has **all** other identities as factors. Indeed, any standard Borel probability space is determined by a sequence  $(t_i)_{i \geq 0}$  of non-negative numbers such that  $\sum_{i \geq 0} t_i = 1$  and  $t_0$  corresponds to the mass of the continuous part and  $t_1, t_2, \dots$  correspond to the masses of atoms. Then, take the corresponding partition of  $[0, 1]$  into intervals  $I_i$  of length  $t_i$  and, for each  $i \geq 1$ , the factor map will glue points in  $I_i$ .

Now, notice that any non-trivial characteristic class  $\mathcal{F}$  contains a non-ergodic automorphism. Indeed, suppose that  $T$  is ergodic, acting on a non-trivial space  $(Y, \nu)$ . Since  $Y$  is non-trivial and so is  $T$ , the graph joinings  $\Delta_{\text{Id}}$  and  $\Delta_T$  are ergodic and different, so any non-trivial convex combination of them yields a non-ergodic member of  $\mathcal{F}$ . It follows that by taking the factor  $\mathcal{I}_\nu$  of (a.e.)  $T$ -invariant sets (which belongs to  $\mathcal{F}$ ), we obtain the identity on a non-trivial standard Borel probability space  $(\bar{Y}, \bar{\nu})$ . But then

the infinite Cartesian product  $(\bar{Y}^{\times\infty}, \bar{\nu}^{\otimes\infty})$  is also in  $\mathcal{F}$  and this infinite product is isomorphic to  $([0, 1], \text{Leb})$ , which completes the proof.  $\square$

In order to prove the existence of the largest characteristic (proper) class, we need to recall some results.

**Theorem 2.4** (non-ergodic Sinai's factor theorem [36, 48]). *Assume that  $R$  is an automorphism of  $(Z, \mathcal{D}, \rho)$  and let  $\rho = \int_{X/\mathcal{I}_\rho} \rho_{\bar{x}} dm(\bar{x})$  stand for the ergodic decomposition of  $\rho$ . Assume that*

$$m - \text{essinf}_{\bar{x} \in X/\mathcal{I}_\rho} h(\rho_{\bar{x}}, R) \geq \alpha > 0.$$

*Then a Bernoulli automorphism of entropy  $\alpha$  is a factor of  $R$ .*

In [48] (see Theorem 4.3 therein) the above result is attributed to Kieffer and Rahe [36], see also [47] p.2 (The non-ergodic factor theorem).

We also need the following well-known result (we include its proof for completeness).

**Proposition 2.5.** *Each automorphism  $R$  is a factor of a self-joining of the (infinite entropy) Bernoulli automorphism  $([0, 1]^{\mathbb{Z}}, \text{Leb}^{\otimes\mathbb{Z}}, S)$ .*

**Remark 2.6.** Before we prove the above result, let us notice that any automorphism  $T$  of  $(X, \mathcal{B}, \mu)$  has an isomorphic copy in the space  $([0, 1]^{\mathbb{Z}}, \kappa, S)$ .<sup>9</sup> Consider first the aperiodic part of  $T$  which is realized on a standard Borel space. This space is isomorphic to  $[0, 1]$ , via a Borel isomorphism  $I$ . It follows that the distribution  $\mu'$  of the process  $(I \circ S^k)_{k \in \mathbb{Z}}$  yields a realization of the aperiodic part. Now,  $\mu'$  takes measure zero on the set of periodic points for the shift. Moreover, the set of periodic points of  $S$  can be identified with a subset of  $[0, 1]$ , points of period 2 with a subset of  $[0, 1]^2$  etc., and we can easily settle an isomorphism of the fixed point subspace for  $T$  with a subset of  $[0, 1]$ , period 2 points with a subset of  $[0, 1]^2$ , etc. Thus, it suffices to take  $\kappa$  equal to the sum of  $\mu'$  and the relevant atomic measures corresponding to the periodic points.

*Proof of Proposition 2.5.* Fix any automorphism  $R$  of  $(Z, \mathcal{D}, \rho)$  and take its isomorphic copy in the space  $([0, 1]^{\mathbb{Z}}, \kappa, S)$ . Take the product space  $([0, 1]^{\mathbb{Z}} \times [0, 1]^{\mathbb{Z}}, \text{Leb}^{\otimes\mathbb{Z}} \otimes \kappa)$  and consider the map  $\psi: [0, 1]^{\mathbb{Z}} \times [0, 1]^{\mathbb{Z}} \rightarrow [0, 1]^{\mathbb{Z}}$  given by

$$(x_n, y_n) \mapsto (x_n + y_n).$$

Then  $\psi_*(\text{Leb}^{\otimes\mathbb{Z}} \otimes \kappa) = \text{Leb}^{\otimes\mathbb{Z}}$  and clearly the join of the  $\sigma$ -algebra of the first coordinate and of  $\psi^{-1}(\mathcal{B}([0, 1]^{\otimes\mathbb{Z}}))$  is the product  $\sigma$ -algebra in  $[0, 1]^{\otimes\mathbb{Z}} \times [0, 1]^{\otimes\mathbb{Z}}$ . The result follows.  $\square$

We now prove the following.

---

<sup>9</sup>The same arguments apply to  $\mathbb{D} = \{z \in \mathbb{C} : |z| \leq 1\}$  instead of  $[0, 1]$ .

**Lemma 2.7.** *Assume that  $\mathcal{F}$  is a characteristic class such that  $\mathcal{F} \setminus \text{ZE} \neq \emptyset$ . Then  $\mathcal{F} = \text{ALL}$ .*

*Proof.* Fix  $T \in \mathcal{F} \setminus \text{ZE}$ . Because of Proposition 2.5, we only need to prove that the infinite entropy Bernoulli automorphism is in  $\mathcal{F}$ . The first step is to consider the factor of  $T$  that arises by gluing together the periodic part and the ergodic components from the aperiodic part whose entropy is smaller than  $\alpha = h(T)$ . Clearly, this factor remains in  $\mathcal{F} \setminus \text{ZE}$ . Moreover, in its ergodic decomposition we have a single point and the remaining part (which may still be non-ergodic) consists of ergodic components of entropy at least  $\alpha$ . By Theorem 2.4, it follows that as a further factor  $R \in \mathcal{F} \setminus \text{ZE}$  we can obtain a non-ergodic automorphism with two ergodic components: one of them is a Bernoulli of entropy  $\alpha$  and the other one is a fixed point. Finally, we take the infinite Cartesian product  $R^{\times\infty}$ . It is not hard to see that a.e. ergodic component of this automorphism is a Bernoulli with infinite entropy. Using once more Sinai's theorem (Theorem 2.4), we obtain that a Bernoulli with infinite entropy belongs to  $\mathcal{F}$  which completes the proof.  $\square$

Now, using the lemma we obtain the following.

**Proposition 2.8.**  *$\text{ZE}$  is the largest proper characteristic class.*  $\square$

### 2.3 Characteristic classes given by ergodic components

Assume that  $\mathcal{F}$  is a characteristic class. By  $\mathcal{F}_{\text{ec}}$  we denote the class of those automorphisms  $R$  such that (a.e.) ergodic components of  $R$  are in  $\mathcal{F}$  (or more precisely, in  $\mathcal{F} \cap \text{Erg}$ , where  $\text{Erg}$  stands for the family of all ergodic automorphisms). Note that we have

$$(20) \quad \mathcal{F} \cap \text{Erg} = \mathcal{F}_{\text{ec}} \cap \text{Erg}.$$

**Lemma 2.9.**  *$\mathcal{F}_{\text{ec}}$  is a characteristic class.*

*Proof.* The proof has two parts: we need to show that  $\mathcal{F}_{\text{ec}}$  is closed under taking factors and joinings.

**Factors** Let  $R$  acting on  $(Z, \mathcal{D}, \kappa)$  belong to  $\mathcal{F}_{\text{ec}}$  and fix a factor  $\mathcal{A} \subset \mathcal{D}$  of  $R$ . Let  $\kappa = \int \kappa_{\bar{x}} dP(\bar{x})$  denote the ergodic decomposition of  $\kappa$ . Since the ergodic components  $\kappa_{\bar{x}}$  are  $R$ -invariant measures,  $\mathcal{A}$  (being an  $R$ -invariant sub- $\sigma$ -algebra) is also a factor of the automorphism  $(Z, \kappa_{\bar{x}}, R)$  and  $\kappa|_{\mathcal{A}} = \int \kappa_{\bar{x}}|_{\mathcal{A}} dP(\bar{x})$  is the ergodic decomposition of  $\kappa|_{\mathcal{A}}$ . It follows that the ergodic components of the factor are factors of ergodic components of  $R$ , and since  $R \in \mathcal{F}_{\text{ec}}$ ,  $(\kappa_{\bar{x}}, R) \in \mathcal{F}$ , so also  $(\kappa_{\bar{x}}|_{\mathcal{A}}, R|_{\mathcal{A}}) \in \mathcal{F}$  for  $P$ -a.e.  $\bar{x}$ .

**Joinings** Take  $(X, \mu, T)$  and  $(Y, \nu, S)$  from  $\mathcal{F}_{ec}$  and let  $\rho \in J(T, S)$  be their joining. Let

$$\rho = \int_0^1 \rho_t dP(t), \quad \mu = \int_0^1 \mu_t dQ(t), \quad \nu = \int_0^1 \nu_t dR(t)$$

be the relevant ergodic decompositions. Then

$$\int_0^1 \mu_t dQ(t) = \mu = \rho|_X = \int_0^1 \rho_t|_X dP(t),$$

so since  $\rho_t|_X$  are also ergodic, these two decompositions are the same. So for a  $P$ -“typical”  $t \in [0, 1]$ , the projection of  $\rho_t$  on  $X$  is an ergodic component of  $T$ . The same argument applies on the coordinate  $Y$  and we see that the ergodic components of  $\rho$  are joinings of ergodic components of  $\mu$  and  $\nu$ . It follows that  $(X \times Y, \rho, T \times S) \in \mathcal{F}_{ec}$ . The argument extends to countable joinings.  $\square$

### 2.3.1 Largest characteristic factor

**ID, ZE, DISP and RIG $_{(q_n)}$**  Given a characteristic class  $\mathcal{F}$ , according to Proposition 2.1, each automorphism  $R$  acting on  $(Z, \mathcal{D}, \kappa)$  has a largest  $\mathcal{F}$ -factor  $\mathcal{D}_{\mathcal{F}} \subset \mathcal{D}$ . Often, its description is classical:

- the  $\sigma$ -algebra of invariant sets for  $\mathcal{F} = \text{ID}$ ,
- the Pinsker factor for  $\mathcal{F} = \text{ZE}$ ,
- the Kronecker factor for  $\mathcal{F} = \text{DISP}$ ,
- the largest factor for which  $(q_n)$  is a rigidity time for  $\mathcal{F} = \text{RIG}_{(q_n)}$ .

**DISP $_{ec}$**  We will comment now on  $\mathcal{D}_{\mathcal{F}_{ec}}$  when  $\mathcal{F} = \text{DISP}$ , cf. Proposition 2.12 (ii) and its connections with the theory of nil-factors. Most of the material presented below is known to aficionados but not necessarily the material is explicitly present in the literature. Our discussion is based on [18], [23], [30] and [31]. We provide some details to explain clearly why the problem of whether  $\mu \perp \text{DISP}_{ec}$  is open, cf. Corollary 1.10, Corollary 2.11 and Remark 5.9.

Recall that according to the Furstenberg-Zimmer theory [23], given  $R$  on  $(Z, \mathcal{D}, \kappa)$  and a factor  $\mathcal{C} \subset \mathcal{D}$ , there exists a certain intermediate factor

$$\mathcal{C} \subset \mathcal{K} = \mathcal{K}(\mathcal{C}) \subset \mathcal{D},$$

called the *relative Kronecker factor* (with respect to  $\mathcal{C}$ ). It is the largest intermediate factor with the following property (see condition C<sub>2</sub> in [23],

p. 131):

there exists a dense set of functions  $F \in L^2(\mathcal{K}, \kappa|_{\mathcal{K}})$  such that for each  $\delta > 0$  there is a finite set  $g_1, \dots, g_k \in L^2(\mathcal{K}, \kappa|_{\mathcal{K}})$  such that for each  $h \in \mathbb{Z}$ ,

$$(21) \quad \min_{1 \leq j \leq k} \|F \circ R^h - g_j\|_{L^2(\kappa_y)} < \delta$$

for a.e.  $y \in Z/\mathcal{C}$ , where

$$(22) \quad \kappa|_{\mathcal{K}} = \int_{Z/\mathcal{C}} \kappa_y d\kappa(y).$$

Whenever condition (21) holds, we speak of *relative compactness* or of *relatively discrete spectrum* of the intermediate factor over  $\mathcal{C}$ .

A particular situation arises when  $\mathcal{C} = \mathcal{I}_\kappa$ , i.e. it is the  $\sigma$ -algebra of invariant sets. Then (22) is nothing but the ergodic decomposition of  $\kappa$  and the conditional measures  $\kappa_y$  are also  $R$ -invariant. In this case condition (21) yields in a.e. fiber  $\pi^{-1}(y)$  (where  $\pi: Z/\mathcal{K} \rightarrow Z/\mathcal{I}_\kappa$  stands for the factor map) a dense set of functions  $F|_{\pi^{-1}(y)}$  in  $L^2(\mathcal{K}, \kappa_y)$  whose orbits under the unitary action of  $R$  are relatively compact. It follows that the (ergodic) automorphism  $(R, \kappa_y)$  has discrete spectrum for a.e.  $y \in Z/\mathcal{I}_\kappa$ . In other words,

$$\mathcal{K}(\mathcal{I}_\kappa) \subset \mathcal{D}_{\text{DISP}_{\text{ec}}}$$

In fact, the opposite inclusion is also true, i.e.

$$(23) \quad \mathcal{D}_{\text{DISP}_{\text{ec}}} = \mathcal{K}(\mathcal{I}_\kappa),$$

that is,  $\mathcal{A} := \mathcal{D}_{\text{DISP}_{\text{ec}}}$  has relatively discrete spectrum over  $\mathcal{I}_\kappa$ . Indeed, by the definition of  $\mathcal{A}$ , a.e. ergodic component of  $R|_{\mathcal{A}}$  has discrete spectrum. Fix  $F \in L^\infty(\mathcal{A}, \kappa|_{\mathcal{A}})$ . Fix also  $\varepsilon, \delta > 0$  and  $k \geq 1$ . Consider the set  $W_k \subset Z/\mathcal{I}_\kappa$  of those  $y$  for which

$$\min_{-k \leq j \leq k} \|F \circ R^n - F \circ R^j\|_{L^2(\kappa_y)} < \varepsilon$$

for each  $n \in \mathbb{Z}$ . Since on each fiber  $R$  is an ergodic automorphism with discrete spectrum, the measure of  $W_k$  goes to 1, when  $k \rightarrow \infty$ , so it will be greater than  $1 - \delta$  for  $k$  large enough. It follows that the function  $F$  is compact as it has been defined in the proof of Theorem 6.15 [23]. Therefore,  $F \in L^2(\mathcal{K}(\mathcal{I}_\kappa))$ , which (by [23]) concludes the proof of (23).

**Remark 2.10.** As a matter of fact, in [5], the Furstenberg-Zimmer theory is developed without assuming ergodicity (cf. e.g. Proposition 5.7 therein to obtain the equality  $\mathcal{D}_{\text{DISP}_{\text{ec}}} = \mathcal{K}(\mathcal{I}_\kappa)$ ).

We will now see that  $\mathcal{D}_{\text{DISP}_{\text{ec}}}$  appears naturally in the classical theory of characteristic nil-factors [30, 31].<sup>10</sup> Recall that if  $R$  acting on  $(Z, \mathcal{D}, \kappa)$  is ergodic then for a function  $f \in L^\infty(Z, \mathcal{D}, \kappa)$  its  $u^s$  norms (in fact, seminorms) are defined in the following way:

$$(24) \quad \|f\|_{u^1} := \left| \int f d\kappa \right|,$$

$$(25) \quad \|f\|_{u^{s+1}}^{2^{s+1}} := \lim_{H \rightarrow \infty} \frac{1}{H} \sum_{h \leq H} \|f \circ R^h \cdot \bar{f}\|_{u^s}^{2^s}.$$

If  $R$  is non-ergodic then instead of (24), we put

$$\|f\|_{u^1}^2 := \lim_{H \rightarrow \infty} \frac{1}{H} \sum_{h \leq H} \int f \circ R^h \cdot \bar{f} d\kappa$$

and (25) remains unchanged. Then, by [30, 31], for each  $s \geq 1$  there is a special factor  $\mathcal{Z}_s = \mathcal{Z}_s(R) \subset \mathcal{B}$ , namely, the largest factor whose

$$(26) \quad \text{a.e. ergodic component is an inverse limit of } s\text{-step nil-systems.}$$

In other words,  $\mathcal{Z}_s(R)$  is the largest (characteristic)  $\text{NIL}_s$ -factor of  $R$ . Moreover (see Proposition 7 (page 138) and Proposition 13 (page 141) in [31]),

$$(27) \quad \|f\|_{u^{s+1}} = 0 \iff f \perp L^2(\mathcal{Z}_s) \iff f \perp L^2(\mathcal{Z}_s(R, \kappa_y)) \text{ for } \kappa\text{-a.e. } y.$$

A special case arises when our measure-preserving systems are Furstenberg systems of a bounded  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$ . As in (for example) [16], see Sections 2.4 and 2.5 therein, one can introduce the uniformity norms (along subsequences of intervals) for  $\mathbf{u}$ . The definitions are given in (14) and (15). They are very similar to those (in the non-ergodic case) to the definitions for functions.

We will now show that

$$(28) \quad \mathcal{Z}_1(R) = \mathcal{K}(\mathcal{I}_\kappa).$$

If  $R$  is ergodic then the above means just that

$$(29) \quad \mathcal{Z}_1 \text{ is the Kronecker factor of } R.$$

To see that (29) indeed holds, notice that (27) for  $s = 1$  yields

$$\|f\|_{\mathbf{u}^2} = 0 \iff f \perp L^2(\mathcal{Z}_1)$$

---

<sup>10</sup>We would like to thank Bryna Kra and Nikos Frantzikinakis for fruitful discussions and useful references on this subject.

and it remains to notice that (using the Wiener lemma)

$$\begin{aligned} \|f\|_{\mathbf{u}^2}^4 &= \lim_{H \rightarrow \infty} \frac{1}{H} \sum_{h \leq H} \|f \circ R^h \cdot \bar{f}\|_{\mathbf{u}^1}^2 = \\ &= \lim_{H \rightarrow \infty} \frac{1}{H} \sum_{h \leq H} \left| \int f \circ R^h \cdot \bar{f} \right|^2 d\kappa \rightarrow \sum_{z \in \mathbb{T}} \sigma_f(\{z\})^2, \end{aligned}$$

where  $\sigma_f$  stands for the spectral measure of  $f$ .

Let us return to a possibly non-ergodic  $R$ . The inclusion  $\mathcal{Z}_1(R) \subset \mathcal{K} := \mathcal{K}(\mathcal{I}_R)$  follows directly by Theorem 5.2 in [18]. To obtain the opposite inclusion, one can argue in the following way. Suppose that  $f \perp L^2(\mathcal{Z}_1(R))$  and  $|f| \leq 1$ . Take  $g \in L^2(\mathcal{K})$ , cf. (21) with  $F = g$ . We want to show that  $\int fg d\kappa = 0$ . Notice that

$$\int fg d\kappa = \int \left( \frac{1}{N} \sum_{n \leq N} \int f \circ T^n \cdot g \circ T^n d\kappa_y \right) d\kappa(y).$$

Let  $g_j$ ,  $1 \leq j \leq k$ , be as in (21). Then

$$\begin{aligned} & \left| \frac{1}{N} \sum_{n \leq N} \int f \circ T^n \cdot g \circ T^n d\kappa_y \right| \\ & \leq \left| \sum_{1 \leq j \leq k} \frac{1}{N} \sum_{n \leq N} \int f \circ T^n \cdot g_j d\kappa_y \right| + \frac{1}{N} \sum_{n \leq N} \min_{1 \leq i \leq k} \int |f \circ T^n (g \circ T^n - g_i)| d\kappa_y. \end{aligned}$$

Each term in the average in the second summand is bounded by  $\delta$ . Moreover,

$$(30) \quad \frac{1}{N} \sum_{n \leq N} \left| \int f \circ T^n \cdot g_j d\kappa_y \right|^2 \rightarrow \sum_{z \in \mathbb{T}} \sigma_{f, g_j, \kappa_y}(\{z\})^2,$$

where  $\sigma_{f, g_j, \kappa_y}$  stands for the spectral measure of the pair  $f, g_j$  (on the ergodic component  $(\pi^{-1}(y), \kappa_y)$  given by  $y$ ). But by (27) and (29), we have

$$\begin{aligned} f \perp L^2(\mathcal{Z}_1) &\iff f \perp L^2(\mathcal{Z}_1(R, \kappa_y)) \text{ for a.e. } y \\ &\iff \sigma_{f, \kappa_y} \text{ is continuous for a.e. } y \end{aligned}$$

( $\sigma_{f, \kappa_y}$  stands for the spectral measure of  $f$  on the ergodic component  $(\pi^{-1}(y), \kappa_y)$  given by  $y$ ). Since  $f \perp L^2(\mathcal{Z}_1)$  and  $\sigma_{f, g_i, \kappa_y} \ll \sigma_{f, \kappa_y}$ , it remains to use the classical equivalence

$$\frac{1}{N} \sum_{n \leq N} a_n \rightarrow 0 \iff \frac{1}{N} \sum_{n \leq N} a_n^2 \rightarrow 0$$

for any bounded sequence  $(a_n) \subset [0, \infty)$ , to conclude that the limit in (30) is equal to zero. Thus  $f \perp L^2(\mathcal{Z}_1) \implies f \perp L^2(\mathcal{K})$ .



Finally, let us compare the above with the notion of relative weak mixing. Recall that relative weak mixing over  $\mathcal{I}_\kappa$  for  $f$  means that

$$\frac{1}{H} \sum_{h \leq H} \int \left| \mathbb{E}(f \circ R^h \cdot \bar{f} | \mathcal{I}_\kappa) \right|^2 d\kappa \rightarrow 0.$$

Moreover,

$$\frac{1}{H} \sum_{h \leq H} \int \left| \mathbb{E}(f \circ R^h \cdot \bar{f} | \mathcal{I}_\kappa) \right|^2 d\kappa = \int \left( \frac{1}{H} \sum_{h \leq H} \left| \int f \circ R^h \cdot \bar{f} d\kappa_y \right|^2 \right) d\kappa,$$

and, once more by the Wiener lemma,

$$\frac{1}{H} \sum_{h \leq H} \left| \int f \circ R^h \cdot \bar{f} d\kappa_y \right|^2 \rightarrow \sum_{z \in \mathbb{T}} \sigma_{f, \kappa_y}(\{z\})^2.$$

It follows immediately that  $\sigma_{f, \kappa_y}$  is continuous for a.e.  $y$  if and only if  $f$  is relatively weakly mixing over  $\mathcal{I}_\kappa$ .

The above discussion can be summarized in the following statement.

**Corollary 2.11.** *Let  $(Z, \mathcal{D}, \kappa, R)$  be a measure-theoretic dynamical system and let  $f \in L^2(Z, \mathcal{D}, \kappa)$ . The following conditions are equivalent:*

- (i)  $f \perp L^2(\mathcal{Z}_1)$ ,
- (ii)  $f \perp L^2(\mathcal{D}_{\text{DISP}_{\text{ec}}})$ ,
- (iii)  $f \perp L^2(\mathcal{K}(\mathcal{I}_\kappa))$ ,
- (iv)  $\sigma_{f, \kappa_y}$  is continuous for  $\kappa$ -a.e.  $y$ ,
- (v)  $f$  is relatively weakly mixing over  $\mathcal{I}_\kappa$ .

### 2.3.2 A class vs. its ec-class

Let us continue our observations on the relations between characteristic classes and the corresponding ec-classes. Note that in general there are no relations between  $\mathcal{F}$  and  $\mathcal{F}_{\text{ec}}$ :

**Proposition 2.12.** *We have:*

- (i)  $\text{ZE} = \text{ZE}_{\text{ec}}$ ,  $\text{ALL} = \text{ALL}_{\text{ec}}$ ,  $\text{ID} = \text{ID}_{\text{ec}}$ ,  $\text{NIL}_s = (\text{NIL}_s)_{\text{ec}}$ ,  $\{*\} \subsetneq \{*\}_{\text{ec}}$ ;
- (ii)  $\text{DISP} \subsetneq \text{DISP}_{\text{ec}}$ ;
- (iii)  $\text{RDISP} = \text{RDISP}_{\text{ec}}$ ;
- (iv)  $\text{DIST} = \text{DIST}_{\text{ec}}$ ;
- (v)  $\left( \text{RIG}_{(q_n)} \right)_{\text{ec}} \subsetneq \text{RIG}_{(q_n)}$ .

*Proof of (i)-(iii).*

(i) The first claim follows from the fact that the entropy function is convex, the other claims are obvious.

(ii) If an automorphism has discrete spectrum then its  $L^2$ -space is generated by eigenfunctions. The restrictions (if non-zero) of these (global) eigenfunctions yield orthonormal bases in  $L^2$ -spaces of ergodic components. The inclusion is strict since  $(x, y) \mapsto (x, x + y)$  (on  $\mathbb{T}^2$ , considered with Lebesgue measure) does not have discrete spectrum while the ergodic components do.

(iii) We want to show that if each ergodic component has rational discrete spectrum then the whole automorphism has. Given  $p/q \in \mathbb{Q}$  and an ergodic component  $c$ , we choose  $f_c$  a modulus 1 eigenfunction corresponding to the eigenvalue  $e^{2\pi ip/q}$ . Since  $f_c$  is unique up to a constant of modulus 1, this choice can be done measurably. In this way, we will create global eigenfunctions.<sup>11</sup>  $\square$

Before we give the proof of (iv), we need to recall some more notions and facts from the relative ergodic theory, e.g. [23, 25]. Given an automorphism  $T$  of  $(X, \mathcal{B}, \mu)$  and its factor  $S$  on  $(Y, \mathcal{C}, \nu)$  with the factor map  $\pi: X \rightarrow Y$ , we say that this extension is *relatively ergodic* (rel. erg.) if each  $f \in L^1(X, \mathcal{B}, \mu)$  satisfying  $f \circ T = f$  ( $\mu$ -a.e.) is  $\pi^{-1}(\mathcal{C})$ -measurable. It follows immediately from the definition that:

- any composition of relatively ergodic extensions remains relatively ergodic;
- an inverse limit of relatively ergodic extensions remains relatively ergodic (as the conditional expectation, with respect to a factor, of an invariant function remains invariant);
- $\bar{\pi}: Y \rightarrow \bar{Y} := Y/\mathcal{I}_\nu$ , where  $(\bar{Y}, \bar{\nu})$  stands for the space of ergodic components (on which acts the identity map), is relatively ergodic.

Let  $\mu = \int_Y \mu_y d\nu(y)$  stand for the disintegration of  $\mu$  over  $\nu$  and let

$$\nu = \int_{\bar{Y}} \bar{\nu}_{\bar{y}} d\bar{\nu}$$

denote the ergodic decomposition of  $\nu$  (which is precisely the disintegration of  $\nu$  over  $\bar{\nu}$ ). Then the ergodic components of  $S$  acting on  $Y$  are of the form

$$(\bar{\pi}^{-1}(\bar{y}), \bar{\nu}_{\bar{y}}, S)$$

(the measures  $\bar{\nu}_{\bar{y}}$  are  $S$ -invariant). Therefore, we have the following lemma.

---

<sup>11</sup>The same argument works if we consider the characteristic class of automorphisms having discrete spectrum contained in a **fixed** countable subgroup of the circle.

**Lemma 2.13.** *If  $T$  is relatively ergodic over  $S$  then the ergodic components of  $T$  are of the form*

$$\left( \pi^{-1}(\bar{\pi}^{-1}(\bar{y})), \int_{\bar{\pi}^{-1}(\bar{y})} \mu_y d\bar{\nu}_{\bar{y}}(y) \right).$$

Note that it follows that the ergodic components of  $T$  have as their factors (via the relevant restriction of  $\pi$ ) ergodic components of  $S$ , and that the spaces of ergodic components of  $T$  and  $S$  are the same (i.e.  $\bar{X} = \bar{Y}$ ).

**Lemma 2.14.** *Let  $T$  be relatively ergodic over  $S$ . Then the following are equivalent:*

- (a)  $T$  over  $S$  has relatively discrete spectrum.
- (b) The ergodic components of  $T$  have relatively discrete spectrum over the ergodic components of  $S$  being their relevant factors.

*Proof.* By Lemma 2.13, we see that the disintegration of an ergodic component  $\pi^{-1}(\bar{\pi}^{-1}(\bar{y}))$  over  $\bar{\pi}^{-1}(\bar{y})$  (which is its factor) consists of the same conditional measures  $\mu_y$  as the total disintegration of  $\mu$  over  $\nu$ . We proceed now as in the proof of the equality  $\mathcal{K}(I_\kappa) = \mathcal{D}_{\text{DISP}_{\text{ec}}}$  (page 21), showing compactness.  $\square$

Recall that an automorphism  $T$  is *distal* if it is a limit of a transfinite (indexed by countable ordinals) sequence of consecutive maximal Kronecker extensions (if an ordinal is not isolated, we pass to the corresponding inverse limit). Note that, by the very definition, the  $\sigma$ -algebra  $\text{Inv}$  is contained in **the** Kronecker factor of  $T$ , so in this transfinite chain of consecutive extensions, all but (perhaps) the first one are relatively ergodic. By applying Lemma 2.14 and transfinite induction, we obtain the following.

**Lemma 2.15.**  *$T$  is distal if and only if all its ergodic components are distal.*

*Proof of (iv)-(v).*

(iv) This follows directly from Lemma 2.15.

(v) It is clear that if  $(q_n)$  is a rigidity time for an a.e. ergodic component, it is also a rigidity time for the whole automorphisms. Not vice versa however (for  $(q_n)$  sufficiently sparsed). We will provide a relevant construction below.  $\square$

**RIG<sub>ec</sub> is a proper subclass of RIG** Let us first notice that we only need to construct a continuous measure  $\sigma$  on the circle such that

$$(31) \quad e^{2\pi i q_n} \rightarrow 1 \text{ in measure } \sigma \text{ but not } \sigma - \text{a.e.}$$

Indeed, suppose (31) holds, and consider on  $\mathbb{T}^2$  the automorphism

$$T(x, y) = (x, y + x) \text{ with measure } \sigma \otimes \text{Leb}.$$

If  $F(x, y) = f(x)e^{2\pi i \ell y}$  then by (31),

$$\int |F(T^{q_n}(x, y)) - F(x, y)| d\sigma(x)dy = \int |f(x)||e^{2\pi i q_n \ell x} - 1| d\sigma(x) \rightarrow 0$$

when  $n \rightarrow \infty$ . On the other hand, the rotation by  $x$  on an ergodic component  $\{x\} \times \mathbb{T}$  has  $(q_n)$  as its rigidity time if and only if  $q_n x \rightarrow 0 \pmod{1}$ . This is not true for  $\sigma$ -a.e.  $x \in \mathbb{T}$  in view of (31).

We now sketch how to construct such a measure assuming that  $(q_n)$  is sufficiently sparsed. Fix  $0 < p_n < 1$  so that  $p_n$  is decreasing to zero and  $\sum_{n \geq 1} p_n = \infty$ . Set  $f_n(x) = \{q_n x\}$ . We intend to construct a Cantor set (together with a Cantor measure  $\sigma$  on it). Let

$$A_n := f_n^{-1}([1/4, 3/4]), \quad B_n = f_n^{-1}([0, p_n]).$$

Our postulates are:

$$\sigma(B_n) = 1 - p_n, \quad \sigma(A_n) = p_n.$$

In fact, we need to be more precise in description of the measure at stage  $n$  to be able to continue its definition. So at stage  $n$  the circle is divided into intervals of the form  $[\frac{j}{q_n}, \frac{j+1}{q_n})$  (many of such intervals are of measure  $\sigma$  equal to zero). We now require that the conditional measures satisfy:

$$(32) \quad \sigma\left(B_n \Big| \left[\frac{j}{q_n}, \frac{j+1}{q_n}\right)\right) = 1 - p_n, \quad \sigma\left(A_n \Big| \left[\frac{j}{q_n}, \frac{j+1}{q_n}\right)\right) = p_n$$

for each  $j = 0, \dots, q_n - 1$ ). Passing to step  $n + 1$ , we require that all the intervals  $[\frac{j}{q_n}, \frac{j+1}{q_n})$  contain at least two intervals of the form  $[\frac{k}{q_{n+1}}, \frac{k+1}{q_{n+1}})$ , we choose two of such (of course only in those  $[\frac{j}{q_n}, \frac{j+1}{q_n})$  which are of positive measure  $\sigma$ ) and apply the rule (32) to  $A_{n+1}, B_{n+1}$  with  $p_n$  replaced with  $p_{n+1}$ .

Note that  $\int e^{2\pi i q_n x} d\sigma(x) = 1 + O(p_n(1 + p_n) + 1 \cdot p_n)$ , so  $e^{2\pi i q_n \cdot} \rightarrow 1$  in measure  $\sigma$ . On the other hand  $\sigma(A_n) = p_n$  and the sets  $A_n$  are almost independent. Since  $\sum_{n \geq 1} p_n = \infty$ , for  $\sigma$ -a.e.  $x$ , we have  $x \in A_n$  for infinitely many  $n$  (by the Borel-Cantelli lemma), so (31) holds.

### 2.3.3 Strong $u$ -MOMO property of systems whose visible measures yield systems in an ec-class

While we have seen rather unclear relations between  $\mathcal{F}$  and  $\mathcal{F}_{\text{ec}}$  (cf. Proposition 2.12), on the topological level we always have the following.

**Proposition 2.16.** *Let  $\mathcal{F}$  be a characteristic class. Then  $\mathcal{C}_{\mathcal{F}} \subset \mathcal{C}_{\mathcal{F}_{ec}}$ .*

*Proof.* This follows immediately from the fact that a homeomorphism  $T$  (acting on a compact metric space  $X$ ) belongs to  $\mathcal{C}_{\mathcal{F}_{ec}}$  if and only if for each  $\kappa \in M^e(X, T)$ ,  $(X, \mathcal{B}(X), \kappa, T) \in \mathcal{F}$ .  $\square$

Note that in view of Proposition 2.16 and Proposition 2.12,

$$(33) \quad \mathcal{C}_{\text{RIG}_{(q_n)}} = \mathcal{C}_{(\text{RIG}_{(q_n)})_{ec}}.$$

The special role of ec-classes stands in the next proposition.

**Proposition 2.17.** *Let  $\mathcal{F}$  be a characteristic class. Then  $\mathbf{u} \perp \mathcal{C}_{\mathcal{F}_{ec}}$  if and only if each element in  $\mathcal{C}_{\mathcal{F}_{ec}}$  satisfies the strong  $\mathbf{u}$ -MOMO property.*

The below proof of Proposition 2.17 is an adaptation of the proof of Corollary 9 in [3]. It uses the following elementary result (see Lemma 18 in [3]).

**Lemma 2.18.** *Assume that  $(c_n) \subset \mathbb{C}$  and  $(m_n) \subset \mathbb{N}$ . Then if the sequence  $(c_n)$  is contained in a closed convex cone which is not a half-plane then*

$$\frac{1}{m_N} \sum_{n \leq N} c_n \rightarrow 0 \iff \frac{1}{m_N} \sum_{n \leq N} |c_n| \rightarrow 0 \text{ as } N \rightarrow \infty.$$

*Proof of Proposition 2.17.* Only one implication needs to be proved. Suppose that  $\mathbf{u} \perp \mathcal{C}_{\mathcal{F}_{ec}}$ , and let  $(Y, S) \in \mathcal{C}_{\mathcal{F}_{ec}}$ . We fix  $f \in C(Y)$ , an increasing sequence  $(b_k)$  in  $\mathbb{N}$ , with  $b_1 = 1$  and  $b_{k+1} - b_k \rightarrow \infty$ , and a sequence  $(y_k)$  of points in  $Y$ . We introduce the finite set  $\mathbb{A} := \{1, e^{2\pi i/3}, e^{4\pi i/3}\}$ , and for each  $k \geq 1$ , we define  $e_k \in \mathbb{A}$  such that the complex number

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

is in the closed convex cone  $\{0\} \cup \{z \in \mathbb{C}^* : \arg(z) \in [-\pi/3, \pi/3]\}$ . Then, by Lemma 2.18, the convergence that we need to prove, i.e.

$$\frac{1}{b_K} \sum_{k < K} \left| \sum_{b_k \leq n < b_{k+1}} f(S^{n-b_k} y_k) \mathbf{u}(n) \right| \xrightarrow{K \rightarrow \infty} 0$$

is equivalent to the convergence

$$(34) \quad \frac{1}{b_K} \sum_{k < K} \sum_{b_k \leq n < b_{k+1}} e_k f(S^{n-b_k} y_k) \mathbf{u}(n) \xrightarrow{K \rightarrow \infty} 0.$$

Consider the dynamical system  $(X, T)$ , where  $X := (Y \times \mathbb{A})^{\mathbb{Z}}$ , and  $T$  is the left shift. Let  $x \in X$  be such that

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

and  $x_n = x_0$  is a fixed arbitrary point of  $Y \times \mathbb{A}$  for  $n \leq 0$ .

By setting  $F := f \otimes \text{Id}$  on  $Y \times \mathbb{A}$ , it easily follows that (34) amounts to

$$\frac{1}{b_K} \sum_{k < K} \sum_{b_k \leq n < b_{k+1}} F \circ \pi_0(T^n x) \mathbf{u}(n) \xrightarrow{K \rightarrow \infty} 0.$$

To prove the convergence above, we define the subspace  $X_x$  as the closure of  $\{T^n x : n \in \mathbb{Z}\}$ . By assumption on  $\mathbf{u}$ , we only have to check that the system  $(X_x, T)$  is in  $\mathcal{C}_{\mathcal{F}_{ec}}$ . So let  $\mu$  be a visible measure in  $(X_x, T)$ , and we first consider the case where  $x$  itself is generic for  $\mu$ , along a sequence  $(N_\ell)$ . Set

$$B := \{(v_j, a_j)_{j \in \mathbb{Z}} \in X : (v_1, a_1) = (Sv_0, a_0)\}$$

Since  $b_{k+1} - b_k \rightarrow \infty$ , we have

$$\frac{1}{N_\ell} \sum_{n < N_\ell} \delta_{T^n x}(B) \xrightarrow{\ell \rightarrow \infty} 1,$$

and since the set  $B$  is closed, by the Portmanteau theorem, it must be of full measure  $\mu$  in  $(X, T)$ . Moreover, such a measure  $\mu$  must be  $T$ -invariant, hence,

$$(35) \quad 1 = \mu \left( \bigcap_{n \in \mathbb{Z}} T^n B \right) = \mu \left( \{(v_j, a_j)_{j \in \mathbb{Z}} \in X : \forall j, (v_j, a_j) = (S^j v_0, a_0)\} \right).$$

Denote by  $\mu^{(0)}$  the restriction of  $\mu$  to the zero-coordinate (that is, with the above notation, the distribution of  $(v_0, a_0)$  under  $\mu$ ). Since  $\mu$  is  $T$ -invariant, it follows that  $\mu^{(0)}$  is  $(S \times \text{Id}_{\mathbb{A}})$ -invariant. Moreover,  $Y \times \mathbb{A}$  consists of three copies of  $Y$ , each of them is invariant under  $S \times \text{Id}_{\mathbb{A}}$ . Thus,

$$\mu^{(0)} = \alpha_0 \mu_1^{(0)} + \alpha_1 \mu_2^{(0)} + \alpha_2 \mu_3^{(0)},$$

where  $\alpha_0 + \alpha_1 + \alpha_2 = 1$ ,  $\alpha_j \geq 0$  and  $\mu_j^{(0)}(Y \times \{e^{2\pi i j/3}\}) = 1$  for  $j = 0, 1, 2$ . It follows that the ergodic components of  $(Y \times \mathbb{A}, \mu^{(0)}, S \times \text{Id}_{\mathbb{A}})$  yield measure-theoretic systems isomorphic to ergodic measures on  $(Y, S)$ , hence in  $\mathcal{F}$  since  $(Y, S) \in \mathcal{C}_{\mathcal{F}_{ec}}$  (this is the moment in our proof where we use the fact that we deal with a characteristic ec-class and not a general characteristic class  $\mathcal{F}$ ). Thus  $(Y \times \mathbb{A}, \mu^{(0)}, S \times \text{Id}_{\mathbb{A}}) \in \mathcal{F}_{ec}$ . Now, using (35), we see that  $(X_x, \mu, T)$  is isomorphic to  $(Y \times \mathbb{A}, \mu^{(0)}, S \times \text{Id}_{\mathbb{A}})$ , thus is also in  $\mathcal{F}_{ec}$ .

Now, suppose that  $\mu \in V_T(x')$  for some point  $x'$  in the orbit closure of  $x$ , say  $x' = \lim_{r \rightarrow \infty} T^{n_r} x$ .

If  $x' = T^n x$  for some  $n \in \mathbb{Z}$ , then  $\mu \in V_T(x)$  and we already know that  $(X_x, \mu, T) \in \{*\} \subset \mathcal{F}_{ec}$  in this case.

If  $n_r \rightarrow -\infty$ , then  $x' = (\dots, x_0, x_0, x_0, \dots)$  is a fixed point, and  $\mu = \delta_{x'}$ . In this case,  $(X_x, \mu, T) \in \{*\} \subset \mathcal{F}_{ec}$ .

If  $n_r \rightarrow +\infty$ , and if we write  $x' = (v_j, a_j)_{j \in \mathbb{Z}} = \lim_{r \rightarrow \infty} T^{n_r} x$ , then as  $b_{k+1} - b_k \rightarrow \infty$ , there exists at most one  $j \in \mathbb{Z}$  such that  $(v_{j+1}, a_j) \neq (Sv_j, a_j)$ . We can then use the same arguments as for  $x$  to show that a measure  $\mu$  for which  $x$  is quasi-generic satisfies  $(X_x, \mu, T) \in \mathcal{F}_{ec}$ .

We conclude that  $(X_x, T)$  is in  $\mathcal{C}_{\mathcal{F}_{ec}}$ .  $\square$

**Remark 2.19.** In general, when instead of  $\mathcal{F}_{ec}$  we consider  $\mathcal{F}$ ,  $\mathbf{u} \perp \mathcal{C}_{\mathcal{F}}$  implies the strong  $\mathbf{u}$ -MOMO property for each for  $(Y, S)$  in which all invariant measures yield systems in  $\mathcal{F}$  (in particular, if  $(Y, S) \in \mathcal{C}_{\mathcal{F}}$  and each invariant measure is visible).

**Question 1.** Is Proposition 2.17 true for **each** characteristic class?

**Remark 2.20.** A straightforward adaptation of the proof shows that the subsequence version of Proposition 2.17 also holds: for each characteristic class  $\mathcal{F}$  and each increasing sequence of integers  $(N_\ell)$ ,  $\mathbf{u}$  is  $(N_\ell)$ -orthogonal to  $\mathcal{C}_{\mathcal{F}_{ec}}$  if and only if each element in  $\mathcal{C}_{\mathcal{F}_{ec}}$  satisfies the strong  $\mathbf{u}$ -MOMO property along  $(N_\ell)$ . See Remarks 1.2 and 1.4.

### 3 Lifting lemma

The purpose of this section is to prove Proposition 3.1, which is an alternative version of Conze-Downarowicz-Serafin lifting lemma from [9] and seems to be of independent interest. It may seem weaker than the original where the genericity was lifted to a single orbit, but the main advantage here is that we do not need assumptions on the nature of the second topological space: it does not have to be a full shift. The second advantage is that the result has its extension to the logarithmic case, see Appendix 3.3, while the lifting lemma of Conze-Downarowicz-Serafin and other results of that type so far have been proved for Cesàro averages.

**Proposition 3.1.** *Let  $(Y, S)$  and  $(X, T)$  be two topological systems and  $u \in Y$  generic along an increasing sequence  $(N_m)$  for some  $S$ -invariant measure  $\kappa$  on  $Y$ . Let  $\rho$  be a joining of  $\kappa$  with a  $T$ -invariant measure  $\nu$  on  $X$ . Then there exist a sequence  $(x_n) \subset X$  and a subsequence  $(N_{m_\ell})$  such that:*

- the sequence  $(S^{n_\ell} u, x_n)$  is generic for  $\rho$  along  $(N_{m_\ell})$ :

$$\frac{1}{N_{m_\ell}} \sum_{0 \leq n < N_{m_\ell}} \delta_{(S^n u, x_n)} \xrightarrow{\ell \rightarrow \infty} \rho,$$

- The sequence  $(x_n)$  is constituted of longer and longer pieces of orbits. More precisely,  $\{n \geq 0 : x_{n+1} \neq T x_n\}$  is of the form  $\{b_1 < b_2 < \dots < b_k < b_{k+1} < \dots\}$ , where  $b_{k+1} - b_k \rightarrow \infty$ .

### 3.1 Good sequences of partitions

We need a convenient tool to estimate the weak\*-convergence of a sequence of probability measures to a given measure.

**Definition 3.1.** Let  $(E, d)$  be a compact metric space, and let  $\nu$  be a Borel probability measure on  $E$ , i.e.  $\nu \in M(E)$ . We consider a sequence  $(\mathcal{P}_\ell)$  of finite partitions of  $E$  into Borel subsets. The sequence  $(\mathcal{P}_\ell)$  is said to be *good for  $(E, \nu)$*  if the following conditions hold:

- for each  $\ell$ ,  $\mathcal{P}_{\ell+1}$  refines  $\mathcal{P}_\ell$ ,
- $\text{diam}(\mathcal{P}_\ell) := \max_{P \text{ atom of } \mathcal{P}_\ell} \text{diam}(P) \xrightarrow{\ell \rightarrow \infty} 0$ ,
- for each  $\ell$  and each atom  $P$  of  $\mathcal{P}_\ell$ ,  $\nu(\partial P) = 0$ .

The motivation for introducing this definition comes from the following result.

**Lemma 3.2.** *If  $(\mathcal{P}_\ell)$  is a good sequence of partitions for  $(E, \nu)$ , then a sequence  $(\nu_n) \subset M(E)$  converges to  $\nu$  in the weak\*-topology if and only if, for each  $\ell$  and each atom  $P$  of  $\mathcal{P}_\ell$ , we have*

$$\nu_n(P) \xrightarrow{n \rightarrow \infty} \nu(P).$$

*Proof.* If  $\nu_n \xrightarrow[n \rightarrow \infty]{w^*} \nu$ , then by the Portmanteau theorem, for each  $P \subset E$  such that  $\nu(\partial P) = 0$ , we have  $\nu_n(P) \rightarrow \nu(P)$ .

Conversely, assume that for each  $\ell$  and each atom  $P$  of  $\mathcal{P}_\ell$  we have  $\nu_n(P) \rightarrow \nu(P)$ . Then any weak\*-limit  $\mu$  of a subsequence of  $(\nu_n)$  satisfies (again by the Portmanteau theorem)  $\mu(P) = \nu(P)$  for each atom  $P$  of  $\mathcal{P}_\ell$ . But since  $\text{diam}(\mathcal{P}_\ell) \rightarrow 0$ , the sequence  $(\mathcal{P}_\ell)$  separates points in  $E$ , hence it generates the Borel  $\sigma$ -algebra of  $E$ . Thus we have  $\mu = \nu$ , and using the compactness of  $M(E)$  for the weak\* topology, we get that  $\nu_n \xrightarrow[n \rightarrow \infty]{w^*} \nu$ .  $\square$

**Lemma 3.3.** *For each  $\nu \in M(E)$  of a compact metric space  $(E, d)$ , there exists a good sequence of partitions for  $(E, \nu)$ .*

*Proof.* We first show that, for each  $\ell \geq 1$ , there exists a finite partition  $\mathcal{Q}_\ell$  in which each atom  $Q$  satisfies

- $\text{diam}(Q) < 1/\ell$ ,
- $\nu(\partial Q) = 0$ .

Indeed, by compactness, there exists a finite set  $\{x_1, \dots, x_k\} \subset E$  such that

$$E \subset \bigcup_{1 \leq i \leq k} B(x_i, \frac{1}{3\ell}).$$



Then, for each  $1 \leq i \leq k$ , there exist at most countably many  $r > 0$  such that

$$\nu(\partial B(x_i, r)) > 0.$$

Therefore, we can find  $r \in (\frac{1}{3\ell}, \frac{1}{2\ell})$  such that

$$\forall 1 \leq i \leq k, \quad \nu(\partial B(x_i, r)) = 0.$$

Then the partition  $\mathcal{Q}_\ell$  generated by the open balls  $B(x_i, r)$ ,  $1 \leq i \leq k$ , satisfies the required conditions.

Once we have  $\mathcal{Q}_\ell$  for each  $\ell \geq 1$ , we set

$$\mathcal{P}_\ell := \mathcal{Q}_1 \vee \cdots \vee \mathcal{Q}_\ell,$$

and we get a good sequence  $(\mathcal{P}_\ell)$  for  $(E, \nu)$ .  $\square$

**Lemma 3.4.** *Let  $(\mathcal{P}_\ell)$  be a good sequence of partitions for  $(E_1, \nu_1)$ , and let  $(\mathcal{Q}_\ell)$  be a good sequence of partitions for  $(E_2, \nu_2)$ . Then for each coupling  $\rho$  of  $\nu_1$  and  $\nu_2$ ,  $(\mathcal{P}_\ell \times \mathcal{Q}_\ell)$  is a good sequence of partitions for  $(E_1 \times E_2, \rho)$ .*

*Proof.* This is obvious, since for each atom  $P$  of  $\mathcal{P}_\ell$  and each atom  $Q$  of  $\mathcal{Q}_\ell$ ,

$$\partial(P \times Q) \subset (\partial P \times E_2) \cup (E_1 \times \partial Q),$$

and the marginals of  $\rho$  are  $\nu_1$  and  $\nu_2$ .  $\square$

### 3.2 Proof of Proposition 3.1

Without loss of generality, we can (and we do) assume that the measure-theoretic dynamical system  $(Y, \kappa, S)$  is aperiodic. Indeed, if this is not the case, we consider any uniquely ergodic topological system  $(Y', S')$  whose unique invariant measure  $\kappa'$  is such that  $(Y', \kappa', S')$  is aperiodic. Then we take any point  $u' \in Y'$ , and we replace  $Y$  by  $Y \times Y'$ ,  $S$  by  $S \times S'$ , and  $u$  by  $(u, u')$ . We also replace  $(N_m)$  by a subsequence of  $(N_m)$  along which  $(u, u')$  is generic, for some measure  $\tilde{\kappa}$  whose marginals have to be  $\kappa$  and  $\kappa'$ . But then the system  $(Y \times Y', \tilde{\kappa}, S \times S')$  is aperiodic, because it is an extension of the aperiodic system  $(Y', \kappa', S')$ .

We fix a good sequence of partitions  $(\mathcal{Q}_\ell)$  for  $(Y, \kappa)$  and a good sequence of partitions  $(\mathcal{P}_\ell)$  for  $(X, \nu)$ . Then by Lemma 3.4,  $(\mathcal{Q}_\ell \times \mathcal{P}_\ell)$  is a good sequence of partitions for  $(Y \times X, \rho)$ .

**Definition 3.2.** Let  $M > 0$ . A subset  $E$  of  $\mathbb{N}$  is said to be  $M$ -separated if for each integers  $n \neq m$ ,  $n, m \in E \implies |n - m| \geq M$ .

The main argument to prove Proposition 3.1 stands in the following proposition.

**Proposition 3.5.** *Under the assumptions of Proposition 3.1, and assuming also that  $(Y, \kappa, S)$  is aperiodic (see above), given  $\ell_0 \geq 1$  and  $\varepsilon \in (0, \frac{1}{2})$ , there exists a sequence  $(x_n)$  of points in  $X$  such that:*

- $\{n \geq 0 : x_{n+1} \neq Tx_n\}$  is  $\frac{1}{\varepsilon}$ -separated,
- for each atom  $A$  of  $\mathcal{Q}_{\ell_0} \times \mathcal{P}_{\ell_0}$ , we have

$$(36) \quad \rho(A) - \varepsilon < \liminf_{m \rightarrow \infty} \frac{1}{N_m} \sum_{0 \leq n < N_m} \mathbb{1}_A(S^n u, x_n),$$

and

$$(37) \quad \limsup_{m \rightarrow \infty} \frac{1}{N_m} \sum_{0 \leq n < N_m} \mathbb{1}_A(S^n u, x_n) < \rho(A) + \varepsilon.$$

*Proof.* Let  $h$  be a natural number such that  $\frac{1}{h} < \varepsilon$ . We claim that for  $\ell$  large enough, we can find a set  $B \subset Y$  which is measurable with respect to  $\bigvee_{0 \leq j \leq h-1} S^j \mathcal{Q}_\ell$ , and such that

- $B, SB, \dots, S^{h-1}B$  are pairwise disjoint,
- $\kappa\left(\bigcup_{0 \leq j \leq h-1} S^j B\right) > 1 - \varepsilon$ .

Indeed, since  $(Y, \kappa, S)$  is assumed to be aperiodic, we can use the Rokhlin lemma to find a Borel subset  $\tilde{B} \subset Y$  such that  $\tilde{B}, S\tilde{B}, \dots, S^{h-1}\tilde{B}$  are pairwise disjoint, and such that

$$\kappa\left(\bigcup_{0 \leq j \leq h-1} S^j \tilde{B}\right) > 1 - \frac{\varepsilon}{2}.$$

Then we use the fact that the good sequence of partitions  $(\mathcal{Q}_\ell)$  generates the Borel  $\sigma$ -algebra: it follows that for  $\ell$  large enough, we can find a  $\mathcal{Q}_\ell$ -measurable set  $B'$  such that

$$\kappa(B' \Delta \tilde{B}) < \frac{\varepsilon}{8h^2}.$$

For each  $1 \leq j \leq h-1$ , we have

$$B' \cap S^j B' \subset (B' \setminus \tilde{B}) \cup (S^j B' \setminus S^j \tilde{B}),$$

hence

$$(38) \quad \kappa(B' \cap S^j B') \leq \frac{\varepsilon}{4h^2}.$$

It remains to define  $B$  by

$$B := B' \setminus \left( \bigcup_{1 \leq j \leq h-1} S^j B' \right).$$

Then, by construction,  $B$  is disjoint from  $S^j B$  for each  $1 \leq j \leq h-1$ , thus  $B, SB, \dots, S^{h-1}B$  are pairwise disjoint. Moreover, from (38), we have

$$\kappa(B) \geq \kappa(B') - \frac{\varepsilon}{4h} \geq \kappa(\tilde{B}) - \frac{\varepsilon}{2h},$$

which implies

$$\kappa \left( \bigcup_{0 \leq j \leq h-1} S^j B \right) = h\kappa(B) \geq h\kappa(\tilde{B}) - \frac{\varepsilon}{2} > 1 - \varepsilon,$$

and our first claim is proved.

Since  $u$  is generic for  $\kappa$  along  $(N_m)$ , and since the set  $\bigcup_{0 \leq j \leq h-1} S^j B$  is measurable with respect to  $\bigvee_{0 \leq j \leq 2h} S^j \mathcal{Q}_\ell$  (in particular, the  $\kappa$ -measure of its boundary vanishes), we have

$$(39) \quad \frac{1}{N_m} \sum_{0 \leq n < N_m} \mathbb{1}_{\bigcup_{0 \leq j \leq h-1} S^j B}(S^n u) \xrightarrow{m \rightarrow \infty} \kappa \left( \bigcup_{0 \leq j \leq h-1} S^j B \right) > 1 - \varepsilon.$$

This implies in particular that the set  $P_B(u) := \{n \geq 0 : S^n u \in B\}$  is infinite. We number in order the elements of this set:

$$P_B(u) = \{b_1 < b_2 < \dots < b_k < \dots\}$$

The integers  $(b_k)$  will correspond to the times when we will be allowed to change the orbit of the desired sequence. As  $B$  is disjoint from  $S^j B$  for each  $1 \leq j \leq h-1$ , the set  $P_B(u)$  is  $h$ -separated, hence  $\frac{1}{\varepsilon}$ -separated.

We consider the partition  $\bigvee_{0 \leq j \leq h-1} (S \times T)^{-j} (\mathcal{Q}_{\ell_0} \times \mathcal{P}_{\ell_0})$  of  $Y \times X$ . Any atom of this partition is of the form  $\bar{Q} \times \bar{P}$ , where  $\bar{Q}$  (respectively  $\bar{P}$ ) is an atom of  $\bigvee_{0 \leq j \leq h-1} S^{-j} \mathcal{Q}_{\ell_0}$  (respectively of  $\bigvee_{0 \leq j \leq h-1} T^{-j} \mathcal{P}_{\ell_0}$ ). For such atoms  $\bar{Q}$  and  $\bar{P}$ , we can write

$$(40) \quad \bar{Q} = Q_0 \cap S^{-1}Q_1 \cap \dots \cap S^{-(h-1)}Q_{h-1},$$

each  $Q_j$  being an atom of  $\mathcal{Q}_{\ell_0}$ , and

$$(41) \quad \bar{P} = P_0 \cap S^{-1}P_1 \cap \dots \cap S^{-(h-1)}P_{h-1},$$

each  $P_j$  being an atom of  $\mathcal{P}_{\ell_0}$ . Since the  $\kappa$ -measure of the boundary of each involved set is always 0, we again have for each atom  $\bar{Q}$  of  $\bigvee_{0 \leq j \leq h-1} S^{-j} \mathcal{Q}_{\ell_0}$

$$(42) \quad \frac{1}{N_m} \sum_{b_k < N_m} \mathbb{1}_{\bar{Q}}(S^{b_k} u) = \frac{1}{N_m} \sum_{0 \leq n < N_m} \mathbb{1}_{B \cap \bar{Q}}(S^n u) \xrightarrow{m \rightarrow \infty} \kappa(B \cap \bar{Q}).$$

If  $C$  is a measurable subset of  $Y$  with  $\kappa(C) > 0$ , we denote by  $\rho_C^Y$  the marginal on  $X$  of the conditional probability measure  $\rho(\cdot | C \times X)$ . Then, for each measurable  $A \subset X$ , we have

$$\begin{aligned}
\rho(C \times A) &= \rho((C \times X) \cap (Y \times A)) \\
(43) \quad &= \rho(C \times X) \rho(Y \times A | C \times X) \\
&= \kappa(C) \rho_C^Y(A).
\end{aligned}$$

On an appropriate probability space, we construct a sequence  $(\xi_k)$  of independent random variables, taking values in  $X$ , such that for each  $k$ ,  $\xi_k$  is distributed according to  $\rho_{B \cap \bar{Q}}^Y$ , where  $\bar{Q}$  is the atom of  $\bigvee_{0 \leq j \leq h-1} S^{-j} \mathcal{Q}_{\ell_0}$  containing  $S^{b_k} u$ .

For each atom  $\bar{Q}$  of  $\bigvee_{0 \leq j \leq h-1} S^{-j} \mathcal{Q}_{\ell_0}$  and each atom  $\bar{P}$  of  $\bigvee_{0 \leq j \leq h-1} T^{-j} \mathcal{P}_{\ell_0}$ , by (42), the law of large numbers and (43), with probability 1, we have

$$(44) \quad \frac{1}{N_m} \sum_{b_k < N_m} \mathbb{1}_{\bar{Q}}(S^{b_k} u) \mathbb{1}_{\bar{P}}(\xi_k) \xrightarrow{m \rightarrow \infty} \kappa(B \cap \bar{Q}) \rho_{B \cap \bar{Q}}^Y(\bar{P}) = \rho((B \cap \bar{Q}) \times \bar{P}).$$

Let us fix a realization of  $(\xi_k)$  which satisfies (44) for each atom  $\bar{Q} \times \bar{P}$  of  $\bigvee_{0 \leq j \leq h-1} (S \times T)^{-j} (\mathcal{Q}_{\ell_0} \times \mathcal{P}_{\ell_0})$ . Then, for each  $n \geq 0$ , we define the point  $x_n \in X$  as follows:

$$x_n := \begin{cases} T^{n-b_1} \xi_1 & \text{if } n < b_1, \\ T^{n-b_k} \xi_k & \text{if } b_k \leq n < b_{k+1} \text{ for some } k \geq 1. \end{cases}$$

The set of integers  $n$  such that  $x_{n+1} \neq T x_n$  is contained in  $P_B(u)$ , therefore, it is  $\frac{1}{\varepsilon}$ -separated.

Now, let  $A = Q \times P$  be a fixed atom of  $\mathcal{Q}_{\ell_0} \times \mathcal{P}_{\ell_0}$ . We set

$$R := \bigcup_{0 \leq j \leq h-1} S^j B \times X,$$

and we observe that

$$(45) \quad \rho(R) = \kappa \left( \bigcup_{0 \leq j \leq h-1} S^j B \right) > 1 - \varepsilon.$$

We also note that for each  $n \geq b_1$ ,  $(S^n u, x_n) \in R$  if and only if there exists  $k$  and  $0 \leq j \leq h-1$  such that  $n = b_k + j$ . In this case,  $(S^n u, x_n) \in A \cap R$  if and only if the atom  $\bar{Q} \times \bar{P}$  of  $\bigvee_{0 \leq j \leq h-1} (S \times T)^{-j} (\mathcal{Q}_{\ell_0} \times \mathcal{P}_{\ell_0})$  containing  $(S^{b_k} u, \xi_k)$  satisfies  $Q_j = Q$  and  $P_j = P$  (using the notations given in (40) and (41), and remembering that  $A = Q \times P$ ).

We can then use (44) to get

$$\begin{aligned}
& \frac{1}{N_m} \sum_{b_1 \leq n < N_m} \mathbb{1}_{A \cap R}(S^n u, x_n) \\
(46) \quad &= \frac{1}{N_m} \sum_{b_k < N_m} \sum_{0 \leq j \leq h-1} \sum_{(\bar{Q}, \bar{P}): Q_j=Q \text{ and } P_j=P} \mathbb{1}_{\bar{Q} \times \bar{P}}(S^{b_k} u, \xi_k) \\
&\xrightarrow{m \rightarrow \infty} \sum_{0 \leq j \leq h-1} \sum_{(\bar{Q}, \bar{P}): Q_j=Q \text{ and } P_j=P} \rho((B \cap \bar{Q}) \times \bar{P}).
\end{aligned}$$

But, on the other hand, we can write

$$\begin{aligned}
\rho(A \cap R) &= \sum_{0 \leq j \leq h-1} \rho(A \cap (S^j B \times X)) \\
(47) \quad &= \sum_{0 \leq j \leq h-1} \rho((S^{-j} Q \times T^{-j} P) \cap (B \times X)) \\
&= \sum_{0 \leq j \leq h-1} \sum_{(\bar{Q}, \bar{P}): Q_j=Q \text{ and } P_j=P} \rho((B \cap \bar{Q}) \times \bar{P}).
\end{aligned}$$

From (46) and (47), it follows that

$$(48) \quad \frac{1}{N_m} \sum_{0 \leq n < N_m} \mathbb{1}_{A \cap R}(S^n u, x_n) \xrightarrow{m \rightarrow \infty} \rho(A \cap R).$$

From (39), we get that

$$\lim_{m \rightarrow \infty} \frac{1}{N_m} \sum_{0 \leq n < N_m} \mathbb{1}_{(Y \times X) \setminus R}(S^n u, x_n) < \varepsilon,$$

and since  $\mathbb{1}_A \leq \mathbb{1}_{A \cap R} + \mathbb{1}_{Y \times X \setminus R}$ , this yields by (48),

$$\begin{aligned}
\limsup_{m \rightarrow \infty} \frac{1}{N_m} \sum_{0 \leq n < N_m} \mathbb{1}_A(S^n u, x_n) &< \lim_{m \rightarrow \infty} \frac{1}{N_m} \sum_{0 \leq n < N_m} \mathbb{1}_{A \cap R}(S^n u, x_n) + \varepsilon \\
&= \rho(A \cap R) + \varepsilon \\
&\leq \rho(A) + \varepsilon,
\end{aligned}$$

and we have (37). On the other hand, using  $\mathbb{1}_A \geq \mathbb{1}_{A \cap R}$ , we get by (45)

$$\begin{aligned}
\liminf_{m \rightarrow \infty} \frac{1}{N_m} \sum_{0 \leq n < N_m} \mathbb{1}_A(S^n u, x_n) &\geq \lim_{m \rightarrow \infty} \frac{1}{N_m} \sum_{0 \leq n < N_m} \mathbb{1}_{A \cap R}(S^n u, x_n) \\
&= \rho(A \cap R) \\
&> \rho(A) - \varepsilon.
\end{aligned}$$

and we have (36).  $\square$

We can now give the proof of Proposition 3.1, in which we use the following obvious fact: if we modify the sequence  $(x_n)$  given by Proposition 3.5 on a finite number of terms, we still get (36) and (37).

*Proof of Proposition (3.1).* We fix a sequence  $(\varepsilon_\ell)_{\ell \geq 1}$  of numbers in  $(0, \frac{1}{2})$ , decreasing to 0, and we construct inductively the desired sequence  $(x_n)$  and the subsequence  $(N_{m_\ell})$  by a repeated use of Proposition 3.5.

We start by applying Proposition 3.5 with  $\varepsilon := \varepsilon_1$  and  $\ell_0 := 1$ . It provides us with an integer  $m_1$ , and a finite sequence  $(x_n)_{0 \leq n < N_{m_1}}$  of points in  $X$  such that

- the set of integers  $n \in \{0, \dots, N_{m_1} - 2\}$  such that  $x_{n+1} \neq Tx_n$  is  $\frac{1}{\varepsilon_1}$ -separated,
- for each atom  $A$  of  $\mathcal{Q}_1 \times \mathcal{P}_1$ , we have

$$\rho(A) - \varepsilon_1 < \frac{1}{N_{m_1}} \sum_{0 \leq n < N_{m_1}} \mathbb{1}_A(S^n u, x_n) < \rho(A) + \varepsilon_1.$$

Now, assume that for some  $\ell \geq 1$  we have already constructed  $m_1 < \dots < m_\ell$  and the sequence  $(x_n)_{0 \leq n < N_{m_\ell}}$  of points in  $X$  such that

- for each  $1 \leq j < \ell$ , the set of integers  $n \in \{N_{m_{j-1}}, \dots, N_{m_j} - 2\}$  such that  $x_{n+1} \neq Tx_n$  is  $\frac{1}{\varepsilon_j}$ -separated (with the convention that  $N_{m_0} = 0$ ),
- for each atom  $A$  of  $\mathcal{Q}_\ell \times \mathcal{P}_\ell$ , we have

$$\rho(A) - \varepsilon_\ell < \frac{1}{N_{m_\ell}} \sum_{0 \leq n < N_{m_\ell}} \mathbb{1}_A(S^n u, x_n) < \rho(A) + \varepsilon_\ell.$$

Then we apply again Proposition 3.5, with  $\varepsilon := \varepsilon_{\ell+1}$  and  $\ell_0 := \ell + 1$ . It provides us with an integer  $m_{\ell+1}$  and a finite sequence of points  $(x_n)_{N_{m_\ell} \leq n < N_{m_{\ell+1}}}$  in  $X$  which satisfy:

- the set of integers  $n \in \{N_{m_\ell}, \dots, N_{m_{\ell+1}} - 2\}$  such that  $x_{n+1} \neq Tx_n$  is  $\frac{1}{\varepsilon_{\ell+1}}$ -separated,
- for each atom  $A$  of  $\mathcal{Q}_{\ell+1} \times \mathcal{P}_{\ell+1}$ , we have

$$(49) \quad \rho(A) - \varepsilon_{\ell+1} < \frac{1}{N_{m_{\ell+1}}} \sum_{0 \leq n < N_{m_{\ell+1}}} \mathbb{1}_A(S^n u, x_n) < \rho(A) + \varepsilon_{\ell+1}.$$

(We keep the points  $(x_n)_{0 \leq n < N_{m_\ell}}$  already provided by the induction hypothesis, referring to the obvious fact stated before the proof.)

Moreover, we can assume that the sequence  $(\varepsilon_\ell)$  decreases sufficiently fast so that the validity of (49) for each atom  $A$  of  $\mathcal{Q}_{\ell+1} \times \mathcal{P}_{\ell+1}$  ensures the

validity of the analog inequalities for each  $A$  which is a finite union of atoms of  $\mathcal{Q}_{\ell+1} \times \mathcal{P}_{\ell+1}$  (in particular, for each  $A$  which is an atom of the previous partitions), but with  $\varepsilon_\ell$  instead of  $\varepsilon_{\ell+1}$ .

The sequence  $(x_n)_{n \geq 0}$  of points in  $X$  and the subsequence  $(N_{m_\ell})$  we construct with the above inductive procedure then satisfy the conditions announced in Proposition 3.1.  $\square$

### 3.3 Logarithmic case

We would like to study the logarithmic version of Proposition 3.1, in which we replace each arithmetic average of the form

$$\frac{1}{N_m} \sum_{0 \leq n < N_m} f(n)$$

by the logarithmic average

$$\frac{1}{L(N_m)} \sum_{1 \leq n \leq N_m} \frac{1}{n} f(n).$$

(Here we use the notation  $L(N) := 1 + \frac{1}{2} + \dots + \frac{1}{N}$ .) In fact, this logarithmic version, whose statement is written below, is also valid, and the arguments to prove it are exactly the same as in the arithmetic average case. We just point out below the few technical changes that need to be made in the proof for the logarithmic case.

**Proposition 3.6.** *Let  $(Y, S)$  and  $(X, T)$  be two topological systems and  $u \in Y$ , logarithmically generic along an increasing sequence  $(N_m)$  for some  $S$ -invariant measure  $\kappa$  on  $Y$ . Let  $\rho$  be a joining of  $\kappa$  with a  $T$ -invariant measure  $\nu$  on  $X$ . Then there exist a sequence  $(x_n) \subset X$  and a subsequence  $(N_{m_\ell})$  such that:*

- *the sequence  $(S^n u, x_n)$  is logarithmically generic for  $\rho$  along  $(N_{m_\ell})$ :*

$$\frac{1}{L(N_{m_\ell})} \sum_{1 \leq n \leq N_{m_\ell}} \frac{1}{n} \delta_{(S^n u, x_n)} \xrightarrow{\ell \rightarrow \infty} \rho,$$

- *the set  $\{n \geq 0 : x_{n+1} \neq T x_n\}$  is of the form  $\{b_1 < b_2 < \dots < b_k < b_{k+1} < \dots\}$ , where  $b_{k+1} - b_k \rightarrow \infty$ .*

The changes that need to be made to the proof are almost all quite obvious, they consist in formally replacing the arithmetic average by the logarithmic average. One point maybe needs some explanations, namely when we arrive at the proof of the logarithmic analog of (44). We put these explanations in the form of a lemma, which we will apply in the following context:  $(d_k)$  is the ordered sequence of positive integers  $n$  such that  $S^n u \in B \cap \bar{Q}$ , and the sequence  $(\rho_k)$  is defined by  $\rho_k := \mathbb{1}_{\bar{P}}(\xi_k)$ .

**Lemma 3.7.** Let  $(d_k)$  be an increasing sequence of positive integers such that

$$\frac{1}{L(N_m)} \sum_{d_k \leq N_m} \frac{1}{d_k} \xrightarrow{m \rightarrow \infty} \kappa \in [0, 1],$$

and let  $(\rho_k)$  be a sequence of real numbers in  $[0, 1]$  such that

$$\frac{1}{K} \sum_{1 \leq k \leq K} \rho_k \xrightarrow{K \rightarrow \infty} \rho \in [0, 1].$$

Then we have

$$\frac{1}{L(N_m)} \sum_{d_k \leq N_m} \frac{\rho_k}{d_k} \xrightarrow{m \rightarrow \infty} \kappa \rho.$$

*Proof.* For each  $m$ , let us denote by  $k_m$  the largest  $k$  such that  $d_k \leq N_m$ . We use the classical identity

$$\sum_{d_k \leq N_m} \frac{\rho_k}{d_k} = \sum_{1 \leq k < k_m} \left( \frac{1}{d_k} - \frac{1}{d_{k+1}} \right) (\rho_1 + \cdots + \rho_k) + \frac{1}{d_{k_m}} (\rho_1 + \cdots + \rho_{k_m}).$$

Given  $\varepsilon > 0$ , let  $K_\varepsilon$  be such that

$$K \geq K_\varepsilon \implies \left| \frac{1}{K} \sum_{1 \leq k \leq K} \rho_k - \rho \right| < \varepsilon.$$

We can then write

$$\begin{aligned} & \left| \frac{1}{L(N_m)} \sum_{d_k \leq N_m} \frac{\rho_k}{d_k} - \frac{1}{L(N_m)} \sum_{d_k \leq N_m} \frac{\rho}{d_k} \right| \\ &= \left| \frac{1}{L(N_m)} \sum_{K_\varepsilon \leq k < k_m} k \left( \frac{1}{d_k} - \frac{1}{d_{k+1}} \right) \left( \frac{1}{k} (\rho_1 + \cdots + \rho_k) - \rho \right) \right| + O\left( \frac{1}{L(N_m)} \right) \\ &< \varepsilon + O\left( \frac{1}{L(N_m)} \right). \end{aligned}$$

But by assumption, we have

$$\frac{1}{L(N_m)} \sum_{d_k \leq N_m} \frac{\rho}{d_k} \xrightarrow{m \rightarrow \infty} \kappa \rho,$$

hence we get

$$\frac{1}{L(N_m)} \sum_{d_k \leq N_m} \frac{\rho_k}{d_k} \xrightarrow{m \rightarrow \infty} \kappa \rho.$$

□



The last place in the proof where a (very easy) correction should be made in the logarithmic case is to get the analog of (46): at some point we have to replace some coefficients  $\frac{1}{b_k+j}$  by  $\frac{1}{b_k}$ , which is of no consequence since  $j$  remains bounded between 0 and  $h-1$  here. To be more precise, (46) becomes

$$\begin{aligned}
(50) \quad & \frac{1}{L(N_m)} \sum_{b_1 \leq n < N_m} \frac{1}{n} \mathbb{1}_{A \cap R}(S^n u, x_n) \\
&= \frac{1}{L(N_m)} \sum_{b_k < N_m} \sum_{0 \leq j \leq h-1} \frac{1}{b_k + j} \sum_{(\bar{Q}, \bar{P}): Q_j=Q \text{ and } P_j=P} \mathbb{1}_{\bar{Q} \times \bar{P}}(S^{b_k} u, \xi_k) \\
&= \frac{1}{L(N_m)} \sum_{b_k < N_m} \sum_{0 \leq j \leq h-1} \frac{1}{b_k} \sum_{(\bar{Q}, \bar{P}): Q_j=Q \text{ and } P_j=P} \mathbb{1}_{\bar{Q} \times \bar{P}}(S^{b_k} u, \xi_k) + o(1) \\
&\xrightarrow{m \rightarrow \infty} \sum_{0 \leq j \leq h-1} \sum_{(\bar{Q}, \bar{P}): Q_j=Q \text{ and } P_j=P} \rho((B \cap \bar{Q}) \times \bar{P}).
\end{aligned}$$

## 4 Proofs of main theorems

### 4.1 Proof of Theorem B

*Proof of Theorem B.* Take any topological system  $(X, T) \in \mathcal{C}_{\mathcal{F}}$  and fix  $f \in C(X)$ ,  $x \in X$ . Take any increasing sequence  $(N_k)$  for which, with no loss of generality, we can assume that  $\frac{1}{N_k} \sum_{n \leq N_k} \delta_{(T^n x, S^n \mathbf{u})} \rightarrow \rho$ . It follows that

$$\lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} f(T^n x) \mathbf{u}(n) = \int f \otimes \pi_0 d\rho.$$

But  $\rho$  is a joining of some  $T$ -invariant measure  $\nu \in V(X, T)$  for which  $x$  is generic along  $(N_k)$ , and some Furstenberg system  $\kappa$  of  $\mathbf{u}$ . Since  $(X, T) \in \mathcal{C}_{\mathcal{F}}$ , the system  $(X, \nu, T)$  is in  $\mathcal{F}$ , and the integral on the right-hand side above vanishes by the Veech condition and Proposition 2.2.

□

### 4.2 Proof of Theorem A

Before we begin the proof, let us make the following remark concerning topological models. Given an automorphism  $(Z, \mathcal{D}, \kappa, R)$ , and a fixed subset of full measure of ergodic components of  $\kappa$ , recall that by a *Hansel model* of  $R$ , we mean a topological system  $(X, T)$  which has a  $T$ -invariant measure  $\nu$  such that, as dynamical systems,  $\kappa$  and  $\nu$  are isomorphic and such that each point  $x \in X$  is generic for one of these chosen ergodic components. In [29], it is proved that each automorphism has a Hansel model.

We assume that  $\mathbf{u} \perp \mathcal{C}_{\mathcal{F}_{ec}}$  for some characteristic class  $\mathcal{F}$ . Take  $\kappa \in V(\mathbf{u})$  and fix  $(N_m)$  so that

$$\frac{1}{N_m} \sum_{n \leq N_m} \delta_{S^n \mathbf{u}} \rightarrow \kappa.$$

Denote by  $\mathcal{A}(\kappa) \subset \mathcal{B}(X_{\mathbf{u}})$  the largest  $\mathcal{F}_{ec}$ -factor of  $(X_{\mathbf{u}}, \kappa, S)$ , i.e.  $\mathcal{A}(\kappa) = \mathcal{B}(X_{\mathbf{u}})_{\mathcal{F}_{ec}}$ .

Consider the factor  $(X_{\mathbf{u}}/\mathcal{A}(\kappa), \mathcal{A}(\kappa), \kappa|_{\mathcal{A}(\kappa)}, S)$  and take a Hansel model  $(X, \nu, T)$  of it (by choosing in the ergodic decomposition of  $\kappa|_{\mathcal{A}(\kappa)}$  only ergodic measures in  $\mathcal{F}$ ). By definition,

$$(51) \quad (X, T) \in \mathcal{C}_{\mathcal{F}_{ec}}.$$

Fix a measure-theoretic factor map  $J: (X_{\mathbf{u}}, \kappa, S) \rightarrow (X, \nu, T)$  such that  $J^{-1}(\mathcal{B}(X)) = \mathcal{A}(\kappa)$ , and let  $\nu_J$  denote the corresponding graph joining (of  $\nu$  and  $\kappa|_{\mathcal{A}(\kappa)}$ ). Let  $\hat{\nu}_J$  be the relatively independent extension of  $\nu_J$  to a joining of  $\nu$  and  $\kappa$ : for  $f \in L^2(\nu)$  and  $g \in L^2(\kappa)$ , we have

$$(52) \quad \int_{X_{\mathbf{u}} \times X} g \otimes f d\hat{\nu}_J = \int_{X_{\mathbf{u}}} \mathbb{E}^{\kappa}(g|\mathcal{A}(\kappa)) f \circ J d\kappa.$$

Now, by applying Proposition 3.1, we can find  $(x_n) \subset Y$  such that

$$(53) \quad ((x_n), \mathbf{u}) \text{ is generic for } \hat{\nu}_J \text{ along some subsequence } (N_{m_\ell}),$$

and the set  $\{n \geq 0 : x_{n+1} \neq Tx_n\}$  is of the form  $\{b_1 < b_2 < \dots\}$  with  $b_{k+1} - b_k \rightarrow \infty$ .

Since  $\mathbf{u} \perp \mathcal{C}_{\mathcal{F}_{ec}}$ , (51) and Proposition 2.17 ensure that the system  $(Y, S)$  satisfies the strong  $\mathbf{u}$ -MOMO property. Therefore, for each  $f \in C(Y)$  we have

$$\begin{aligned} & \lim_{m \rightarrow \infty} \frac{1}{N_m} \sum_{n \leq N_m} f(x_n) \mathbf{u}(n) = \\ & \lim_{\ell \rightarrow \infty} \frac{1}{b_{K_m}} \sum_{k < K_m} \left( \sum_{b_k \leq n < b_{k+1}} f(T^{n-b_k} x_{b_k}) \mathbf{u}(n) \right) = 0, \end{aligned}$$

and it follows from (53) that  $\int f \otimes \pi_0 d\hat{\nu}_J = 0$ . Using (52), we get

$$\int_{X_{\mathbf{u}}} \mathbb{E}^{\kappa}(\pi_0|\mathcal{A}(\kappa)) f \circ J d\kappa = 0.$$

But  $\{f \circ J : f \in C(X)\}$  is dense in  $L^2(\mathcal{A}(\kappa))$  and therefore  $\pi_0 \perp L^2(\mathcal{A}(\kappa))$ , which is the Veech condition for  $\mathbf{u}$  with respect to the characteristic class  $\mathcal{F}_{ec}$ .

## 5 Combinatorics

### 5.1 Orthogonality to zero entropy systems

#### 5.1.1 Cancellations. Proof of Corollaries C and C'

We need the following interpretation of the Veech condition in terms of relative uniform mixing (K-mixing) of the function  $\pi_0$ . For  $n \in \mathbb{N}$ , let  $\pi_n := \pi_0 \circ S^n$ .

**Proposition 5.1.** *Let  $\kappa \in V_S(\mathbf{u})$  and  $\int \pi_0 d\kappa = 0$ . Then the following conditions are equivalent:*

- (a)  $\pi_0 \perp L^2(\Pi(\kappa))$ ,
- (b)  $\pi_0$  is relatively K-mixing, i.e. for each  $\varepsilon > 0$ , there exists  $N$  such that

$$\left| \int \pi_0 \mathbf{1}_C d\kappa - \int \pi_0 d\kappa \int \mathbf{1}_C d\kappa \right| = \left| \int \pi_0 \mathbf{1}_C d\kappa \right| < \varepsilon$$

for each set  $C \in \sigma(\pi_n, \pi_{n+1}, \dots)$  and  $n \geq N$ .

If we additionally assume that  $\mathbf{u}$  takes values in a finite set  $E \subset \mathbb{C}$  and  $(M_k)$  is a sequence along which we have a Furstenberg system  $\kappa$  then the above conditions are equivalent to

- (c) for each  $\varepsilon > 0$  there exists  $N \geq 1$  such that for any  $s \geq 1$  and any function  $f$  depending on coordinates  $N \leq n, n+1, \dots, n+s$ ,  $\|f\|_{C(X_{\mathbf{u}})} \leq 1$ , we have

$$\limsup_{k \rightarrow \infty} \left| \frac{1}{M_k} \sum_{m \leq M_k} \mathbf{u}(m) f(S^m \mathbf{u}) \right| < \varepsilon.$$

*Proof.* (a)  $\implies$  (b). Assume that  $\mathbb{E}(\pi_0 | \Pi(\kappa)) = 0$ . Let  $C \in \sigma(\pi_n, \pi_{n+1}, \dots)$ . We have

$$\begin{aligned} \left| \int \pi_0 \mathbf{1}_C d\kappa \right| &= \left| \int \mathbf{1}_C \mathbb{E}(\pi_0 | \sigma(\pi_n, \pi_{n+1}, \dots)) d\kappa \right| \\ &\leq \int |\mathbb{E}(\pi_0 | \sigma(\pi_n, \pi_{n+1}, \dots))| d\kappa. \end{aligned}$$

Hence, we have an upper bound which does not depend on  $C$ . Since

$$\mathbb{E}(\pi_0 | \sigma(\pi_n, \pi_{n+1}, \dots)) \rightarrow \mathbb{E}(\pi_0 | \Pi(\kappa)) = 0$$

$\kappa$ -a.e. and thus also in  $L^1$ , which is precisely the relative K-mixing for  $\pi_0$ .

(b)  $\implies$  (a). Suppose that  $\pi_0$  is relatively K-mixing. Then, in particular, we have (5.1) for each  $C \in \Pi(\kappa)$ . In fact, since  $\varepsilon > 0$  is arbitrary,  $\int \pi_0 \mathbf{1}_C d\kappa = 0$  for each  $C \in \Pi(\kappa)$ . Whence  $\mathbb{E}(\pi_0 | \Pi(\kappa)) = 0$ .

(a)  $\implies$  (c). Since

$$\frac{1}{M_k} \sum_{m \leq M_k} \mathbf{u}(m) f(S^m \mathbf{u}) = \frac{1}{M_k} \sum_{m \leq M_k} (\pi_0 f)(S^m \mathbf{u}) \rightarrow \int_{X_{\mathbf{u}}} \pi_0 f d\kappa,$$

we can repeat the same argument as was used to prove (a)  $\implies$  (b) (replacing  $\mathbb{1}_C$  by  $f$ ).

(c)  $\implies$  (b). Suppose that  $\left| \int_{X_{\mathbf{u}}} \pi_0 f d\kappa \right| < \varepsilon$ , whenever  $f$  depending on coordinates  $n, n+1, \dots, n+s$  with  $n \geq N$  is bounded by 1. Consider all blocks on coordinates  $n, n+1, \dots, n+s$  that is all

$$B = \{x \in X_{\mathbf{u}} : x_n = b_0, \dots, x_{n+s} = b_s\}$$

with  $b_j \in E$ . Let  $C$  be any union of such blocks. Then  $\mathbb{1}_C$  is a (continuous) function depending on coordinates  $n, \dots, n+s$  and is bounded by 1 and, by assumption,

$$\left| \int_{X_{\mathbf{u}}} \pi_0 \mathbb{1}_C d\kappa \right| < \varepsilon.$$

Note that with  $N$  fixed and  $s$  arbitrary, the family of  $C$  defined above is dense in the  $\sigma$ -algebra  $\sigma(\pi_N, \pi_{N+1}, \dots)$ . Hence, given  $D \in \sigma(\pi_N, \pi_{N+1}, \dots)$  and  $\varepsilon > 0$ , we first find  $s \geq 0$  and then  $C$  as above (a union of blocks “sitting” on coordinates  $N, \dots, N+s$ ) such that  $\kappa(C \Delta D) < \varepsilon$  and find that

$$\left| \int \pi_0 \mathbb{1}_D d\kappa \right| \leq \left| \int \pi_0 \mathbb{1}_C d\kappa \right| + \kappa(C \Delta D) < 2\varepsilon.$$

□

Now, since each clopen set is a finite union of cylinders of a fixed length, Corollary C' follows directly by the above proposition. Corollary C' is a special case of Corollary C.

### 5.1.2 Conditional cancellations. Remark 1.6

The “cancellation law” of the values of  $\mathbf{u}$  along large shifts of the return times to a block (for most of the blocks) claimed in Remark 1.6 is a consequence of a refinement of Proposition 5.1.

**Proposition 5.2.** *Let  $\kappa \in V_S(\mathbf{u})$  and  $\int \pi_0 d\kappa = 0$ . Then the following conditions are equivalent:*

(a)  $\pi_0 \perp L^2(\Pi(\kappa))$ ,

(d) *for each  $\varepsilon > 0$  there exist  $N \geq 1$  and  $L \geq 1$  such that for each  $\ell \geq L$  there is a family of “good”  $\ell$ -blocks  $C \in \sigma(\pi_N, \pi_{N+1}, \dots)$ , i.e. of blocks satisfying*

$$\left| \int \mathbb{1}_C \cdot \pi_0 d\kappa \right| \leq \kappa(C)\varepsilon,$$

whose measure is  $> 1 - \varepsilon$ . In other words, for a “good”  $\ell$ -block  $C$ ,  $|\int \pi_0 d\kappa_C| < \varepsilon$ , where  $\kappa_C$  stands for the conditional measure on  $C$ .

*Proof.* (a)  $\implies$  (d). Fix  $\varepsilon > 0$  and note that  $\mathbb{E}^\kappa(\pi_0 | \sigma(\pi_N, \pi_{N+1}, \dots)) \rightarrow 0$   $\kappa$ -a.e. This implies convergence in measure, i.e., we can find a set  $A_\varepsilon$  of measure at least  $1 - \varepsilon$  such that for  $N$  large enough,

$$|\mathbb{E}^\kappa(\pi_0 | \sigma(\pi_N, \pi_{N+1}, \dots))(x)| < \varepsilon \text{ for all } x \in A_\varepsilon.$$

Fix such an  $N$ . There is  $M \geq 1$  large enough such that

$$\kappa(A_\varepsilon \Delta A_\varepsilon^{(M)}) < \varepsilon,$$

where  $A_\varepsilon^{(M)} \in \sigma(\pi_{-M}, \pi_{-M+1}, \dots)$  and note that  $S^{N+M} A_\varepsilon^{(M)} \in \sigma(\pi_N, \pi_{N+1}, \dots)$ . Now, for  $\ell$  large enough, we can approximate  $S^{N+M} A_\varepsilon^{(M)}$  by a (disjoint) union of  $\ell$ -blocks belonging to  $\sigma(\pi_N, \pi_{N+1}, \dots)$ ,

$$\kappa(S^{N+M} A_\varepsilon^{(M)} \Delta \bigcup_{j \in J} C_j^{(\ell)}) < \varepsilon.$$

But  $\kappa(S^{N+M} A_\varepsilon^{(M)} \Delta A_\varepsilon) < 2\varepsilon$ , so

$$\kappa(\bigcup_{j \in J} C_j^{(\ell)} \setminus A_\varepsilon) \leq \kappa(A_\varepsilon \Delta \bigcup_{j \in J} C_j^{(\ell)}) < 3\varepsilon.$$

Consider those  $j \in J$  for which  $\kappa(C_j^{(\ell)} \setminus A_\varepsilon) \geq \sqrt{\varepsilon} \kappa(C_j^{(\ell)})$ . Then the measure  $m$  of the union of such blocks has to satisfy  $\sqrt{\varepsilon} m < 3\varepsilon$ , so  $m < 3\sqrt{\varepsilon}$ . In other words “most” (in measure) of the  $C_j^{(\ell)}$ ’s are “good”, i.e. they satisfy  $\kappa(C_j^{(\ell)} \cap A_\varepsilon) > (1 - 3\sqrt{\varepsilon}) \kappa(C_j^{(\ell)})$ . Take such a “good”  $C$ . We have

$$\begin{aligned} \left| \int \mathbf{1}_C \cdot \pi_0 d\kappa \right| &\leq \int \mathbf{1}_C |\mathbb{E}^\kappa(\pi_0 | \sigma(\pi_N, \pi_{N+1}, \dots))| d\kappa \\ &= \int_{A_\varepsilon} \mathbf{1}_C |\mathbb{E}^\kappa(\pi_0 | \sigma(\pi_N, \pi_{N+1}, \dots))| d\kappa + \int_{A_\varepsilon^c} \mathbf{1}_C |\mathbb{E}^\kappa(\pi_0 | \sigma(\pi_N, \pi_{N+1}, \dots))| d\kappa \\ &\leq \varepsilon \kappa(C) + 3\sqrt{\varepsilon} \kappa(C). \end{aligned}$$

(d)  $\implies$  (a). Fix  $A \in \Pi(\kappa)$  of positive measure  $\kappa$ . Then for  $\varepsilon > 0$ , we can find  $N$  such that for all  $\ell$  large enough “most” of the  $\ell$ -blocks in  $\sigma(\pi_N, \pi_{N+1}, \dots)$  is “good” in the sense that

$$\left| \int \mathbf{1}_C \cdot \pi_0 d\kappa \right| \leq \kappa(C) \varepsilon.$$

Since  $A \in \sigma(\pi_N, \pi_{N+1}, \dots)$ , we can approximate it by unions of  $\ell$ -blocks (for  $\ell$  sufficiently large) and most of the used blocks is “good”. Whence

$$\left| \int \mathbf{1}_A \cdot \pi_0 d\kappa \right| \leq 2\varepsilon,$$

and since  $\varepsilon > 0$  was arbitrary,  $\pi_0 \perp L^2(\Pi(\kappa))$ .  $\square$

## 5.2 Orthogonality to $\mathcal{C}_{\text{ID}}$ . Proof of Corollary 1.8

We recall that (Proposition 2.12)

$$\mathcal{C}_{\text{ID}} = \mathcal{C}_{\text{ID}_{\text{ec}}}.$$

Since the characteristic factor is represented by the  $\sigma$ -algebra of invariant sets, by Theorems B and A, we obtain immediately that:

**Corollary 5.3.**  $\mathbf{u} \perp \mathcal{C}_{\text{ID}}$  if and only if for each Furstenberg system  $\kappa$  of  $\mathbf{u}$ ,  $\pi_0 \perp L^2(\mathcal{I}_\kappa)$ .

Let us now pass to a combinatorial characterization of the Veech condition. Assume that  $\kappa$  is given as the limit of  $\frac{1}{N_k} \sum_{n \leq N_k} \delta_{S^n \mathbf{u}}$ . In view of Corollary 5.3, we need to decipher  $\mathbb{E}(\pi_0 | \mathcal{I}_\kappa) = 0$ . By the von Neumann theorem, we have

$$\frac{1}{H} \sum_{h \leq H} \pi_0 \circ S^h \rightarrow \mathbb{E}(\pi_0 | \mathcal{I}_\kappa) \text{ in } L^2,$$

i.e.

$$\lim_{H \rightarrow \infty} \int_{X_{\mathbf{u}}} \left| \frac{1}{H} \sum_{h \leq H} \pi_0 \circ S^h \right|^2 d\kappa = 0$$

as  $\mathbb{E}(\pi_0 | \mathcal{I}_\kappa) = 0$ . So, given  $\varepsilon > 0$ ,

$$\int_{X_{\mathbf{u}}} \left| \frac{1}{H} \sum_{h \leq H} \pi_0 \circ S^h \right|^2 d\kappa < \varepsilon \text{ for all } H \geq H_\varepsilon.$$

The latter is equivalent to

$$\lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} \left| \frac{1}{H} \sum_{h \leq H} \pi_0 \circ S^h \right|^2 (S^n \mathbf{u}) < \varepsilon,$$

that is,

$$\lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} \left| \frac{1}{H} \sum_{h \leq H} \mathbf{u}(n+h) \right|^2 < \varepsilon.$$

The proof of Corollary 1.8 follows immediately.  $\square$

**Remark 5.4.** Hence, the Matomäki-Radziwiłł theorem [41] on the behaviour of a strongly aperiodic multiplicative function  $\mathbf{u}$  on a typical short interval implies  $\mathbf{u} \perp \mathcal{C}_{\text{ID}}$ . However, as shown in [28], the aperiodic multiplicative functions defined in [42] do not satisfy the assertion of Corollary 1.8.

In Corollary 5.3, the Veech condition (for  $\mathbf{u}$ ) equivalent to  $\mathbf{u} \perp \mathcal{C}_{\text{ID}}$  is written as  $\pi_0 \perp L^2(\mathcal{I}_\kappa)$ . If we look at it more spectrally, we obtain immediately that

$$(54) \quad \mathbf{u} \perp \mathcal{C}_{\text{ID}} \text{ if and only if } \sigma_{\pi_0, \kappa}(\{1\}) = 0$$

for all  $\kappa \in V_S(\mathbf{u})$ , i.e. the spectral measure of  $\pi_0$  (with respect to each Furstenberg system) has no atom at 1. Classically (by a simple computation), we have:

**Lemma 5.5.** *If  $\sigma$  is a measure on the circle  $\mathbb{S}^1$  then*

$$\frac{1}{H} \sum_{h=0}^{H-1} \hat{\sigma}(h) \rightarrow \sigma(\{1\}).$$

Hence, the Veech condition is equivalent to

$$\frac{1}{H} \sum_{h=0}^{H-1} \int \pi_0 \cdot \overline{\pi_0} \circ S^h d\kappa \rightarrow 0.$$

Combinatorially, we obtain

$$(55) \quad \frac{1}{H} \sum_{h=0}^{H-1} \lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} \mathbf{u}(n) \overline{\mathbf{u}}(n+h) \rightarrow 0$$

for each sequence  $(N_k)$  defining a Furstenberg system  $\kappa$ .

It follows that (55) is equivalent to the short interval behaviour (13). In other words, condition

$$\lim_{H \rightarrow \infty} \left( \lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} \left| \frac{1}{H} \sum_{h \leq H} \mathbf{u}(n+h) \right|^2 \right) = 0$$

is equivalent to

$$\lim_{H \rightarrow \infty} \frac{1}{H} \sum_{h \leq H} \left( \lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} \mathbf{u}(n) \overline{\mathbf{u}}(n+h) = 0 \right).$$

### 5.3 Orthogonality to $\mathcal{C}_{\text{DISP}(G)}$ with $G$ countable

Let  $G \subset \mathbb{S}^1$  be a countable subgroup and recall that  $\text{DISP}(G)$  stands for the (characteristic) class of discrete spectrum automorphisms whose groups of eigenvalues are contained in  $G$ . Since  $z \in \mathbb{S}^1$  is an eigenvalue of  $(Z, \mathcal{D}, \kappa, R)$  if and only if it is an eigenvalues of a subset of positive measure of ergodic components, it is not hard to see that

$$\mathcal{F}_{\text{DISP}(G)} = (\mathcal{F}_{\text{DISP}(G)})_{\text{ec}}.$$

It follows that

$$\mathbf{u} \perp \mathcal{C}_{\mathcal{F}_{\text{DISP}(G)}} \text{ if and only if } \sigma_{\pi_0, \kappa}(G) = 0,$$

i.e. the spectral measure of  $\pi_0$  has no atoms belonging to  $G$  (for each Furstenberg system  $\kappa \in V_S(\mathbf{u})$ ).

Suppose that  $e^{2\pi i \alpha} \in G$ . Consider  $\mathbf{v}(n) := e^{2\pi i n \alpha} \mathbf{u}(n)$  for  $n \in \mathbb{N}$ . Note that

$$\frac{1}{N_k} \sum_{n \leq N_k} \mathbf{v}(n) \overline{\mathbf{v}}(n+h) = e^{-2\pi i h \alpha} \frac{1}{N_k} \sum_{n \leq N_k} \mathbf{u}(n) \overline{\mathbf{u}}(n+h).$$

So, if we have a subsequence  $(N_k)$  along which both  $\frac{1}{N_k} \sum_{n \leq N_k} \delta_{S^n} \mathbf{w}$  with  $\mathbf{w} = \mathbf{u}, \mathbf{v}$  converge to  $\kappa, \kappa'$  respectively,<sup>12</sup> then

$$\sigma_{\pi_0, \kappa} = \delta_{e^{2\pi i \alpha}} * \sigma_{\pi_0, \kappa'},$$

whence

$$\sigma_{\pi_0, \kappa}(\{e^{2\pi i \alpha}\}) = 0 \text{ if and only if } \sigma_{\pi_0, \kappa'}(\{1\}).$$

By our previous subsection it follows that the latter condition is equivalent to:

$$\lim_{H \rightarrow \infty} \left( \lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} \left| \frac{1}{H} \sum_{h \leq H} \mathbf{v}(n+h) \right|^2 \right) = 0,$$

that is,

$$\lim_{H \rightarrow \infty} \left( \lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} \left| \frac{1}{H} \sum_{h \leq H} e^{2\pi i (n+h)\alpha} \mathbf{u}(n+h) \right|^2 \right) = 0$$

which is the strong  $\mathbf{u}$ -MOMO condition for the irrational rotation by  $\alpha$ .<sup>13</sup>

## 5.4 Furstenberg systems and the strong $\mathbf{u}$ -MOMO property

The following proposition helps us to exclude some measure-theoretic systems from the list of Furstenberg systems of an arithmetic function.

**Proposition 5.6.** *Let  $\mathbf{u} : \mathbb{N} \rightarrow \mathbb{C}$  be a bounded arithmetic function. Then no Furstenberg system  $\kappa \in V_S(\mathbf{u})$  has a topological model which is strongly  $\mathbf{u}$ -MOMO.*

<sup>12</sup>Note that these common sequences yield **all** Furstenberg systems for both  $\mathbf{u}$  and  $\mathbf{v}$ .

<sup>13</sup>Note that if  $f(t) = e^{2\pi i t}$  then

$$\frac{1}{b_K} \sum_{k < K} \left| \sum_{b_k \leq n < b_{k+1}} f(R_\alpha^n x_k) \mathbf{u}(n) \right| = \frac{1}{b_K} \sum_{k < K} \left| \sum_{b_k \leq n < b_{k+1}} e^{2\pi i n \alpha} \mathbf{u}(n) \right|.$$



*Proof.* Suppose  $(X_{\mathbf{u}}, \kappa, S)$  has a topological model  $(Z, \nu, R)$  which satisfies the strong  $\mathbf{u}$ -MOMO property. Let  $J : Z \rightarrow X_{\mathbf{u}}$  settles a measure-theoretic isomorphism and let  $\nu_J$  be the corresponding graph joining. We assume that  $\frac{1}{N_j} \sum_{n \leq N_j} \delta_{S^n \mathbf{u}} \rightarrow \kappa$ . From Proposition 3.1 we can find a sequence  $(z_n) \subset Z$  consisting of pieces of orbits of different points:  $\{n : Rz_n \neq z_{n+1}\} = \{b_k : k \geq 1\}$  with  $b_{k+1} - b_k \rightarrow \infty$ , and a subsequence  $(N_{j_\ell})$  such that

$$\frac{1}{N_{j_\ell}} \sum_{n \leq N_{j_\ell}} \delta_{(z_n, S^n \mathbf{u})} \rightarrow \nu_J.$$

Then, by the strong  $\mathbf{u}$ -MOMO property of  $(Z, R)$ ,

$$\int f \otimes \pi_0 d\nu_J = \lim_{\ell \rightarrow \infty} \frac{1}{b_{K_\ell}} \sum_{k < K_\ell} \left( \sum_{b_k \leq n < b_{k+1}} f(R^{n-b_k} z_{b_k}) \mathbf{u}(n) \right) = 0.$$

Hence,  $\int \mathbb{E}^{\nu_J}(f|X_{\mathbf{u}}) \pi_0 d\kappa = 0$  for each continuous  $f$  on  $Z$ , and we obtain a contradiction since  $\mathbb{E}^{\nu_J}(L^2(\nu)|X_{\mathbf{u}}) = L^2(\kappa)$ .  $\square$

**Corollary 5.7.** *Assume that for each  $(b_k)$  with  $b_{k+1} - b_k \rightarrow \infty$ ,*

$$(56) \quad \lim_{K \rightarrow \infty} \frac{1}{b_K} \sum_{k < K} \sup_{\alpha \in \mathbb{R}} \left| \sum_{b_k \leq n < b_{k+1}} \mathbf{u}(n) e^{2\pi i \alpha n} \right| = 0.$$

*Then the unipotent system  $(x, y) \mapsto (x, y + x)$  (on  $\mathbb{T}^2$ ) is not a Furstenberg system of  $\mathbf{u}$ .*

*Proof.* Since condition (56) is the strong  $\mathbf{u}$ -MOMO property of the unipotent system, the result follows from Proposition 5.6.  $\square$

**Remark 5.8.** Corollary 5.7 brings a better understanding of Problem 3.1 (due to Frantzikinakis) of the workshop [4]:

*The system  $(x, y) \mapsto (x, y + x)$  is not a Furstenberg system of the Liouville function*

*(see also slide no 6 in [17]).*

We recall that in [42] (see Theorem 1.3 therein), it is proved that

$$\lim_{M \rightarrow \infty} \limsup_{N \rightarrow \infty} \sup_{\alpha \in \mathbb{R}} \frac{1}{N} \sum_{n \leq N} \left| \frac{1}{M} \sum_{n \leq m < n+M} \lambda(m) e^{2\pi i m \alpha} \right| = 0,$$

so the sup has changed the place! The strong  $\lambda$ -MOMO property for the unipotent system remains hence open.

For the equivalence of

$$\lim_{M \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \sup_{\alpha \in \mathbb{R}} \left| \frac{1}{M} \sum_{n \leq m < n+M} \mathbf{u}(m) e^{2\pi i m \alpha} \right| = 0,$$

with (56) see the appendix in [34] - only in the arXiv version of the paper.

## 5.5 Orthogonality to $\mathcal{C}_{\text{DISP}_{ec}}$ . Proof of Corollary 1.10

In view of Corollary 2.11 (see also (23)) and Theorem A, in order to obtain  $\mathbf{u} \perp \mathcal{C}_{\text{DISP}_{ec}}$  it is sufficient and necessary to have  $\mathbf{u} \perp L^2(\mathcal{K}(\mathcal{I}_\kappa))$  for each Furstenberg system  $\kappa$  of  $\mathbf{u}$ .

By our previous results, for the class of all topological systems whose all **ergodic** measures yield discrete spectra, Sarnak and Veech conditions are equivalent. We now write the Veech condition combinatorially, i.e., we provide the proof of Corollary 1.10.

*Proof of Corollary 1.10.* By Corollary 2.11, we need to show that for each  $\kappa$  being a Furstenberg system of  $\mathbf{u}$ , we have

$$(57) \quad \int \frac{1}{H} \sum_{h \leq H} |\mathbb{E}(\pi_0 \circ S^h \cdot \bar{\pi}_0 | \mathcal{I}_\kappa)|^2 d\kappa \rightarrow 0.$$

By the von Neumann theorem,

$$\int |\mathbb{E}(\pi_0 \circ S^h \cdot \bar{\pi}_0 | \mathcal{I}_\kappa)|^2 d\kappa = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \int (\pi_0 \circ S^h \cdot \bar{\pi}_0) \circ S^n \cdot \overline{\pi_0 \circ S^h \cdot \bar{\pi}_0} d\kappa.$$

Therefore, (57) is equivalent to

$$\lim_{H \rightarrow \infty} \frac{1}{H} \sum_{h \leq H} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \int (\pi_0 \circ S^h \cdot \bar{\pi}_0) \circ S^n \cdot \overline{\pi_0 \circ S^h \cdot \bar{\pi}_0} d\kappa = 0.$$

Let  $(M_k)$  be such that  $\kappa = \lim_{k \rightarrow \infty} \frac{1}{M_k} \sum_{m \leq M_k} \delta_{S^m \mathbf{u}}$ . It follows immediately that (57) is equivalent to

$$\lim_{H \rightarrow \infty} \frac{1}{H} \sum_{h \leq H} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \lim_{k \rightarrow \infty} \frac{1}{M_k} \sum_{m \leq M_k} \mathbf{u}(m+n+h) \overline{\mathbf{u}(m+n) \mathbf{u}(m+h)} \mathbf{u}(m) = 0$$

which is precisely  $\|\mathbf{u}\|_{u^2((M_k))} = 0$ . Now, it suffices to use (27).  $\square$

**Remark 5.9.** In fact, already Frantzikinakis [17] (see slide no 10) showed that if  $\mathbf{u}$  is generic then  $\|\mathbf{u}\|_{u^2} = 0$  if and only if

$$\lim_{M \rightarrow \infty} \limsup_{N \rightarrow \infty} \frac{1}{N} \sum_{n \leq N} \sup_{\alpha \in \mathbb{R}} \left| \frac{1}{M} \sum_{n \leq m < n+M} \mathbf{u}(m) e^{2\pi i m \alpha} \right| = 0.$$

We recall that this condition is equivalent to the strong  $\mathbf{u}$ -MOMO property of the unipotent system  $(x, y) \mapsto (x, y + x)$ .

**Remark 5.10.** Note that for each (bounded)  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  satisfying  $\|\mathbf{u}\|_{u^2} = 0$  the system

$$(58) \quad (x, y) \mapsto (x, x + y) \text{ on } (\mathbb{T}^2, \text{Leb} \otimes \text{Leb})$$

cannot appear (up to isomorphism) as a Furstenberg system of  $\mathbf{u}$  (because  $\pi_0$  is orthogonal to the  $L^2(\mathcal{K}(\mathcal{I}_\kappa))$  but for the unipotent system (58) the whole system is relative Kronecker over the  $\sigma$ -algebra of invariant sets).

In particular, if  $\|\boldsymbol{\lambda}\|_{u^2} = 0$  holds for the Liouville function then (58) is not its Furstenberg system – this would answer a question by N. Frantzikinakis asked in 2016 (it is an official Problem 3.1 in [4]). However, the problem of whether  $\|\boldsymbol{\lambda}\|_{u^2} = 0$  (or more generally  $\|\boldsymbol{\lambda}\|_{u^s} = 0$ ) seems to be difficult. The best known results [43, 44] require a quantitative dependence between the parameters  $M$  and  $N$ , i.e.  $M = N^\theta$ , for arbitrary small, but fixed  $\theta > 0$ .

If  $\|\boldsymbol{\lambda}\|_{u^2} = 0$  holds then Sarnak’s conjecture holds for all (zero entropy) systems whose **ergodic** measures yield discrete spectrum. So far it is only known that Sarnak’s conjecture holds for systems whose **all** invariant measure yield discrete spectrum [15, 32, 33].

Ruling out (58) (or, more generally, nilpotent type systems) from the list of potential Furstenberg systems of  $\boldsymbol{\lambda}$  is important in view of Frantzikinakis and Host’s results [20, 21] concerning the structure of Furstenberg systems of multiplicative functions (although, for the moment, this structure is known only for the logarithmic case).

In the light of [42], it would be also interesting to know whether  $\|\mathbf{u}\|_{u^2} = 0$  holds for some classical multiplicative functions. Note that this is **not** the case for the class of aperiodic multiplicative functions defined in [42] since as shown in [28] they have the unipotent system as a Furstenberg system<sup>14</sup> (see also Remark 5.4).

## 5.6 Orthogonality to $\mathcal{C}_{\text{DISP}}$ . Averaged Chowla property for multiplicative functions

The assertion “iff” of Theorem A cannot be applied to the class  $\mathcal{C}_{\text{DISP}}$ . In this section we will show however that the assertion of this theorem holds whenever  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  satisfies the following additional property:

(\*) all rotations on the circle satisfy the strong  $\mathbf{u}$ -MOMO property.

We will need the following fact (see, e.g., [14]):

(59) each discrete spectrum automorphism is a factor of  $R_\alpha \times \text{Id}_{[0,1]}$ ,

for some ergodic rotation by  $\alpha \in G$  on a compact (Abelian) metric group  $G$ . Our key tool will be the following lemma.

**Lemma 5.11.** *Suppose that (\*) holds. Then  $R_\alpha \times \text{Id}_{[0,1]}$  satisfies the strong  $\mathbf{u}$ -MOMO property.*

<sup>14</sup>In fact, for such functions  $\mathbf{u}$  we have already  $\|\mathbf{u}\|_{u^1((N_k))} > 0$  (for some  $(N_k)$ ), see Corollary 6.5 in [28].

*Proof.* It is enough to check the strong  $\mathbf{u}$ -MOMO for functions  $F$  of the form  $\chi \otimes f$ , where  $\chi \in \widehat{G}$  and  $f \in C([0, 1])$ . We have

$$\begin{aligned} & \frac{1}{b_K} \sum_{k < K} \left| \sum_{b_k \leq n < b_{k+1}} F((R_\alpha \times \text{Id})^n(h_k, u_k)) \mathbf{u}(n) \right| \\ &= \frac{1}{b_K} \sum_{k < K} \left| \sum_{b_k \leq n < b_{k+1}} \chi((R_\alpha^n(h_k))f(u_k)) \mathbf{u}(n) \right| \\ &= O \left( \frac{1}{b_K} \sum_{k < K} \left| \sum_{b_k \leq n < b_{k+1}} \chi(n\alpha) \mathbf{u}(n) \right| \right). \end{aligned}$$

Our claim follows from (\*).  $\square$

**Theorem 5.12.** *Assume that  $\mathbf{u}$  enjoys the property (\*). Then  $\mathbf{u} \perp \mathcal{C}_{DISP}$  if and only if  $\pi_0 \perp L^2(\mathcal{K}(\kappa))$  for each  $\kappa \in V_S(\mathbf{u})$  (iff the spectral measure  $\sigma_{\pi_0}$  is continuous for each Furstenberg system  $\kappa$ ).*

*Proof.* We only need to show that  $\mathbf{u} \perp \mathcal{C}_{DISP}$  implies  $\pi_0 \perp L^2(\mathcal{K}(\kappa))$  for each  $\kappa \in V_S(\mathbf{u})$ .<sup>15</sup> Using (59), let  $p$  settle a factor map from  $R_\alpha \times \text{Id}_{[0,1]}$  acting on  $(G \times [0, 1], m_G \otimes \text{Leb})$  and  $(X_{\mathbf{u}}/\mathcal{K}(\kappa), \mathcal{K}(\kappa), \kappa|_{\mathcal{K}(\kappa)})$ . Let  $(m_G \otimes \text{Leb})_p$  stand for the corresponding graph joining and  $\rho$  for the relatively independent extension of it to a joining of  $(G \times [0, 1], m_G \otimes \text{Leb}, R_\alpha \times \text{Id})$  with  $(X_{\mathbf{u}}, \kappa, S)$ .

Now, by Proposition 3.1, the integral  $\int F \otimes \pi_0 d\rho$  can be computed using a quasi-generic sequence  $((g_n), (S^n \mathbf{u}))$ . Since, by Lemma 5.11, our topological system  $R_\alpha \times \text{Id}$  satisfies the strong  $\mathbf{u}$ -MOMO property, this integral vanishes. On the other hand, for each  $F \in C(G \times [0, 1])$ ,

$$\int F \otimes \pi_0 d\rho = \int \mathbb{E}(F|X_{\mathbf{u}}) \pi_0 d\kappa$$

and since  $\mathbb{E}^\rho(C(G \times [0, 1])|X_{\mathbf{u}})$  is dense in  $L^2(\mathcal{K}, \kappa|_{\mathcal{K}})$  (in view of the definition of  $\rho$ ), it follows that  $\pi_0 \perp L^2(\mathcal{K}(\kappa))$ .  $\square$

*Proof of Corollary 1.15.* Note that in the proof of Theorem 5.12, we have shown that our original assumption (\*) already implies the Veech condition. In particular, the Sarnak and the Veech properties are equivalent. Condition (17) is just rewriting the Wiener condition combinatorially. Finally, the last part (18) is proved in Appendix A.  $\square$

*Proof of Corollary 1.16.* By Corollary 1.15, we only need to show that irrational rotations satisfy the strong  $\mathbf{u}$ -MOMO property. This follows from the fact that irrational rotations satisfy the AOP property [1] and that the AOP property implies the strong  $\mathbf{u}$ -MOMO property [3].  $\square$

<sup>15</sup> $\mathcal{K}(\kappa)$  stands for the Kronecker factor of  $(X_{\mathbf{u}}, \kappa, S)$ .

## 6 No strong $\mathbf{u}$ -MOMO in positive entropy

In this section we discuss the problem of orthogonality to  $\mathcal{C}_{ZE}$  and the reversed problem of the **absence** of orthogonality to an **arbitrary** positive entropy systems, following some ideas from [3]. Recall that the following has been proved in [3].

**Proposition 6.1** ([3]). *Let  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  be a bounded arithmetic function. The following are equivalent:*

- (a)  $\mathbf{u} \perp \mathcal{C}_{ZE}$ .
- (b) For each  $(X, T)$  of zero entropy and  $f \in C(X)$ , (1) holds uniformly in  $x \in X$ .
- (c) Each zero entropy  $(X, T)$  satisfies the strong  $\mathbf{u}$ -MOMO property.

On the other hand, it follows from the results of Downarowicz and Serafin [11, 12] that for each  $\mathbf{u} \perp \mathcal{C}_{ZE}$  **there exists**  $(X, T)$  such that

$$(60) \quad \mathbf{u} \perp (X, T) \text{ and } (X, T) \notin \text{ZE}.$$

In fact, one can get a positive entropy system  $(X, T)$  in which for every  $f \in C(X)$  (1) holds uniformly in  $x \in X$ .

We prove however that (60) fails if orthogonality is replaced by the strong  $\mathbf{u}$ -MOMO property. To avoid technical details, we restrict ourselves to the case of an arithmetic function  $\mathbf{u}$  taking finitely many values.

**Theorem D.** *Let  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$  be an arithmetic function taking finitely many values. Assume that  $\mathbf{u} \perp \mathcal{C}_{ZE}$ . Then no positive entropy topological dynamical system satisfies the strong  $\mathbf{u}$ -MOMO property.*

### 6.1 Proof of Theorem D

We fix a bounded arithmetic function  $\mathbf{u}: \mathbb{N} \rightarrow \mathbb{C}$ . We need a series of results from [3] in some modified forms. In [3], the equivalence of certain three properties (P1), (P2) and (P3) of an ergodic measure-theoretic dynamical system  $(Z, \mathcal{B}(Z), \kappa, R)$  was proved. Condition (P1) was nothing but the strong  $\mathbf{u}$ -MOMO for **some** topological system being a model of the system given by  $\kappa$ . Instead of recalling (P2), let us formulate red its subsequence version:

- (P2') Assume that  $(X, T)$  is any topological system and let  $x \in X$ . If  $x$  is generic along  $(N_k)$  for a measure which is isomorphic (as dynamical systems) to  $\kappa$  then
- $$\lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} f(T^n x) \mathbf{u}(n) = 0 \text{ for each } f \in C(X).$$

The proof of the implication (P1)  $\implies$  (P2') is a repetition of the proof of (P1) implies (P2). In Lemma 17 in [3], we need to consider the sequence  $(N_k)$  instead of  $\mathbb{N}$  and start with  $\limsup$  along this sequence.

As a consequence of the above, we obtain the following version of Corollary 12 from [3].

**Corollary 6.2.** *Assume that  $\kappa$  is an ergodic shift-invariant measure on  $\mathbb{D}_L^{\mathbb{Z}}$ , and that there exists  $y \in \mathbb{D}_L^{\mathbb{Z}}$ , generic along  $(N_k)$  for  $\kappa$ , correlating with  $\mathbf{u}$  along  $(N_k)$ , i.e. the sequence  $(\frac{1}{N_k} \sum_{n \leq N_k} y(n)\mathbf{u}(n))$  does not go to zero. Then the strong  $\mathbf{u}$ -MOMO property fails for any uniquely ergodic model of  $(\mathbb{D}_L^{\mathbb{Z}}, \kappa, S)$ .*

Then, by repeating the proof from [3], we obtain the following form of Corollary 14 in [3].

**Corollary 6.3.** *Assume that  $y$  is generic along  $(N_k)$  for a Bernoulli measure  $\nu$ , and that  $y$  and  $\mathbf{u}$  correlate along  $(N_k)$ . Then the strong  $\mathbf{u}$ -MOMO property fails for any  $(X, T)$  with  $h(X, T) > h(\nu)$ .*

We also need the following crucial probabilistic lemma whose proof we postpone to the next subsection.

**Lemma 6.4.** *Assume that  $X = (X_n)_{n \in \mathbb{Z}}$  is a stationary process of positive entropy, taking finitely many complex values. Then for any non-trivial probability distribution  $\beta$  concentrated on a finite subset of  $\mathbb{R}$ , there exists a stationary coupling of  $X$  with a Bernoulli process  $Y = (Y_n)_{n \in \mathbb{Z}}$  of distribution  $\beta^{\otimes \mathbb{Z}}$  such that  $\mathbb{E}[X_0 Y_0] \neq \mathbb{E}[X_0] \mathbb{E}[Y_0]$ .*

We now assume that  $\mathbf{u}$  takes finitely many values and satisfies the Veech condition:  $\pi_0 \perp L^2(\Pi(\kappa))$  for each Furstenberg system  $\kappa$  of  $\mathbf{u}$ .

**Lemma 6.5.** *For each  $h > 0$  there exists a sequence  $y$ , generic for a Bernoulli measure of entropy  $h$  along some increasing sequence  $(N_k)$ , and correlating with  $\mathbf{u}$  along  $(N_k)$ .*

*Proof.* Let  $\kappa$  be a Furstenberg system of  $\mathbf{u}$ , and  $(M_\ell)$  such that  $\mathbf{u}$  is generic for  $\kappa$  along  $(M_\ell)$ . By assumption, the entropy of the stationary process defined by  $\pi_0$  under  $\kappa$  is positive. Take a real-valued Bernoulli shift of entropy  $h$  (Bernoulli measure denoted by  $\nu$ ). Using Lemma 6.4, find a joining of  $\kappa$  and  $\nu$  for which  $\pi_0$  (in  $L^2(X_{\mathbf{u}}, \kappa)$ ) is not orthogonal to  $\pi_0$  in  $L^2(\nu)$ :  $\int \pi_0 \otimes \pi_0 d\rho \neq 0$ . Now, use a subsequence version of the lifting lemma (Theorem 5.16 in [6]) to find  $y$  in the subshift defining the Bernoulli automorphism such that  $(\mathbf{u}, y)$  is generic, along a subsequence  $(N_k) = (M_{\ell_k})$ , for  $\rho$ . Then

$$0 \neq \int \pi_0 \otimes \pi_0 d\rho = \lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} \pi_0(S^n \mathbf{u}) \pi_0(S^n y) = \lim_{k \rightarrow \infty} \frac{1}{N_k} \sum_{n \leq N_k} \mathbf{u}(n) y_n$$

which means that  $\mathbf{u}$  and  $y$  correlate along  $(N_k)$ .  $\square$

Now the proof of Theorem D is a straightforward consequence of Lemma 6.5 and Corollary 6.3.

## 6.2 Proof of Lemma 6.4

Let  $X = (X_n)_{n \in \mathbb{Z}}$  be a positive entropy stationary process as in the statement of the lemma. Without loss of generality (considering its real or imaginary part), we can assume that this process takes its values in a finite subset  $\{x_1 < x_2 < \dots < x_r\}$  of  $\mathbb{R}$ . We also consider a given probability measure  $\beta$  supported on a possibly different finite subset of  $\mathbb{R}$   $\{y_1 < y_2 < \dots < y_s\}$ , which is supposed to be non trivial (i.e. not reduced to a Dirac measure). Thus we can assume that  $s \geq 2$ , and  $\beta(y_j) > 0$  for each  $1 \leq j \leq s$ . The purpose of this section is to show how we can construct a stationary coupling of  $X$  with a Bernoulli process  $Y$  whose distribution is  $\beta^{\otimes \mathbb{Z}}$ , in such a way that for each  $n \in \mathbb{Z}$ ,

$$(61) \quad \mathbb{E}[X_n Y_n] > \mathbb{E}[X_n] \mathbb{E}[Y_n].$$

We observe that the validity of the preceding inequality is unchanged if we replace  $Y_n$  by  $Y_n + C$  for a fixed  $C$ . Thus we can and we do assume without loss of generality that the probability  $\beta$  is such that  $\mathbb{E}[Y_n] = 0$ .

To construct the announced coupling, we just assume that, on the probability space where the process  $X$  is defined, we also have an i.i.d. process  $V = (V_n)_{n \in \mathbb{Z}}$  such that

- each  $V_n$  is uniformly distributed on  $[0, 1]$ ,
- $V$  is independent of  $X$ .

The construction will be divided into two steps: first we construct an auxiliary (uniform i.i.d.) process  $U$  and then we use it to construct  $Y$  which satisfies the assertion of Lemma 6.4.

### Step 1: uniform i.i.d. process $U$

For  $n \in \mathbb{Z}$  and  $j \in \{1, \dots, r\}$ , we consider the random variable  $P_{j,n}$  defined by

$$P_{j,n} := \mathbb{P}(X_n = x_j \mid (X_m)_{m \leq n-1}).$$

When  $j$  is fixed,  $(P_{j,n})_{n \in \mathbb{Z}}$  is a stationary process. On the other hand, if we fix  $n$ , then  $(P_{1,n}, \dots, P_{r,n})$  is the conditional distribution of  $X_n$  given  $(X_m)_{m \leq n-1}$ , in particular we have almost surely  $0 \leq P_{j,n} \leq 1$ , and

$$\sum_{j=1}^r P_{j,n} = 1.$$

This allows us to define a random partition of  $[0, 1[$  into disjoint subintervals  $I_{1,n}, \dots, I_{r,n}$  where for each  $j$ ,  $I_{j,n}$  is the interval of length  $P_{j,n}$  defined by

$$I_{j,n} := \left[ \sum_{1 \leq i \leq j-1} P_{i,n} ; \sum_{1 \leq i \leq j} P_{i,n} \right[.$$

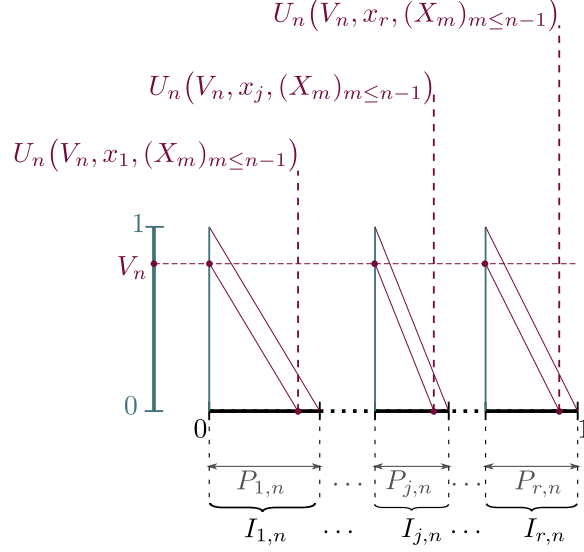


Figure 1: Definition of  $U_n$

Then we can define the random variable  $U_n$  by

$$U_n := \sum_{j=1}^r \mathbf{1}_{X_n=x_j} \left( \sum_{1 \leq i \leq j-1} P_{i,n} + V_n P_{j,n} \right).$$

Informally, if  $X_n = x_j$ , we pick  $U_n$  uniformly at random (using  $V_n$ ) inside  $I_{j,n}$  (see Figure 1). Therefore,

$$\mathcal{L}(U_n | (X_m)_{m \leq n-1}, (V_m)_{m \leq n-1}) = \mathcal{U}_{[0,1]},$$

i.e., it is uniform on  $[0, 1]$ . But all  $U_m$ ,  $m \leq n-1$ , are measurable with respect to  $(X_m)_{m \leq n-1}$  and  $(V_m)_{m \leq n-1}$ , thus we also have

$$(62) \quad \mathcal{L}(U_n | (U_m)_{m \leq n-1}) = \mathcal{U}_{[0,1]} \text{ and } \mathcal{L}(U_n) = \mathcal{U}_{[0,1]}.$$

Indeed, this is just the application of the tower property of conditional expectations: to obtain the left equality, notice that for any measurable  $A \subset [0, 1]$ , we have

$$\begin{aligned} & \mathbb{P}(U_n \in A | (U_m)_{m \leq n-1}) \\ &= \mathbb{E} \left[ \underbrace{\mathbb{P}(U_n \in A | (X_m)_{m \leq n-1}, (V_m)_{m \leq n-1})}_{\text{Leb}(A)} | (U_m)_{m \leq n-1} \right] \\ &= \text{Leb}(A). \end{aligned}$$

Moreover, it also follows from (62) that  $U$  is i.i.d.



Note that by construction,  $U_n$  is a measurable function of  $V_n$ ,  $X_n$  and  $(X_m)_{m \leq n-1}$ , which we abusively write as

$$U_n = U_n(V_n, X_n, (X_m)_{m \leq n-1}).$$

Moreover, whenever we fix realizations  $\xi$  of  $(X_m)_{m \leq n}$  and  $v$  of  $V_n$  then  $U_n$  as a function of its second argument is *increasing*:

$$(63) \quad U_n(v, x_{j_1}, \xi) < U_n(v, x_{j_2}, \xi), \text{ whenever } x_{j_1} < x_{j_2}.$$

**Step 2: process  $Y$  as a function of  $U$**

We want to define  $Y_n$  for a given  $n \in \mathbb{Z}$ . We use another partition of  $[0, 1[$  into subintervals, according to the probability distribution  $\beta$  intended for  $Y_n$ : for  $1 \leq k \leq s$ , set  $\beta_k := \beta(y_k)$  and define the interval  $J_k := [\beta_1 + \dots + \beta_{k-1}; \beta_1 + \dots + \beta_k[$ . Then we simply define  $Y_n$  as a function of  $U_n$  by setting

$$Y_n := \sum_{k=1}^s y_k \mathbb{1}_{J_k}(U_n).$$

It follows by the choice of the intervals  $J_k$  and by  $\mathcal{L}(U_n) = \mathcal{U}_{[0,1]}$  that  $Y_n$  is distributed according to  $\beta$ . Moreover, by the independence of  $U$ , we have the independence of  $Y$ . Thus,  $Y$  is a Bernoulli process with distribution  $\beta^{\otimes \mathbb{Z}}$ .

It remains to prove the announced inequality (61). Observe that  $Y_n$  is, like  $U_n$ , constructed as a measurable function of  $V_n$ ,  $X_n$  and  $(X_m)_{m \leq n-1}$ , which we also abusively write as

$$Y_n = Y_n(V_n, X_n, (X_m)_{m \leq n-1}).$$

Since  $Y_n$  is a non-decreasing function of  $U_n$ , we get from (63) that for a fixed realization  $\xi$  of  $(X_m)_{m \leq n-1}$  and  $v$  of  $V_n$ , we have for  $1 \leq j_1 < j_2 \leq r$

$$(64) \quad Y_n(v, x_{j_1}, \xi) < Y_n(v, x_{j_2}, \xi)$$

and it follows that the map

$$x \in A \mapsto \mathbb{E}[Y_n \mid (X_m)_{m \leq n-1} = \xi, X_n = x]$$

is non-decreasing. Moreover, by the construction of  $Y$ , we have

$$\mathcal{L}(Y_n \mid (X_m)_{m \leq n-1}) = \mathcal{L}(Y_n) = \beta,$$

whence

$$(65) \quad \mathbb{E}[Y_n \mid (X_m)_{m \leq n-1}] = \mathbb{E}[Y_n] = 0.$$

Thus there exists  $j_0 \in \{1, \dots, r\}$  (depending on  $\xi$ ) such that

$$(66) \quad \begin{aligned} \mathbb{E}[Y_n \mid (X_m)_{m \leq n-1} = \xi, X_n = x_j] &\leq 0 \text{ for } 1 \leq j \leq j_0, \\ \mathbb{E}[Y_n \mid (X_m)_{m \leq n-1} = \xi, X_n = x_j] &> 0 \text{ for } j_0 + 1 \leq j \leq r. \end{aligned}$$

We then have, using (65) and (66),

$$\begin{aligned}
\mathbb{E}[X_n Y_n | (X_m)_{m \leq n-1} = \xi] &= \mathbb{E}[(X_n - x_{j_0}) Y_n | (X_m)_{m \leq n-1} = \xi] \\
&= \sum_{j=1}^{j_0} (x_j - x_{j_0}) \mathbb{E}[Y_n | (X_m)_{m \leq n-1} = \xi, X_n = x_j] \\
(67) \quad &+ \sum_{j=j_0+1}^r (x_j - x_{j_0}) \mathbb{E}[Y_n | (X_m)_{m \leq n-1} = \xi, X_n = x_j] \\
&\geq 0.
\end{aligned}$$

Now, we claim that the announced result is a consequence of the following lemma.

**Lemma 6.6.** *If the realization  $\xi$  of  $(X_m)_{m \leq n-1}$  is such that the conditional distribution  $\mathcal{L}(X_n | (X_m)_{m \leq n-1} = \xi)$  is non-trivial, then*

$$\mathbb{E}[X_n Y_n | (X_m)_{m \leq n-1} = \xi] > 0.$$

Indeed, since  $X$  has positive entropy,  $\mathcal{L}(X_n | (X_m)_{m \leq n-1})$  is non-trivial with positive probability, and thus we can conclude that

$$\mathbb{E}[X_n Y_n] = \mathbb{E}[\mathbb{E}[X_n Y_n | (X_m)_{m \leq n-1}]] > 0.$$

*Proof of Lemma 6.6.* We fix a realization  $\xi$  of  $(X_m)_{m \leq n-1}$  such that the conditional distribution  $\mathcal{L}(X_n | (X_m)_{m \leq n-1} = \xi)$  is non-trivial. Then the random variables  $P_{j,n}$  and the intervals  $I_{j,n}$  are fixed, because their values only depend on  $\xi$ . Setting

$$\begin{aligned}
j_1 &:= \min\{j \in \{1, \dots, r\} : P_{j,n} > 0\}, \\
\text{and } j_2 &:= \max\{j \in \{1, \dots, r\} : P_{j,n} > 0\},
\end{aligned}$$

we have  $j_1 < j_2$ . Moreover the intervals  $I_{j_1,n}$  and  $I_{j_2,n}$  are respectively of the form  $[0, P_{j_1,n}[$  and  $[1 - P_{j_2,n}, 1[$ , with  $0 < P_{j_1,n} \leq 1 - P_{j_2,n} < 1$ . We now discuss according to the relative position of the interval  $I_{j_2,n}$  with respect to the interval  $J_1$  (used to define  $Y_n$ ).

**Case 1:**  $J_1 \cap I_{j_2,n} = \emptyset$  (see Figure 2). Then we have

$$(68) \quad \mathbb{P}(Y_n = y_1 | (X_m)_{m \leq n-1} = \xi, X_n = x_{j_2}) = 0,$$

whereas

$$(69) \quad \mathbb{P}(Y_n = y_1 | (X_m)_{m \leq n-1} = \xi, X_n = x_{j_1}) > 0.$$

Moreover, notice that (68) is equivalent to

$$(70) \quad \mathbb{P}(Y_n > y_1 | (X_m)_{m \leq n-1} = \xi, X_n = x_{j_2}) = 1,$$

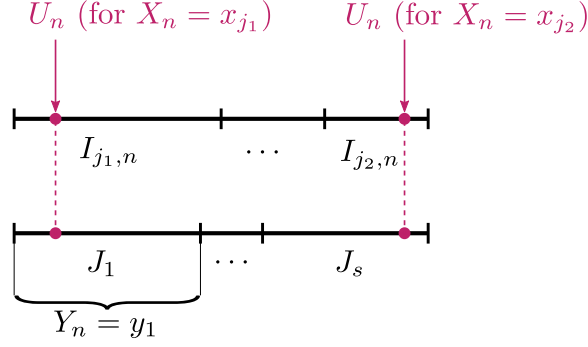


Figure 2: Case 1 ( $J_1 \cap I_{j_2, n} = \emptyset$ )

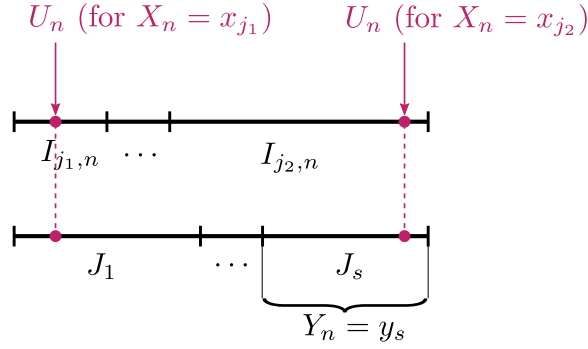


Figure 3: Case 2 ( $J_1 \cap I_{j_2, n} \neq \emptyset$ )

It follows from (69) and (70) that there exists a  $V_n$ -measurable event  $A$  of positive probability such that, on  $A$ ,

$$Y_n(V_n, x_{j_2}, \xi) > y_1 = Y_n(V_n, x_{j_1}, \xi).$$

Remembering (64), we get

$$(71) \quad \mathbb{E}[Y_n | (X_m)_{m \leq n-1} = \xi, X_n = x_{j_2}] > \mathbb{E}[Y_n | (X_m)_{m \leq n-1} = \xi, X_n = x_{j_1}].$$

**Case 2:**  $J_1 \cap I_{j_2, n} \neq \emptyset$  (see Figure 3). Then  $I_{j_1, n} \subset J_1$  and  $I_{j_1, n} \cap J_s = \emptyset$ . It follows that

$$\mathbb{P}(Y_n = y_s | (X_m)_{m \leq n-1} = \xi, X_n = x_{j_1}) = 0,$$

whereas

$$\mathbb{P}(Y_n = y_s | (X_m)_{m \leq n-1} = \xi, X_n = x_{j_2}) > 0.$$

In this case, we get a  $V_n$ -measurable event  $A$  of positive probability such that, on  $A$ ,

$$Y_n(V_n, x_{j_2}, \xi) = y_s > Y_n(V_n, x_{j_1}, \xi),$$

and as before we conclude that (71) holds.

Now, since (71) always holds, and since

$$\mathbb{E}[Y_n | (X_m)_{m \leq n-1} = \xi] = 0 = \sum_{j=1}^r P_{j,n} \mathbb{E}[Y_n | (X_m)_{m \leq n-1} = \xi, X_n = x_j],$$

we deduce that

$$\mathbb{E}[Y_n | (X_m)_{m \leq n-1} = \xi, X_n = x_{j_2}] > 0.$$

It follows that in the sum (67), at least the term corresponding to  $j = j_2$  is positive, and this yields

$$\mathbb{E}[X_n Y_n | (X_m)_{m \leq n-1} = \xi] > 0.$$

□

# Appendix

## A From averaged double to averaged multiple correlations

This section follows some arguments from [42].

**Remark A.1.** In the proof below we will use the following standard fact: let  $(x(n))$  be a sequence of complex number bounded by 1. Then

$$\sum_{m \leq M} |x(m)| = o(M)$$

is equivalent to

$$\sum_{m \leq M} |x(m)|^2 = o(M).$$

The little “o” is uniform with respect to  $M$ .<sup>16</sup>

We have the following general lemma:

**Lemma A.2.** *Let  $(N_\ell)_{\ell \in \mathbb{N}}$  be a sequence of natural numbers. For  $k \in \mathbb{N}$  let  $a, b_1, \dots, b_k: \mathbb{N} \rightarrow \mathbb{C}$  be sequences bounded by 1. Assume that  $a$  satisfies*

$$(72) \quad \lim_{H \rightarrow \infty} \frac{1}{H} \sum_{h \leq H} \lim_{\ell \rightarrow \infty} \frac{1}{N_\ell} \left| \sum_{n \leq N_\ell} a(n) a(n+h) \right| = 0.$$

Then

$$(73) \quad \lim_{H \rightarrow \infty} \frac{1}{H^k} \sum_{h_1, \dots, h_k \leq H} \lim_{\ell \rightarrow \infty} \frac{1}{N_\ell} \left| \sum_{n \leq N_\ell} a(n) \prod_{i=1}^k b_i(n+h_i) \right| = 0.$$

*Proof.* Notice first that (73) can be rewritten as the following: for every  $\varepsilon > 0$ , there exists  $H_\varepsilon$  such that for  $H > H_\varepsilon$  and all  $\ell$  sufficiently large (depending on  $H$ ), we have

$$A := \frac{1}{H^k} \sum_{h_1, \dots, h_k \leq H} \frac{1}{N_\ell} \left| \sum_{n \leq N_\ell} a(n) \prod_{i=1}^k b_i(n+h_i) \right| < \varepsilon.$$

<sup>16</sup>If  $\varepsilon := \frac{1}{M} \sum_{m \leq M} |x(m)|^2$  then by Markov's inequality

$$\frac{1}{M} |\{m \leq M : |c_m|^2 \geq \varepsilon^{1/2}\}| \leq \frac{1}{\varepsilon^{1/2}} \cdot \varepsilon = \varepsilon^{1/2}$$

and then

$$\frac{1}{M} \sum_{m \leq M} |x(m)| = \frac{1}{M} \sum_{m \leq M, |x(m)| \geq \varepsilon^{1/4}} |x(m)| + \frac{1}{M} \sum_{m \leq M, |x(m)| < \varepsilon^{1/4}} |x(m)| \leq \varepsilon^{1/2} + \varepsilon^{1/4}.$$

Now, notice that for any  $H, N_\ell, H'$  and any  $h' \leq H'$ , by shifting the summation over  $n \leq N_\ell$  by  $h'$  (for every fixed choice of  $h_1, \dots, h_k$ ), we have

$$\begin{aligned} & \sum_{h_1, \dots, h_k \leq H} \left| \sum_{n \leq N_\ell} a(n) \prod_{i=1}^k b_i(n + h_i) \right| = \\ & \sum_{h_1, \dots, h_k \leq H} \left| \sum_{h' \leq n \leq N_\ell + h'} a(n) \prod_{i=1}^k b_i(n + h_i) \right| + O(H^k \cdot h') = \\ & \sum_{h_1, \dots, h_k \leq H} \left| \sum_{n \leq N_\ell} a(n + h') \prod_{i=1}^k b_i(n + h_i + h') \right| + O(H^k \cdot h') \end{aligned}$$

and  $O(H^k \cdot h') = O(H^k \cdot H')$ . Notice that as  $h_i$  is taken from  $[0, H]$ , then  $h_i + h'$  is taken from  $[h', H + h']$  (which is a small shift of  $[0, H]$  if  $h'$  is much smaller than  $H$ ). So putting  $h'$  to the summation over  $h_i$ , we get

$$\begin{aligned} & \sum_{h_1, \dots, h_k \leq H} \left| \sum_{n \leq N_\ell} a(n + h') \prod_{i=1}^k b_i(n + h_i + h') \right| = \\ & \sum_{h' \leq h_1, \dots, h_k \leq H + h'} \left| \sum_{n \leq N_\ell} a(n + h') \prod_{i=1}^k b_i(n + h_i) \right| = \\ & \sum_{h_1, \dots, h_k \leq H} \left| \sum_{n \leq N_\ell} a(n + h') \prod_{i=1}^k b_i(n + h_i) \right| + O\left((h'H^{k-1} + (h')^2 H^{k-2} + \dots + (h')^k) N_\ell\right) \end{aligned}$$

and  $O((h'H^{k-1} + (h')^2 H^{k-2} + \dots + (h')^k) N) = O((H')^k H^{k-1} N_\ell)$ . Putting the two displayed equations together we get that for every  $h' \leq H'$ ,

$$A = \frac{1}{H^k} \sum_{h_1, \dots, h_k \leq H} \frac{1}{N_\ell} \left| \sum_{n \leq N_\ell} a(n + h') \prod_{i=1}^k b_i(n + h_i) \right| + O\left(\frac{H'}{N_\ell}\right) + O\left(\frac{(H')^k}{H}\right).$$

Averaging the above equation over all  $h' \leq H'$ , we get that

$$A = \frac{1}{H^k} \sum_{h_1, \dots, h_k \leq H} \frac{1}{H'} \sum_{h' \leq H'} \frac{1}{N_\ell} \left| \sum_{n \leq N_\ell} a(n + h') G(n) \right| + O\left(\frac{H'}{N_\ell}\right) + O\left(\frac{(H')^k}{H}\right),$$

where  $G(n) = \prod_{i=1}^k b_i(n + h_i)$ .

We will now estimate

$$(74) \quad \frac{1}{H'} \sum_{h' \leq H'} \frac{1}{N_\ell^2} \left| \sum_{n \leq N_\ell} a(n + h') G(n) \right|^2 = \frac{1}{N_\ell^2} \sum_{n, n' \leq N_\ell} \left( \frac{1}{H'} \sum_{h' \leq H'} a(n + h') \overline{a(n' + h')} \right) G(n) \overline{G(n')},$$

which will be easier to handle than the above expression for  $A$  (and then use Remark A.1 to get rid of the squares). Clearly, to obtain an upper bound for (74), it suffices to obtain an upper bound for

$$(75) \quad \frac{1}{N_\ell^2} \sum_{n, n' \leq N_\ell} \frac{1}{H'} \left| \sum_{h' \leq H'} a(n+h') \overline{a(n'+h')} \right|.$$

Again, it will be easier to deal with

$$(76) \quad \frac{1}{N_\ell^2} \sum_{n, n' \leq N_\ell} \frac{1}{H'^2} \left| \sum_{h' \leq H'} a(n+h') \overline{a(n'+h')} \right|^2$$

(and use Remark A.1 to get rid of the squares). Expanding the square again we get that (76) is equal to

$$\begin{aligned} & \frac{1}{N_\ell^2} \sum_{n, n' \leq N_\ell} \frac{1}{H'^2} \sum_{h', h'' \leq H'} a(n+h') \overline{a(n'+h')} \overline{a(n+h'')} \cdot \overline{a(n'+h'')} = \\ & \frac{1}{N_\ell^2} \sum_{n, n' \leq N_\ell} \frac{1}{H'^2} \sum_{h', h'' \leq H'} a(n+h') \overline{a(n+h'')} a(n'+h'') \overline{a(n'+h')}. \end{aligned}$$

The sum in the last term by exchanging the order of summation is equal to

$$\begin{aligned} & \frac{1}{H'^2} \sum_{h', h'' \leq H'} \left| \frac{1}{N_\ell} \sum_{n \leq N_\ell} a(n+h') \overline{a(n+h'')} \right|^2 = \\ & \frac{1}{H'^2} \sum_{h', h'' \leq H'} \left| \frac{1}{N_\ell} \sum_{n \leq N_\ell} a(n) \overline{a(n+h''-h')} \right|^2 + O\left(\frac{H'^2}{N_\ell^2}\right). \end{aligned}$$

Finally, grouping according to  $h = h'' - h'$ , we get that the above is equal to

$$\begin{aligned} & \frac{1}{H'^2} \sum_{|h| \leq H'} |H' - h| \cdot \left| \frac{1}{N_\ell} \sum_{n \leq N_\ell} a(n) \overline{a(n+h)} \right|^2 + O\left(\frac{H'^2}{N_\ell^2}\right) \leq \\ & \frac{1}{H'} \cdot \sum_{|h| \leq H'} \left| \frac{1}{N_\ell} \sum_{n \leq N_\ell} a(n) \overline{a(n+h)} \right|^2 + O\left(\frac{H'^2}{N_\ell^2}\right). \end{aligned}$$

That is, the expression from (76) equals

$$(77) \quad \frac{1}{H'} \sum_{|h| \leq H'} \left| \frac{\sum_{n \leq N_\ell} a(n) \overline{a(n+h)}}{N_\ell} \right|^2 + O\left(\frac{(H')^2}{N_\ell^2}\right).$$

Now, by the assumption of our lemma, it follows that

$$\frac{1}{H'} \sum_{|h| \leq H'} \left| \frac{\sum_{n \leq N_\ell} a(n) \overline{a(n+h)}}{N_\ell} \right| = o(1),$$

which, by Remark A.1, is equivalent to

$$\frac{1}{H'} \sum_{|h| \leq H'} \left| \frac{\sum_{n \leq N_\ell} a(n) \overline{a(n+h)}}{N_\ell} \right|^2 = o(1).$$

Therefore, (77) (and, thus, also (76)) is of the order of  $o(1) + O\left(\frac{(H')^2}{N_\ell^2}\right)$ . Using again Remark A.1, we conclude that also (75) is of the same order. It follows immediately that also the order of (74) is the same.

Thus, we have proved that

$$A = o(1) + O\left(\frac{(H')^2}{N_\ell^2}\right) + O\left(\frac{H'}{N_\ell}\right) + O\left(\frac{(H')^k}{H}\right).$$

□

**Acknowledgments:** We would like to thank Tomasz Downarowicz, Nikos Frantzikinakis and Krzysztof Frączek for useful discussions on the paper. Research of the second and third authors supported by Narodowe Centrum Nauki grant UMO-2019/33/B/ST1/00364.

## References

- [1] H. El Abdalaoui, M. Lemańczyk, T. de la Rue, *Automorphisms with quasi-discrete spectrum, multiplicative functions and average orthogonality along short intervals*, Int. Math. Res. Not. IMRN 2017, no. 14, 4350-4368.
- [2] H. El Abdalaoui, J. Kułaga-Przymus, M. Lemańczyk, T. de la Rue, *The Chowla and the Sarnak conjectures from ergodic theory point of view*, Discrete and Continuous Dynam. Systems **37** (2017), 2899-2944.
- [3] H. El Abdalaoui, J. Kułaga-Przymus, M. Lemańczyk, T. de la Rue, *Möbius disjointness for models of an ergodic system and beyond*, Israel J. Math. **228** (2018), 707-751.
- [4] American Institute of Mathematics, workshop Sarnak’s Conjecture, December 2018, <http://aimpl.org/sarnakconjecture/3/>
- [5] T. Austin, *Extensions of probability-preserving systems by measurably-varying homogeneous spaces and applications*, Fundamenta Math. **210** (2010), 133–206.
- [6] V. Bergelson, T. Downarowicz, J. Vandehey, *Deterministic functions on amenable semigroups and a generalization of the Kamae-Weiss theorem on normality preservation*, arXiv 2004.02811
- [7] J. Bourgain, P. Sarnak, T. Ziegler, *Disjointness of Möbius from horocycle flows*, in “From Fourier and Number Theory to Radon Transforms and Geometry”, in memory of Leon Ehrenpreis, *Developments in Mathematics*, **28** (2012), 67–83, Springer Verlag.
- [8] J. Chinis, *Siegel zeros and Sarnak’s conjecture*, preprint May 2021.



- [9] J.-P. Conze, T. Downarowicz, J. Serafin, *Correlation of sequences and of measures, generic points for joinings and ergodicity of certain cocycles*, Trans. Amer. Math. Soc. **369** (2017), 3421–3441.
- [10] H. Davenport, *On some infinite series involving arithmetical functions. II*, Quart. J. Math. Oxford **8** (1937), 313–320.
- [11] T. Downarowicz, J. Serafin, *Almost Full Entropy Subshifts Uncorrelated to the Möbius Function*, Int. Math. Res. Not. IMRN 2019, no. 11, 3459–3472.
- [12] T. Downarowicz, J. Serafin, *A strictly ergodic, positive entropy subshift uniformly uncorrelated to the Möbius function*, Studia Math. **251** (2020), 195–206.
- [13] T. Downarowicz, B. Weiss, *Pure strictly uniform models of non-ergodic measure automorphisms*, arxiv.org/abs/2104.09604
- [14] N. Edeko, *On the isomorphism problem for non-minimal transformations with discrete spectrum*, Discrete Continuous Dynam. Systems **39** (2019), 6001–6021
- [15] S. Ferenczi, J. Kułaga-Przymus, M. Lemańczyk, *Sarnak’s Conjecture – what’s new*, in: Ergodic Theory and Dynamical Systems in their Interactions with Arithmetics and Combinatorics, CIRM Jean-Morlet Chair, Fall 2016, Editors: S. Ferenczi, J. Kułaga-Przymus, M. Lemańczyk, Lecture Notes in Mathematics **2213**, Springer International Publishing, pp. 418.
- [16] N. Frantzikinakis, *Ergodicity of the Liouville system implies the Chowla conjecture*, Discrete Analysis 2017, paper no 8, pp. 23.
- [17] N. Frantzikinakis, <https://www.cirm-math.fr/ProgWeebly/Renc1553/Frantzikinakis.pdf>,
- [18] N. Frantzikinakis, B. Host, *Weighted multiple ergodic averages and correlation sequences*, Ergodic Theory Dynam. Systems **38** (2018), 81–142.
- [19] N. Frantzikinakis, *An averaged Chowla and Elliott conjecture along independent polynomials*, Int. Math. Res. Not. IMRN 2018, no. 12, 3721–3743.
- [20] N. Frantzikinakis, B. Host, *The logarithmic Sarnak conjecture for ergodic weights*, Annals Math. (2), **187**(3) (2018), 869–931.
- [21] N. Frantzikinakis, B. Host, *Furstenberg systems of bounded multiplicative functions and applications*, Int. Math. Res. Not. IMRN 2021, no. 8, 6077–6107.
- [22] H. Furstenberg, *Disjointness in ergodic theory, minimal sets, and a problem in Diophantine approximation*, Math. Systems Theory **1**, 1–49.
- [23] H. Furstenberg, *Recurrence in ergodic theory and combinatorial number theory*, Princeton University Press, Princeton, N.J., 1981, M. B. Porter Lectures.
- [24] E. Glasner, *On the multipliers of  $W^\perp$* , Ergodic Theory Dynam. Systems **14** (1994), 129–140.
- [25] E. Glasner, *Ergodic Theory via Joinings*, American Mathematical Society, Providence, 2003.
- [26] A. Gomilko, D. Kwietniak, M. Lemańczyk, *Sarnak’s conjecture implies the Chowla conjecture along a subsequence*, in: Ergodic theory and dynamical systems in their interactions with arithmetics and combinatorics, 237–247, Lecture Notes in Math., **2213**, Springer, Cham, 2018.

- [27] A. Gomilko, M. Lemańczyk, T. de la Rue, *Möbius orthogonality in density for zero entropy dynamical systems*, Pure and Applied Functional Analysis **5** (2020), 1357-1376.
- [28] A. Gomilko, M. Lemańczyk, T. de la Rue, *On Furstenberg systems of aperiodic multiplicative functions of Matomäki, Radziwiłł and Tao*, arXiv:2006.09958
- [29] G. Hansel, *Strict Uniformity in Ergodic Theory*, Math. Z. **135** (1974), 221-248.
- [30] B. Host, B. Kra, *Nonconventional ergodic averages and nilmanifolds*, Annals Math. (2) **161** (2005), 397-488.
- [31] B. Host, B. Kra, *Nilpotent structures in ergodic theory*, Mathematical Surveys and Monographs, 236. American Mathematical Society, Providence, RI, 2018.
- [32] W. Huang, Z. Wang, X. Ye, *Measure complexity and Möbius disjointness*, Advances Math. **347** (2019), 827 – 858.
- [33] W. Huang, Z. Wang, and G. Zhang, *Möbius disjointness for topological models of ergodic systems with discrete spectrum*, J. Modern Dyn. **14** (2019), 277-290.
- [34] A. Kanigowski, M. Lemańczyk, C. Ulcigrai, *On disjointness of some parabolic flows*, Inventiones Math. **221** vol. 1 (2020), 1-111; see the version arXiv:1810.11576
- [35] I. Kátai, *A remark on a theorem of H. Daboussi*, Acta Math. Hungar. **47** (1- 2) (1986), 223-225.
- [36] J.C. Kieffer, M. Rahe, *Selecting universal partitions in ergodic theory*, Annals of Probability **9** (1981), 705-709.
- [37] J. Kułaga-Przymus, M. Lemańczyk, *Sarnak’s conjecture from the ergodic theory point of view*, to appear in Encyclopedia for Complexity, arXiv:2009.04757
- [38] M. Lemańczyk, F. Parreau, *Rokhlin extensions and lifting disjointness*, Ergodic Theory Dynam. Systems **23** (2003), 1525-1550.
- [39] M. Lemańczyk, F. Parreau, J.-P. Thouvenot, *Gaussian automorphisms whose ergodic self-joinings are Gaussian*, Fundamenta Math. **164** (2000), 253-293.
- [40] E. Lesigne, B. Rittaud, T. de la Rue. *Weak disjointness of measure preserving dynamical systems*, Ergodic Theory Dynam. Systems **23** (2003), 1173-1198.
- [41] K. Matomäki, M. Radziwiłł, *Multiplicative functions in short intervals*, Annals Math. (2) **183** (2016), 1015-1056.
- [42] K. Matomäki, M. Radziwiłł, T. Tao, *An averaged form of Chowla’s conjecture*, Algebra & Number Theory **9** (2015), 2167-2196.
- [43] K. Matomäki, M. Radziwiłł, T. Tao, *Fourier uniformity of bounded multiplicative functions in short intervals on average*, Invent. Math. **220** (1), 1-58, 2020.
- [44] K. Matomäki, M. Radziwiłł, T. Tao, J. Teravainen, T. Ziegler, *Higher uniformity of bounded multiplicative functions in short intervals on average*, arXiv:2007.15644.
- [45] T. de la Rue, *Notes on Austin’s multiple ergodic theorem*, <http://hal.archives-ouvertes.fr/hal-00400975/>

- [46] P. Sarnak, *Three lectures on the Möbius function, randomness and dynamics*, Web-page, <http://publications.ias.edu/sarnak/>.
- [47] B. Seward, *Positive entropy actions of countable groups factor onto Bernoulli shifts*, arXiv arXiv:1804.05269
- [48] F. Takens, E. Verbitsky, *Renyi entropies of aperiodic dynamical systems*, Israel J. Math. **127** (2002), 279–302.
- [49] T. Tao, *The Chowla and the Sarnak conjecture*, What’s new, <http://terrytao.wordpress.com/2012/10/14/the-chowla-conjecture-and-the-sarnak-conjecture/>.
- [50] T. Tao, *Equivalence of the logarithmically averaged Chowla and Sarnak conjectures*, Number Theory – Diophantine Problems, Uniform Distribution and Applications: Festschrift in Honour of Robert F. Tichy’s 60th Birthday (C. Elsholtz and P. Grabner, eds.), Springer International Publishing, Cham, 2017, pp. 391–421.
- [51] T. Tao, *The logarithmically averaged and non-logarithmically averaged Chowla conjectures*, <https://terrytao.wordpress.com/2017/10/20/the-logarithmically-averaged-and-non-logarithmically-averaged-chowla-conjectures/>
- [52] W.A. Veech, *Möbius Dynamics*, unpublished lecture notes, Spring semester 2016.

Department of Mathematics, The Maryland University  
adkanigowski@gmail.com

Faculty of Mathematics and Computer Science  
Nicolaus Copernicus University, Toruń, Poland  
joasiak@mat.umk.pl, mlem@mat.umk.pl

Laboratoire de Mathématiques Raphaël Salem, CNRS – Université de Rouen Normandie  
Avenue de l’Université – 76801 Saint Étienne du Rouvray, France  
Thierry.de-la-Rue@univ-rouen.fr