Annals of Mathematics **187** (2018), 869–931
https://doi.org/10.4007/annals.2018.187.3.6

The logarithmic Sarnak conjecture for ergodic weights

By Nikos Frantzikinakis and Bernard Host

Abstract

The Möbius disjointness conjecture of Sarnak states that the Möbius function does not correlate with any bounded sequence of complex numbers arising from a topological dynamical system with zero topological entropy. We verify the logarithmically averaged variant of this conjecture for a large class of systems, which includes all uniquely ergodic systems with zero entropy. One consequence of our results is that the Liouville function has super-linear block growth. Our proof uses a disjointness argument, and the key ingredient is a structural result for measure preserving systems naturally associated with the Möbius and the Liouville function. We prove that such systems have no irrational spectrum and their building blocks are infinite-step nilsystems and Bernoulli systems. To establish this structural result we make a connection with a problem of purely ergodic nature via some identities recently obtained by Tao. In addition to an ergodic structural result of Host and Kra, our analysis is guided by the notion of strong stationarity that was introduced by Furstenberg and Katznelson in the early 90's and naturally plays a central role in the structural analysis of measure preserving systems associated with multiplicative functions.

1. Introduction and main results

1.1. *Main results related to the Sarnak conjecture*. Let $\boldsymbol{\lambda}\colon \mathbb{N} \to \{-1,1\}$ be the Liouville function, which is defined to be 1 on positive integers with an even number of prime factors, counted with multiplicity, and -1 elsewhere. We extend $\boldsymbol{\lambda}$ to the integers in an arbitrary way, for example by letting $\boldsymbol{\lambda}(-n) = \boldsymbol{\lambda}(n)$ for negative $n \in \mathbb{Z}$ and $\boldsymbol{\lambda}(0) = 0$. The Möbius function $\boldsymbol{\mu}$ is equal to $\boldsymbol{\lambda}$ on integers that are not divisible by any square number and is 0 otherwise.

It is widely believed that the values of the Liouville function and the non-zero values of the Möbius function fluctuate between -1 and 1 in such a random way that forces non-correlation with any "reasonable" sequence of

Keywords: Möbius function, Liouville function, Sarnak conjecture, Chowla conjecture
AMS Classification: Primary: 11N37; Secondary: 37A45.

complex numbers. This rather vague principle is referred to as the "Möbius randomness law" (see [41, §13.1]) and is often used to give heuristic asymptotics for various sums over primes (see [61] for examples). The class of "reasonable" sequences is expected to include all bounded "low complexity" sequences, and in this direction a precise conjecture that uses the language of dynamical systems was formulated by Sarnak in [60], [59]:

Conjecture (Sarnak). *Let (Y, R) be a topological dynamical system*[1] *with zero topological entropy. Then for every $g \in C(Y)$ and $y \in Y$, we have*

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} g(R^n y)\, \boldsymbol{\mu}(n) = 0.$$

This is a fundamental and difficult problem, and there is a long list of partial results that cover a variety of dynamical systems (see Section 1.3). The goal of this article is to verify the conjecture of Sarnak for a large class of dynamical systems (Y, R), by exploiting mostly the structure of measure preserving dynamical systems generated by the Möbius and the Liouville function rather than the structure of the topological dynamical system (Y, R) for which we have limited information. The price to pay is that we have to restrict to logarithmic averages rather than the more standard Cesàro averages.

We give two variants of our main result; the first imposes a global condition on the topological dynamical system:

Theorem 1.1. *Let (Y, R) be a topological dynamical system with zero topological entropy and at most countably many ergodic invariant measures. Then for every $y \in Y$ and every $g \in C(Y)$, we have*

$$\lim_{N\to\infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{g(R^n y)\, \boldsymbol{\mu}(n)}{n} = 0. \tag{1}$$

Moreover, a similar statement holds with the Liouville function $\boldsymbol{\lambda}$ in place of $\boldsymbol{\mu}$.

Remark. In particular, our result applies if the system (Y, R) has zero topological entropy and is uniquely ergodic.

A rather surprising consequence of the previous result is a seemingly unrelated statement about the block complexity $P_{\boldsymbol{\lambda}}(n)$ of the Liouville function, which is defined to be the number of sign patterns of size n that are taken by consecutive values of the Liouville function. (See Section 1.2 for a more formal definition.) Since the Liouville function is not periodic (because $\lambda(2n) = -\lambda(n)$), it follows from [53] that $P_{\boldsymbol{\lambda}}(n) \geq n + 1$ for every $n \in \mathbb{N}$. Moreover, in [51, Prop. 2.9] it was shown that $P_{\boldsymbol{\lambda}}(n) \geq n + 5$ for every $n \geq 3$

[1] Meaning that Y is a compact metric space and $R\colon Y \to Y$ is a homeomorphism.

and that these $n+5$ sign patterns are taken on a set of positive upper density of starting points. The Chowla conjecture predicts that $P_{\boldsymbol{\lambda}}(n) = 2^n$ for every $n \in \mathbb{N}$; equivalently, all possible sign patterns of size n are taken by the Liouville function. But we are far from being able to verify this. In fact, it was not even known that $P_{\boldsymbol{\lambda}}(n)$ has super-linear growth, meaning, $\lim_{n\to\infty} P_{\boldsymbol{\lambda}}(n)/n = \infty$. We verify this property:

THEOREM 1.2. *The Liouville function has super-linear block growth.*

Remark. In fact, we prove something stronger. If $a\colon \mathbb{N} \to \mathbb{C}$ takes finitely many values and has linear block growth, then the logarithmic averages of $a(n)\,\boldsymbol{\lambda}(n)$ are 0. It follows that even if we modify the values of $\boldsymbol{\lambda}$ on a set of logarithmic density 0, using values taken from a finite set of real numbers, then the new sequence still has super-linear block growth.

Theorem 1.2 is deduced from Theorem 1.1 in Section 7.

Another variant of our main result assumes genericity of the point defining the weight sequence for a zero entropy system that has at most countably many ergodic components:

THEOREM 1.3. *Let (Y, R) be a topological dynamical system and $y \in Y$ be generic for a measure with zero entropy and at most countably many ergodic components. Then for every $g \in C(Y)$, we have*

$$\lim_{N\to\infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{g(R^n y)\,\boldsymbol{\mu}(n)}{n} = 0. \tag{2}$$

Moreover, a similar statement holds with the Liouville function $\boldsymbol{\lambda}$ in place of $\boldsymbol{\mu}$.

Genericity of $y \in Y$ for a Borel probability measure ν on Y means that for every $f \in C(Y)$, we have $\lim_{N\to\infty} \frac{1}{N}\sum_{n=1}^{N} f(R^n y) = \int f\,d\nu$. Our assumption is that the induced system (Y, ν, R) has zero entropy and at most countably many ergodic components.

Remarks. • A straightforward adaptation of our argument shows that the conclusion of Theorem 1.3 holds for those $y \in Y$ that satisfy the following property: for any sequence $(N_k)_{k\in\mathbb{N}}$ with $N_k \to \infty$ along which y is quasi-generic for logarithmic averages for some measure ν (meaning, $\lim_{k\to\infty} \frac{1}{\log N_k}\sum_{n=1}^{N_k} \frac{f(R^n y)}{n} = \int f\,d\nu$ for every $f \in C(Y)$), the system (Y, ν, R) has zero entropy and countably many ergodic components.

• See Section 1.4 for an example of a topological system and a point that is generic for a zero entropy system with uncountably many ergodic components; in this case our result does not apply.

If the ergodic components of the measure in the statement of Theorem 1.3 are assumed to be totally ergodic, then we get a much stronger conclusion:

THEOREM 1.4. *Let (Y,R) be a topological dynamical system and $y \in Y$ be generic for a measure ν with zero entropy and at most countably many ergodic components, all of which are totally ergodic. Then for every $g \in C(Y)$ with $\int g\,d\nu = 0$, we have*

$$\lim_{N\to\infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{g(R^n y)\,\prod_{j=1}^{\ell} \boldsymbol{\mu}(n+h_j)}{n} = 0 \tag{3}$$

for all $\ell \in \mathbb{N}$ and $h_1,\dots,h_\ell \in \mathbb{Z}$. Moreover, a similar statement holds with the Liouville function $\boldsymbol{\lambda}$ in place of $\boldsymbol{\mu}$.

Remarks. • For $\ell = 2$ and all odd values of ℓ, the conclusion holds even if we omit the hypothesis $\int g\,d\nu = 0$ assuming that $h_1 \neq h_2$ when $\ell = 2$. Indeed, if g is constant, then (3) holds for $\ell = 2$ by [62] and for odd values of ℓ by [64]. By adding and subtracting a constant we can thus reduce to the zero integral case.

• A variant similar to Theorem 1.1 can be proved in the same way: the conclusion of Theorem 1.4 holds for every $y \in Y$ if (Y,R) has zero entropy and at most countably many ergodic invariant measures assuming in addition that they are all totally ergodic.

• The remark following Theorem 1.3 is also valid in this case if we assume in addition that the ergodic components of (Y,ν,R) are totally ergodic.

Theorem 1.4 is new even in the case where R is given by an irrational rotation on $\mathbb{T}$ and $g(t) := \mathrm{e}^{2\pi i t}$, $t \in \mathbb{T}$. In this case we have $g(R^n 0) = \mathrm{e}^{2\pi i n\alpha}$, $n \in \mathbb{N}$, for some irrational α, and we get the following result as a consequence:

COROLLARY 1.5. *Let $\alpha \in \mathbb{R}$ be irrational. Then*

$$\lim_{N\to\infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{\mathrm{e}^{2\pi i n\alpha}\,\prod_{j=1}^{\ell} \boldsymbol{\mu}(n+h_j)}{n} = 0 \tag{4}$$

for all $\ell \in \mathbb{N}$ and $h_1,\dots,h_\ell \in \mathbb{Z}$. Moreover, a similar statement holds with the Liouville function $\boldsymbol{\lambda}$ in place of $\boldsymbol{\mu}$.

Remarks. • For $\ell = 1$, the result is well known and follows from classical methods of Vinogradov. But even for $\ell = 2$, the result is new.

• More generally, if we apply Theorem 1.4 for R given by appropriate totally ergodic affine transformations on a torus with the Haar measure (as in [25, §3.3]), we get that (4) holds with $(\mathrm{e}^{2\pi i n\alpha})_{n\in\mathbb{N}}$ replaced by any sequence of the form $(\mathrm{e}^{2\pi i P(n)})_{n\in\mathbb{N}}$, where $P \in \mathbb{R}[t]$ has an irrational non-constant coefficient.

It is straightforward to adapt our arguments in order to strengthen the conclusion in Theorems 1.1, 1.3, and 1.4 replacing $\lim_{N\to\infty} \frac{1}{\log N} \sum_{n=1}^{N}$ by

$$\lim_{N/M\to\infty} \frac{1}{\log(N/M)} \sum_{n=M}^{N}.$$

1.2. *Proof strategy and a key structural result.* A brief description of the proof strategy of Theorem 1.4 is as follows (Theorems 1.1 and 1.3 are proved similarly): In the case where the system (Y, ν, R) is totally ergodic (the more general case can be treated similarly), we first reinterpret the result as a statement in ergodic theory about the disjointness of two measure preserving systems. The first is what we call a Furstenberg system of the Möbius (or the Liouville) function. Roughly speaking, it is defined on the sequence space $X := \{-1, 0, 1\}^{\mathbb{Z}}$ with the shift transformation, by a measure that assigns to each cylinder set $\{x \in X : x(j) = \epsilon_j, j = -m, \dots, m\}$ value equal to the logarithmic density of the set $\{n \in \mathbb{N} : \boldsymbol{\mu}(n + j) = \epsilon_j, j = -m \dots, m\}$, where $\epsilon_{-m}, \dots, \epsilon_m \in \{-1, 0, 1\}$ and $m \in \mathbb{N}$. (We restrict to sequences of intervals along which all these densities exist.) The precise definition is given in Section 3.2 and is motivated by analogous constructions made by Furstenberg in [26]. The second system is an arbitrary totally ergodic system with zero entropy. In order to prove that these two systems are disjoint, we have to understand in some fine detail the structure of all possible Furstenberg systems of the Möbius and the Liouville function. Our main structural result is the following (see Sections 2 and 3.2 and Appendix A.3 for the definition of the notions involved):

THEOREM 1.6 (Structural result). *A Furstenberg system of the Möbius or the Liouville function is a factor of a system that*

(i) *has no irrational spectrum*;

(ii) *has ergodic components isomorphic to direct products of infinite-step nilsystems and Bernoulli systems.*

Remarks. • We allow the Bernoulli systems and the infinite-step nilsystems to be trivial; in other words, a direct product of a Bernoulli system and an infinite-step nilsystem is either a Bernoulli system, an infinite-step nilsystem, or a direct product of both.

• The product decomposition depends on the ergodic component; in particular, the infinite-step nilsystem depends on the ergodic component. On the other hand, our argument allows us to take the Bernoulli system to be the same on every ergodic component; we are not going to use this property though.

• A related result in a complementary direction was recently obtained in [22]; it states that if a Furstenberg system of the Möbius or the Liouville function is ergodic, then it is isomorphic to a Bernoulli system. The tools and the underlying ideas used in the proof of this result are very different and apply to a larger class of multiplicative functions.

• It is not clear to us how to adapt our argument in order to deal with more general bounded multiplicative functions. One would have to find a suitable variant of Proposition 3.9 below and to also modify significantly the subsequent analysis.

Using ergodic theory machinery we prove (see part (ii) of Proposition 3.12) that any system satisfying properties (i) and (ii) of Theorem 1.6 is necessarily disjoint from every totally ergodic system with zero entropy, leading to a proof of Theorem 1.4. The argument used in the proof of Theorems 1.1 and 1.3 depends on a different disjointness result (see part (i) of Proposition 3.12), and this necessitates the use of some additional input from number theory that is contained in [62] in order to verify its hypothesis.

To prove properties (i) and (ii) of Theorem 1.6, we combine tools from analytic number theory and ergodic theory. Our starting point is an identity of Tao (Theorem 3.6) that is implicit in [62] and enables us to express the self-correlations of the Möbius and the Liouville function as an average of its dilated self-correlations with prime dilates. (This step necessitates the use of logarithmic averages.) We use this identity in order to reduce our problem to a result of purely ergodic context. Roughly speaking, it asserts that if we average the correlations of an arbitrary measure preserving system (X, μ, T) over all prime dilates of its iterates, then the resulting system $(\widetilde{X}, \widetilde{\mu}, \widetilde{T})$ (see Definition 3.8), which we call the "system of arithmetic progressions with prime steps," necessarily possesses properties (i) and(ii) (see Theorem 3.10). Our motivation for establishing this property comes from the case where the ergodic components of the system (X, μ, T) are totally ergodic. It can then be shown that the resulting system $(\widetilde{X}, \widetilde{\mu}, \widetilde{T})$ has additional structure; namely, it is strongly stationary (see Definition 5.1). The structure of strongly stationary systems was completely determined in [42] and [21], where it was shown that they satisfy properties (i) and (ii) of Theorem 1.6. Unfortunately, we do not know how to establish total ergodicity of the ergodic components of Furstenberg systems of the Liouville function. (For the Möbius function this property is not even true.) In order to overcome this obstacle we use a more complicated line of arguing, which we briefly describe next.

To prove that the system $(\widetilde{X}, \widetilde{\mu}, \widetilde{T})$ enjoys property (ii) we initially use a structural result of Host and Kra (see Theorem 4.1 and Corollary A.6) and an ergodic theorem (see Theorem 4.3) in order to reduce the problem to the case where the system (X, μ, T) is an ergodic infinite-step nilsystem (see Lemma 4.11). In this case, we show (see Proposition 4.8) that the ergodic components of the system $(\widetilde{X}, \widetilde{\mu}, \widetilde{T})$ are infinite-step nilsystems. The essential role in this part of the argument plays the theory of arithmetic progressions on nilmanifolds, which we briefly review in Appendix B. The details are given in Section 4.

The key ingredient in the proof of property (i) is to establish that the system $(\widetilde{X}, \widetilde{\mu}, \widetilde{T})$ satisfies a somewhat weaker property than strong stationarity; roughly speaking, it is an inverse limit of partially strongly stationary systems (a notion defined in Definition 5.1). We then adjust an argument of Jenvey

[42] in order to show that such systems do not have irrational spectrum. The details are given in Section 5.

Finally, we briefly record the input from analytic number theory needed to carry out our analysis: The structural result of Theorem 1.6 uses some identities of Tao for the Möbius and the Liouville function that are implicit in [62] and were obtained from first principles using techniques from probabilistic number theory. It also uses indirectly (via the use of Theorems 4.3 and 4.4 in various places) the Gowers uniformity of the W-tricked von Mangoldt function that was established in [31], [32], [34]. Theorem 1.4 does not use any other tools from number theory. Theorems 1.1, 1.2, and 1.3 use, in addition to the previous number theoretic tools, a recent result of Tao [62] on the two-point correlations of the Liouville function, which in turn depends upon a recent result of Matomäki and Radziwiłł [50] on averages of the Möbius and the Liouville function on short intervals. This additional input from number theory is used in order to verify that on any Furstenberg system of the Möbius (resp. Liouville) function, a function naturally associated to $\boldsymbol{\mu}$ (or $\boldsymbol{\lambda}$) is orthogonal to the rational Kronecker factor of the system; this is needed in order to verify the hypothesis of the disjointness result stated in part (i) of Proposition 3.12.

1.3. *Comparison with existing results*. We say that a topological dynamical system (Y, R) *satisfies the Sarnak conjecture* if for every continuous function g on Y and every $y \in Y$, the Cesàro averages

$$\frac{1}{N}\sum_{n=1}^{N} g(R^n y)\,\boldsymbol{\mu}(n)$$

tend to 0 as $N \to \infty$. We say that (Y, R) *satisfies the logarithmic Sarnak conjecture* if the same property holds with the logarithmic averages

$$\frac{1}{\log N}\sum_{n=1}^{N} \frac{g(R^n y)\,\boldsymbol{\mu}(n)}{n}$$

in place of the Cesàro averages. Note that the Sarnak conjecture for a system implies the logarithmic Sarnak conjecture for the same system.

The Sarnak conjecture has been proved for a variety of systems, for example nilsystems [32], some horocycle flows [6] and more general zero entropy systems arising from homogeneous dynamics [56], certain distal systems, in particular, some extensions of a rotation by a torus [45], [49], [67], a large class of rank one transformations [16], [5], [19], systems generated by various substitutions [14], [11], [18], [52], all automatic sequences [54], some interval exchange transformations [5], [9], [19], some systems of number theoretic origin [4], [29], and more. The survey article [17] contains an up-to-date list of relevant references. In most cases the systems under consideration are uniquely ergodic. The proof techniques vary a lot since they make essential use of special properties of the system at hand. However, in many cases, the proof is based upon a lemma of Kátai [43], in a way introduced in [6], and our method is completely different.

Theorems 1.1 and 1.3 in this article allow one to deal with the vastly more general class of zero entropy topological dynamical systems that are uniquely ergodic or have at most countably many ergodic invariant measures. The price to pay is that we cover only the logarithmic variant of Sarnak's conjecture. Modulo this shortcoming, Theorems 1.1 and 1.3 cover most of the systems cited above and can be used to handle a wide variety of new systems. We briefly give a non-exhaustive list of examples covered by our main results:

Systems with countable support. If Y is a countable set, then the system (Y,R) has at most countably many ergodic invariant measures, all of them giving rise to periodic systems. Hence, Theorem 1.1 applies and shows that the system (Y,R) satisfies the logarithmic Sarnak conjecture. (The same conclusion can also be obtained using [40, Th. 1.4], which deals with Cesàro averages.) In particular, this implies that the support of the subshift generated by the Liouville function is an uncountable set, and this is true even if we change the values of the Liouville function on a set of logarithmic density 0.

Homogeneous dynamics. Nilsystems and several horocycle flows have zero entropy and every point is generic for an ergodic measure; hence Theorem 1.3 applies. The same holds for more general unipotent actions on homogeneous spaces of connected Lie groups.

Some distal systems. Our result applies for a wide family of topological distal systems. For example, suppose that (W,T) is a uniquely ergodic system and (Y,R) is built from (W,T) by a sequence of compact group extensions in the topological sense. Then the transformation R admits a "natural" invariant measure ν, and if (Y,ν,R) is ergodic, then (Y,R) is uniquely ergodic [26, Prop. 3.10], and Theorem 1.1 applies.

Rank one transformations. Strictly speaking, rank one systems are defined in a pure measure theoretical setting, but they have a natural topological model. Most of these models (including those considered in the bibliography cited above) are uniquely ergodic and have zero topological entropy; hence Theorem 1.1 applies.

Subshifts with linear block growth. Let (Y,R) be a transitive subshift with linear block growth (see Section 7). Then (Y,R) has zero topological entropy, and by Proposition 7.1 it admits only finitely many ergodic invariant measures (for minimal subshifts this result was already known [3]). Hence, Theorem 1.1 applies and shows that it satisfies the logarithmic Sarnak conjecture. We use this fact in the proof of Theorem 1.2.

Substitution dynamical systems. Theorem 1.1 applies to all systems of *primitive substitutions* [58] with not necessarily constant length, because they have zero topological entropy and are uniquely ergodic.

Interval exchange transformations. All interval exchange transformations have zero entropy, and minimality of the interval exchange (which is equivalent to the non-existence of a point with a finite orbit) implies that it has a finite number of ergodic invariant measures [44], [65]. Hence, Theorem 1.1 applies and shows that all minimal interval exchange transformations satisfy the logarithmic Sarnak conjecture.

Finite rank Bratteli-Vershik dynamical systems. Theorem 1.1 applies to all finite rank Bratteli-Vershik dynamical systems [7] (minimality is part of their defining properties) because they have zero entropy and finitely many ergodic invariant measures. This class contains all the examples mentioned in the previous two classes.

Although the class of topological dynamical systems to which Theorem 1.4 applies is more restrictive (due to our total ergodicity assumption), it is still large. For instance, totally ergodic nilsystems, several horocycle flows, several distal systems as the ones described above, some classical rank one transformations (for example, the Chacon system), and typical interval exchange transformations, have zero topological entropy and are uniquely ergodic and totally ergodic; hence Theorem 1.4 applies.

1.4. *Further comments and some conjectures*. Theorems 1.1, 1.3, and 1.4 deal with logarithmic averages rather than the more standard Cesàro averages. This is necessary for our proof since on the first step of our argument we use the identities of Tao stated in Theorem 3.6, and these are only known in a form useful to us for logarithmic averages.

If one shows that Furstenberg systems of the Liouville function have no rational spectrum except 1, then Theorem 1.4 can be proved for the Liouville function for any $y \in Y$ that is generic for a measure ν such that the system (Y, ν, S) has zero entropy and at most countably many ergodic components and all $g \in C(Y)$ with $\int g \, d\nu = 0$.

Theorem 1.3 handles the case where a point $y \in Y$ is generic (or quasi-generic) for a measure ν such that the system (Y, ν, S) has zero entropy and at most countably many ergodic components. But if (Y, ν, S) has uncountably many ergodic components, our argument falls apart. A particular instance is the following one: Let $(\alpha_k)_{k\in\mathbb{N}}$ be a sequence that is equidistributed in $\mathbb{T}$, and suppose that the finite sequences $(n\alpha_k)_{n\in[k^2,(k+1)^2)}$, $k \in \mathbb{N}$, are asymptotically equidistributed in $\mathbb{T}$ as $k \to \infty$, meaning, $\lim_{k\to\infty} \frac{1}{2k+1} \sum_{k^2\leq n<(k+1)^2} f(n\alpha_k) =$

$\int f\, dm_{\mathbb{T}}$ for every $f \in C(\mathbb{T})$. We let

$$y_0(n) := \sum_{k=1}^{\infty} \mathrm{e}^{2\pi i n \alpha_k}\, \mathbf{1}_{[k^2,(k+1)^2)}(n), \quad n \in \mathbb{N},$$

and $y_0(n) := 1$ for $n \leq 0$. Let $\mathbb{S}$ be the unit circle, $Y = \mathbb{S}^{\mathbb{Z}}$, $R\colon Y \to Y$ be the shift transformation, and let $g \in C(Y)$ be defined by $g(y) := y(0)$ for $y \in Y$. Note that $y_0(n) = g(R^n y_0)$ for every $n \in \mathbb{Z}$. We claim that the point $y_0 \in Y$ is generic for some invariant measure ν on Y and that the system (Y, ν, R) is measure-theoretically isomorphic to the system $(\mathbb{T}^2, m_{\mathbb{T}^2}, T)$, where $m_{\mathbb{T}^2}$ is the Haar measure of $\mathbb{T}^2$ and $T\colon \mathbb{T}^2 \to \mathbb{T}^2$ is defined by

$$T(s,t) := (s, t+s), \quad s,t \in \mathbb{T}.$$

Assuming the claim for the moment, we easily conclude that the system (Y, ν, R) has zero entropy, no eigenvalue other than 1, uncountably many ergodic components, and is disjoint from every ergodic system. Our methods do not allow us to prove that this system is disjoint from Furstenberg systems of the Möbius or the Liouville function or that the logarithmic averages of $y_0(n)\, \boldsymbol{\mu}(n)$ or $y_0(n)\, \boldsymbol{\lambda}(n)$ are 0.

To prove the claim, define the map $\phi\colon \mathbb{T}^2 \to \mathbb{S}$ by $\phi(s,t) := \mathrm{e}^{2\pi i t}$, for $s, t \in \mathbb{T}$, and the map $\Phi\colon \mathbb{T}^2 \to Y$ by $(\Phi(s,t))(n) = \phi(T^n(s,t)) := \mathrm{e}^{2\pi i (t+ns)}$ for $n \in \mathbb{Z}$, $s,t \in \mathbb{T}$. We have $\Phi \circ T = R \circ \Phi$, and the image ν of the measure $m_{\mathbb{T}^2}$ under Φ is invariant under R. Moreover, $\phi(T(s,t))\, \overline{\phi(s,t)} = \mathrm{e}^{2\pi i s}$, and it follows that Φ is one-to-one and thus is an isomorphism from $(\mathbb{T}^2, m_{\mathbb{T}^2}, T)$ to (Y, ν, R). It remains to show that the point y_0 is generic for the measure ν. For $m \in \mathbb{N}$, let $\ell_{-m}, \dots, \ell_m \in \mathbb{Z}$, and define

$$F(y) := \prod_{j=-m}^{m} y(j)^{\ell_j} \quad \text{for} \quad y = (y(n))_{n \in \mathbb{Z}} \in Y.$$

Then by a direct computation it is not hard to verify that

$$\begin{aligned} \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} F(R^n y_0) &= \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{j=-m}^{m} y_0(n+j)^{\ell_j} \\ &= \int_{\mathbb{T}^2} \prod_{j=-m}^{m} \mathrm{e}^{2\pi i (t+js)\ell_j}\, ds\, dt = \int_{\mathbb{T}^2} F \circ \Phi\, dm_{\mathbb{T}^2} = \int_Y F\, d\nu. \end{aligned}$$

By linearity and density, the same formula holds for every continuous function F on Y, and the claim follows.

We would also like to remark that it is consistent with existing knowledge (though highly unlikely) that some Furstenberg system of the Liouville function is isomorphic to the low complexity system $(\mathbb{T}^2, m_{\mathbb{T}^2}, T)$ described above. Here is a related problem:

PROBLEM. *Let* $\phi\colon \mathbb{T} \to \{-1,1\}$ *be the function defined by* $\phi(t) := \mathbf{1}_{[0,1/2)}(t) - \mathbf{1}_{[1/2,1)}(t)$. *Show that the following identity cannot hold*:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} \prod_{j=1}^{\ell} \boldsymbol{\lambda}(n+h_j) = \int_{\mathbb{T}^2} \prod_{j=1}^{\ell} \phi(t+h_j s)\, dt\, ds$$

for all $\ell \in \mathbb{N}$ *and* $h_1, \dots, h_\ell \in \mathbb{Z}$.

In the initial step of our argument (Proposition 3.9) we make essential use of the fact that $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$ are equal to -1 on the primes. But we expect the conclusion of Theorem 1.6 to remain valid even when one uses an arbitrary multiplicative function $f\colon \mathbb{N} \to [-1,1]$ in place of $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$. In fact, we expect ergodicity in all cases, and we conjecture the following:

CONJECTURE 1. *Every multiplicative function* $f\colon \mathbb{N} \to [-1,1]$ *has a unique Furstenberg system.*[2] *This system is ergodic and isomorphic to the direct product of a Bernoulli system and an ergodic odometer.*[3]

Note that all three possibilities can occur; for example, it is known that the Furstenberg system of $\boldsymbol{\mu}^2$ (called the square-free system) is an ergodic odometer [8], and conditional to the Chowla conjecture it is known that the Furstenberg system of the Liouville function $\boldsymbol{\lambda}$ is isomorphic to a Bernoulli system and the Furstenberg system of the Möbius function $\boldsymbol{\mu}$ is a relatively Bernoulli extension over the procyclic factor induced by $\boldsymbol{\mu}^2$ (see [15, Lemma 4.6]).

How do we then distinguish (at least conjecturally) between the possible structures of the Furstenberg system of a multiplicative function $f\colon \mathbb{N} \to [-1,1]$? It seems easier to do this when f takes values in $\{-1,1\}$, in which case we expect the following dichotomy:

CONJECTURE 2. *The Furstenberg system of a multiplicative function* $f\colon \mathbb{N} \to \{-1,1\}$ *is either a Bernoulli system or an ergodic odometer. Moreover*, *it is a Bernoulli system if and only if* f *is aperiodic.*

Aperiodicity, which is also often referred to as non-pretentiousness, means that the averages $\frac{1}{N}\sum_{n=1}^{N} f(an+b)$ converge to 0 as $N \to \infty$ for all $a, b \in \mathbb{N}$. It can be shown that the Furstenberg system of a zero mean multiplicative function $f\colon \mathbb{N} \to \{-1,1\}$ is Bernoulli if and only if all multiple correlations of distinct shifts of f vanish. When one works with logarithmic averages, Tao showed in [63] (when $f = \boldsymbol{\lambda}$, but his argument applies with some modifications

[2]Equivalently, the point $(f(n))_{n\in\mathbb{N}}$ is generic for some measure on the sequence space $[-1,1]^{\mathbb{N}}$.

[3]An ergodic odometer is an ergodic inverse limit of periodic systems, or equivalently, an ergodic system (X,μ,T) for which the rational eigenfunctions span a dense subspace of $L^2(\mu)$.

for general multiplicative $f\colon \mathbb{N} \to \{-1,1\}$; see [22, Th. 1.8]) that this is equivalent to asserting that f satisfies the Sarnak conjecture. So for multiplicative functions $f\colon \mathbb{N} \to \{-1,1\}$,

- aperiodicity,
- Bernoullicity of the corresponding Furstenberg system,
- f satisfies the logarithmic Chowla conjecture, and
- f satisfies the logarithmic Sarnak conjecture

are expected to be equivalent properties. Of course, none of the last three properties is known unconditionally even for the Liouville function. (Only aperiodicity is known.)

1.5. *Notation and conventions*. For the reader's convenience, we gather here some notation used throughout the article.

We write $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ and $\mathbb{S}$ for the unit circle. For $t \in \mathbb{R}$ or $\mathbb{T}$, we write $\mathrm{e}(t) := \mathrm{e}^{2\pi i t}$.

We denote by $\mathbb{N}$ the set of positive integers and by $\mathbb{P}$ the set of prime numbers. For $N \in \mathbb{N}$, we denote by $[N]$ the set $\{1, \ldots, N\}$. Whenever we write $\mathbf{N}$ we mean a sequence of intervals of integer $([N_k])_{k\in\mathbb{N}}$ with $N_k \to \infty$.

Unless otherwise specified, with $\ell^\infty(\mathbb{Z})$ we denote the space of all bounded, real valued, doubly infinite sequences.

If A is a finite non-empty set, we let $\mathbb{E}_{n\in A} := \frac{1}{|A|}\sum_{n\in A}$.

With (Y, R) we denote the topological dynamical system used to define the weight in the formulation of Theorems 1.1, 1.3, and 1.4; it sometimes comes equipped with an R-invariant measure ν.

With (X, μ, T) we denote a Furstenberg system of the Möbius or the Liouville function, and we also use the same notation when we study properties of abstract measure preserving systems.

With $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ we denote the system of arithmetic progressions with prime steps associated with a system (X, μ, T).

1.6. *Acknowledgement*. We would like to thank F. Durand, B. Kra, M. Lemańczyk, and P. Sarnak for useful remarks. The second author thanks the CMM – Universidad de Chile for its hospitality and support.

2. **Background in ergodic theory**

We gather here some basic background in ergodic theory and related notation used throughout the article.

Topological dynamical systems. A *topological dynamical system* (X, T) is a compact metric space endowed with a homeomorphism $T\colon X \to X$. It is *topologically transitive* if it has at least one dense orbit under T, and it is *minimal* if each orbit is dense.

If (X,T) and (Y,S) are two topological dynamical systems, then the second system is a *factor* of the first if there exists a map $\pi\colon X\to Y$, continuous and onto, such that $S\circ\pi(x)=\pi\circ T(x)$ for every $x\in X$. If the factor map π is injective, then the two systems are *isomorphic*.

Measure preserving systems. Throughout the article, we make the standard assumption that all probability spaces $(X,\mathcal{X},\mu)$ considered are Lebesgue, meaning, X can be given the structure of a compact metric space and $\mathcal{X}$ is its Borel σ-algebra. A *measure preserving system*, or simply *a system*, is a quadruple $(X,\mathcal{X},\mu,T)$, where $(X,\mathcal{X},\mu)$ is a probability space and $T\colon X\to X$ is an invertible, measurable, measure preserving transformation. We often omit the σ-algebra $\mathcal{X}$ and write (X,μ,T). Throughout, for $n\in\mathbb{N}$, we denote with T^n the composition $T\circ\cdots\circ T$ (n times), and let $T^{-n}:=(T^n)^{-1}$ and $T^0:=\mathrm{id}_X$. Also, for $f\in L^1(\mu)$ and $n\in\mathbb{Z}$, we denote by T^nf the function $f\circ T^n$.

Factors and isomorphisms. A *homomorphism*, also called a *factor map*, from a system $(X,\mathcal{X},\mu,T)$ onto a system $(Y,\mathcal{Y},\nu,S)$ is a measurable map $\pi\colon X\to Y$, such that $\mu\circ\pi^{-1}=\nu$ and with $S\circ\pi=\pi\circ T$ valid μ-almost everywhere. When we have such a homomorphism we say that the system $(Y,\mathcal{Y},\nu,S)$ is a *factor* of the system $(X,\mathcal{X},\mu,T)$. If the factor map $\pi\colon X\to Y$ is invertible,[4] we say that π is an *isomorphism* and that the systems $(X,\mathcal{X},\mu,T)$ and $(Y,\mathcal{Y},\nu,S)$ are *isomorphic*.

If $\pi\colon(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$ is a factor map and $\phi\in L^1(\mu)$, the function $\mathbb{E}_\mu(\phi\mid Y)$ in $L^1(\nu)$ is determined by the property $\int_A\mathbb{E}_\mu(\phi\mid Y)\,d\nu=\int_{\pi^{-1}(A)}\phi\,d\mu$ for every $A\in\mathcal{Y}$.

If $\pi\colon(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$ is a factor map, then $\pi^{-1}(\mathcal{Y})$ is a T-invariant sub-σ-algebra of $\mathcal{X}$. Conversely, for any T-invariant sub-σ-algebra $\mathcal{Y}'$ of $\mathcal{X}$, there exists a factor map $\pi\colon(X,\mathcal{X},\mu,T)\to(Y,\mathcal{Y},\nu,S)$ with $\mathcal{Y}'=\pi^{-1}(\mathcal{Y})$ up to μ-null sets. This factor is unique up to isomorphism, and we call it *the factor associated with* (*or induced by*) $\mathcal{Y}'$. See [66, §2.3] or [13, §6.2] for details. When there is no danger of confusion, we may abuse notation and denote the transformation S on Y by T. We pass constantly from invariant sub-σ-algebras to factors, the convention being that the factors associated to the σ-algebras $\mathcal{Y},\mathcal{Z},\ldots$, are written $Y,Z,\ldots$.

We will sometimes abuse notation and use the sub-σ-algebra $\mathcal{Y}$ in place of the subspace $L^2(X,\mathcal{Y},\mu)$. For example, if we write that a function is orthogonal to $\mathcal{Y}$, we mean that it is orthogonal to the subspace $L^2(X,\mathcal{Y},\mu)$.

Spectrum. Let (X,μ,T) be a system. For $t\in\mathbb{T}$, we say that $\mathrm{e}(t)$ is an *eigenvalue* of the system if there exists a non-identically zero function $f\in$

[4]Meaning that there exists a factor map $Y\to X$, written π^{-1}, with $\pi^{-1}\circ\pi=\mathrm{id}_X$ valid μ-almost everywhere. (This implies that $\pi\circ\pi^{-1}=\mathrm{id}_Y$ holds ν-almost everywhere.)

$L^2(\mu)$ such that $Tf = \mathrm{e}(t)f$, in which case we say that f is an *eigenfunction* associated to the eigenvalue $\mathrm{e}(t)$. We call the eigenvalue $\mathrm{e}(t)$ *rational* if t is rational and *irrational* otherwise. The *spectrum* of the system is the subset of $\mathbb{T}$ consisting of all eigenvalues, and we define the *rational* and the *irrational spectrum* to be the subset of the spectrum consisting of all rational (resp. irrational) eigenvalues. With $\mathcal{K}_{\mathrm{rat}}(T)$ we denote the *rational Kronecker factor* of $(X, \mathcal{X}, \mu, T)$; it is the smallest T-invariant sub-σ-algebra of $\mathcal{X}$ with respect to which all eigenfunctions with rational eigenvalues are measurable. The linear span of these eigenfunctions is dense in $L^2(X, \mathcal{K}_{\mathrm{rat}}(T), \mu)$.

Ergodicity and ergodic decomposition. A system (X, μ, T) is *ergodic* if all functions $f \in L^1(\mu)$ that satisfy $Tf = f$ are constant. It is *totally ergodic* if (X, μ, T^d) is ergodic for every $d \in \mathbb{N}$, equivalently, if it is ergodic and has no rational spectrum except 1.

Let $(X, \mathcal{X}, \mu, T)$ be a system, and let $\pi\colon (X, \mathcal{X}, \mu, T) \to (\Omega, \mathcal{O}, P, T)$ be the factor map associated to the σ-algebra of T-invariant sets of X. Then the disintegration of μ over P,

$$\mu = \int_\Omega \mu_\omega \, dP(\omega), \tag{5}$$

is called the *ergodic decomposition* of μ under T (see [28, Th. 3.22]). The following properties hold:

- T acts as the identity on Ω;
- the map $\omega \mapsto \mu_\omega$ is a measurable map from Ω to the set of ergodic T-invariant measures on X;
- the decomposition (5) is unique in the following sense: if $(Y, \mathcal{Y}, \nu)$ is a probability space and $y \mapsto \mu'_y$ is a measurable map from Y into the set of ergodic measures on X such that $\mu = \int_Y \mu'_y \, d\nu(y)$, then there exists a measurable map $\phi\colon Y \to \Omega$, mapping the measure ν to the measure P, such that $\mu_{\phi(y)} = \mu'_y$ for ν-almost every $y \in Y$.

We call the systems $(X, \mathcal{X}, \mu_\omega, T)$, $\omega \in \Omega$, the *ergodic components* of $(X, \mathcal{X}, \mu, T)$.

Unique ergodicity. A topological dynamical system (X, T) is *uniquely ergodic* if there is a unique T-invariant Borel probability measure on X.

Bernoulli systems. For the purposes of this article, a *Bernoulli system* has the form $(X^{\mathbb{Z}}, \mathcal{B}_{X^{\mathbb{Z}}}, \nu, S)$, where $(X, \mathcal{X}, \rho)$ is a probability space, S is the shift transformation on $X^{\mathbb{Z}}$, $\mathcal{B}_{X^{\mathbb{Z}}}$ is the product σ-algebra of $X^{\mathbb{Z}}$, and ν is the product measure $\rho^{\mathbb{Z}}$.

Nilsystems. Let $s \in \mathbb{N}$, G be an s-step nilpotent Lie group and Γ be a discrete cocompact subgroup of G. Then the quotient space $X = G/\Gamma$ is called an *s-step nilmanifold.* We denote the elements of X as points $x, y, \ldots$, not as cosets. The point e_X is the image in X of the unit element of G. The natural action of G on X is written $(g,x) \mapsto g \cdot x$, and the unique Borel measure on X that is invariant under this action is called the *Haar measure* of X and is denoted by μ_X. If $a \in G$, then the transformation $T\colon X \to X$ defined by $Tx = ax$, $x \in X$, is called a *nilrotation of* X, and the system $(X, \mathcal{X}, \mu_X, T)$, where $\mathcal{X}$ is the Borel-σ-algebra of X, is called an *s-step nilsystem*. When we do not care about the degree of nilpotency s, we simply call it a *nilsystem*. It is well known that if T is a nilrotation on X, then the statements (X,T) is topologically transitive, (X,T) is minimal, (X,μ_X,T) is ergodic, and (X,T) is uniquely ergodic, are equivalent. Moreover, an ergodic nilsystem (X,μ_X,T) is totally ergodic if and only if the nilmanifold X is connected.

Joinings and disjoint systems. Let $(X,\mathcal{X},\mu,T)$ and $(Y,\mathcal{Y},\nu,S)$ be two systems. We call a measure ρ on $(X \times Y, \mathcal{X} \times \mathcal{Y})$ a *joining* of the two systems if it is $T \times S$ invariant and its projection onto the X and Y coordinates are the measures μ and ν respectively. We say that the systems on X and on Y are *disjoint* if the only joining of the systems is the product measure $\mu \times \nu$. If two systems are disjoint, then they have no non-trivial common factor, but the converse is not true. It is well known that every Bernoulli system is disjoint from every zero-entropy system; we will use the zero entropy assumption in the proofs of our main results only via this property.

3. Overview of the proof and reduction to an ergodic statement

In this section we give an overview of the proof of our main results and eventually reduce to some statements of purely ergodic context, which we establish in Sections 4–6. In Section 3.2 we define the notion of a Furstenberg system of an arbitrary bounded sequence. In Section 3.4 we reproduce some striking identities of Tao that are implicit in [62] and we use them in Section 3.5 in order to show that a Furstenberg system of the Liouville function is a factor of a measure preserving system of purely ergodic origin; we call it the "system of arithmetic progressions with prime steps." In Section 3.6 we state our main structural results for such systems, and we use them in Section 3.7 in order to get similar structural results for Furstenberg systems of the Möbius and the Liouville function, thus proving Theorem 1.6. In Section 3.8 we state a disjointness result that we use in Section 3.9 in order to prove Theorems 1.1, 1.3, and 1.4.

3.1. *Notation regarding averages*. For $N \in \mathbb{N}$, we let $[N] = \{1, \dots, N\}$. For an arbitrary bounded sequence $a = (a(n))_{n \in \mathbb{N}}$, we write

$$\mathbb{E}_{n \in [N]} a(n) := \frac{1}{N} \sum_{n=1}^{N} a(n) \quad \text{and} \quad \mathbb{E}_{n \in \mathbb{N}} a(n) := \lim_{N \to \infty} \mathbb{E}_{n \in [N]} a(n)$$

if this limit exists. Let $\mathbf{N} = ([N_k])_{k\in\mathbb{N}}$ be a sequence of intervals with $N_k \to \infty$. For an arbitrary bounded sequence $a = (a(n))_{n\in\mathbb{N}}$, we write

$$\mathbb{E}_{n\in\mathbf{N}}\, a(n) := \lim_{k\to\infty} \mathbb{E}_{n\in[N_k]}\, a(n)$$

if this limit exists and

$$\mathbb{E}^{\log}_{n\in[N_k]} := \frac{1}{\log N_k} \sum_{n=1}^{N_k} \frac{a(n)}{n}, \qquad \mathbb{E}^{\log}_{n\in\mathbf{N}}\, a(n) := \lim_{k\to\infty} \mathbb{E}^{\log}_{n\in[N_k]}\, a(n)$$

if this limit exists. If $(a(p))_{p\in\mathbb{P}}$ is a sequence indexed by the primes, we write

$$\mathbb{E}_{p\in\mathbb{P}}\, a(p) := \lim_{N\to\infty} \frac{1}{\pi(N)} \sum_{p\leq N} a(p),$$

where $\pi(N)$ denotes the number of prime numbers less than N, if this limit exists.

Using partial summation one easily verifies that for a bounded sequence $(a(n))_{n\in\mathbb{N}}$, convergence of the Cesàro averages $\mathbb{E}_{n\in[N]}\, a(n)$ implies convergence of the logarithmic averages $\mathbb{E}^{\log}_{n\in[N]}\, a(n)$ as $N \to \infty$, but the converse does not hold. Moreover, the direct implication does not hold if we average over subsequences of intervals.

3.2. *Furstenberg systems of bounded sequences*. To each bounded sequence that is distributed "regularly" along a sequence of intervals with lengths increasing to infinity, we associate a measure preserving system. For the purposes of this article, all averages in the definition of Furstenberg systems of bounded sequences are taken to be logarithmic and we restrict to real valued bounded sequences.

Definition 3.1. Let $\mathbf{N} := ([N_k])_{k\in\mathbb{N}}$ be a sequence of intervals with $N_k \to \infty$. We say that the real valued sequence $a \in \ell^\infty(\mathbb{Z})$ *admits log-correlations on* $\mathbf{N}$ if the following limits exist:

$$\lim_{k\to\infty} \mathbb{E}^{\log}_{n\in[N_k]}\, a(n+h_1)\cdots a(n+h_\ell)$$

for every $\ell \in \mathbb{N}$ and $h_1, \ldots, h_\ell \in \mathbb{Z}$ (not necessarily distinct).

Remarks. • If $a \in \ell^\infty(\mathbb{Z})$, then using a diagonal argument we get that every sequence of intervals $\mathbf{N} = ([N_k])_{k\in\mathbb{N}}$ has a subsequence $\mathbf{N}' = ([N'_k])_{k\in\mathbb{N}}$, such that the sequence $a \in \ell^\infty(\mathbb{Z})$ admits log-correlations on $\mathbf{N}'$.

• If $a(n)$ is only defined for $n \in \mathbb{N}$, we extend it in an arbitrary way to $\mathbb{Z}$ and define the analogous notion. Then all the limits above do not depend on the choice of the extension.

The correspondence principle of Furstenberg was originally used in [25] in order to restate Szemerédi's theorem on arithmetic progressions in ergodic terms. We will use the following variant of this principle, which applies to general real valued bounded sequences:

PROPOSITION 3.2. *Let $a \in \ell^\infty(\mathbb{Z})$ be a real valued sequence that admits log-correlations on $\mathbf{N} := ([N_k])_{k\in\mathbb{N}}$. Then there exist a topological system (X, T), a T-invariant Borel probability measure μ, and a real valued T-generating function $F_0 \in C(X)$,*[5] *such that*

$$\mathbb{E}^{\log}_{n\in\mathbf{N}} \prod_{j=1}^{\ell} a(n+h_j) = \int \prod_{j=1}^{\ell} T^{h_j} F_0 \, d\mu \tag{6}$$

for every $\ell \in \mathbb{N}$ and $h_1, \ldots, h_\ell \in \mathbb{Z}$.

Definition 3.3. Let $a \in \ell^\infty(\mathbb{N})$ be a real valued sequence that admits log-correlations on $\mathbf{N} := (N_k)_{k\in\mathbb{N}}$. We call the system (or the measure μ) defined in Proposition 3.2 the *Furstenberg system* (*or measure*) *associated with a and* $\mathbf{N}$.

Remarks. • Given $a \in \ell^\infty(\mathbb{Z})$ and $\mathbf{N}$, the measure μ is uniquely determined by (6) since this identity determines the values of $\int f\, d\mu$ for all real valued $f \in C(X)$.

• A priori a sequence $a \in \ell^\infty(\mathbb{Z})$ may have several, perhaps uncountably many, non-isomorphic Furstenberg systems depending on which sequence of intervals $\mathbf{N}$ we use in the evaluation of the log-correlations of the sequence $a \in \ell^\infty(\mathbb{Z})$. When we write that a Furstenberg measure or system of a sequence has a certain property, we mean that any of these measures or systems has the asserted property.

In the construction of the Furstenberg system $(X, \mathcal{X}, \mu, T)$ we can take X to be the compact metric space $I^{\mathbb{Z}}$ (with the product topology), where I is any closed and bounded interval containing the range of $(a(n))_{n\in\mathbb{Z}}$, $\mathcal{X}$ to be the Borel-σ-algebra of $I^{\mathbb{Z}}$, and T to be the shift transformation on $I^{\mathbb{Z}}$. Points of X are written as $x = (x(n))_{n\in\mathbb{Z}}$, and we let $F_0(x) := x(0)$, $x \in X$. Then $F_0 \in C(X)$ and F_0 is T-generating. We consider the sequence $a = (a(n))_{n\in\mathbb{Z}}$ as a point of X. Our hypothesis implies that the measures

$$\mathbb{E}^{\log}_{n\in[N_k]} \delta_{T^n a}, \quad k \in \mathbb{N}, \tag{7}$$

converge weak-star as $k \to \infty$ to a measure μ on X, and this measure is clearly T-invariant and satisfies (6). Indeed, if $F = \prod_{j=1}^{\ell} T^{h_j} F_0$, then $F \in C(X)$

[5] A real valued function $F_0 \in C(X)$ is T-generating if the functions $T^n F_0$, $n \in \mathbb{Z}$, separate points of X. By the Stone-Weierstrass theorem, this holds if and only if the T-invariant subalgebra generated by F_0 is dense in $C(X)$ (we restrict to real valued functions) with the uniform topology.

and $F(T^n a) = \prod_{j=1}^{\ell} a(n+h_j)$, $n \in \mathbb{N}$, and the weak-star convergence of the measures in (7) to μ gives identity (6).

In this article we are mostly interested in applying the previous result when $a = \boldsymbol{\mu}$, in which case we take $X := \{-1,0,1\}^{\mathbb{Z}}$. For every $h \in \mathbb{Z}$, we write $F_h \colon X \to \{-1,0,1\}$ for the function given by

$$F_h(x) := x(h), \quad x \in X.$$

Then for every $h \in \mathbb{Z}$, we have $F_h = T^h F_0$. If $(X, \mathcal{X}, \mu, T)$ is the Furstenberg system associated with the Möbius function and the sequence $\mathbf{N}$, by Proposition 3.2 we have

$$\int \prod_{j=1}^{\ell} F_{h_j}(x)\, d\mu(x) = \int \prod_{j=1}^{\ell} T^{h_j} F_0\, d\mu = \mathbb{E}^{\log}_{n\in\mathbf{N}} \prod_{j=1}^{\ell} \boldsymbol{\mu}(n+h_j)$$

for every $\ell \in \mathbb{N}$ and $h_1, \dots, h_\ell \in \mathbb{Z}$.

3.3. *A convergence result for multiple correlation sequences*. We will make use of the following consequence of Theorem 4.3 below:

Proposition 3.4. *Suppose that the sequence $a \in \ell^\infty(\mathbb{Z})$ admits log-correlations on the sequence of intervals $\mathbf{N}$. Then the limit*

$$\mathbb{E}_{p\in\mathbb{P}} \Big(\mathbb{E}^{\log}_{n\in\mathbf{N}} \prod_{j=1}^{\ell} a(n+ph_j) \Big)$$

exists for all $\ell \in \mathbb{N}$ and $h_1, \dots, h_\ell \in \mathbb{Z}$.

Proof. Let $(X, \mathcal{X}, \mu, T)$ be the Furstenberg system associated with $a \in \ell^\infty(\mathbb{Z})$ and $\mathbf{N}$, and also let $F_0 \in L^\infty(\mu)$ be as in Proposition 3.2. Using Theorem 4.3 in Section 4.1.2 we get that for every $\ell \in \mathbb{N}$ and $h_1, \dots, h_\ell \in \mathbb{Z}$, the limit

$$\mathbb{E}_{p\in\mathbb{P}} \int \prod_{j=1}^{\ell} T^{ph_j} F_0\, d\mu$$

exists. By (6) we can replace $\int \prod_{j=1}^{\ell} T^{ph_j} F_0\, d\mu$ by $\mathbb{E}^{\log}_{n\in\mathbf{N}} \prod_{j=1}^{\ell} a(n+ph_j)$, and we arrive to the asserted conclusion. $\square$

3.4. *Tao's identities*. A key tool in our argument is the following rather amazing identity, which is implicit in [62]:

Theorem 3.5 (Tao's identity for general sequences). *Let $\mathbf{N} = ([N_k])_{k\in\mathbb{N}}$ be a sequence of intervals with $N_k \to \infty$, $a \in \ell^\infty(\mathbb{Z})$ be a sequence (perhaps complex valued), and $\ell \in \mathbb{N}$, $h_1, \dots, h_\ell \in \mathbb{Z}$. If we assume that on the left- and right-hand sides below the limits $\mathbb{E}^{\log}_{n\in\mathbf{N}}$ exist for every $p \in \mathbb{P}$ and the limit $\mathbb{E}_{p\in\mathbb{P}}$*

exists, then we have the identity

$$\mathbb{E}_{p\in\mathbb{P}}\Big(\mathbb{E}^{\log}_{n\in\mathbf{N}}\prod_{j=1}^{\ell} a(pn+ph_j)\Big)=\mathbb{E}_{p\in\mathbb{P}}\Big(\mathbb{E}^{\log}_{n\in\mathbf{N}}\prod_{j=1}^{\ell} a(n+ph_j)\Big).$$

We give a sketch of the proof of a more general identity in Appendix C; the argument is almost entirely based on the argument given by Tao in [62].

Using the previous result we verify the following identities for the Möbius and the Liouville function:

THEOREM 3.6 (Tao's identity for $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$). *Suppose that the Möbius function $\boldsymbol{\mu}$ admits log-correlations on the sequence of intervals $\mathbf{N}$. Then we have*

$$\mathbb{E}^{\log}_{n\in\mathbf{N}}\prod_{j=1}^{\ell}\boldsymbol{\mu}(n+h_j)=(-1)^{\ell}\,\mathbb{E}_{p\in\mathbb{P}}\Big(\mathbb{E}^{\log}_{n\in\mathbf{N}}\prod_{j=1}^{\ell}\boldsymbol{\mu}(n+ph_j)\Big)$$

for all $\ell\in\mathbb{N}$ and $h_1,\dots,h_\ell\in\mathbb{Z}$; in particular, the limit $\mathbb{E}_{p\in\mathbb{P}}$ on the right-hand side exists. A similar statement holds for the Liouville function $\boldsymbol{\lambda}$.

Proof. We first check the identity for the Liouville function. We verify that the hypothesis of Theorem 3.5 applies for $a:=\boldsymbol{\lambda}$. The limit $\mathbb{E}^{\log}_{n\in\mathbf{N}}$ on the left- and right-hand sides exists for every $p\in\mathbb{P}$ since $\boldsymbol{\lambda}$ admits log-correlations on $\mathbf{N}$ and it is completely multiplicative. Moreover, using complete multiplicativity, the left-hand side becomes $(-1)^{\ell}\,\mathbb{E}^{\log}_{n\in\mathbf{N}}\prod_{j=1}^{\ell}\boldsymbol{\lambda}(n+h_j)$. The right-hand side is $\mathbb{E}_{p\in\mathbb{P}}\Big(\mathbb{E}^{\log}_{n\in\mathbf{N}}\prod_{j=1}^{\ell}\boldsymbol{\lambda}(n+ph_j)\Big)$; note that the existence of the limit $\mathbb{E}_{p\in\mathbb{P}}$ follows from Proposition 3.4. So Theorem 3.5 applies for $a:=\boldsymbol{\lambda}$ and gives the asserted identity.

The argument is slightly more complicated for the Möbius function because in this case we lose complete multiplicativity. Arguing by contradiction, suppose that the asserted estimate fails. Then there exist a subsequence $\mathbf{N'}:=([N'_k])_{k\in\mathbb{N}}$ of $\mathbf{N}:=([N_k])_{k\in\mathbb{N}}$ and $\ell\in\mathbb{N}$, $h_1,\dots,h_\ell\in\mathbb{Z}$, such that the limit $\mathbb{E}^{\log}_{n\in\mathbf{N'}}\prod_{j=1}^{\ell}\boldsymbol{\mu}(pn+ph_j)$ exists for every $p\in\mathbb{P}$, and we have

$$\mathbb{E}^{\log}_{n\in\mathbf{N'}}\prod_{j=1}^{\ell}\boldsymbol{\mu}(n+h_j)\neq(-1)^{\ell}\,\mathbb{E}_{p\in\mathbb{P}}\Big(\mathbb{E}^{\log}_{n\in\mathbf{N'}}\prod_{j=1}^{\ell}\boldsymbol{\mu}(n+ph_j)\Big).\tag{8}$$

Note that the existence of the limit $\mathbb{E}_{p\in\mathbb{P}}$ on the right-hand side follows again from Proposition 3.4. For $j=1,\dots,\ell$ and $p\in\mathbb{P}$, we have $\boldsymbol{\mu}(pn+ph_j)=-\boldsymbol{\mu}(n+h_j)$ unless $n+h_j\equiv 0 \pmod p$. For $p\in\mathbb{P}$, this leads to the identity

$$\mathbb{E}^{\log}_{n\in\mathbf{N'}}\prod_{j=1}^{\ell}\boldsymbol{\mu}(pn+ph_j)=(-1)^{\ell}\,\mathbb{E}^{\log}_{n\in\mathbf{N'}}\prod_{j=1}^{\ell}\boldsymbol{\mu}(n+h_j)+O(1/p),$$

where the implicit constant depends only on ℓ. Averaging over $p \in \mathbb{P}$ we get

$$\mathbb{E}_{p\in\mathbb{P}}\Big(\mathbb{E}^{\log}_{n\in\mathbf{N}'} \prod_{j=1}^{\ell} \boldsymbol{\mu}(pn+ph_j)\Big) = (-1)^{\ell}\, \mathbb{E}^{\log}_{n\in\mathbf{N}'} \prod_{j=1}^{\ell} \boldsymbol{\mu}(n+h_j); \tag{9}$$

in particular, the limit $\mathbb{E}_{p\in\mathbb{P}}$ on the left-hand side exists. So Theorem 3.5 applies for $a := \boldsymbol{\mu}$ and the sequence of intervals $\mathbf{I}'$, and it gives that

$$\mathbb{E}_{p\in\mathbb{P}}\Big(\mathbb{E}^{\log}_{n\in\mathbf{N}'} \prod_{j=1}^{\ell} \boldsymbol{\mu}(pn+ph_j)\Big) = \mathbb{E}_{p\in\mathbb{P}}\ \Big(\mathbb{E}^{\log}_{n\in\mathbf{N}'} \prod_{j=1}^{\ell} \boldsymbol{\mu}(n+ph_j)\Big).$$

Combining this identity with (9) we get an identity that contradicts (8). This completes the proof. □

Using Theorem 3.6 we immediately deduce the following identities for Furstenberg systems of the Möbius and the Liouville function:

Theorem 3.7 (Ergodic form of Tao's identities for $\boldsymbol{\mu}$ and $\boldsymbol{\lambda}$). *Suppose that $(X, \mathcal{X}, \mu, T)$ is a Furstenberg system of the Möbius or the Liouville function, and let F_0 be as in Proposition* 3.2. *Then the limit in the right-hand side below exists and we have*

$$\int \prod_{j=1}^{\ell} T^{h_j} F_0\, d\mu = (-1)^{\ell}\, \mathbb{E}_{p\in\mathbb{P}} \int \prod_{j=1}^{\ell} T^{ph_j} F_0\, d\mu \tag{10}$$

for all $\ell \in \mathbb{N}$ and $h_1, \ldots, h_\ell \in \mathbb{Z}$.

Henceforth, our goal is to describe the structure of measure preserving systems that satisfy the identities in (10) for some T-generating function $F_0 \in C(X)$. For technical reasons, it is essential for us to work with suitable extensions of such systems, which we describe in the next subsection. Our main task will then be to get structural results for these extended systems.

3.5. *The system of arithmetic progressions with prime steps.* Motivated by Theorem 3.7, given a system (X, μ, T), we are going to construct a new system on the space $X^{\mathbb{Z}}$ by averaging the prime dilates of correlations of the system on the space X. Since in some cases X is itself a sequence space with elements denoted by $x = (x(n))_{n\in\mathbb{Z}}$, we denote elements of $X^{\mathbb{Z}}$ by $\underline{x} = (x_n)_{n\in\mathbb{Z}}$.

Definition 3.8. Let $(X, \mathcal{X}, \mu, T)$ be a system, and let $X^{\mathbb{Z}}$ be endowed with the product σ-algebra. We write $\widetilde{\mu}$ for the measure on $X^{\mathbb{Z}}$ characterized as follows: For every $m \in \mathbb{N}$ and all $f_{-m}, \ldots, f_m \in L^\infty(\mu)$, we define

$$\int_{X^{\mathbb{Z}}} \prod_{j=-m}^{m} f_j(x_j)\, d\widetilde{\mu}(\underline{x}) := \mathbb{E}_{p\in\mathbb{P}} \int_X \prod_{j=-m}^{m} T^{pj} f_j\, d\mu. \tag{11}$$

Note that the limit above exists by Theorem 4.3 in Section 4.1.2. Using the identity $\int_X \prod_{j=-m}^{m} T^{p(j+1)} f_j\, d\mu = \int_X \prod_{j=-m}^{m} T^{pj} f_j\, d\mu$, we get that the measure

$\widetilde{\mu}$ is invariant under the shift transformation S on $X^{\mathbb{Z}}$. We say that $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ is the *system of arithmetic progressions with prime steps* associated with the system (X, μ, T).

We return now to the case where (X, μ, T) is a Furstenberg system of the Liouville function and make the following key observation:

PROPOSITION 3.9. *A Furstenberg system (X, μ, T) of the Möbius or the Liouville function is a factor of the associated system $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ of arithmetic progressions with prime steps.*

Remark. The fact that the Möbius and the Liouville function is -1 on primes is crucial for the proof of this result and is used via the identity (10). In fact, our argument also works for all bounded multiplicative functions that take the value -1 on a subset of the primes with relative density 1.

Proof. We can take $X = \{-1, 0, 1\}^{\mathbb{Z}}$. We define the map $\pi\colon X^{\mathbb{Z}} \to X$ as follows: For $\underline{x} = (x_n)_{n\in\mathbb{Z}} \in X^{\mathbb{Z}}$, let

$$(\pi(\underline{x}))(n) := -x_n(0) = -F_0(x_n), \quad n \in \mathbb{Z},$$

where, as usual, $F_h(x) = x(h)$, $x \in X$, $h \in \mathbb{Z}$. For $n \in \mathbb{Z}$, we then have

$$(\pi(S\underline{x}))(n) = -F_0((S\underline{x})_n) = -F_0(x_{n+1}) = (\pi(\underline{x}))(n+1) = (T\pi(\underline{x}))(n).$$

Thus

$$\pi \circ S = T \circ \pi.$$

Next, we claim that $\widetilde{\mu} \circ \pi^{-1} = \mu$. Indeed, for every $\ell \in \mathbb{N}$ and $h_1, \dots, h_\ell \in \mathbb{Z}$, by identity (10) in Theorem 3.5 and the definition (11) of $\widetilde{\mu}$, we have

$$\begin{aligned}
\int_X \prod_{j=1}^{\ell} F_{h_j}(x)\, d\mu(x) &= \int_X \prod_{j=1}^{\ell} F_0(T^{h_j}x)\, d\mu(x) \\
&= (-1)^\ell\, \mathbb{E}_{p\in\mathbb{P}} \int_X \prod_{j=1}^{\ell} F_0(T^{ph_j}x)\, d\mu(x) \\
&= (-1)^\ell \int_{X^{\mathbb{Z}}} \prod_{j=1}^{\ell} F_0(x_{h_j})\, d\widetilde{\mu}(\underline{x}) \\
&= \int_{X^{\mathbb{Z}}} \prod_{j=1}^{\ell} \big(-F_0(x_{h_j})\big)\, d\widetilde{\mu}(\underline{x}) = \int_{X^{\mathbb{Z}}} \prod_{j=1}^{\ell} (F_{h_j} \circ \pi)(\underline{x})\, d\widetilde{\mu}(\underline{x}).
\end{aligned}$$

Since the algebra generated by the functions F_h, $h \in \mathbb{Z}$, is dense in $C(X)$ with the uniform topology, the claim follows.

Therefore, $\pi\colon (X^{\mathbb{Z}}, \widetilde{\mu}, S) \to (X, \mu, T)$ is a factor map and the proof is complete. $\square$

From this point on we work with abstract systems of arithmetic progressions with prime steps and use Proposition 3.9 in order to transfer any structural result we get to a structural result for Furstenberg systems of the Möbius and the Liouville function.

3.6. *Structure of systems of arithmetic progressions with prime steps*. We state our main structural results for abstract systems of arithmetic progressions with prime steps. In Section 4 we show

THEOREM 3.10. *Let (X,μ,T) be a system. Then almost every ergodic component of the system $(X^{\mathbb{Z}},\widetilde{\mu},S)$, of arithmetic progressions with prime steps, is isomorphic to a direct product of an infinite-step nilsystem and a Bernoulli system.*

In Section 5 we show

THEOREM 3.11. *Let (X,μ,T) be a system. Then the system $(X^{\mathbb{Z}},\widetilde{\mu},S)$, of arithmetic progressions with prime steps, has no irrational spectrum.*

We also establish similar results for systems of arithmetic progressions with integer steps (see Definition 4.2).

3.7. *Proof of Theorem* 1.6 *assuming the preceding material*. Combining Proposition 3.9 and Theorem 3.11, we get that any Furstenberg system of the Möbius or the Liouville function is a factor of a system with no irrational spectrum (and hence has no irrational spectrum), thus establishing property (i) of Theorem 1.6. Combining Proposition 3.9 and Theorem 3.10, we get property (ii) of Theorem 1.6. □

3.8. *Disjointness*. As we previously remarked, our proof strategy for Theorems 1.1, 1.3, and 1.4 is to study the structure of Furstenberg systems of the Möbius and the Liouville function in enough detail to enable us to prove a useful disjointness result. The relevant disjointness result is the following one and is proved in Section 6:

PROPOSITION 3.12. *Let (X,μ,T) be a system with ergodic components isomorphic to direct products of infinite-step nilsystems and Bernoulli systems. Let (Y,ν,R) be an ergodic system of zero entropy.*

(i) *If the two systems have disjoint irrational spectrum, then for every joining σ of the two systems and function $f\in L^\infty(\mu)$ orthogonal to $\mathcal{K}_{\mathrm{rat}}(T)$, we have*

$$\int f(x)\,g(y)\,d\sigma(x,y)=0$$

for every $g\in L^\infty(\nu)$.

(ii) *If the two systems have no common eigenvalue except* 1, *then they are disjoint.*

We will use the following direct consequence:

COROLLARY 3.13. *Proposition* 3.12 *holds under the weaker assumption that* (Y, ν, R) *is a zero entropy system with at most countably many ergodic components.*

Proof. Let $\nu = \sum_{j\in J} c_j\nu_j$ be the ergodic decomposition of ν under R, where J is a finite or an infinite countable set, $c_j > 0$, $\sum_{j\in J} c_j = 1$, and ν_j, $j \in J$, are ergodic R-invariant measures. Let $Y = \cup_{j\in J} Y_j$ be a partition of Y into R-invariant subsets such that for every $j \in J$, we have $\nu_j(Y_j) = 1$.

Let σ be a joining of the systems (X, μ, T) and (Y, ν, R). For $j \in J$, we let $\sigma_j := \frac{1}{c_j}\mathbf{1}_{X\times Y_j} \cdot \sigma$ and μ_j be the image of σ_j under the projection of $X \times Y$ on X. Then for $j \in J$, we have that μ_j is a T-invariant probability measure on X, the image of σ_j under the projection of $X \times Y$ onto Y is ν_j, and σ_j is a joining of the systems (X, μ_j, T) and (Y, ν_j, R).

For $j \in J$, the measure ν_j is absolutely continuous with respect to ν and thus the spectrum of (Y, ν_j, R) is contained in the spectrum of (Y, ν, R). Similarly, for $j \in J$, the measure μ_j is absolutely continuous with respect to μ and thus the spectrum of (X, μ_j, T) is contained in the spectrum of (X, μ, T). Moreover, every ergodic component of μ_j is an ergodic component of μ and thus is isomorphic to the direct product of an infinite-step nilsystem and a Bernoulli system.

In case (i), suppose that $f \in L^\infty(\mu)$ is orthogonal to $\mathcal{K}_{\mathrm{rat}}(X, \mu, T)$. This means that f is orthogonal in $L^2(\mu)$ to every eigenfunction of (X, μ, T) corresponding to a rational eigenvalue. It follows that for every $j \in J$, the function f is orthogonal in $L^2(\mu_j)$ to every eigenfunction of (X, μ_j, T) corresponding to a rational eigenvalue, and by part (i) of Proposition 3.12 we have $\int f(x)\, g(y)\, d\sigma_j(x, y) = 0$ for every $g \in L^\infty(\nu_j)$. Summing up, we obtain $\int f(x)\, g(y)\, d\sigma(x, y) = 0$ for every $g \in L^\infty(\nu)$.

In case (ii), for every $j \in J$ the systems (X, μ_j, T) and (Y, ν_j, R) have no common eigenvalue except 1, and thus they are disjoint by part (ii) of Proposition 3.12. Therefore, for every $j \in J$, the measure σ_j defined above is equal to $\mu_j \times \nu_j$. Summing up, we obtain $\sigma = \mu \times \nu$. This completes the proof. □

3.9. *Proof of Theorem* 1.3 *assuming the preceding material.* We consider only the case of the Möbius function; the proof for the Liouville function is identical.

Arguing by contradiction, suppose that the conclusion of Theorem 1.3 fails. Then there exist a topological dynamical system (Y, R), a point $y_0 \in Y$ generic for a measure ν such that the system (Y,ν,R) has zero entropy and at most countably many ergodic components, and a function $g_0 \in C(Y)$ such that the averages

$$\mathbb{E}^{\log}_{n\in[N]} g_0(R^n y_0)\,\boldsymbol{\mu}(n) \tag{12}$$

do not converge to 0 as $N \to \infty$. Hence, there exists a sequence $\mathbf{N} = (N_k)_{k\in\mathbb{N}}$ of intervals with $N_k \to \infty$ such that the limit

$$\mathbb{E}^{\log}_{n\in\mathbf{N}}\, g_0(R^n y_0)\,\boldsymbol{\mu}(n) = \lim_{k\to\infty} \mathbb{E}^{\log}_{n\in[N_k]}\, g_0(R^n y_0)\,\boldsymbol{\mu}(n) \tag{13}$$

exists and is non-zero. After passing to a subsequence, which we also denote by $\mathbf{N}$, we can further assume that the limit

$$\mathbb{E}^{\log}_{n\in\mathbf{N}}\, g(R^n y_0) \prod_{j=1}^{\ell} \boldsymbol{\mu}(n+h_j) \tag{14}$$

exists for every $\ell \in \mathbb{N}$, $h_1,\dots,h_\ell \in \mathbb{Z}$, and $g \in C(Y)$.

Let $X := \{-1,0,1\}^{\mathbb{Z}}$, $T\colon X \to X$ be the shift transformation, and let $x_0 \in X$ be defined by $x_0(n) = \boldsymbol{\mu}(n)$, $n \in \mathbb{Z}$. Then the convergence (14) implies that for every $\ell \in \mathbb{N}$, $h_1,\dots,h_\ell \in \mathbb{Z}$, and every $g \in C(Y)$, the limit

$$\mathbb{E}^{\log}_{n\in\mathbf{N}}\, g(R^n y_0) \Big(\prod_{j=1}^{\ell} F_{h_j}\Big)(T^n x_0)$$

exists. (Recall that $F_h(x) = x(h)$, $x \in X$, $h \in \mathbb{Z}$.) Since the algebra generated by the functions F_h, $h \in \mathbb{Z}$, is dense in $C(X)$ with the uniform topology, we deduce that the sequence of measures

$$\mathbb{E}^{\log}_{n\in[N_k]} \delta_{(T^n x_0, R^n y_0)}, \quad k \in \mathbb{N},$$

converges weak-star to some probability measure σ on $X \times Y$ that satisfies

$$\mathbb{E}^{\log}_{n\in\mathbf{N}}\, g(R^n y_0) \prod_{j=1}^{\ell} \boldsymbol{\mu}(n+h_j) = \int \prod_{j=1}^{\ell} F_{h_j}(x)\, g(y)\, d\sigma(x,y) \tag{15}$$

for every $\ell \in \mathbb{N}$, $h_1,\dots,h_\ell \in \mathbb{Z}$, and $g \in C(Y)$. By construction, σ is invariant under $T \times R$.

The projection of σ on Y is the weak-star limit of the sequence of measures $\mathbb{E}^{\log}_{n\in[N_k]} \delta_{R^n y_0}$, $k \in \mathbb{N}$, and since the point y_0 is generic for ν, this measure is equal to ν, and thus the corresponding measure preserving system has zero entropy and at most countably many ergodic components.

The projection of σ on X is the weak-star limit of the sequence of measures $\mathbb{E}^{\log}_{n\in[N_k]} \delta_{T^n x_0}$, $k \in \mathbb{N}$. It is thus a T-invariant measure μ that is the Furstenberg measure associated with $\boldsymbol{\mu}$ and $\mathbf{N}$ by Proposition 3.2, and σ is a joining of the systems (X,μ,T) and (Y,ν,R).

By Proposition 3.9 and its proof, the system (X, μ, T) is a factor of the system $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$, with factor map $\pi \colon X^{\mathbb{Z}} \to X$ given by

$$(\pi(\underline{x}))(n) = -x_n(0), \quad \underline{x} \in X^{\mathbb{Z}},\ n \in \mathbb{Z}.$$

We define the joining $\widetilde{\sigma}$ of the systems $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ and (Y, ν, R) by

$$\int_{X^{\mathbb{Z}} \times Y} f(\underline{x}) \cdot g(y)\, d\widetilde{\sigma}(\underline{x}, y) = \int_{X \times Y} \mathbb{E}_{\widetilde{\mu}}(f \mid X)(x) \cdot g(y)\, d\sigma(x, y) \tag{16}$$

for every $f \in L^\infty(\widetilde{\mu})$ and $g \in L^\infty(\nu)$.

By Theorems 3.10 and 3.11, the system $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ has no irrational spectrum and its ergodic components are isomorphic to direct products of infinite-step nilsystems and Bernoulli systems.

We now verify that the function $\widetilde{F_0} := F_0 \circ \pi$ is orthogonal to the rational Kronecker factor of the system $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$. In fact we will show that $\widetilde{F_0}$ is orthogonal to the Kronecker factor of this system. By a well-known consequence of the spectral theorem for unitary operators, this property is equivalent to establishing that

$$\mathbb{E}_{n \in \mathbb{N}} \Big| \int \widetilde{F_0} \cdot S^n \widetilde{F_0}\, d\widetilde{\mu} \Big| = 0. \tag{17}$$

By the definition of the measure $\widetilde{\mu}$ (see (11)) and since for $h \in \mathbb{N}$ we have $\widetilde{F_0}(\underline{x})\, \widetilde{F_0}(S^h \underline{x}) = (-F_0(x_0))\,(-F_0(x_h))$, we get for every $n \in \mathbb{N}$ that

$$\int \widetilde{F_0} \cdot S^n \widetilde{F_0}\, d\widetilde{\mu} = \mathbb{E}_{p \in \mathbb{P}} \int F_0 \cdot T^{pn} F_0\, d\mu.$$

By (6), for every $h \in \mathbb{N}$ we have

$$\int F_0 \cdot T^h F_0\, d\mu = \mathbb{E}^{\log}_{n \in \mathbf{N}}\, \boldsymbol{\mu}(n)\, \boldsymbol{\mu}(n+h) = 0,$$

where the vanishing of the average follows from the main result of Tao in [62]. Combining the above identities we get (17).

By Corollary 3.13, we have

$$0 = \int \widetilde{F_0}(\underline{x}) \cdot g_0(y)\, d\widetilde{\sigma}(\underline{x}, y) = \int F_0(x) \cdot g_0(y)\, d\sigma(x, y) = \mathbb{E}^{\log}_{n \in \mathbf{N}}\, g_0(R^n y_0)\, \boldsymbol{\mu}(n)$$

by (15), contradicting our assumption that the limit in (13) is non-zero. This completes the proof. $\square$

3.10. *Proof of Theorem* 1.1 *assuming the preceding material*. We proceed exactly as in the proof of Theorem 1.3 in Section 3.9. Arguing by contradiction, we assume that there exist a topological dynamical system (Y, R), a point $y_0 \in Y$, and a continuous function g_0 on Y such that the logarithmic averages (12) do not converge to 0. We construct a sequence of intervals $\mathbf{N} = (N_k)_{k \in \mathbb{N}}$, a system (X, T), and a measure σ on $X \times Y$, as in the proof of Theorem 1.3 in Section 3.9. The projection ν of σ on Y is an R-invariant measure, and since (Y, R) has at most countably many ergodic invariant measures,

ν has at most countably many ergodic components. Since the system (Y, R) has zero topological entropy, all these components have zero entropy and the system (Y, ν, R) has zero entropy. We conclude as in the proof of Theorem 1.3 in Section 3.9. □

3.11. *Proof of Theorem* 1.4 *assuming the preceding material.* We consider only the case of the Möbius function; the proof for the Liouville function is identical.

Arguing by contradiction, suppose that the conclusion of Theorem 1.4 fails. Then there exist a topological dynamical system (Y, R), a point $y_0 \in Y$ that is generic for a measure ν such that the system (Y, ν, R) has zero entropy and at most countably many ergodic components, all of which are totally ergodic, and a function $g_0 \in C(Y)$ such that for some $\ell_0 \in \mathbb{N}$ and some $h_{0,1}, \dots, h_{0,\ell_0} \in \mathbb{Z}$, the identity (3) fails; namely, the averages

$$\mathbb{E}^{\log}_{n\in[N]}\, g_0(R^n y_0) \prod_{j=1}^{\ell_0} \boldsymbol{\mu}(n + h_{0,j})$$

do not converge to 0 as $N \to \infty$.

As in the proof of Theorem 1.3 in Section 3.9, we define a sequence of intervals $\mathbf{N} = (N_k)_{k\in\mathbb{N}}$ such that the above averages converge to some non-zero number, a system (X, T), and a measure σ on $X \times Y$ such that (15) holds. By construction, σ is invariant under $T \times R$. By assumption and the definition of genericity, the projection of σ on Y is the measure ν, and thus the system (Y, ν, R) has zero entropy, at most countably many ergodic components, and no rational eigenvalue except 1.

The projection of σ on X is a T-invariant measure μ that by (15) is the Furstenberg measure associated with $\boldsymbol{\mu}$ and $\mathbf{N}$ by Proposition 3.2. Hence, by Proposition 3.9, the system (X, μ, T) is a factor of the system $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$. By Theorems 3.10 and 3.11, the system $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ has no irrational spectrum and its ergodic components are isomorphic to direct products of infinite-step nilsystems and Bernoulli systems.

From the previous discussion it follows that the systems $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ and (Y, ν, R) satisfy the hypothesis of part (ii) of Corollary 3.13, hence, they are disjoint. Since the system (X, μ, T) is a factor of $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$, the systems (X, μ, T) and (Y, ν, R) are also disjoint. Since σ is a joining of the systems (X, μ, T) and (Y, ν, R), it is the product measure $\mu \times \nu$. It follows that

$$\mathbb{E}^{\log}_{n\in\mathbf{N}}\, g_0(R^n y_0) \prod_{j=1}^{\ell_0} \boldsymbol{\mu}(n + h_{0,j}) = \int_{X\times Y} \prod_{j=1}^{\ell_0} F_{h_{0,j}}(x) \cdot g_0(y)\, d\sigma(x, y)$$

$$= \int \prod_{j=1}^{\ell_0} F_{h_{0,j}}(x)\, d\mu \cdot \int g(y)\, d\nu = \mathbb{E}^{\log}_{n\in\mathbf{N}} \Big(\prod_{j=1}^{\ell_0} F_{h_{0,j}}\Big)(T^n x_0) \cdot \mathbb{E}^{\log}_{n\in\mathbf{N}}\, g(R^n y_0).$$

The last limit is zero since $\mathbb{E}^{\log}_{n\in\mathbf{N}}\, g(R^n y_0) = \int g\, d\nu = 0$. This contradicts our assumption that $\mathbb{E}^{\log}_{n\in\mathbf{N}}\, g_0(R^n y_0) \prod_{j=1}^{\ell_0} \boldsymbol{\mu}(n + h_{0,j}) \neq 0$ and completes the proof of Theorem 1.4. $\square$

4. **The structure of systems of arithmetic progressions**

The goal of this section is to prove Theorem 3.10, which gives information about the structure of systems of arithmetic progressions with prime steps associated with a system (X, μ, T). We will work progressively with systems of increasing complexity starting from the case where (X, μ, T) is a nilsystem. This important case will be dealt using the theory of arithmetic progressions on nilmanifolds, which is summarized in Appendix B.

4.1. *Systems of arithmetic progressions*. We start with the definition of systems of arithmetic progressions with integer steps. These systems are a stepping stone towards understanding the structure of the systems of arithmetic progressions with prime steps.

4.1.1. *The system of arithmetic progressions with integer steps*. We will use the following result from [36] (convergence was also established in [70]):

THEOREM 4.1. *Let (X, μ, T) be a system. Then for every $\ell \in \mathbb{N}$ and $f_1, \dots, f_\ell \in L^\infty(\mu)$, the following limit exists in $L^2(\mu)$:*

$$\mathbb{E}_{n\in\mathbb{N}} \prod_{j=1}^{\ell} T^{nj} f_j. \tag{18}$$

Furthermore, if the system is ergodic, then Z_∞ is the infinite-step nilfactor of the system (see Appendix A.4*), and if $\mathbb{E}_\mu(f_j \mid Z_\infty) = 0$ for some $j \in \{1, \dots, \ell\}$, then the limit* (18) *is* 0.

In accordance to the system of arithmetic progressions with prime steps (see Definition 3.8) we define systems of arithmetic progressions with integer steps as follows:

Definition 4.2. Let (X, μ, T) be a system. We write $\underline{\mu}$ for the measure on $X^{\mathbb{Z}}$ characterized as follows: For every $m \in \mathbb{N}$ and all $f_{-m}, \dots, f_m \in L^\infty(\mu)$, we define

$$\int_{X^{\mathbb{Z}}} \prod_{j=-m}^{m} f_j(x_j)\, d\underline{\mu}(\underline{x}) := \mathbb{E}_{n\in\mathbb{N}} \int_X \prod_{j=-m}^{m} T^{nj} f_j\, d\mu. \tag{19}$$

Note that the limit above exists by Theorem 4.1 and the measure $\underline{\mu}$ is invariant under the shift S of $X^{\mathbb{Z}}$. We say that $(X^{\mathbb{Z}}, \underline{\mu}, S)$ is the *system of arithmetic progressions with integer steps* associated with the system (X, μ, T).

4.1.2. *The system of arithmetic progressions with prime steps*. The system of arithmetic progressions with prime steps $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ was defined in Section 3.5. We recall here the defining property of the measure $\widetilde{\mu}$: For every $m \in \mathbb{N}$ and $f_{-m}, \ldots, f_m \in L^\infty(\mu)$, we have

$$\int_{X^{\mathbb{Z}}} \prod_{j=-m}^{m} f_j(x_j)\, d\widetilde{\mu}(\underline{x}) = \mathbb{E}_{p\in\mathbb{P}} \int_X \prod_{j=-m}^{m} T^{pj} f_j \, d\mu.$$

Note that convergence of the averages on the right-hand side follows from the next result, which was proved in [23] conditional to some conjectures obtained later in [32], [34], and the convergence part was also proved in [68]:

THEOREM 4.3. *Let (X, μ, T) be a system. Then for every $\ell \in \mathbb{N}$ and $f_1, \ldots, f_\ell \in L^\infty(\mu)$, the following limit exists in $L^2(\mu)$:*

$$\mathbb{E}_{p\in\mathbb{P}} \prod_{j=1}^{\ell} T^{pj} f_j. \tag{20}$$

Furthermore, if the system is ergodic, Z_∞ is the infinite-step nilfactor of the system (see Appendix A.4), and if $\mathbb{E}_\mu(f_j \mid Z_\infty) = 0$ for some $j \in \{1, \ldots, \ell\}$, then the limit (20) is 0.

Remark. This result is not stated explicitly in [23], but follows from the argument in [23, §5], using Theorem 4.1 and $U_{\ell+1}$-uniformity of the W-tricked von Mangoldt function (established in [31], [32], [34]) in place of U_3-uniformity.

In order to determine the support of the measure $\widetilde{\mu}$ we will use the following multiple ergodic theorem:

THEOREM 4.4. *Let (X, μ, T) be a system, and suppose that for some $d \in \mathbb{N}$, the ergodic components of the system (X, μ, T^d) are totally ergodic. Then*

$$\mathbb{E}_{p\in\mathbb{P}} \prod_{j=1}^{\ell} T^{pj} f_j = \mathbb{E}_{(k,d)=1} \mathbb{E}_{n\in\mathbb{N}} \prod_{j=1}^{\ell} T^{(nd+k)j} f_j \tag{21}$$

for all $\ell \in \mathbb{N}$ and $f_1, \ldots, f_\ell \in L^\infty(\mu)$, where convergence takes place in $L^2(\mu)$ and the average $\mathbb{E}_{(k,d)=1}$ is taken over those $k \in \{1, \ldots, d-1\}$ such that $(k, d) = 1$.

Remark. The existence of the limits on the left- and right-hand sides follows from Theorems 4.3 and 4.1 respectively.

Proof. For $w \in \mathbb{N}$, let W denote the product of the first w primes that are relatively prime to d. Following the proof of [24, Th. 1.3] we get that the limit on the left-hand side of (21) is equal to the following limit:[6]

$$\lim_{W\to\infty} \mathbb{E}_{(k,dW)=1}\, \mathbb{E}_{n\in\mathbb{N}} \prod_{j=1}^{\ell} T^{(ndW+k)j} f_j,$$

where the average $\mathbb{E}_{(k,dW)=1}$ is taken over those $k \in \{1,\dots,dW-1\}$ such that $(k,dW)=1$. Since the ergodic components of T^d are totally ergodic, we get by [21, Th. 6.4] (see also Theorem 5.4 below) that

$$\mathbb{E}_{n\in\mathbb{N}} \prod_{j=1}^{\ell} T^{(ndW+k)j} f_j = \mathbb{E}_{n\in\mathbb{N}} \prod_{j=1}^{\ell} T^{(nd+k)j} f_j$$

holds for every $W \in \mathbb{N}$. Hence, the limit we want to compute is

$$\lim_{W\to\infty} \mathbb{E}_{(k,dW)=1} \mathbb{E}_{n\in\mathbb{N}} \prod_{j=1}^{\ell} T^{(nd+k)j} f_j. \tag{22}$$

We claim that for general d-periodic sequences $(a(k))_{k\in\mathbb{N}}$, for every $W \in \mathbb{N}$ with $(d,W)=1$, we have

$$\mathbb{E}_{(k,dW)=1} a(k) = \mathbb{E}_{(k,d)=1} a(k). \tag{23}$$

To see this, for $j \in \{0,\dots,d-1\}$ consider the set

$$A_j := \{k \in \{1,\dots dW\}\colon k \equiv j \pmod d \ \text{ and } \ (k,Wd)=1\}.$$

If $(j,d)>1$, then $A_j=\emptyset$. If $(j,d)=1$, then $(k,d)=1$ and

$$A_j = \{k \in \{1,\dots dW\}\colon k \equiv j \pmod d \ \text{ and } \ (k,W)=1\}.$$

Since $(W,d)=1$, we have $|A_j| = \phi(W)$ if $(j,d)=1$. It follows from these simple facts and our assumption of d-periodicity of $(a(k))_{k\in\mathbb{N}}$ that (23) holds.

Applying (23) for $a(k) := \mathbb{E}_{n\in\mathbb{N}} \prod_{j=1}^{\ell} T^{(nd+k)j} f_j$, $k\in\mathbb{N}$, which is d-periodic, we see that the limit in (22) is equal to the expression on the right-hand side of (21). This completes the proof. $\square$

4.2. *The case of a nilsystem.* We start with the following intermediate result, which establishes Theorem 3.10 in the case where (X,μ,T) is a (finite-step) nilsystem:

PROPOSITION 4.5. *If (X,μ,T) is an ergodic nilsystem, then the ergodic components of the systems $(X^{\mathbb{Z}},\underline{\mu},S)$ and $(X^{\mathbb{Z}},\widetilde{\mu},S)$ are isomorphic to nilsystems.*

The proof is given in Section 4.2.3. We start with some preliminaries.

[6]This is established in [24] only for $d=1$, but the same argument works for every $d\in\mathbb{N}$ using the Gowers uniformity (as $N\to\infty$ and then $W\to\infty$) of the W-tricked von Mangoldt function $(\frac{\phi(dW)}{dW}\Lambda(dWn+k)-1)_{n\in[N]}$ for $k\in\mathbb{N}$ relatively prime to dW.

Notation. If T is a transformation on X, we write $\overline{T}$ and $\overrightarrow{T}$ for the transformations of $X^{\mathbb{Z}}$ given by

$$(\overline{T}\underline{x})_j = Tx_j \quad \text{and} \quad (\overrightarrow{T}\underline{x})_j = T^j x_j, \quad j \in \mathbb{Z},$$

where $\underline{x} = (x_k)_{k\in\mathbb{Z}} \in X^{\mathbb{Z}}$. We call $\overline{T}$ the *diagonal transformation.* As usual, with S we denote the shift transformation on $X^{\mathbb{Z}}$.

We remark that $\overline{T}$ commutes with $\overrightarrow{T}$ and with S, and that $[S, \overrightarrow{T}] = \overline{T}$.

4.2.1. *Integer steps*. We use the same hypothesis and notation as in the preceding sections, and now we assume in addition that $X = G/\Gamma$ is a nilmanifold, $\mu = \mu_X$ is the Haar measure on X, and T is an ergodic translation by some $\tau \in G$. Arguing as in [48, §2.1] we can and will assume that G is spanned by the connected component G^0 of e_G and τ. This condition implies that the groups G_s are connected for every $s \geq 2$ (see [2, Th. 4.1]). The transformations $\overline{T}$ and $\overrightarrow{T}$ of $X^{\mathbb{Z}}$ are the translations by $\overline{\tau} = (\dots, \tau, \tau, \tau \dots)$ and $\overrightarrow{\tau} = (\dots, \tau^{-2}, \tau^{-1}, e_G, \tau, \tau^2, \dots)$, respectively.

The Hall-Petresco group $\underline{G}$ and the nilmanifold of arithmetic progressions $\underline{X}$ are defined in the Appendices B.1 and B.2. It is immediate from the definition of $\underline{G}$ that $\overline{\tau}, \overrightarrow{\tau} \in \underline{G}$. Therefore, $\overline{T}$ and $\overrightarrow{T}$ are nilrotations of $\underline{X}$. The next result was established in [2, Lemma 5.2]:

Lemma 4.6. *If (X, T) is a minimal nilsystem, then*

$$\underline{X} = \overline{\left\{\overrightarrow{T}^n \overline{T}^m e_{\underline{X}} \colon m, n \in \mathbb{Z}\right\}}.$$

The next result was established in the form stated in [2, Th. 5.4] and previously in a slightly different form in [69]:

Proposition 4.7. *Let (X, T, μ) be an ergodic nilsystem. Then for every $m \in \mathbb{N}$ and all $f_{-m}, \dots, f_m \in L^\infty(\mu)$, we have*

$$\int_{\underline{X}} \prod_{j=-m}^{m} f_j(x_j)\, d\mu_{\underline{X}}(\underline{x}) = \mathbb{E}_{n\in\mathbb{N}} \int_X \prod_{j=-m}^{m} T^{nj} f_j \, d\mu.$$

In other words, the Haar measure $\mu_{\underline{X}}$ of $\underline{X}$ coincides with the measure $\underline{\mu}$ on $\underline{X}$ defined in Definition 4.2.

4.2.2. *Prime steps*. Let (X, μ, T) be an ergodic nilsystem. It is a known and easy to prove fact that this system is totally ergodic if and only if X is connected. In general, let X_0 be the connected component of e_X and μ_0 be its Haar measure. Then there exists $d \in \mathbb{N}$ such that the sets $T^l X_0$, $l \in \{0, \dots, d-1\}$, form a partition of X and we have

$$\mu = \mathbb{E}_{0\leq l\leq d-1} T^l \mu_0. \tag{24}$$

Moreover, the system (X_0, μ_0, T^d) and the other ergodic components of the system (X, μ, T^d) are totally ergodic. We call d the *index* of X_0.

Let $\underline{X}_0 \subset X_0^{\mathbb{Z}}$ and the measure $\underline{\mu}_0$ on $\underline{X}_0$ be defined as $\underline{X}$ and $\underline{\mu}$ are defined in Definition 4.2, with the system (X_0, μ_0, T^d) in place of (X, μ, T). Then $\underline{X}_0$ and $\underline{\mu}_0$ are invariant under $\overline{T}^d$, $\overrightarrow{T}^d$, and S. Applying Theorem 4.4 for the nilsystem (X, μ, T) that has index d, we get that for every $m \in \mathbb{N}$ and $f_{-m}, \ldots, f_m \in L^\infty(\mu)$, we have

$$\mathbb{E}_{p\in\mathbb{P}} \int_X \prod_{j=-m}^{m} T^{pj} f_j \, d\mu = \mathbb{E}_{(k,d)=1} \mathbb{E}_{n\in\mathbb{N}} \int_X \prod_{j=-m}^{m} T^{(nd+k)j} f_j \, d\mu, \tag{25}$$

where the average $\mathbb{E}_{(k,d)=1}$ is taken over those $k \in \{1, \ldots, d-1\}$ such that $(k, d) = 1$. Combining (11), (24), and (25), we get for every $m \in \mathbb{N}$ and $f_{-m}, \ldots, f_m \in L^\infty(\mu)$ that

$$\int_{\underline{X}} \prod_{j=-m}^{m} f_j(x_j) \, d\widetilde{\mu}(\underline{x}) = \mathbb{E}_{0\leq l\leq d-1} \mathbb{E}_{(k,d)=1} \mathbb{E}_{n\in\mathbb{N}} \int_X \prod_{j=-m}^{m} T^{(nd+k)j+l} f_j \, d\mu_0.$$

Moreover, applying (19) for the system (X_0, μ_0, T^d) we get

$$\int_{\underline{X}} \prod_{j=-m}^{m} f_j(x_j) \, d\underline{\mu}_0(\underline{x}) = \mathbb{E}_{n\in\mathbb{N}} \int_X \prod_{j=-m}^{m} T^{ndj} f_j \, d\mu_0.$$

Combining the last two identities we deduce that

$$\widetilde{\mu} = \mathbb{E}_{0\leq l\leq d-1} \mathbb{E}_{(k,d)=1} \overline{T}^l \overrightarrow{T}^k \underline{\mu}_0. \tag{26}$$

Since the support of $\underline{\mu}_0$ is $\underline{X}_0$, it follows that the measure $\widetilde{\mu}$ is supported on the set

$$\widetilde{X} := \bigcup_{l=0}^{d-1} \bigcup_{k\colon (k,d)=1} \overline{T}^l \overrightarrow{T}^k \underline{X}_0.$$

The precise form of $\widetilde{X}$ is not important; the crucial point is that $\widetilde{X} \subset \underline{X}$. To see this, note that Lemma 4.6 implies that the set $\underline{X}$ is $\overline{T}$ and $\overrightarrow{T}$ invariant and

$$\underline{X}_0 = \overline{\left\{ \overrightarrow{T}^{dn} \overline{T}^{dm} e_{\underline{X}_0} \colon m, n \in \mathbb{Z} \right\}} \subset \underline{X}.$$

4.2.3. *Proof of Proposition* 4.5. Let $\widetilde{\mu} = \int \widetilde{\mu}_\omega \, dP(\omega)$ be the ergodic decomposition of the measure $\widetilde{\mu}$ with respect to the transformation S acting on $X^{\mathbb{Z}}$. Since, as established above, $\widetilde{\mu}$ is supported on the S-invariant set $\underline{X}$, almost every ergodic component $\widetilde{\mu}_\omega$ admits a generic point in $\underline{X}$. For these ω, we have that $\widetilde{\mu}_\omega$ is supported on a closed S-orbit in $\underline{X}$, which we denote by $\widetilde{X}_\omega$. By Proposition B.4, the system $(\widetilde{X}_\omega, S)$ is topologically isomorphic to a uniquely ergodic nilsystem. Thus, $\widetilde{\mu}_\omega$ is the unique invariant measure for the action of S on $\widetilde{X}_\omega$ and the system $(\widetilde{X}_\omega, \widetilde{\mu}_\omega, S)$ is (measure theoretically) isomorphic to an ergodic nilsystem.

A similar argument applies to the system $(\underline{X}, \underline{\mu}, S)$. ∎

4.3. *The case of an infinite-step nilsystem*. Our next goal is to treat the case where (X,μ,T) is an ergodic infinite-step nilsystem and prove the following intermediate result:

PROPOSITION 4.8. *If (X,μ,T) is an ergodic infinite-step nilsystem, then the ergodic components of the systems $(X^{\mathbb{Z}},\underline{\mu},S)$ and $(X^{\mathbb{Z}},\widetilde{\mu},S)$ are isomorphic to infinite-step nilsystems.*

The proof is given in Section 4.3.3. We start with some preliminaries.

Our setup is as follows (see Appendix A for definitions and properties of inverse limits): We have $(X,\mu,T)=\varprojlim(X_j,\mu_j,T)$ where for $j\in\mathbb{N}$, the system (X_j,μ_j,T) is an ergodic nilsystem with base point e_{X_j}. For $j\in\mathbb{N}$, the factor maps are written $\pi_{j,j+1}\colon X_{j+1}\to X_j$ and $\pi_j\colon X\to X_j$ and, as explained in Appendix A.3, $\pi_{j,j+1}$ and π_j are also topological factor maps. Thus, we also have $(X,T)=\varprojlim(X_j,T)$ in the topological sense (see Appendix A.3).

The sequence $(X_j^{\mathbb{Z}},\overline{T},\overrightarrow{T})$, $j\in\mathbb{N}$, with factor maps $\pi_{j,j+1}^{\mathbb{Z}}\colon X_{j+1}^{\mathbb{Z}}\to X_j^{\mathbb{Z}}$, $j\in\mathbb{N}$, is an inverse system. By the characterization of inverse limits stated in (i) and (ii) of Appendix A.2, we get that $(X^{\mathbb{Z}},\overline{T},\overrightarrow{T})$, endowed with the factor maps $\pi_j^{\mathbb{Z}}\colon X^{\mathbb{Z}}\to X_j^{\mathbb{Z}}$, $j\in\mathbb{N}$, is the inverse limit of the sequence $(X_j^{\mathbb{Z}},\overline{T},\overrightarrow{T})$, $j\in\mathbb{N}$.

4.3.1. *Integer steps*. Let $\underline{X}$ be the orbit closure in $X^{\mathbb{Z}}$ of

$$e_{\underline{X}}:=(\dots,e_X,e_X,e_X,\dots)$$

under the transformations $\overline{T}$ and $\overrightarrow{T}$. Since $\pi_j^{\mathbb{Z}}(e_{\underline{X}})=e_{\underline{X}_j}$ for every $j\in\mathbb{N}$, it follows from Lemma 4.6 and part (i) of Lemma A.2 that $\pi_j^{\mathbb{Z}}(\underline{X})=\underline{X}_j$, $j\in\mathbb{N}$, and $(\underline{X},\overline{T},\overrightarrow{T})$ is the inverse limit of the systems $(\underline{X}_j,\overline{T},\overrightarrow{T})$, $j\in\mathbb{N}$. In particular, we have

$$\underline{X}=\Big\{\underline{x}\in X^{\mathbb{Z}}\colon \pi_j^{\mathbb{Z}}(\underline{x})\in\underline{X}_j\ \text{ for every } j\in\mathbb{N}\Big\}. \tag{27}$$

Note that for $j\in\mathbb{N}$, the maps $\pi_{j,j+1}^{\mathbb{Z}}\colon\underline{X}_{j+1}\to\underline{X}_j$ and $\pi_j^{\mathbb{Z}}\colon\underline{X}\to\underline{X}_j$ commute with the shift transformation S, and thus are factor maps from $(\underline{X}_{j+1},S)$ and $(\underline{X},S)$ to $(\underline{X}_j,S)$, respectively. It follows from the characterization of topological inverse limits stated in (i) and (ii) of Appendix A.2 that

$$(\underline{X},S)=\varprojlim(\underline{X}_j,S)$$

with factor maps $\pi_{j,j+1}^{\mathbb{Z}}\colon\underline{X}_{j+1}\to\underline{X}_j$ and $\pi_j^{\mathbb{Z}}\colon\underline{X}\to\underline{X}_j$, $j\in\mathbb{N}$. By Proposition B.4, for every $j\in\mathbb{N}$ we have that $(\underline{X}_j,S)$ is topologically isomorphic to a nilsystem. Hence the action of S on each closed orbit under S in $\underline{X}_j$ induces a uniquely ergodic nilsystem. From Lemma A.2 we deduce the following:

Proposition 4.9. *Let $\underline{X}$ be as above. For $\underline{x} \in \underline{X}$, let $\underline{X}' := \overline{\{S^n \underline{x} \colon n \in \mathbb{Z}\}}$ be the closed orbit of $\underline{x}$ under S. Then the system $(\underline{X}', S)$ is topologically isomorphic to a uniquely ergodic infinite-step nilsystem.*

4.3.2. *Prime steps*. From Definition 3.8 it follows that for every $j \in \mathbb{N}$, the image of the measure $\widetilde{\mu}$ under the maps $\pi_j^{\mathbb{Z}}$ is equal to $\widetilde{\mu}_j$ and that the image of $\widetilde{\mu}_{j+1}$ under $\pi_{j,j+1}^{\mathbb{Z}}$ is equal to $\widetilde{\mu}_j$. These maps commute with S. Hence it follows from the characterization of inverse limits (i) and (ii) given in Appendix A.1 that

$$(X^{\mathbb{Z}}, \widetilde{\mu}, S) = \varprojlim (X_j^{\mathbb{Z}}, \widetilde{\mu}_j, S). \tag{28}$$

Furthermore, we saw in Section 4.2.2 that for every $j \in \mathbb{N}$, the measure $\widetilde{\mu}_j$ is supported inside $\underline{X}_j$ and thus

$$\widetilde{\mu}\Big(\big\{\underline{x} \in X^{\mathbb{Z}} \colon \pi_j^{\mathbb{Z}}(\underline{x}) \notin \underline{X}_j\big\}\Big) = 0.$$

It follows from this and (27) that $\widetilde{\mu}$ is supported inside the subset $\underline{X}$ of $X^{\mathbb{Z}}$.

4.3.3. *Proof of Proposition* 4.8. In the previous subsection we established that the measure $\widetilde{\mu}$ is supported inside the S-invariant set $\underline{X}$. Using this and Proposition 4.9 we deduce that almost every ergodic component of the system $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ is isomorphic to an infinite-step nilsystem; the argument is identical to the one used in the last step of the proof of Proposition 4.5 (see Section 4.2.3).

A similar argument applies to the system $(X^{\mathbb{Z}}, \underline{\mu}, S)$. □

4.4. *General ergodic systems*. Our next goal is to prove the following result, which comes very close to establishing Theorem 3.10:

Proposition 4.10. *If (X, μ, T) is an ergodic system, then almost every ergodic component of the systems $(X^{\mathbb{Z}}, \underline{\mu}, S)$ and $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ is isomorphic to a direct product of an infinite-step nilsystem and a Bernoulli system.*

This result is proved in Section 4.4.1. First we make some preparatory work.

Let (X, μ, T) be an ergodic system. The infinite-step nilfactor of the system is defined in Section A.4 and is denoted by $(Z_\infty, \mu_\infty, T)$; in Corollary A.6 we show that it is isomorphic to an infinite-step nilsystem. Let $p_\infty \colon X \to Z_\infty$ be the corresponding factor map, and let the measures $\underline{\mu}_\infty$ and $\widetilde{\mu}_\infty$ on $Z_\infty^{\mathbb{Z}}$ be associated with the system $(Z_\infty, \mu_\infty, T)$ as in Definitions 3.8 and 4.2 respectively. Then $\underline{\mu}_\infty$ and $\widetilde{\mu}_\infty$ are respectively the images of $\underline{\mu}$ and $\widetilde{\mu}$ under $p_\infty^{\mathbb{Z}} \colon X^{\mathbb{Z}} \to Z_\infty^{\mathbb{Z}}$. Combining the second part of Theorems 4.1 and 4.3 with the definitions of the measures $\underline{\mu}$ and $\widetilde{\mu}$, we get for every $m \in \mathbb{N}$ and $f_{-m}, \ldots, f_m \in L^\infty(\mu)$ that

$$\int_{X^{\mathbb{Z}}} \prod_{j=-m}^{m} f_j(x_j)\, d\underline{\mu}(\underline{x}) = \int_{Z_\infty^{\mathbb{Z}}} \prod_{j=-m}^{m} \mathbb{E}_\mu(f_j \mid Z_\infty)(z_j)\, d\underline{\mu}_\infty(\underline{z})$$

and

$$\int_{X^{\mathbb{Z}}} \prod_{j=-m}^{m} f_j(x_j)\, d\widetilde{\mu}(\underline{x}) = \int_{Z_\infty^{\mathbb{Z}}} \prod_{j=-m}^{m} \mathbb{E}_\mu(f_j \mid Z_\infty)(z_j)\, d\widetilde{\mu}_\infty(\underline{z}). \tag{29}$$

LEMMA 4.11. *Let (X,μ,T) be an ergodic system and (Z_∞,μ_∞,T) be its infinite-step nilfactor. Then the system $(X^{\mathbb{Z}},\widetilde{\mu},S)$ is isomorphic to the direct product of the system $(Z_\infty^{\mathbb{Z}},\widetilde{\mu}_\infty,S)$ and a Bernoulli system (that can be trivial). A similar statement also holds for the system $(X^{\mathbb{Z}},\underline{\mu},S)$.*

Proof of Lemma 4.11. We give the argument for the system $(X^{\mathbb{Z}},\widetilde{\mu},S)$; an analogous argument works for the system $(X^{\mathbb{Z}},\underline{\mu},S)$.

Since the system (X,μ,T) is ergodic (and it is our working assumption that it is Lebesgue), it is a classical result of Rohlin (see, for example, [28, Th. 3.18]) that there exists a (Lebesgue) probability space (U,ρ) such that the (Lebesgue) probability spaces (X,μ) and $(Z_\infty,\mu_\infty)\times(U,\rho)$ are isomorphic, the factor map $p_\infty\colon X\to Z_\infty$ corresponds to the first coordinate projection $Z_\infty\times U\to Z_\infty$, and the conditional expectation $f\mapsto \mathbb{E}(f\mid Z_\infty)$ corresponds to the map $f\mapsto \int f(\cdot,u)\,d\rho(u)$ from $L^1(\mu_\infty\times\rho)$ to $L^1(\mu_\infty)$. We identify x with (z,u) and $\underline{x}$ with $(\underline{z},\underline{u})$; then identity (29) becomes

$$\begin{aligned}\int_{X^{\mathbb{Z}}} \prod_{j=-m}^{m} f_j(x_j)\, d\widetilde{\mu}(\underline{x}) &= \int_{Z_\infty^{\mathbb{Z}}} \prod_{j=-m}^{m} \Big(\int_U f_j(z_j,u_j)\,d\rho(u_j)\Big) d\widetilde{\mu}_\infty(\underline{z})\\ &= \int_{Z_\infty^{\mathbb{Z}}\times U^{\mathbb{Z}}} \prod_{j=-m}^{m} f_j(z_j,u_j)\, d(\widetilde{\mu}_\infty\times\rho^{\mathbb{Z}})(\underline{z},\underline{u}),\end{aligned}$$

where $\rho^{\mathbb{Z}}$ is the measure $\cdots\times\rho\times\rho\times\rho\times\cdots$ on $U^{\mathbb{Z}}$.

Since the algebra generated by functions of the form $\underline{x}\mapsto f(x_j)$, $j\in\mathbb{Z}$, $f\in C(X)$, is dense in $C(X^{\mathbb{Z}})$ with the uniform topology, we deduce that $\widetilde{\mu}=\widetilde{\mu}_\infty\times\rho^{\mathbb{Z}}$. Let S_1, S_2 denote the shift transformations on the spaces $Z_\infty^{\mathbb{Z}}$ and U^Z respectively. Then the system $(X^{\mathbb{Z}},\widetilde{\mu},S)$ is the direct product of the system $(Z_\infty^{\mathbb{Z}},\widetilde{\mu}_\infty,S_1)$ and the Bernoulli system $(U^{\mathbb{Z}},\rho^{\mathbb{Z}},S_2)$. This completes the proof. □

4.4.1. *Proof of Proposition* 4.10. We give the argument for the system $(X^{\mathbb{Z}},\widetilde{\mu},S)$; an analogous argument works for the system $(X^{\mathbb{Z}},\underline{\mu},S)$.

By Lemma 4.11, the system $(X^{\mathbb{Z}},\widetilde{\mu},S)$ is isomorphic to the direct product of the system $(Z_\infty^{\mathbb{Z}},\widetilde{\mu}_\infty,S)$ and a Bernoulli system. Since Bernoulli systems are weakly mixing, almost every ergodic component of $(X^{\mathbb{Z}},\widetilde{\mu},S)$ is a direct product of an ergodic component of the system $(Z_\infty^{\mathbb{Z}},\widetilde{\mu}_\infty,S)$ and the Bernoulli system given by Lemma 4.11. (We used the uniqueness property of the ergodic decomposition here.) As explained in Section A.4, the system (Z_∞,μ_∞,T) is isomorphic to an ergodic infinite-step nilsystem. Hence Proposition 4.8 applies and gives that the ergodic components of the system $(Z_\infty^{\mathbb{Z}}, \widetilde{\mu}_\infty, S)$ are isomorphic to infinite-step nilsystems. This completes the proof of Proposition 4.10. $\square$

4.5. *General systems — Proof of Theorem* 3.10. Let (X, μ, T) be a system, and let $\mu = \int \mu_\omega \, dP(\omega)$ be the ergodic decomposition of μ under T. It follows from Definition 3.8 that

$$\widetilde{\mu} = \int \widetilde{\mu}_\omega \, dP(\omega).$$

As a consequence, by the uniqueness property of the ergodic decomposition, almost every ergodic component of the system $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ is an ergodic component of the system $(X^{\mathbb{Z}}, \widetilde{\mu}_\omega, S)$ for some $\omega \in \Omega$. We can therefore restrict to the case where the system (X, μ, T) is ergodic. In this case the result follows from Proposition 4.10. This completes the proof of Theorem 3.10.

A similar argument applies for the system $(X^{\mathbb{Z}}, \underline{\mu}, S)$. $\square$

5. Strong stationarity and systems of arithmetic progressions

The goal of this section is to introduce the notion of strong stationarity and variants of it that turn out to be linked to structural properties of systems of arithmetic progressions. We then use this connection in order to prove that systems of arithmetic progressions have no irrational spectrum, thus establishing Theorem 3.11, which in turn gives the first part of Theorem 1.6 (via Proposition 3.9).

5.1. *Strong stationarity*. Throughout this section we continue to denote by X a compact metric space, and we equip the sequence space $X^{\mathbb{Z}}$ with the product topology and the Borel σ-algebra. With S we denote the shift transformation on $X^{\mathbb{Z}}$. With $\mathcal{B}_0$ we denote all Borel subsets of $X^{\mathbb{Z}}$ that depend only on the 0-th coordinate of elements of $X^{\mathbb{Z}}$. Equivalently, $\mathcal{B}_0$ consists of sets of the form $\{x \in X^{\mathbb{Z}} \colon x(0) \in A\}$, where A is a Borel subset of X. We also denote by $\mathcal{F}_0$ the algebra of $\mathcal{B}_0$-measurable functions.

For $r \in \mathbb{N}$, we define the map $\tau_r \colon X^{\mathbb{Z}} \to X^{\mathbb{Z}}$ by

$$(\tau_r(\underline{x}))(j) := x(rj) \quad \text{for } \underline{x} \in X^{\mathbb{Z}} \text{ and } j \in \mathbb{Z}.$$

We remark that the maps S and τ_r satisfy the following commutation relation

$$S \circ \tau_r = \tau_r \circ S^r. \tag{30}$$

The notion of strong stationarity was introduced in a rather abstract setting by Furstenberg and Katznelson in [27]; here we use a variant adapted to our purposes:

Definition 5.1. If X is as above, we say that an S-invariant Borel measure ν on $X^{\mathbb{Z}}$ is *strongly stationary* if it is invariant under τ_r for every $r \in \mathbb{N}$, and *partially strongly stationary* if for some $d \in \mathbb{N}$, it is invariant under τ_r for every $r \in d\mathbb{N}+1$. Respectively, we say that the system $(X^{\mathbb{Z}}, \nu, S)$ is *strongly stationary* and *partially strongly stationary.*

Remark. Equivalently, we have strong stationarity if and only if

$$\int \prod_{j=-m}^{m} S^j f_j \, d\nu = \int \prod_{j=-m}^{m} S^{rj} f_j \, d\nu$$

for all $m, r \in \mathbb{N}$ and $f_{-m}, \ldots, f_m \in C(X^{\mathbb{Z}}) \cap \mathcal{F}_0$. A similar equivalent condition holds for partial strong stationarity.

In the next subsection we explain why the notion of partial strong stationarity is linked to structural properties of systems of arithmetic progressions.

5.2. *Systems of arithmetic progressions and partial strong stationarity*. If a system is totally ergodic, then it can be shown that the associated system of arithmetic progressions with prime and integer steps is strongly stationary. The notion of total ergodicity turns out to be too restrictive, so we introduce a somewhat weaker notion that is better adapted to our purposes.

Definition 5.2. We say that a system (X, μ, T) has *finite rational spectrum* if the set of eigenvalues of the system of the form $\mathrm{e}(t)$ with $t \in \mathbb{Q}$ is finite.

Remark. Equivalently, (X, μ, T) has finite rational spectrum if there exists $d \in \mathbb{N}$ such that the ergodic components of the system (X, μ, T^d) are totally ergodic.

The link between strong stationarity and systems of arithmetic progressions is given by the next result, which is proved in Section 5.2.2 and forms an essential part of the proof of Theorem 3.11:

Proposition 5.3. *Let (X, μ, T) be a system with finite rational spectrum. Then the systems $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ and $(X^{\mathbb{Z}}, \underline{\mu}, S)$ are partially strongly stationary.*

Remark. Our argument shows that we get full strong stationarity if the ergodic components of the system (X, μ, T) are totally ergodic. We do not use this fact though because we are not able to verify this hypothesis for Furstenberg systems of the Liouville function.

5.2.1. *Some multiple ergodic theorems*. The proof of Proposition 5.3 is rather simple but is based on some highly non-trivial known identities involving multiple ergodic averages that we use as a black box. Note that we implicitly assume convergence in $L^2(\mu)$ for all the multiple ergodic averages in this subsection; this is guaranteed to be the case by Theorems 4.1 and 4.3.

The first identity we use was proved in [21, Th. 6.4]:

THEOREM 5.4. *Suppose that the ergodic components of the system (X, μ, T) are totally ergodic. Then for every $r \in \mathbb{N}$, we have*

$$\mathbb{E}_{n\in\mathbb{N}} \prod_{j=1}^{\ell} T^{nj} f_j = \mathbb{E}_{n\in\mathbb{N}} \prod_{j=1}^{\ell} T^{rnj} f_j$$

for all $\ell \in \mathbb{N}$ and $f_1, \dots, f_\ell \in L^\infty(\mu)$, where convergence takes place in $L^2(\mu)$.

Combining this result with Theorem 4.4 we get the following ergodic theorem that is better adapted to our purposes:

COROLLARY 5.5. *Let $d \in \mathbb{N}$ and (X, μ, T) be a system such that the ergodic components of the system (X, μ, T^d) are totally ergodic. Then for every $r \in \mathbb{N}$ with $(r, d) = 1$, we have*

$$\mathbb{E}_{n\in\mathbb{N}} \prod_{j=1}^{\ell} T^{nj} f_j = \mathbb{E}_{n\in\mathbb{N}} \prod_{j=1}^{\ell} T^{rnj} f_j \quad \text{and} \quad \mathbb{E}_{p\in\mathbb{P}} \prod_{j=1}^{\ell} T^{pj} f_j = \mathbb{E}_{p\in\mathbb{P}} \prod_{j=1}^{\ell} T^{rpj} f_j$$

for all $\ell \in \mathbb{N}$ and $f_1, \dots, f_\ell \in L^\infty(\mu)$, where convergence takes place in $L^2(\mu)$.

Proof. We prove the second identity; the proof of the first is similar. (Simply replace below $p \in \mathbb{P}$ with $n \in \mathbb{N}$ and $\mathbb{E}_{(k,d)=1}$ with $\mathbb{E}_{k\in[d]}$.) Our assumption gives that the ergodic components of $(T^r)^d$ are also totally ergodic. By Theorem 4.4 (applied for T^r in place of T), we get the identity

$$\mathbb{E}_{p\in\mathbb{P}} \prod_{j=1}^{\ell} T^{rpj} f_j = \mathbb{E}_{(k,d)=1} \mathbb{E}_{n\in\mathbb{N}} \prod_{j=1}^{\ell} T^{(dn+k)rj} f_j,$$

where the average $\mathbb{E}_{(k,d)=1}$ is taken over those $k \in \{1, \dots, d-1\}$ such that $(k, d) = 1$. Using Theorem 5.4, we get that the average on the right-hand side is equal to

$$\mathbb{E}_{(k,d)=1} \mathbb{E}_{n\in\mathbb{N}} \prod_{j=1}^{\ell} T^{(dn+kr)j} f_j = \mathbb{E}_{(k,d)=1} \mathbb{E}_{n\in\mathbb{N}} \prod_{j=1}^{\ell} T^{(dn+k)j} f_j = \mathbb{E}_{p\in\mathbb{P}} \prod_{j=1}^{\ell} T^{pj} f_j,$$

where the first identity follows since $(r, d) = 1$ and the second from Theorem 4.4. Combining the above we get the asserted identity. $\square$

5.2.2. *Proof of Proposition* 5.3. Our assumption gives that there exists $d \in \mathbb{N}$ such that the ergodic components of the system (X, μ, T^d) are totally ergodic. Let $m \in \mathbb{N}$ and $f_{-m}, \ldots, f_m \in C(X^{\mathbb{Z}}) \cap \mathcal{F}_0$. We have

$$\begin{aligned}\int_{X^{\mathbb{Z}}} \prod_{j=-m}^{m} S^{(dn+1)j} f_j \, d\widetilde{\mu} &= \mathbb{E}_{p \in \mathbb{P}} \int_X \prod_{j=-m}^{m} T^{(dn+1)pj} f_j \, d\mu \\ &= \mathbb{E}_{p \in \mathbb{P}} \int_X \prod_{j=-m}^{m} T^{pj} f_j \, d\mu = \int_{X^{\mathbb{Z}}} \prod_{j=-m}^{m} S^j f_j \, d\widetilde{\mu},\end{aligned}$$

where we used the defining property of the measure $\widetilde{\mu}$ (see Definition 3.8) to get the first and third identity and the second identity of Corollary 5.5 (for $r := dn+1$) to get the middle identity. This proves that the system $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ is partially strongly stationary.

A similar argument shows that the system $(X^{\mathbb{Z}}, \underline{\mu}, S)$ is partially strongly stationary; the only difference is that one uses the first identity of Corollary 5.5 instead of the second.

5.3. *Spectrum of partially strongly stationary systems*. The next result was obtained in [42, §3] for ergodic strongly stationary systems, but the same argument also works with minor modifications for partially strongly stationary systems that are not necessarily ergodic. We will summarize its proof for completeness. Note also that a somewhat more complicated argument can be used to show that a strongly stationary system can only have 1 in its spectrum (see [42, §4]); but unfortunately a similar result fails for partially strongly stationary systems that can have a rational spectrum different than 1.

Proposition 5.6. *Let $(X^{\mathbb{Z}}, \nu, S)$ be a partially strongly stationary system. Then the system has no irrational spectrum.*

In the proof of Proposition 5.6 we will use the following key property of the maps τ_r:

Lemma 5.7 (Lemma 2.3 in [42]). *Let χ be an eigenfunction of the system $(X^{\mathbb{Z}}, \nu, S)$ with eigenvalue $\mathrm{e}(t)$, and suppose that for some $r \in \mathbb{N}$, the measure ν is invariant under τ_r. Then $\chi \circ \tau_r$ is a finite linear combination of eigenfunctions for eigenvalues of the form $\mathrm{e}((j+t)/r)$ for $j = 0, \ldots, r-1$.*

Proof. For $j = 0, \ldots, r-1$, let $g_j := \sum_{k=0}^{r-1} \mathrm{e}(-k(j+t)/r)\, \chi \circ \tau_r \circ S^k$. Then direct computation shows that $g_j \circ S = \mathrm{e}((j+t)/r)\, g_j$, $j = 0, \ldots, r-1$ and that $\chi = \sum_{j=0}^{r-1} g_j$. □

We will also use the following classical variant of van der Corput's fundamental lemma (the stated version is from [1]):

Lemma 5.8 (Van der Corput). *Let $(v_n)_{n \in \mathbb{N}}$ be a bounded sequence of vectors in a Hilbert space. Suppose that for each $h \in \mathbb{N}$, we have*

$$\mathbb{E}_{n \in \mathbb{N}} \langle v_{n+h}, v_n \rangle = 0.$$

Then

$$\mathbb{E}_{n\in\mathbb{N}}\, v_n = 0,$$

where convergence takes place in norm.

We are now ready to prove Proposition 5.6.

Proof of Proposition 5.6. By our assumption, there exists $d \in \mathbb{N}$ such that the measure ν is τ_r-invariant for every $r \in d\mathbb{N}+1$.

Let $\chi \in L^\infty(\mu)$ be such that $S\chi = \lambda \cdot \chi$, where $\lambda = \mathrm{e}(\alpha)$ with α irrational. We will show that $\chi = 0$. To do this we follow closely the argument of Jenvey in [42, §3].

Since for $r \in d\mathbb{N}+1$ the maps τ_r leave the 0-th coordinate of $x \in X^{\mathbb{Z}}$ unchanged, we have $f = f \circ \tau_r$ for every $f \in \mathcal{F}_0$. Since linear combinations of functions of the form $\prod_{j=-m}^m S^j f_j$ with $f_{-m}, \dots, f_m \in C(X^{\mathbb{Z}}) \cap \mathcal{F}_0$, $m \in \mathbb{N}$, are dense in the space $C(X^{\mathbb{Z}})$ with the uniform topology, it suffices to show that

$$\int \chi \cdot \prod_{j=-m}^{m} S^j f_j \, d\nu = 0$$

for all $m \in \mathbb{N}$ and $f_{-m}, \dots, f_m \in C(X^{\mathbb{Z}}) \cap \mathcal{F}_0$. Composing with the ν-preserving maps S^m for $m \in \mathbb{N}$, we see that it suffices to show that

$$\int \chi \cdot \prod_{j=0}^{m} S^j f_j \, d\nu = 0 \tag{31}$$

for all $m \in \mathbb{N}$ and $f_0, \dots, f_m \in C(X^{\mathbb{Z}}) \cap \mathcal{F}_0$.

For $r \in d\mathbb{N}+1$, we compose the integrand with the ν-preserving maps τ_r and then use the commutation relations (30) and the fact that $f \circ \tau_r = f$ for $f \in \mathcal{F}_0$. We deduce that the integral in (31) is equal to

$$\int \chi \circ \tau_r \cdot \prod_{j=0}^{m} S^{rj} f_j \, d\nu$$

for every $r \in d\mathbb{N}+1$. Averaging over $r \in d\mathbb{N}+1$ gives the identity

$$\int \chi \cdot \prod_{j=0}^{m} S^j f_j \, d\nu = \mathbb{E}_{n\in\mathbb{N}} \int \chi \circ \tau_{dn+1} \cdot \prod_{j=0}^{m} S^{(dn+1)j} f_j \, d\nu.$$

Hence, it suffices to show that for every $m \in \mathbb{N}$ and $f_1, \dots, f_m \in L^\infty(\nu)$, we have

$$\mathbb{E}_{n\in\mathbb{N}}\, \chi \circ \tau_{dn+1} \cdot \prod_{j=1}^{m} S^{dnj} f_j = 0, \tag{32}$$

where the limit is taken in $L^2(\nu)$. Note that from this point on we work with general functions $f_j \in L^\infty(\nu)$, $j = 1, \dots, m$, not just those in $C(X^{\mathbb{Z}}) \cap \mathcal{F}_0$.

Our first goal is to successively apply van der Corput's lemma and the Cauchy-Schwarz inequality in order to reduce our problem to establishing convergence to zero for an expression that does not depend on the functions $f_1, \dots, f_m$. In our first iteration, we apply Lemma 5.8, compose the integrand with S^{-dn}, and use the Cauchy-Schwarz inequality; we see that in order to establish (32) it suffices to show that for every $h_1 \in \mathbb{N}$, we have

$$\mathbb{E}_{n\in\mathbb{N}}\, S^{-dn}(\chi \circ \tau_{d(n+h_1)+1} \cdot \overline{\chi} \circ \tau_{dn+1}) \prod_{j=1}^{m-1} S^{dnj} f_j = 0$$

for all $f_1, \dots, f_{m-1} \in L^\infty(\nu)$. Note that the number of functions f_j has decreased by one. Note also that by Lemma 5.7 the function

$$F_{h_1,n} := S^{-dn}(\chi \circ \tau_{d(n+h_1)+1} \cdot \overline{\chi} \circ \tau_{dn+1}) \tag{33}$$

is a finite linear combination of eigenfunctions for S with eigenvalue some root of unity times

$$\mathrm{e}\Big(\alpha \cdot (\phi(n+h_1) - \phi(n)\Big),$$

where

$$\phi(n) := \frac{1}{dn+1}, \quad n \in \mathbb{N}.$$

We define inductively the functions $F_{h_1,\dots,h_k,n}$, $h_1, \dots, h_k, n \in \mathbb{N}$ as follows: For $k = 1$ and $h_1, n \in \mathbb{N}$, we let $F_{h_1,n}$ be as in (33), and for $k \geq 2$ and $h_1, \dots, h_k, n \in \mathbb{N}$, we let

$$F_{h_1,\dots,h_k,n} := S^{-dn}\Big(F_{h_1,\dots,h_{k-1},n+h_k} \cdot \overline{F_{h_1,\dots,h_{k-1},n}}\Big).$$

After successively applying Lemma 5.8 ($m+1$ times) and using the Cauchy-Schwarz inequality (m times) we are left with showing that we have, for every $h_1, \dots, h_{m+1} \in \mathbb{N}$,

$$\mathbb{E}_{n\in\mathbb{N}} \int F_{h_1,\dots,h_{m+1},n}\, d\nu = 0. \tag{34}$$

Using Lemma 5.7 and the inductive definition of the functions $F_{h_1,\dots,h_{m+1},n}$, we get that for every $h_1, \dots, h_{m+1}, n \in \mathbb{N}$, the function $F_{h_1,\dots,h_{m+1},n}$ is a finite linear combination of eigenfunctions with eigenvalue equal to some root of unity times the number

$$\mathrm{e}\bigg(\alpha \cdot \sum_{\epsilon\in\{0,1\}^{m+1}} (-1)^{|\epsilon|} \phi(n + \epsilon \cdot h)\bigg),$$

where $h := (h_1, \dots, h_{m+1})$, $|\epsilon| := \epsilon_1 + \dots + \epsilon_{m+1}$, and $\epsilon \cdot h := \epsilon_1 h_1 + \dots + \epsilon_{m+1} h_{m+1}$. Hence,

$$\int F_{h_1,\dots,h_{m+1},n}\, d\nu = 0 \tag{35}$$

unless some of the eigenvalues of the eigenfunctions composing the function $F_{h_1,\dots,h_{m+1},n}$ is 1. Since α is irrational and ϕ takes rational values, this can only happen if

$$\sum_{\epsilon\in\{0,1\}^{m+1}} (-1)^{|\epsilon|}\phi(n+\epsilon\cdot h) = 0. \tag{36}$$

Note that for fixed $h = (h_1,\dots,h_{m+1}) \in \mathbb{N}^{m+1}$, the left-hand side in (36) is a rational function in the variable n and has a pole at $n = 0$; hence it is not identically zero. After clearing denominators, (36) becomes a non-trivial polynomial identity in n; hence it can only have finitely many solutions in n. We deduce that (35) holds for all large enough $n \in \mathbb{N}$. As a consequence, (34) holds for all $h_1,\dots,h_{m+1} \in \mathbb{N}$. As remarked above, this proves that $\chi = 0$ and completes the proof. □

5.4. *Proof of Theorem* 3.11. Let (X,μ,T) be a system with ergodic decomposition $\mu = \int \mu_\omega \, dP(\omega)$. It follows from (11) that

$$\widetilde{\mu} = \int \widetilde{\mu}_\omega \, dP(\omega).$$

If $\alpha \in \mathbb{T}$ is irrational and $\mathrm{e}(\alpha)$ is an eigenvalue of $(X^{\mathbb{Z}},\widetilde{\mu},S)$, then for ω in a set of positive P-measure, the number $\mathrm{e}(\alpha)$ is an eigenvalue of $(X^{\mathbb{Z}},\widetilde{\mu}_\omega,S)$. It thus suffices to prove the theorem in the case where (X,μ,T) is ergodic, and we restrict to this case.

Let (Z_∞,μ_∞,T) be the infinite-step nilfactor of (X,μ,T). By Lemma 4.11, the system $(X^{\mathbb{Z}},\widetilde{\mu},S)$ is isomorphic to the direct product of a Bernoulli system and the system $(Z_\infty^{\mathbb{Z}},\widetilde{\mu}_\infty,S)$. Since Bernoulli systems are weakly mixing, the system $(X^{\mathbb{Z}},\widetilde{\mu},S)$ has the same eigenvalues as the system $(Z_\infty^{\mathbb{Z}},\widetilde{\mu}_\infty,S)$. We can therefore restrict to the case where (X,μ,T) is an ergodic infinite-step nilsystem.

If $(X,\mu,T) = \varprojlim(X_j,\mu_j,T)$ where for $j\in\mathbb{N}$ each system (X_j,μ_j,T) is an ergodic nilsystem, then we get by (28) that

$$(X^{\mathbb{Z}},\widetilde{\mu},S) = \varprojlim(X_j^{\mathbb{Z}},\widetilde{\mu}_j,S).$$

Suppose that α is irrational and $\mathrm{e}(\alpha)$ is an eigenvalue of $(X^{\mathbb{Z}},\widetilde{\mu},S)$ with eigenfunction f. Then for every large enough $j\in\mathbb{N}$, the conditional expectation of f with respect to $X_j^{\mathbb{Z}}$ is non-zero, and this function is an eigenfunction of $(X_j^{\mathbb{Z}},\widetilde{\mu}_j,S)$ with eigenvalue $\mathrm{e}(\alpha)$ as well. Therefore, we can and will restrict to the case where (X,μ,T) is an ergodic nilsystem.

If (X,μ,T) is an ergodic nilsystem, then it has finite rational spectrum. Hence, Proposition 5.3 applies and gives that the system $(X^{\mathbb{Z}},\widetilde{\mu},S)$ is partially strongly stationary. Proposition 5.6 then shows that the system $(X^{\mathbb{Z}},\widetilde{\mu},S)$ has no irrational spectrum. This finishes the proof of the absence of irrational spectrum for the system $(X^{\mathbb{Z}},\widetilde{\mu},S)$. □

We remark that a similar argument also shows that the system $(X^{\mathbb{Z}}, \underline{\mu}, S)$ has no irrational spectrum.

5.5. *An alternate approach to Theorem* 3.10. In [21] it is shown that almost every ergodic component of a strongly stationary system is isomorphic to a direct product of an infinite-step nilsystem and a Bernoulli system. A similar statement with exactly the same proof is valid under the weaker assumption of partial strong stationarity. If (X, μ, T) is an ergodic nilsystem, then it has finite rational spectrum and Proposition 5.3 shows that the system $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ is partially strongly stationary. By combining these results we get a different proof for a weaker version of Proposition 4.5, which states that in the case where (X, μ, T) is an ergodic nilsystem, the ergodic components of the system $(X^{\mathbb{Z}}, \widetilde{\mu}, S)$ are direct products of infinite-step nilsystems and Bernoulli systems. (Note that Proposition 4.5 shows that the Bernoulli systems are superfluous.) One could use this result as a starting point for an alternate proof of Theorems 3.10 and 3.11. The disadvantage of this approach is that we get an unwanted Bernoulli component at a very early stage in the argument, which causes some delicate technical problems in the subsequent analysis.

6. **Disjointness result**

The goal of this section is to prove the disjointness result of Proposition 3.12. We start with the following simpler result:

LEMMA 6.1. *Let (X, μ, T) be an ergodic infinite-step nilsystem, and let (Y, ν, R) be an ergodic system.*

(i) *If the two systems have disjoint irrational spectrum, then for every joining σ of the two systems and function $f \in L^\infty(\mu)$ orthogonal to $\mathcal{K}_{\mathrm{rat}}(T)$, we have*

$$\int f(x)\, g(y)\, d\sigma(x, y) = 0$$

for every $g \in L^\infty(\nu)$.

(ii) *If the two systems have disjoint spectrum different than* 1, *then they are disjoint.*

Proof. We prove part (i). We write $(X, \mu, T) = \varprojlim (X_j, \mu_j, T)$, where (X_j, μ_j, T), $j \in \mathbb{N}$, are ergodic (finite-step) nilsystems, and we let $\pi_j \colon X \to X_j$, $j \in \mathbb{N}$, be the factor maps. Then for every $j \in \mathbb{N}$, the image σ_j of σ under $\pi_j \times \mathrm{id} \colon X \times Y \to X_j \times Y$ is a joining of X_j and Y, and for every $f \in L^\infty(\mu)$ and $g \in L^\infty(\nu)$, we have

$$\int f(x)\, g(y)\, d\sigma(x, y) = \lim_{j\to\infty} \int (f \circ \pi_j)(x)\, g(y)\, d\sigma_j(x, y).$$

Since the function f is orthogonal to $\mathcal{K}_{\mathrm{rat}}(X,T)$, the function $f\circ\pi_j$ is orthogonal to $\mathcal{K}_{\mathrm{rat}}(X_j,T)$ for every $j\in\mathbb{N}$. We can therefore restrict to the case where (X,μ,T) is an ergodic nilsystem.

Suppose that (X,μ,T) is an ergodic s-step nilsystem for some $s\in\mathbb{N}$. The eigenfunctions of X associated to rational eigenvalues are constant on the connected components of X. Therefore, we can approximate in $L^2(\mu)$ the function f that is orthogonal to $\mathcal{K}_{\mathrm{rat}}(X,T)$ by a function in $C^\infty(X)$, still orthogonal to $\mathcal{K}_{\mathrm{rat}}(X,T)$, thus reducing to the case where $f\in C^\infty(X)$. Let $g\in L^\infty(\nu)$. Since σ is $(T\times R)$-invariant, we have

$$\int f(x)\,g(y)\,d\sigma(x,y)=\int f(T^nx)\,g(R^ny)\,d\sigma(x,y)$$

for every $n\in\mathbb{N}$. We average over $n\in\mathbb{N}$ and reduce to showing that

$$\lim_{N\to\infty}\mathbb{E}_{n\in[N]}\int f(T^nx)\cdot g(R^ny)\,d\sigma(x,y)=0. \tag{37}$$

Since (X,T) is an s-step nilsystem and $f\in C^\infty(X)$, it follows from [37, Th. 2.13] and the property characterizing the factors $\mathcal{Z}_s$ given in (44) of Appendix A.4 that if g is orthogonal to the factor $\mathcal{Z}_s(R)$, then there exists a set Y_0 with $\nu(Y_0)=1$ such that for every $y\in Y_0$, we have

$$\lim_{N\to\infty}\mathbb{E}_{n\in[N]}f(T^nx)\cdot g(R^ny)=0$$

for every $x\in X$. This implies that the last identity holds for σ-almost every $(x,y)\in X\times Y$, and the bounded convergence theorem gives (37).

Hence, we have reduced the problem to verifying (37) when $g\in\mathcal{Z}_s(R)$. By Theorem A.5, the factor $(Z_s,\mathcal{Z}_s,\nu_s,R)$ associated with $\mathcal{Z}_s$ is an inverse limit of ergodic s-step nilsystems. Thus, by $L^2(\nu)$-approximation, in order to verify (37), we can assume that the system on Y is an ergodic s-step nilsystem and $g\in C(Y)$.

Let X_0 be the connected components of e_X in X and let μ_0 be the Haar measure of this nilmanifold. Then μ_0 is the normalized restriction of μ to X_0. It is a general fact about nilsystems that there exists $k\in\mathbb{N}$ such that the sets T^jY_0, $0\le j<k$, form a partition of X and that (X_0,μ_0,T^k) is totally ergodic. The rational eigenvalues of (X,μ,T) are $\mathrm{e}(i/k)$ for $i=0,\dots,k-1$. Let Y_0, ν_0 and ℓ be defined in the same way as X_0,μ_0,k was defined with Y substituted for X, and let d be the least common multiple of k and ℓ. Then (X_0,μ_0,T^d) and (Y_0,ν_0,R^d) are totally ergodic and thus have no rational spectrum except 1. Moreover, if for some irrational t we have that $e(t)$ is a common eigenvalue for (X_0,μ_0,T^d) and (Y_0,ν_0,S^d), then $e(t)$ is a common eigenvalue for the systems (X,μ,T^d) and (Y,ν,S^d). It is then an easy consequence that the systems (X,μ,T) and (Y,ν,S) have a common eigenvalue of the form $e(s)$ with s irrational (which can be chosen to satisfy $ds=t \bmod 1$), contradicting our assumption that these systems have disjoint irrational spectrum.

We conclude from the previous analysis that the systems (X_0, μ_0, T^d) and (Y_0, ν_0, S^d) have disjoint spectrum different than 1. As a consequence, the product system $(X_0 \times Y_0, \mu_0 \times \nu_0, T^d \times R^d)$ is ergodic, and since it is a nilsystem, it is uniquely ergodic. Let $x \in X, y \in Y$. There exist $i, j \in \{0, \ldots, d-1\}$ such that $x' := T^{-i}x \in X_0$ and $y' := R^{-j}y \in Y_0$. Since the action of $T^d \times R^d$ on $X_0 \times Y_0$ is uniquely ergodic, we have

$$\begin{aligned}\mathbb{E}_{n\in[N]} f(T^{dn}x) \cdot g(R^{dn}y) &= \mathbb{E}_{n\in[N]} f(T^{dn+i}x') \cdot g(R^{dn+j}y') \\ &\to \int T^i f \, d\mu_0 \cdot \int T^j g \, d\nu_0 = 0,\end{aligned}$$

where the last identity follows since our assumption that f is orthogonal to $\mathcal{K}_{\mathrm{rat}}(T)$ implies that $\int T^i f \, d\mu_0 = 0$ for every $i \in \mathbb{N}$. Applying the last identity for $T^q x, R^r y$ where $q, r \in \{0, \ldots, d-1\}$, in place of x, y, we deduce that

$$\lim_{N\to\infty} \mathbb{E}_{n\in[N]} f(T^n x) \cdot g(R^n y) = 0$$

holds for every $x \in X, y \in Y$, and the bounded convergence theorem gives (37). This completes the proof of part (i).

We prove part (ii). In order to show that the systems are disjoint, it suffices to show that for all $f \in C^\infty(X)$ and $g \in L^\infty(\nu)$, with $\int g \, d\nu = 0$, we have

$$\int f(x) \cdot g(y) \, d\sigma(x, y) = 0. \tag{38}$$

As in the proof of part (i) we reduce to the case where the system (X, μ, T) is a nilsystem. Composing with $(T \times R)^n$ and averaging over $n \in \mathbb{N}$, it thus suffices to show that

$$\lim_{N\to\infty} \mathbb{E}_{n\in[N]} \int f(T^n x) \cdot g(R^n y) \, d\sigma(x, y) = 0. \tag{39}$$

As in the proof of part (i) we reduce to the case where the system (Y, ν, R) is also a nilsystem, so now the systems on X and on Y are ergodic nilsystems with disjoint spectrum other than 1. Then the product system $(X \times Y, \mu \times \nu, T \times R)$ is ergodic, and since it is a nilsystem, it is uniquely ergodic. Hence, for every $x \in X$ and $y \in Y$, we have

$$\lim_{N\to\infty} \mathbb{E}_{n\in[N]} f(T^n x) \cdot g(R^n y) = \int f \, d\mu \cdot \int g \, d\nu = 0, \tag{40}$$

where the second identity follows since by assumption, $\int g \, d\nu = 0$. Finally, using (40) and the bounded convergence theorem we get (39). This completes the proof of part (ii). □

LEMMA 6.2. *Proposition* 3.12 *holds under the additional assumption that the system* (X, μ, T) *is ergodic.*

Proof. By assumption, (X, μ, T) is the direct product of an ergodic infinite-step nilsystem (X', μ', T') and a Bernoulli system (W, λ, S).

We prove part (i). After identifying X with $X' \times W$, we have to show that

$$\int f(x', w)\, g(y)\, d\sigma(x', w, y) = 0 \tag{41}$$

for every $g \in L^\infty(\nu)$.

Using $L^2(\mu' \times \lambda)$-approximation on the orthocomplement of $\mathcal{K}_{\mathrm{rat}}(T' \times S)$, we get that it suffices to verify (41) when $f(x', w) = f_1(x')\, f_2(w)$ for some $f_1 \in L^\infty(\mu')$ and $f_2 \in L^\infty(\lambda)$. Since Bernoulli systems are weakly mixing, we get that $\mathcal{K}_{\mathrm{rat}}(T' \times S) = \mathcal{K}_{\mathrm{rat}}(T')$. Hence, our assumption on f translates to the fact that either $\int f_2\, d\lambda = 0$, or f_1 is orthogonal to $\mathcal{K}_{\mathrm{rat}}(T')$.

Suppose that $\int f_2\, d\lambda = 0$. Let τ be the image of σ under the projection of $X' \times W \times Y$ onto $X' \times Y$. Then σ defines a joining of the zero entropy system $(X' \times Y, \tau, T' \times R)$ and the Bernoulli system (W, λ, S). Since these systems are disjoint, we have $\sigma = \tau \times \lambda$. Hence,

$$\int f_1(x')\, f_2(w)\, g(y)\, d\sigma(x', w, y) = \int f_1(x')\, g(y)\, d\tau(x', y) \int f_2(w)\, d\lambda(w) = 0,$$

establishing that (41) holds in this case.

Suppose now that f_1 is orthogonal to $\mathcal{K}_{\mathrm{rat}}(T')$. Let ρ be the image of σ under the projection of $X' \times W \times Y$ onto $W \times Y$. Then ρ defines a joining of the Bernoulli system (W, λ, S) and the zero entropy system (Y, ν, R). Since the systems are disjoint, we have $\rho = \lambda \times \nu$. Hence, we can consider σ as a joining of the system (X', μ', T') and the system $(W \times Y, \lambda \times \nu, S \times R)$. Since Bernoulli systems are weakly mixing, the system on $W \times Y$ is ergodic and has the same eigenvalues as the system (Y, ν, R) and hence no common irrational eigenvalue with the system (X', μ', T'). It follows that the assumptions of part (i) of Lemma 6.1 are satisfied, and we conclude that (41) holds in this case as well, completing the proof.

We prove part (ii). Let σ be a joining of the systems on $X' \times W$ and on Y. As in the proof of part (i) we get that σ is a joining of the ergodic infinite-step nilsystem (X', μ', T') and the ergodic system $(W \times Y, \lambda \times \nu, S \times R)$ and that these systems have disjoint spectrum other than 1. It follows that the assumptions of part (ii) of Lemma 6.1 are satisfied, and we conclude that $\sigma = \mu' \times \lambda \times \nu$. Hence, the systems on X and on Y are disjoint, completing the proof. □

We are now ready to complete the proof of Proposition 3.12.

Proof of Proposition 3.12. We write

$$\sigma = \int \sigma_\omega\, dP(\omega) \tag{42}$$

for the ergodic decomposition of the joining σ under $T \times R$. Since the system on Y is ergodic, for almost every $\omega \in \Omega$, the projection of σ_ω onto Y is equal to ν. We write μ_ω for the projection of σ_ω on X. Then by the uniqueness property of the ergodic decomposition, we get that for almost every $\omega \in \Omega$, the measure μ_ω is T-invariant and ergodic, the measure σ_ω is an ergodic joining of the systems (X, μ_ω, T) and (Y, ν, R), and the following identity holds:

$$\mu = \int \mu_\omega \, dP(\omega). \tag{43}$$

We prove part (i). Let λ be an irrational eigenvalue of (Y, ν, R). By assumption, λ is not an eigenvalue of (X, μ, T), hence

$$P\Big(\big\{\omega \colon \lambda \text{ is an eigenvalue of } (X, \mu_\omega, T)\big\}\Big) = 0.$$

Since (Y, ν, R) has at most countably many eigenvalues, it follows that there exists a subset Ω_1 of Ω with $P(\Omega_1) = 1$ and such that for every $\omega \in \Omega_1$, the systems (Y, ν, T) and (X, μ_ω, T) do not have any irrational eigenvalue in common. Moreover, since f is orthogonal to $\mathcal{K}_{\mathrm{rat}}(\mu, T)$, there exists $X_1 \subset X$ with $\mu(X_1) = 1$ and such that

$$\mathbb{E}_{n \in \mathbb{N}} \, \mathrm{e}(n\alpha) \, f(T^n x) \to 0 \text{ for every } \alpha \in \mathbb{Q} \text{ and every } x \in X_1.$$

By (43), there exists a subset Ω_2 of Ω_1 with $P(\Omega_2) = 1$ and such that for every $\omega \in \Omega_2$, we have $\mu_\omega(X_1) = 1$, and the convergence above holds for μ_ω almost every $x \in X$. We conclude that for every $\omega \in \Omega_2$, the function f is orthogonal to $\mathcal{K}_{\mathrm{rat}}(\mu_\omega, T)$.

From the above discussion we have that for every $\omega \in \Omega_2$, the hypothesis of part (i) of Lemma 6.2 is satisfied for the function f and the joining σ_ω of the systems (X, μ_ω, T) and (Y, ν, S). We deduce that for every $\omega \in \Omega_2$, we have

$$\int f(x) \, g(y) \, d\sigma_\omega(x, y) = 0$$

for every $g \in L^\infty(\nu)$. Since $P(\Omega_2) = 1$, it follows from (43) that

$$\int f(x) \, g(y) \, d\sigma(x, y) = 0$$

for every $g \in L^\infty(\nu)$. This completes the proof of part (i).

We prove part (ii). As in the first part we show that for P-almost every $\omega \in \Omega$, the systems (Y, ν, T) and (X, μ_ω, T) have disjoint spectrum other than 1. Hence, part (ii) of Lemma 6.2 applies and gives that these two systems are disjoint and thus $\sigma_\omega = \mu_\omega \times \nu$ for almost every $\omega \in \Omega$. Therefore, by (42) and (43) we get $\sigma = \mu \times \nu$. This completes the proof of part (ii). □

7. Subshifts with linear block growth and proof of Theorem 1.2

The goal of this section is to deduce Theorem 1.2 from Theorem 1.1 and some facts about invariant measures of subshifts with linear block growth.

7.1. *Measures on a subshift with linear block growth.* We start with some definitions. Let A be a non-empty finite set whose elements are called *letters*. A is endowed with the discrete topology and $A^{\mathbb{Z}}$ with the product topology and with the shift T. For $n \in \mathbb{N}$, a *word of length* n is a sequence $u = u_1 \cdots u_n$ of n letters (we omit the commas), and we write $[u] = \{x \in A^{\mathbb{Z}} \colon x_1 \cdots x_n = u_1 \cdots u_n\}$.

A *subshift*, also called a *symbolic system*, is a closed non-empty T-invariant subset X of $A^{\mathbb{Z}}$. Recall that X is transitive if it has at least one dense orbit under T.

Let (X,T) be a transitive subshift, equal to the closed orbit of some point $\omega \in A^{\mathbb{Z}}$. For every $n \in \mathbb{N}$, we let $L_n(X)$ denote the set of words u of length n such that $[u] \cap X \neq \emptyset$. Then $L_n(X)$ is also the set of words of length n that occur (as consecutive values) in ω. Note that the set $L(X) := \bigcup_{n\in\mathbb{N}} L_n(X)$ determines X. The *block complexity* of X or of ω is defined by $p_X(n) = |L_n(X)|$ for $n \in \mathbb{N}$. We say that the subshift (X,T) (or the sequence ω) *has linear block growth* if $\liminf_{n\to\infty} p_X(n)/n < \infty$.

Proposition 7.1. *Let (X,T) be a transitive subshift with linear block growth. Then (X,T) admits only finitely many ergodic invariant measures.*

This result was proved in [3] under the stronger hypothesis that (X,T) is minimal. In order to replace this hypothesis with transitivity we will use a result from [10] (alternatively we could use [20, Th. 7.3.7]) that treats the case of non-atomic invariant measures.

Proof of Proposition 7.1. Let X be the closed orbit under T of some $\omega \in A^{\mathbb{Z}}$, and suppose that the subshift (X,T) has linear block growth. If ω is periodic, then X is a finite orbit and the shift transformation on X admits only one invariant measure; hence, we can restrict to the case where ω is not periodic. Let K be an integer such that $\liminf_{n\to\infty} p_X(n)/n \leq K$. Then for infinitely many $n \in \mathbb{N}$, we have $p_X(n+1) - p_X(n) \leq K$.

We say that a word $u \in L_n(X)$ is right special if there exist two different letters $a, b \in A$ such that ua and ub belong to $L_{n+1}(X)$. The number of right special words of length n is clearly bounded by $p_X(n+1) - p_X(n)$. The left special words of length n are defined in a similar way, and their number is also bounded by $p_X(n+1) - p_X(n)$. By a *special word* of length n we mean a left or right special word. Then for infinitely many values of $n \in \mathbb{N}$, there are at most $2K$ special words of length n.

We claim that for every finite orbit Y in X and every $n \in \mathbb{N}$, the set $L_n(Y)$ contains a special word. Suppose that this is not the case. Let $x \in Y$. Since the orbit of ω is dense in X, there exists $k \in \mathbb{Z}$ such that the words $\omega_{k+1} \cdots \omega_{k+n}$ and $x_1 \cdots x_n$ are equal. We show that $T^k\omega = x$. We claim first that for $\ell \geq 1$ we have $\omega_{k+\ell} = x_\ell$. For $1 \leq \ell \leq n$ there is nothing to prove. Suppose that this property holds until some $\ell \geq n$. Then the words $\omega_{k+\ell-n+1} \cdots \omega_{k+\ell}$ and $x_{\ell-n+1} \ldots x_\ell$ are equal, and since $x \in Y$, this word belongs to $L_n(Y)$ and thus is not right special. Since $\omega_{k+\ell-n+1} \cdots \omega_{k+\ell}\omega_{k+\ell+1}$ and $x_{\ell-n+1} \ldots x_\ell x_{\ell+1}$ belong to $L_{n+1}(X)$, we have $\omega_{k+\ell+1} = x_{\ell+1}$, and the claim is proved. In the same way, using now the fact that $L_n(Y)$ does not contain any left special word, we obtain that $\omega_{k+\ell} = x_\ell$ for $\ell \leq 0$, and we conclude that $T^k\omega = x$. Since the orbit of ω is dense, we deduce that $X = Y$ and thus ω is periodic. This contradicts our assumption and proves the claim.

We claim now that X contains at most $2K$ distinct finite orbits. Suppose that this is not the case and that $Y_1, \ldots, Y_{2K+1}$ are distinct finite orbits. Then the sets Y_j, $j = 1, \ldots, 2K+1$, are closed, invariant, pairwise disjoint, and it follows that for every sufficiently large $n \in \mathbb{N}$, the sets $L_n(Y_1), \ldots, L_n(Y_{2K+1})$ are pairwise disjoint. Let $n \in \mathbb{N}$ be chosen so that there are at most $2K$ special words of length n. By the preceding step, each set $L_n(Y_j)$ contains a special word, and since these words are distinct, we have a contradiction and the claim is proved.

By [10], the subshift (X, T) has only finitely many non-atomic ergodic measures. Each atomic ergodic invariant measure is the uniform measure of a finite orbit, and we previously showed that there are at most $2K$ such orbits; hence there are at most $2K$ such measures. This completes the proof. $\square$

7.2. *Proof of Theorem* 1.2. Suppose that $\boldsymbol{\lambda}$ has linear block growth. We extend $\boldsymbol{\lambda}$ to a two-sided sequence, written also $\boldsymbol{\lambda} \in \{-1, 1\}^{\mathbb{Z}}$, by letting $\boldsymbol{\lambda}(n) = 1$ for non-positive $n \in \mathbb{Z}$; then the extended sequence still has linear block growth. Let Y be the closed orbit of $\boldsymbol{\lambda}$ in $\{-1, 1\}^{\mathbb{Z}}$, and let R be the shift on Y. Then (Y, R) is a transitive subshift, and since it has linear block growth, it has zero topological entropy. Moreover, by Proposition 7.1 this system admits only finitely many ergodic invariant measures. Note that for every $n \in \mathbb{N}$, we have $\boldsymbol{\lambda}(n) = F_0(R^n\boldsymbol{\lambda})$, where $F_0 \colon \{-1, 1\}^{\mathbb{Z}} \to \mathbb{R}$ is the map $x \mapsto x_0$. By Theorem 1.1 we get

$$0 = \lim_{N\to\infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{F_0(R^n\boldsymbol{\lambda})\, \boldsymbol{\lambda}(n)}{n} = \lim_{N\to\infty} \frac{1}{\log N} \sum_{n=1}^{N} \frac{\boldsymbol{\lambda}(n)^2}{n} = 1,$$

a contradiction. $\square$

Appendix A. Inverse limits and infinite-step nilsystems

A.1. *Inverse limits in ergodic theory.* Let $(X_j, \mathcal{X}_j, \mu_j, T_j)$, $j \in \mathbb{N}$, be measure preserving systems, and let $\pi_{j,j+1} \colon X_{j+1} \to X_j$, $j \in \mathbb{N}$, be factor maps. We say that $(X_j, \pi_{j,j+1} \colon j \in \mathbb{N})$ is an *inverse sequence* of systems. An *inverse limit* of this inverse sequence is defined to be a system $(X, \mathcal{X}, \mu, T)$ endowed with factor maps $\pi_j \colon X \to X_j$, $j \in \mathbb{N}$, satisfying the following two properties:

(i) $\pi_j = \pi_{j,j+1} \circ \pi_{j+1}$ for every $j \in \mathbb{N}$;
(ii) $\mathcal{X} = \vee_{j\in\mathbb{N}} \pi_j^{-1}(\mathcal{X}_j)$.

For a given inverse sequence of systems, the existence of an inverse limit can be shown by an explicit construction. Properties (i) and (ii) characterize the system (X, μ, T) up to isomorphism. Thus we can say that (X, μ, T), endowed with the factor maps π_j, $j \in \mathbb{N}$, is *the* inverse limit instead of *an* inverse limit, and write

$$(X, \mu, T) = \varprojlim (X_j, \mu_j, T_j)$$

when the factor maps are clear from the context.

A typical example is when a system $(X, \mathcal{X}, \mu, T)$ is given and for $j \in \mathbb{N}$ the systems on X_j are the ones associated to an increasing sequence $\mathcal{X}_j$ of T-invariant sub-σ-algebras of $\mathcal{X}$. Then the inverse limit of this inverse sequence can be defined as the factor of $\mathcal{X}$ associated with the T-invariant sub-σ-algebra $\mathcal{X}' := \vee_{j\in\mathbb{N}} \pi_j^{-1}(\mathcal{X}_j)$.

We record some easy but important properties of inverse limits:

LEMMA A.1. *Suppose that $(X, \mu, T) = \varprojlim (X_j, \mu_j, T_j)$. Then*

(i) *(X, μ, T) is ergodic if and only if (X_j, μ_j, T_j) is ergodic for every $j \in \mathbb{N}$;*
(ii) *a complex number of modulus 1 is an eigenvalue of (X, μ, T) if and only if it is an eigenvalue of (X_j, μ_j, T_j) for every sufficiently large $j \in \mathbb{N}$.*

A.2. *Inverse limits of topological dynamical systems*. Let (X_j, T_j), $j \in \mathbb{N}$, be topological dynamical systems and $\pi_{j,j+1} \colon X_{j+1} \to X_j$, $j \in \mathbb{N}$, be factor maps. We say that $(X_j, \pi_{j,j+1} \colon j \in \mathbb{N})$ is an *inverse sequence* of topological dynamical systems. An *inverse limit* of this inverse sequence is defined to be a topological dynamical system (X, T) endowed with factor maps $\pi_j \colon X \to X_j$, $j \in \mathbb{N}$, satisfying

(i) $\pi_j = \pi_{j,j+1} \circ \pi_{j+1}$ for every $j \in \mathbb{N}$;
(ii) if $x, x' \in X$ are distinct, then $\pi_j(x) \neq \pi_j(x')$ for some $j \in \mathbb{N}$.

Again, for a given inverse sequence of topological systems the existence of an inverse limit can be established by an explicit construction. Properties (i) and (ii) characterize the system (X, T) up to isomorphism. We state the following easy but important properties:

LEMMA A.2. *Suppose $(X, T) = \varprojlim (X_j, T_j)$ with factor maps $\pi_j \colon X \to X_j$, $j \in \mathbb{N}$. Then*

(i) *Let $x \in X$ and Y be the orbit closure of x under T. Then for every $j \in \mathbb{N}$, $\pi_j(Y)$ is the orbit closure of $\pi_j(x)$ under T and $(Y, T) = \varprojlim (\pi_j(Y), T_j)$.*
(ii) *If (X_j, T_j) is minimal for every $j \in \mathbb{N}$, then (X, T) is minimal.*
(iii) *If (X_j, T_j) is uniquely ergodic for every $j \in \mathbb{N}$, then (X, T) is uniquely ergodic.*

We verify the third property only. Let μ, μ' be two T-invariant measures on X. For every $j \in \mathbb{N}$, the system (X_j, T_j) is uniquely ergodic with invariant measure μ_j. Hence, for every $j \in \mathbb{N}$, the images of μ and μ' under π_j are equal to μ_j, and $\int f \circ \pi_j \, d\mu = \int f \circ \pi_j \, d\mu'$ for every $f \in C(X_j)$. It follows from property (ii) of topological inverse limits and the Stone-Weierstrass theorem that the collection of functions $f \circ \pi_j$, where $f \in C(X_j)$ and $j \in \mathbb{N}$, is dense in $C(X)$ with the uniform norm. We conclude that $\mu = \mu'$. Hence, the system (X, T) is uniquely ergodic.

Up to notational changes, all definitions and results of Sections A.1 and A.2 remain valid for systems with several commuting transformations.

A.3. *Infinite step nilsystems*. Let (X_j, μ_j, T_j), $j \in \mathbb{N}$, be ergodic nilsystems and $\pi_{j,j+1} \colon X_{j+1} \to X_j$, $j \in \mathbb{N}$, be factor maps. By [55, Th. 3.3],[7] for every $j \in \mathbb{N}$, the measure theoretic factor map $\pi_{j,j+1} \colon X_{j+1} \to X_j$ agrees almost everywhere with a topological factor map, which we also denote by $\pi_{j,j+1}$. Therefore, the topological dynamical systems (X_j, T_j), with factor maps $\pi_{j,j+1}$, $j \in \mathbb{N}$, form an inverse system. Let (X, T) be the inverse limit of this sequence, and let $\pi_j \colon X \to X_j$ be the associated factor maps. By part (iii) of Lemma (A.2), the system (X, T) is uniquely ergodic. Let μ be the unique invariant measure of (X, T). Then properties (i) and (ii) of Section A.1 are satisfied and $(X, \mu, T) = \varprojlim (X_j, \mu_j, T_j)$.

We use the following terminology from [12]:

Definition A.3. We say that a measure preserving system (X, μ, T) is an *ergodic infinite-step nilsystem* if it is the inverse limit of a sequence (X_j, μ_j, T_j), $j \in \mathbb{N}$, of ergodic nilsystems. By the preceding discussion, the topological dynamical system (X, T) is then the inverse limit of the minimal nilsystems (X_j, T_j), $j \in \mathbb{N}$, and we say that (X, T) is a *minimal infinite-step nilsystem*. We often abuse notation and denote the transformation T_j on X_j by T.

We caution the reader that if s_j is the degree of nilpotency of the nilmanifolds X_j, $j \in \mathbb{N}$, then the sequence $(s_j)_{j \in \mathbb{N}}$ may be unbounded.

From property (iii) of Lemma A.2 and the well-known fact that minimal (finite-step) nilsystems are uniquely ergodic, it follows that minimal infinite-step nilsystems are uniquely ergodic.

Lemma A.4. *An ergodic joining of two ergodic finite or infinite-step nilsystems is a finite or an infinite-step nilsystem respectively.*

[7] In [55] the result is given only when the groups defining the nilmanifolds are connected, but the proof extends to the general case. Another proof is implicit in [39, §6]; see also [38, Ch. XII].

Proof. We give the argument for infinite-step nilsystems only; the other case is similar. Let σ be an ergodic joining of the ergodic infinite-step nilsystems (X,μ,T) and (X',μ',T'). We write $(X,\mu,T) = \varprojlim_j (X_j,\mu_j,T_j)$ and $(X',\mu',T') = \varprojlim_j (X'_j,\mu'_j,T'_j)$ where the systems on X_j and X'_j are ergodic nilsystems for every $j\in\mathbb{N}$. For $j\in\mathbb{N}$, let σ_j be the projection of σ on $X_j\times X'_j$; then σ_j is an ergodic joining of the systems on X_j and X'_j. By [48, Ths. 2.19 and 2.21], for $j\in\mathbb{N}$, the measure σ_j is the Haar measure on some sub-nilmanifold of the product nilmanifold $X_j\times X'_j$. Hence $(X_j\times X'_j,\sigma_j,T_j\times T'_j)$ is an ergodic nilsystem. Since $(X\times X',\sigma,T\times T') = \varprojlim_j (X_j\times X'_j,\sigma_j,T_j\times T'_j)$, the result follows. $\square$

A.4. *The infinite-step nilfactor of a system*. Let (X,μ,T) be an ergodic system, and for $k\in\mathbb{N}$, let $(Z_k,\mathcal{Z}_k,\mu_k,T)$ be the factor of order k of X as defined in [36]. In [36] it is shown that Z_k is characterized by the following property:

$$\text{(44)}\qquad \text{for } f\in L^\infty(\mu),\quad \mathbb{E}(f|Z_k)=0 \quad \text{if and only if} \quad \|\!|f|\!\|_{k+1}=0,$$

where the seminorms $\|\!|\cdot|\!\|_k$ are defined inductively as follows: for $f\in L^\infty(\mu)$, we let $\|\!|f|\!\|_1 := \left|\int f\; d\mu\right|$ and $\|\!|f|\!\|_{k+1}^{2^{k+1}} := \mathbb{E}_{n\in\mathbb{N}} \|\!|\bar{f}\cdot T^n f|\!\|_k^{2^k}$ for $k\in\mathbb{N}$, where all limits can be shown to exist.

The following result was proved in [36]:

THEOREM A.5. *If* (X,μ,T) *is an ergodic system*, *the system* $(Z_k,\mathcal{Z}_k,\mu_k,T)$ *is an inverse limit of ergodic* k*-step nilsystems*.

The factors $\mathcal{Z}_k$, $k\in\mathbb{N}$, form an increasing sequence of T-invariant sub-σ-algebras of $\mathcal{X}$, and we let $\mathcal{Z}_\infty := \vee_{k\in\mathbb{N}}\mathcal{Z}_k$ and $(Z_\infty,\mathcal{Z}_\infty,\mu_\infty,T)$ be the factor system associated with the $\mathcal{Z}_\infty$. Then, this system is the inverse limit of the systems $(Z_k,\mathcal{Z}_k,\mu_k,T)$, $k\in\mathbb{N}$.

COROLLARY A.6. *If* (X,μ,T) *is an ergodic system*, *then* (Z_∞,μ_∞,T) *is an ergodic infinite-step nilsystem*.

Proof. For $k\in\mathbb{N}$, we write $(Z_k,\mu_k,T) = \varprojlim_j (Z_{k,j},\mu_{k,j},T_j)$, where the systems on $Z_{k,j}$ are ergodic k-step ergodic nilsystems for every $j\in\mathbb{N}$. For $\ell\in\mathbb{N}$, let (Y_ℓ,ν_ℓ,T) be the factor of X associated with the σ-algebra

$$\mathcal{Y}_\ell := \bigvee_{k,j\in\mathbb{N},\ k+j\le\ell} \mathcal{Z}_{k,j}.$$

Then the system on Y_ℓ is an ergodic joining of the nilsystems on $Z_{k,j}$ with $k+j\le\ell$. Hence, Lemma A.4 gives that (Y_ℓ,ν_ℓ,T) is an ergodic nilsystem. Moreover, for every $\ell\in\mathbb{N}$ and for all $k,j\in\mathbb{N}$ with $k+j\le\ell$, we have $\mathcal{Z}_{k,j}\subset\mathcal{Z}_\ell$ and thus $\mathcal{Y}_\ell\subset\mathcal{Z}_\ell$ and $\vee_\ell\mathcal{Y}_\ell\subset\mathcal{Z}_\infty$. Conversely, for every $k\in\mathbb{N}$, we have $\mathcal{Y}_{k+j}\supset\mathcal{Z}_{k,j}$ for every $j\in\mathbb{N}$, and hence $\vee_\ell\mathcal{Y}_\ell = \vee_j\mathcal{Y}_{k+j}\supset\vee_j\mathcal{Z}_{k,j} = \mathcal{Z}_k$.

Therefore, $\vee_\ell \mathcal{Y}_\ell \supset \mathcal{Z}_\infty$, and we have equality $\vee_\ell \mathcal{Y}_\ell = \mathcal{Z}_\infty$. By characterizations (i) and (ii) of inverse limits, we deduce that $(Z_\infty, \mu_\infty, T) = \varprojlim_\ell (Y_\ell, \nu_\ell, T)$ and thus $(Z_\infty, \mu_\infty, T)$ is an infinite-step nilsystem. $\square$

Appendix B. The nilmanifold and nilsystem of arithmetic progressions

A key step in the proof of Theorem 1.6 is to determine the structure of the system of arithmetic progressions with integer steps (see Definition 4.2) in the case where the base system is a nilsystem. We are thus naturally led to study configurations defined by arithmetic progressions on $G^{\mathbb{Z}}$, where G is some nilpotent group, of the form $(\ldots, h^{-2}g, h^{-1}g, g, hg, h^2g, \ldots)$, where $g, h \in G$. It turns out that such configurations are not closed under pointwise multiplication and the smallest closed subgroup of $G^{\mathbb{Z}}$ that contains these "arithmetic progressions" is the Hall-Petresco group that we define next. An extensive study of arithmetic progressions in a nilpotent group and in a nilmanifold can be found in [38, Ch. XIV] and in [38].

B.1. *The group of arithmetic progressions*. Let $s \in \mathbb{N}$, and let $X = G/\Gamma$ be an s-step nilmanifold. We write

$$G = G_0 = G_1 \supset G_2 \supset \cdots \supset G_s \supset G_{s+1} = \{e_G\}$$

for the lower central series of G. We denote by μ_X the Haar measure of X and by e_X the image of e_G in X. The action of G on X is written $(g, x) \mapsto g \cdot x$.

We use the following convention for binomial coefficients with negative entries:

$$\binom{n}{m} = \frac{n(n-1)\cdots(n-m+1)}{m!}, \quad n \in \mathbb{Z},\, m \geq 0,$$

where the empty product is equal to 1 by convention.

We write $\underline{G}$ for the set of sequences $\underline{g} = (g_j)_{j\in\mathbb{Z}}$ given by

$$g_j = a_0 a_1^{\binom{j}{1}} a_2^{\binom{j}{2}} \cdots a_s^{\binom{j}{s}}, \quad j \in \mathbb{Z}, \tag{45}$$

where $a_m \in G_m$ for $m = 0, 1, \ldots, s$.

It is known since the work of Hall [35] and Petresco [57] that $\underline{G}$ forms a group with respect to pointwise multiplication. This group is called the *Hall-Petresco group of* G and was extensively studied by Leibman [47] and later by Green and Tao [30], [33].

Elements of $\underline{G}$ have the following useful equivalent characterization: For $\underline{g} = (g_j)_{j\in\mathbb{Z}}$ in $G^{\mathbb{Z}}$, let $\partial\underline{g} \in G^{\mathbb{Z}}$ be defined by

$$(\partial\underline{g})_j := g_{j+1} g_j^{-1}, \quad j \in \mathbb{Z}.$$

In other words, $\partial\underline{g} = \sigma\underline{g} \cdot \underline{g}^{-1}$, where $\sigma\colon G^{\mathbb{Z}} \to G^{\mathbb{Z}}$ is the shift defined by

$$(\sigma(\underline{g}))_j := g_{j+1}, \quad \underline{g} \in G^{\mathbb{Z}},\ j \in \mathbb{Z}.$$

For $m \in \mathbb{N}$, we let $\partial^{\circ m} := \partial \circ \cdots \circ \partial$ (m times). The next result was proved in [47, Prop. 3.1] and also in [46]:

LEMMA B.1. *An element* $\underline{g} \in G^{\mathbb{Z}}$ *belongs to* $\underline{G}$ *if and only if for every* $m \in \mathbb{N}$, *we have* $\partial^{\circ m}\underline{g} \in G_m^{\mathbb{Z}}$.

We immediately deduce from Lemma B.1 the following basic properties:

- $\underline{G}$ is invariant under the shift $\sigma\colon G^{\mathbb{Z}} \to G^{\mathbb{Z}}$;
- $\partial^{\circ(s+1)}\underline{g} = e_{\underline{G}}$ for every $\underline{g} \in \underline{G}$; that is, σ is a unipotent automorphism of $\underline{G}$;
- $\underline{G}$ is a closed subgroup of $G^{\mathbb{Z}}$.

B.2. *The nilmanifold of arithmetic progressions*. Let $X^{\mathbb{Z}}$ be endowed with the action of $\underline{G}$ given by $(\underline{g} \cdot \underline{x})_j = g_j \cdot x_j$ for $\underline{g} \in \underline{G}$, $\underline{x} \in X^{\mathbb{Z}}$, and $j \in \mathbb{Z}$. If $e_{\underline{X}} = (\dots, e_X, e_X, e_X, \dots)$, we define

$$\underline{X} := \underline{G} \cdot e_{\underline{X}} = \big\{(g_j \cdot e_X)_{j\in\mathbb{Z}}\colon (g_j)_{j\in\mathbb{Z}} \in \underline{G}\big\}.$$

The stabilizer of $e_{\underline{X}}$ in $\underline{G}$ is the subgroup $\underline{\Gamma} := \underline{G} \cap \Gamma^{\mathbb{Z}}$, and thus we have

$$\underline{X} = \underline{G} \,/\, \underline{\Gamma}.$$

A priori, $\underline{X}$ is an infinite dimensional object, but it will be convenient for us to represent it as a nilmanifold, in order to be able to apply the machinery of nilmanifolds. To this end, we show that $\underline{G}$ can be represented as a subgroup of G^{s+1} and $\underline{X}$ as a sub-nilmanifold of X^{s+1}. We make use of the next lemma, which follows from Lemma B.1 and was established by Green and Tao in the course of proving Lemma 14.2 in [30].

LEMMA B.2. *The projection homomorphism*

$$p\colon \underline{G} \to G^{s+1} \text{ given by } p(\underline{g}) := (g_0, g_1, \dots, g_s)$$

is one-to-one and satisfies $p^{-1}(\Gamma^{s+1}) = \underline{\Gamma}$. *Furthermore, the projection*

$$q\colon \underline{X} \to X^{s+1} \text{ given by } q(\underline{x}) := (x_0, x_1, \dots, x_s)$$

is one-to-one.

We let

$$\underline{G}' := p(\underline{G}), \quad \underline{\Gamma}' := p(\underline{\Gamma}) = \underline{G} \cap \Gamma^{s+1}, \quad \underline{X}' := q(\underline{X}).$$

Writing $\underline{e}'_X := (e_X, e_X, \dots, e_X) \in X^{s+1}$, we have $\underline{X}' = \underline{G}' \cdot \underline{e}'_X$ by construction, and we can identify $\underline{X}'$ with $\underline{G}'/\underline{\Gamma}'$.

By [2, §5] (see also [69]), $\underline{G}'$ is a closed subgroup of G^{s+1}, hence a nilpotent Lie group, and the discrete subgroup $\underline{\Gamma}'$ of $\underline{G}'$ is cocompact. Therefore, $\underline{X}'$ is compact and can be identified with the nilmanifold $\underline{G}'/\underline{\Gamma}'$.

Since $\underline{G}$ and $\underline{G}'$ are Polish groups and $p\colon \underline{G} \to \underline{G}'$ is a continuous bijective homomorphism, the inverse homomorphism is also continuous. Since $\underline{\Gamma}'$ is cocompact in $\underline{G}'$, it follows that $\underline{\Gamma}$ is cocompact in $\underline{G}$, hence $\underline{X}$ is compact, and thus $q\colon \underline{X} \to \underline{X}'$ is a homeomorphism.

Convention. In the sequel, we use the isomorphism p to identify $\underline{G}$ with $\underline{G}'$ and $\underline{\Gamma}$ with $\underline{\Gamma}'$. We use the homeomorphism q to identify $\underline{X} = \underline{G}/\underline{\Gamma}$ with the nilmanifold $\underline{X}' = \underline{G}'/\underline{\Gamma}'$. We write $\mu_{\underline{X}}$ for the Haar measure of $\underline{X}$.

Definition B.3. $\underline{X} = \underline{G}/\underline{\Gamma}$ is called the *nilmanifold of arithmetic progressions* in X.

B.3. *The nilsystem of arithmetic progressions.* Since $\underline{G}$ is invariant under the shift σ of $G^{\mathbb{Z}}$, we get that $\underline{X}$ is invariant under the shift S of $X^{\mathbb{Z}}$. We have

$$S(\underline{g} \cdot \underline{x}) = \sigma(\underline{g}) \cdot S\underline{x}, \quad \underline{x} \in \underline{X},\ \underline{g} \in \underline{G}. \tag{46}$$

By (46) the image of the measure $\mu_{\underline{X}}$ under S is invariant under translation by elements of $\underline{G}$, and hence it is equal to $\mu_{\underline{X}}$. We have thus established that $(\underline{X}, \mu_{\underline{X}}, S)$ is a measure preserving system and our next goal is to give $(\underline{X}, S)$ the structure of a nilsystem, called the *nilsystem of arithmetic progressions in* X.

We define the group $\hat{\underline{G}}$ to be the semidirect product $\hat{\underline{G}} = G \rtimes_\phi \mathbb{Z}$, where $\phi\colon \mathbb{Z} \to \mathrm{Aut}(G)$ is the homomorphism $n \mapsto \overrightarrow{\sigma}^{\circ n}$, where $\sigma^{\circ n} = \sigma \circ \cdots \circ \sigma$ (n times). More explicitly, as a set we have $\hat{\underline{G}} = \underline{G} \times \mathbb{Z}$, and the multiplication is given by

$$(\underline{g}, m) \cdot (\underline{h}, n) = (\underline{g} \cdot \sigma^{\circ m}(\underline{h}), m + n), \quad \underline{g}, \underline{h} \in \underline{G},\ m, n \in \mathbb{Z}.$$

Then $\underline{G} \times \{0\}$ is a normal subgroup of $\hat{\underline{G}}$ that we identify with $\underline{G}$. Since $\underline{G}$ is nilpotent and the automorphism σ of $\underline{G}$ is unipotent, it follows that $\hat{\underline{G}}$ is nilpotent [48, Prop. 3.9]. We give $\hat{\underline{G}}$ the structure of a Lie group by letting $\underline{G}$ be an open subgroup of $\hat{\underline{G}}$.

The group $\hat{\underline{G}}$ acts on $\underline{X}$ by $(\underline{g}, m) \cdot \underline{x} = \underline{g} \cdot S^m \underline{x}$, and this action preserves the Haar measure of $\underline{X}$. Moreover, the stabilizer of $e_{\underline{X}}$ is the discrete cocompact subgroup $\hat{\underline{\Gamma}} := \underline{\Gamma} \rtimes_\phi \mathbb{Z}$ of $\hat{\underline{G}}$, and we can identify $\underline{X}$ with the nilmanifold $\hat{\underline{G}}/\hat{\underline{\Gamma}}$. Since the measure $\underline{\mu}$ is invariant under S and the action of $\underline{G}$, it is invariant under the action of $\hat{\underline{G}}$ and thus coincides with the Haar measure of $\underline{X}$ when identified with $\hat{\underline{G}}/\hat{\underline{\Gamma}}$. Finally, with the above identifications, the transformation S is the translation by the element $(\underline{e_G}, 1)$ of $\hat{\underline{G}}$, and thus $(\underline{X}, \mu_{\underline{X}}, S)$ is a nilsystem. The previous discussion leads to the following basic result:

Proposition B.4. *If X is a nilmanifold, then the system $(\underline{X}, S)$ is topologically isomorphic to a nilsystem. As a consequence, if $Y = \overline{\{S^n \underline{x} \colon n \in \mathbb{Z}\}}$ for some $\underline{x} \in \underline{X}$, then the system (Y, S) is topologically isomorphic to a uniquely ergodic nilsystem.*

The first claim was established in the previous discussion. The consequence follows, for example, from [48, Ths. 2.19 and 2.21].

Appendix C. Sketch of proof of Tao's identity

We recall the statement of Theorem 3.5 and briefly sketch its proof almost entirely following [62]. The only difference in our presentation is that our assumption of existence of certain limits allows us to perform a partial summation at the beginning of the argument in order to connect the averages we are interested in to the averages treated in [62].

Proposition C.1. *Let $\mathbf{N} = ([N_k])_{k\in\mathbb{N}}$ be a sequence of intervals, $\ell \in \mathbb{N}$, $a_1, \dots, a_\ell$ be bounded sequences of complex numbers, and $h_1, \dots, h_\ell \in \mathbb{Z}$. Also let $(c_p)_{p\in\mathbb{P}}$ be a bounded sequence of complex numbers. Then, assuming that on the left- and right-hand sides below the limits $\mathbb{E}^{\log}_{n\in\mathbf{N}}$ exist for every $p \in \mathbb{P}$ and the limit $\mathbb{E}_{p\in\mathbb{P}}$ exists, we have the identity*

$$\mathbb{E}_{p\in\mathbb{P}}\, c_p \Big(\mathbb{E}^{\log}_{n\in\mathbf{N}} \prod_{j=1}^{\ell} a_j(pn + ph_j)\Big) = \mathbb{E}_{p\in\mathbb{P}}\, c_p \Big(\mathbb{E}^{\log}_{n\in\mathbf{N}} \prod_{j=1}^{\ell} a_j(n + ph_j)\Big). \tag{47}$$

Sketch of Proof. For $H \in \mathbb{N}$, let[8]

$$\mathcal{P}_H := \{p \in \mathbb{P}\colon H/2 \le p < H\}, \qquad W_H := \sum_{p\in\mathcal{P}_H} \frac{1}{p} \sim \frac{1}{\log H},$$

where the last asymptotic means that the quotient of the two quantities involved converges to a non-zero constant as $H \to \infty$ and follows from the prime number theorem using partial summation.

We first claim that the limits on the left- and right-hand sides of (47) are equal to

$$\lim_{H\to\infty} \frac{1}{W_H} \sum_{p\in\mathcal{P}_H} \frac{c_p}{p}\, \mathbb{E}^{\log}_{n\in\mathbf{N}} \prod_{j=1}^{\ell} a_j(pn + ph_j) \tag{48}$$

and

$$\lim_{H\to\infty} \frac{1}{W_H} \sum_{p\in\mathcal{P}_H} \frac{c_p}{p}\, \mathbb{E}^{\log}_{n\in\mathbf{N}} \prod_{j=1}^{\ell} a_j(n + ph_j)$$

respectively. To see this, let

$$A(p) := c_p\, \mathbb{E}^{\log}_{n\in\mathbf{N}} \prod_{j=1}^{\ell} a_j(pn + ph_j).$$

Our assumptions give that the limit $L := \mathbb{E}_{p\in\mathbb{P}} A(p)$ exists and we want to show that

$$B(H) := \frac{1}{W_H} \sum_{p\in\mathcal{P}_H} \frac{A(p)}{p} \to L \quad \text{as} \quad H \to \infty.$$

[8]In [62] the respective set $\mathcal{P}_H$ consists of primes on the interval $[\delta H/2, \delta H)$ for a sufficiently small δ, but for our purposes we can take $\delta = 1$.

(In a similar manner we treat the second average.) Let $\varepsilon > 0$. If $S(x) := \sum_{p\le x}(A(p)-L)$, where $x \in \mathbb{N}$, our hypothesis gives that $|S(x)| \le \varepsilon \frac{x}{\log x}$ for all sufficiently large x. Since $S(x) - S(x-1)$ is equal to $A(x) - L$ if x is a prime and is 0 otherwise, we get that for every $H \in \mathbb{N}$, we have

$$B(H) - L = \frac{1}{W_H} \sum_{H/2\le n<H} \frac{S(n)-S(n-1)}{n}.$$

Using partial summation we get that $|B(H)-L|$ is bounded by a sum of terms of the form $S(H)/(HW_H)$ and $\frac{1}{W_H}\sum_{H/2\le n<H}\frac{S(n)}{n^2}$. For sufficiently large $H \in \mathbb{N}$, the first term is bounded by ε and the second by $\varepsilon H \sum_{H/2\le n<H}\frac{1}{n^2} \le 2\varepsilon$. This completes the proof of the claim.

Next note the simple but important fact that if $b \in \ell^\infty(\mathbb{Z})$, then for every $r \in \mathbb{N}$, we have[9]

$$\mathbb{E}^{\log}_{n\in\mathbf{N}}(b(rn) - b(n)\, r\, \mathbf{1}_{r\mathbb{Z}}(n)) = 0.$$

Using this for $r = p$ and for the sequence b_p, $p \in \mathbb{P}$, defined by

$$b_p(n) := c_p \prod_{j=1}^{\ell} a_j(n + ph_j), \quad n \in \mathbb{N},$$

we can rewrite the limit in (48) as

$$\lim_{H\to\infty} \frac{1}{W_H} \sum_{p\in\mathcal{P}_H} c_p\, \mathbb{E}^{\log}_{n\in\mathbf{N}} \prod_{j=1}^{\ell} a_j(n+ph_j)\cdot \mathbf{1}_{p\mathbb{Z}}(n).$$

Hence, in order to establish (47) and because all relevant limits exist, it suffices to show that

$$\liminf_{H\to\infty} \Big|\mathbb{E}^{\log}_{n\in\mathbf{N}} \frac{1}{W_H}\sum_{p\in\mathcal{P}_H} c_p \prod_{j=1}^{\ell} a_j(n+ph_j)\cdot\big(\mathbf{1}_{p\mathbb{Z}}(n) - p^{-1}\big)\Big| = 0. \tag{49}$$

We argue by contradiction. Suppose (49) fails for some $h_1, \dots, h_\ell \in \mathbb{Z}$. Since $W_H \sim \frac{1}{\log H}$, there exists $\varepsilon > 0$ such that for $\delta := \varepsilon^2$ (we can choose it any function of ε we like), we have (the argument is similar if $\le -\varepsilon\,\frac{1}{\log H}$)

$$\mathbb{E}^{\log}_{n\in\mathbf{N}} \sum_{p\in\mathcal{P}_H} c_p \prod_{j=1}^{\ell} a_j(n+ph_j)\cdot\big(\mathbf{1}_{p\mathbb{Z}}(n) - p^{-1}\big) \ge \varepsilon\,\frac{1}{\log H} \tag{50}$$

for all large enough $H \in \mathbb{N}$. Using the translation invariance of the average $\mathbb{E}^{\log}_{n\in\mathbf{N}}$ we shift n by h and sum over $h \in [H]$. We get that

[9] This identity holds for logarithmic averages and fails in general for Cesàro averages, which is the main reason why we cannot treat Cesàro averages in this article.

$$\mathbb{E}^{\log}_{n\in[N_k]} \sum_{p\in\mathcal{P}_H} \sum_{h\in[H]} c_p \prod_{j=1}^{\ell} a_j(n+h+ph_j)\cdot\big(\mathbf{1}_{p\mathbb{Z}}(n+h)-p^{-1}\big) \geq \varepsilon\,\frac{H}{\log H} \tag{51}$$

for all large enough $H\in\mathbb{N}$ depending on ε and all large enough k depending on ε and H. Furthermore, after approximating the sequences a_j, $j=1,\dots,\ell$, to the nearest element of the lattice $\varepsilon^2\mathbb{Z}[i]$, we can assume that they take values on a finite set $A=A_\varepsilon$ and (51) continues to hold (with $\varepsilon/2$ in place of ε). For details, see [62, §2].

For $k\in\mathbb{N}$, on the space $\mathbb{N}$ we define the (non-shift invariant) probability measure $\mathbb{P}_k$ on all subsets of $\mathbb{N}$ by letting

$$\mathbb{P}_k(E):=\mathbb{E}^{\log}_{n\in[N_k]}\mathbf{1}_E(n), \qquad E\subset\mathbb{N}.$$

We also define the vector valued random variables $\mathbf{X}_H\colon\mathbb{N}\to\mathbb{C}^{\ell H}$ and $\mathbf{Y}_H\colon\mathbb{N}\to\prod_{p\leq H}\mathbb{Z}/p\mathbb{Z}$ as follows:

$$\mathbf{X}_H(n):=(a_{j,h}(n))_{j\in[\ell],h\in[H]},\ n\in\mathbb{N}, \quad \text{where } \ a_{j,h}(n):=a_j(n+h),$$
$$\mathbf{Y}_H(n):=\big(n\ (p)\big)_{p\leq H},\ n\in\mathbb{N},$$

where $\big(n\ (p)\big)_{p\leq H}$ denotes the reductions of n modulo the primes p that are less than H. Furthermore, for $H\in\mathbb{N}$, we let $F_H\colon A^{\ell H}\times\prod_{p\leq H}\mathbb{Z}/p\mathbb{Z}\to\mathbb{R}$ be defined by

$$F_H((x_{j,h})_{j\in[\ell],h\in[LH]},(r_p)_{p\leq H}):=\sum_{p\in\mathcal{P}_H}\sum_{h\in[H]} c_p\prod_{j=1}^{\ell} x_{j,h+ph_j}\,(\mathbf{1}_{p\mathbb{Z}}(r_p+h)-p^{-1}), \tag{52}$$

where $L:=\max_{j=1,\dots,\ell}(h_j)+1$. Also let $\mathbb{E}_kF$ denote the expectation of a function $F\colon\mathbb{N}\to\mathbb{C}$ with respect to the probability measure $\mathbb{P}_k$. Then (51) gives that

$$|\mathbb{E}_kF_H(\mathbf{X}_H(n),\mathbf{Y}_H(n))|\geq\varepsilon\,\frac{H}{\log H} \tag{53}$$

for all large enough H depending on ε and all large enough k depending on ε and H.

Using the entropy decrement argument as in [62, Lemma 3.2], we get that there exist a positive integer $H_-=H_-(\varepsilon)$ (which can be chosen suitably large depending on ε), a larger positive integer $H^+=H^+(\varepsilon)$, and for $k\in\mathbb{N}$, there exist $H_k\in[H_-,H_+]$ such that

$$\mathbb{I}_k(\mathbf{X}_{H_k},\mathbf{Y}_{H_k})\leq\frac{H_k}{\log H_k\log\log H_k}$$

for every $k\in\mathbb{N}$ where $\mathbb{I}_k$ is the mutual information function (defined in [62, §3]) with respect to the probability measure $\mathbb{P}_k$. Since the integers H_k belong to the finite interval $[H_-, H_+]$ for every $k \in \mathbb{N}$, there exists a fixed integer $H_0 \in [H_-, H_+]$ such that

$$\mathbb{I}_k(\mathbf{X}_{H_0}, \mathbf{Y}_{H_0}) \leq \frac{H_0}{\log H_0 \log\log H_0} \tag{54}$$

for infinitely many $k \in \mathbb{N}$. We deduce that for $H := H_0$, (53) and (54) hold simultaneously for infinitely many $k \in \mathbb{N}$.

Using (54) one gets, as in [62] (using the Pinsker type inequality [62, Lemma 3.3] and then the Hoeffding inequality as in [62, Lemma 3.5]), the following estimate (it corresponds to [62, eq. (3.16)]):

$$\mathbb{E}_{(r_p)_{p\leq H} \in \prod_{p\leq H_0} \mathbb{Z}/p\mathbb{Z}}\, \mathbb{E}_k F_{H_0}(\mathbf{X}_{H_0}(n), (r_p)_{p\leq H_0}) \geq C\varepsilon \frac{H_0}{\log H_0} \tag{55}$$

for some $C > 0$ and for infinitely many $k \in \mathbb{N}$. But by (52), we have

$$\mathbb{E}_{(r_p)_{p\leq H} \in \prod_{p\leq H} \mathbb{Z}/p\mathbb{Z}}\, F_H(\mathbf{X}_H(n), (r_p)_{p\leq H}) = 0$$

for every $n, H \in \mathbb{N}$. This contradicts (55) and completes the proof. □

References

[1] V. Bergelson, Weakly mixing PET, *Ergodic Theory Dynam. Systems* **7** no. 3 (1987), 337–349. MR 0912373. Zbl 0645.28012. https://doi.org/10.1017/S0143385700004090.

[2] V. Bergelson, B. Host, and B. Kra, Multiple recurrence and nilsequences, *Invent. Math.* **160** no. 2 (2005), 261–303, With an appendix by Imre Ruzsa. MR 2138068. Zbl 1087.28007. https://doi.org/10.1007/s00222-004-0428-6.

[3] M. Boshernitzan, A unique ergodicity of minimal symbolic flows with linear block growth, *J. Analyse Math.* **44** (1984/85), 77–96. MR 0801288. Zbl 0602.28008. https://doi.org/10.1007/BF02790191.

[4] J. Bourgain, Möbius-Walsh correlation bounds and an estimate of Mauduit and Rivat, *J. Anal. Math.* **119** (2013), 147–163. MR 3043150. Zbl 1357.11092. https://doi.org/10.1007/s11854-013-0005-2.

[5] J. Bourgain, On the correlation of the Moebius function with rank-one systems, *J. Anal. Math.* **120** (2013), 105–130. MR 3095150. Zbl 1358.37017. https://doi.org/10.1007/s11854-013-0016-z.

[6] J. Bourgain, P. Sarnak, and T. Ziegler, Disjointness of Moebius from horocycle flows, in *From Fourier Analysis and Number Theory to Radon Transforms and Geometry*, *Dev. Math.* **28**, Springer, New York, 2013, pp. 67–83. MR 2986954. Zbl 1336.37030. https://doi.org/10.1007/978-1-4614-4075-8_5.

[7] X. Bressaud, F. Durand, and A. Maass, On the eigenvalues of finite rank Bratteli-Vershik dynamical systems, *Ergodic Theory Dynam. Systems* **30** no. 3 (2010), 639–664. MR 2643706. Zbl 1204.37008. https://doi.org/10.1017/S0143385709000236.

[8] F. Cellarosi and Y. G. Sinai, Ergodic properties of square-free numbers, *J. Eur. Math. Soc.* (*JEMS*) **15** no. 4 (2013), 1343–1374. MR 3055764. Zbl 1325.37010. https://doi.org/10.4171/JEMS/394.

[9] J. Chaika, *Interval Exchange Transformations: Applications of Keane's Construction and Disjointness*, ProQuest LLC, Ann Arbor, MI, 2010, thesis (Ph.D.)–Rice University. MR 2782362. Available at http://gateway.proquest.com/openurl?url_ver=Z39.88-2004&rft_val_fmt=info:ofi/fmt:kev:mtx:dissertation&res_dat=xri:pqdiss&rft_dat=xri:pqdiss:3421391.

[10] V. Cyr and B. Kra, Counting generic measures for a subshift of linear growth, 2015, to appear in *J. Eur. Math. Soc.* arXiv 1505.02748.

[11] J.-M. Deshouillers, M. Drmota, and C. Müllner, Automatic sequences generated by synchronizing automata fulfill the Sarnak conjecture, *Studia Math.* **231** no. 1 (2015), 83–95. MR 3460628. Zbl 1371.11066. https://doi.org/10.4064/sm8479-2-2016.

[12] P. Dong, S. Donoso, A. Maass, S. Shao, and X. Ye, Infinite-step nilsystems, independence and complexity, *Ergodic Theory Dynam. Systems* **33** no. 1 (2013), 118–143. MR 3009105. Zbl 1268.37019. https://doi.org/10.1017/S0143385711000861.

[13] M. Einsiedler and T. Ward, *Ergodic Theory with a View Towards Number Theory*, *Grad. Texts in Math.* **259**, Springer-Verlag London, Ltd., London, 2011. MR 2723325. Zbl 1206.37001. https://doi.org/10.1007/978-0-85729-021-2.

[14] E. H. El Abdalaoui, S. Kasjan, and M. Lemańczyk, 0-1 sequences of the Thue-Morse type and Sarnak's conjecture, *Proc. Amer. Math. Soc.* **144** no. 1 (2016), 161–176. MR 3415586. Zbl 1352.37031. https://doi.org/10.1090/proc/12683.

[15] E. H. El Abdalaoui, J. Kułaga-Przymus, M. Lemańczyk, and T. de la Rue, The Chowla and the Sarnak conjectures from ergodic theory point of view, *Discrete Contin. Dyn. Syst.* **37** no. 6 (2017), 2899–2944. MR 3622068. Zbl 1360.37019. https://doi.org/10.3934/dcds.2017125.

[16] E. H. El Abdalaoui, M. Lemańczyk, and T. de la Rue, On spectral disjointness of powers for rank-one transformations and Möbius orthogonality, *J. Funct. Anal.* **266** no. 1 (2014), 284–317. MR 3121731. Zbl 1351.37048. https://doi.org/10.1016/j.jfa.2013.09.005.

[17] S. Ferenczi, J. Kułaga-Przymus, and M. Lemańczyk, Sarnak's conjecture - what's new, 2017. arXiv 1710.04039.

[18] S. Ferenczi, J. Kułaga-Przymus, M. Lemańczyk, and C. Mauduit, Substitutions and Möbius disjointness, in *Ergodic Theory, Dynamical Systems, and the Continuing Influence of John C. Oxtoby*, *Contemp. Math.* **678**, Amer. Math. Soc., Providence, RI, 2016, pp. 151–173. MR 3589820. Zbl 1364.37027. https://doi.org/10.1090/conm/678/13645.

[19] S. Ferenczi and C. Mauduit, On Sarnak's conjecture and Veech's question for interval exchanges, to appear in *J. Analyse Math.* Available at https://hal.archives-ouvertes.fr/hal-01263088.

[20] S. Ferenczi and T. Monteil, Infinite words with uniform frequencies, and invariant measures, in *Combinatorics, Automata and Number Theory*, *Encyclopedia Math. Appl.* **135**, Cambridge Univ. Press, Cambridge, 2010, pp. 373–409. MR 2759110. Zbl 1217.68168. https://doi.org/10.1017/CBO9780511777653.008.

[21] N. Frantzikinakis, The structure of strongly stationary systems, *J. Anal. Math.* **93** (2004), 359–388. MR 2110334. Zbl 1073.60040. https://doi.org/10.1007/BF02789313.

[22] N. Frantzikinakis, Ergodicity of the Liouville system implies the Chowla conjecture, *Discrete Anal.* **19** (2017), 41pp. arXiv 1611.09338.

[23] N. Frantzikinakis, B. Host, and B. Kra, Multiple recurrence and convergence for sequences related to the prime numbers, *J. Reine Angew. Math.* **611** (2007), 131–144. MR 2361086. Zbl 1126.37005. https://doi.org/10.1515/CRELLE.2007.076.

[24] N. Frantzikinakis, B. Host, and B. Kra, The polynomial multidimensional Szemerédi theorem along shifted primes, *Israel J. Math.* **194** no. 1 (2013), 331–348. MR 3047073. Zbl 1301.37008. https://doi.org/10.1007/s11856-012-0132-y.

[25] H. Furstenberg, Ergodic behavior of diagonal measures and a theorem of Szemerédi on arithmetic progressions, *J. Analyse Math.* **31** (1977), 204–256. MR 0498471. Zbl 0347.28016. https://doi.org/10.1007/BF02813304.

[26] H. Furstenberg, *Recurrence in Ergodic Theory and Combinatorial Number Theory*, Princeton University Press, Princeton, N.J., 1981, M. B. Porter Lectures. MR 0603625. Zbl 0459.28023.

[27] H. Furstenberg and Y. Katznelson, A density version of the Hales-Jewett theorem, *J. Anal. Math.* **57** (1991), 64–119. MR 1191743. Zbl 0770.05097. https://doi.org/10.1007/BF03041066.

[28] E. Glasner, *Ergodic theory via joinings*, *Mathematical Surveys and Monographs* **101**, American Mathematical Society, Providence, RI, 2003. MR 1958753. Zbl 1038.37002. https://doi.org/10.1090/surv/101.

[29] B. Green, On (not) computing the Möbius function using bounded depth circuits, *Combin. Probab. Comput.* **21** no. 6 (2012), 942–951. MR 2981162. Zbl 1279.11095. https://doi.org/10.1017/S0963548312000284.

[30] B. Green and T. Tao, An inverse theorem for the Gowers $U^3(G)$ norm, *Proc. Edinb. Math. Soc.* (2) **51** no. 1 (2008), 73–153. MR 2391635. Zbl 1202.11013. https://doi.org/10.1017/S0013091505000325.

[31] B. Green and T. Tao, Linear equations in primes, *Ann. of Math.* (2) **171** no. 3 (2010), 1753–1850. MR 2680398. Zbl 1242.11071. https://doi.org/10.4007/annals.2010.171.1753.

[32] B. Green and T. Tao, The Möbius function is strongly orthogonal to nilsequences, *Ann. of Math.* (2) **175** no. 2 (2012), 541–566. MR 2877066. Zbl 1347.37019. https://doi.org/10.4007/annals.2012.175.2.3.

[33] B. Green and T. Tao, The quantitative behaviour of polynomial orbits on nilmanifolds, *Ann. of Math.* (2) **175** no. 2 (2012), 465–540. MR 2877065. Zbl 1251.37012. https://doi.org/10.4007/annals.2012.175.2.2.

[34] B. Green, T. Tao, and T. Ziegler, An inverse theorem for the Gowers $U^{s+1}[N]$-norm, *Ann. of Math.* (2) **176** no. 2 (2012), 1231–1372. MR 2950773. Zbl 1282.11007. https://doi.org/10.4007/annals.2012.176.2.11.

[35] M. Hall, Jr., *The Theory of Groups*, The Macmillan Co., New York, N.Y., 1959. MR 0103215. Zbl 0084.02202.

[36] B. Host and B. Kra, Nonconventional ergodic averages and nilmanifolds, *Ann. of Math.* (2) **161** no. 1 (2005), 397–488. MR 2150389. Zbl 1077.37002. https://doi.org/10.4007/annals.2005.161.397.

[37] B. Host and B. Kra, Uniformity seminorms on ℓ^∞ and applications, *J. Anal. Math.* **108** (2009), 219–276. MR 2544760. Zbl 1183.37011. https://doi.org/10.1007/s11854-009-0024-1.

[38] B. Host and B. Kra, *Nilpotent Structures in Ergodic Theory*, book in preparation.

[39] B. Host, B. Kra, and A. Maass, Nilsequences and a structure theorem for topological dynamical systems, *Adv. Math.* **224** no. 1 (2010), 103–129. MR 2600993. Zbl 1203.37022. https://doi.org/10.1016/j.aim.2009.11.009.

[40] W. Huang, Z. Wang, and G. Zhang, Möbius disjointness for topological models of ergodic systems with discrete spectrum, 2016. arXiv 1608.08289.

[41] H. Iwaniec and E. Kowalski, *Analytic Number Theory*, *Amer. Math. Soc. Colloq. Publ.* **53**, American Mathematical Society, Providence, RI, 2004. MR 2061214. Zbl 1059.11001. https://doi.org/10.1090/coll/053.

[42] E. Jenvey, Strong stationarity and de Finetti's theorem, *J. Anal. Math.* **73** (1997), 1–18. MR 1616457. Zbl 0901.28011. https://doi.org/10.1007/BF02788136.

[43] I. Kátai, A remark on a theorem of H. Daboussi, *Acta Math. Hungar.* **47** no. 1-2 (1986), 223–225. MR 0836415. Zbl 0607.10034. https://doi.org/10.1007/BF01949145.

[44] A. B. Katok, Invariant measures of flows on orientable surfaces, *Dokl. Akad. Nauk SSSR* **211** (1973), 775–778. MR 0331438. Zbl 0298.28013.

[45] J. Kułaga-Przymus and M. Lemańczyk, The Möbius function and continuous extensions of rotations, *Monatsh. Math.* **178** no. 4 (2015), 553–582. MR 3422903. Zbl 1359.37019. https://doi.org/10.1007/s00605-015-0808-6.

[46] M. Lazard, Sur les groupes nilpotents et les anneaux de Lie, *Ann. Sci. Ecole Norm. Sup.* (3) **71** (1954), 101–190. MR 0088496. Zbl 0055.25103. https://doi.org/10.24033/asens.1021.

[47] A. Leibman, Polynomial sequences in groups, *J. Algebra* **201** no. 1 (1998), 189–206. MR 1608723. Zbl 0908.20029. https://doi.org/10.1006/jabr.1997.7269.

[48] A. Leibman, Pointwise convergence of ergodic averages for polynomial sequences of translations on a nilmanifold, *Ergodic Theory Dynam. Systems* **25** no. 1 (2005), 201–213. MR 2122919. Zbl 1080.37003. https://doi.org/10.1017/S0143385704000215.

[49] J. Liu and P. Sarnak, The Möbius function and distal flows, *Duke Math. J.* **164** no. 7 (2015), 1353–1399. MR 3347317. Zbl 06455743. https://doi.org/10.1215/00127094-2916213.

[50] K. Matomäki and M. Radziwiłł, Multiplicative functions in short intervals, *Ann. of Math.* (2) **183** no. 3 (2016), 1015–1056. MR 3488742. Zbl 1339.11084. https://doi.org/10.4007/annals.2016.183.3.6.

[51] K. Matomäki, M. Radziwiłł, and T. Tao, Sign patterns of the Liouville and Möbius functions, *Forum Math. Sigma* **4** (2016), e14, 44. MR 3513734. Zbl 06601295. https://doi.org/10.1017/fms.2016.6.

[52] C. Mauduit and J. Rivat, Prime numbers along Rudin-Shapiro sequences, *J. Eur. Math. Soc.* (*JEMS*) **17** no. 10 (2015), 2595–2642. MR 3420517. Zbl 06509927. https://doi.org/10.4171/JEMS/566.

[53] M. Morse and G. A. Hedlund, Symbolic dynamics II. Sturmian trajectories, *Amer. J. Math.* **62** (1940), 1–42. MR 0000745. Zbl 0022.34003. https://doi.org/10.2307/2371431.

[54] C. Müllner, Automatic sequences fulfill the Sarnak conjecture, *Duke Math. J.* **166** no. 17 (2017), 3219–3290. MR 3724218. Zbl 06825580. https://doi.org/10.1215/00127094-2017-0024.

[55] W. Parry, Dynamical representations in nilmanifolds, *Compositio Math.* **26** (1973), 159–174. MR 0320277. Zbl 0256.22013.

[56] R. Peckner, Möbius disjointness for homogeneous dynamics, 2015. arXiv 1506.07778.

[57] J. Petresco, Sur les commutateurs, *Math. Z.* **61** (1954), 348–356. MR 0066380. Zbl 0056.25505. https://doi.org/10.1007/BF01181351.

[58] M. Queffélec, *Substitution Dynamical Systems—Spectral Analysis*, second ed., *Lecture Notes in Math.* **1294**, Springer-Verlag, New York, 2010. MR 2590264. Zbl 1225.11001. https://doi.org/10.1007/978-3-642-11212-6.

[59] P. Sarnak, Mobius randomness and dynamics, *Not. S. Afr. Math. Soc.* **43** no. 2 (2012), 89–97. MR 3014544. Available at https://drive.google.com/file/d/0B9_NIFunxShBa3JKVVB2bDlrUjA/edit.

[60] P. Sarnak, Three lectures on the Möbius function randomness and dynamics. Available at https://publications.ias.edu/sites/default/files/MobiusFunctionsLectures(2).pdf.

[61] T. Tao, Probabilistic models and heuristics for the primes-Möbius pseudorandomness, blog entry. Available at https://terrytao.wordpress.com/2015/01/04/.

[62] T. Tao, The logarithmically averaged Chowla and Elliott conjectures for two-point correlations, *Forum Math. Pi* **4** (2016), e8, 36. MR 3569059. Zbl 06700128. https://doi.org/10.1017/fmp.2016.6.

[63] T. Tao, Equivalence of the logarithmically averaged Chowla and Sarnak conjectures, in *Number Theory—Diophantine Problems, Uniform Distribution and Applications*, Springer, Cham, 2017, pp. 391–421. MR 3676413. https://doi.org/10.1007/978-3-319-55357-3_21.

[64] T. Tao and J. Teräväinen, The structure of logarithmically averaged correlations of multiplicative functions, with applications to the Chowla and Elliott conjectures, 2017. arXiv 1708.02610.

[65] W. A. Veech, Interval exchange transformations, *J. Analyse Math.* **33** (1978), 222–272. MR 0516048. Zbl 0455.28006. https://doi.org/10.1007/BF02790174.

[66] P. Walters, *An Introduction to Ergodic Theory*, *Grad. Texts in Math.* **79**, Springer-Verlag, New York, 1982. MR 0648108. Zbl 0958.28011.

[67] Z. Wang, Möbius disjointness for analytic skew products, *Invent. Math.* **209** no. 1 (2017), 175–196. MR 3660308. Zbl 1377.37069. https://doi.org/10.1007/s00222-016-0707-z.

[68] T. D. Wooley and T. D. Ziegler, Multiple recurrence and convergence along the primes, *Amer. J. Math.* **134** no. 6 (2012), 1705–1732. MR 2999293. Zbl 1355.11005. https://doi.org/10.1353/ajm.2012.0048.

[69] T. Ziegler, A non-conventional ergodic theorem for a nilsystem, *Ergodic Theory Dynam. Systems* **25** no. 4 (2005), 1357–1370. MR 2158410. Zbl 1099.37002. https://doi.org/10.1017/S0143385703000518.

[70] T. Ziegler, Universal characteristic factors and Furstenberg averages, *J. Amer. Math. Soc.* **20** no. 1 (2007), 53–97. MR 2257397. Zbl 1198.37014. https://doi.org/10.1090/S0894-0347-06-00532-7.

(Received: September 19, 2017)

University of Crete, Department of Mathematics, Voutes University Campus, Heraklion 71003, Greece
E-mail: frantzikinakis@gmail.com

Université Paris-Est Marne-la-Vallée, Laboratoire d'analyse et de mathématiques appliquées, UMR CNRS 8050, 77454 Marne la Vallée Cedex, France
E-mail: bernard.host@u-pem.fr