Ergod. Th. & Dynam. Sys. (200?), **??**, 1–11
Printed in the United Kingdom

Approximate Transitivity for Zero Entropy Systems

A. Dooley†† and Anthony Quas‡†

†*School of Mathematics, University of New South Wales, Sydney, NSW 2052, Australia (e-mail: a.dooley@unsw.edu.au)*

‡*Department of Mathematical Sciences, University of Memphis, Memphis, TN 38152, USA (e-mail: aquas@memphis.edu)*

(*Received ?????*)

Abstract. We show that the Morse system is AT and that a system of positive entropy cannot be AT. We give examples of zero entropy systems that are not AT, one of which is a two-point extension of a system that is AT.

1. *Introduction and Statement of Results*

A dynamical system $(X, \mathcal{B}, \mu, T)$ is said to be *approximately transitive* (or to have the *AT property*) if any finite collection of positive L^1 functions may be simultaneously approximated by a positive linear combination of iterates of a single L^1 function, that is given $\epsilon > 0, f_1, \ldots, f_n \in L^1_+(X)$, there exist $f \in L^1_+(X)$, $n_j : j = 1, \ldots, m$ and $\alpha_{i,j}$, $i = 1, \ldots, n$, $j = 1, \ldots, m \in \mathbb{R}^+$ such that

$$\left\| f_i - \sum_{i,j} \alpha_{i,j} \mathcal{L}_{n_j} f \right\| \leq \epsilon$$

where $\mathcal{L}_n f(x) = \frac{d\mu \circ T^n}{d\mu}(x) f(T^n x)$ and $\| \cdot \|$ denotes the L^1 norm. In this paper, we consider exclusively invertible probability measure-preserving transformations and so $\mathcal{L}_n f(x)$ reduces to $f(T^n x)$.

This property was introduced by Connes and Woods [**1**], who used it in their study of von Neumann algebras. Indeed, in the Connes-Krieger "dictionary" between von Neumann factors, ergodic dynamical systems and flows, the algebra is of ITPFI type (infinite product of finite type I) if and only if the associated dynamical system is of product type, if and only if the associated flow has the

† This research was partially funded by the ARC
† AQ was partially supported by NSF grant DMS-0200703 and wishes to thank the University of New South Wales for its hospitality while this paper was being written

AT property. While their original proof of these equivalences used deep facts from the theory of von Neumann algebras, later proofs of the second equivalence were given using only ergodic theoretic methods by Hawkins and Hamachi [**14, 13**]. Subsequent work on the structure of ITPFI factors and AT flows was carried out by Giordano and Skandalis [**10, 11, 12**].

The AT property appears to be related to various rank properties as well as entropy. For a survey on rank of transformations, the reader is referred to Ferenczi's work [**7**]. A transformation has *rank one* if there exist a sequence of subsets B_n and integers k_n with the following properties:

1. The sets $T^i(B_n)$ for $0 \le i < k_n$ are pairwise disjoint;
2. $\mu(\bigcup_{0\le i<k_n} T^i(B_n)) \to 1$ as $n \to \infty$;
3. the algebras $\mathcal{A}_n$ generated by the sets $T^i(B_n)$ for $0 \le i < k_n$ have the property that for all $A \in \mathcal{B}$ and all $\epsilon > 0$, there exists an n_0 such that for $n \ge n_0$, there exists a set $B \in \mathcal{A}_n$ such that $\mu(B \triangle A) < \epsilon$.

A system has *funny rank one* (a concept introduced by Thouvenot and studied by Ferenczi in [**5**]) if there exist a sequence of subsets B_n, and a sequence of sequences $a_0^n, \dots, a_{k_n-1}^n$ with the following properties:

1. The sets $T^{a_i^n}(B_n)$ are pairwise disjoint;
2. $\mu(\bigcup_{0\le i<k_n} T^{a_i^n}(B_n)) \to 1$ as $n \to \infty$;
3. the algebras $\mathcal{A}_n$ generated by the sets $T^{a_i^n}(B_n)$ for $0 \le i < k_n$ have the property that for all $A \in \mathcal{B}$ and all $\epsilon > 0$, there exists an n_0 such that for $n \ge n_0$, there exists a set $B \in \mathcal{A}_n$ such that $\mu(B \triangle A) < \epsilon$.

It is not hard to see that a funny rank one transformation has the AT property: taking $f = \mathbf{1}_{B_n}$ for a suitably large n, we are able to take positive linear combinations of iterates to approximate arbitrary finite collections of positive functions.

The Morse system is a well-known and much-studied dynamical system, which has been the source of many counter-examples. It is a two-point extension of an odometer. Del Junco [**4**] showed that it has simple spectrum, but is not of rank one. We note also work of Ferenczi [**6**], which studies the tiling properties of the Morse system. As it is not rank one, it cannot be tiled by translates of a single word: Ferenczi shows more, that the maximum density of the translates of a single word is 2/3. We show in Theorem 1 that the Morse system has the AT property.

The AT property implies [**2**] a rather strong lower bound for the measures of certain "funny $\bar{d}$-cylinders", which we reprove in Theorem 2. A corollary of this result is that a positive entropy system cannot have the AT property. Since we then have the implications rank one implies funny rank one implies approximately transitive implies zero entropy, it is natural to ask which of the reverse implications hold.

Ferenczi [**5**] gave an example of a system that was of funny rank one, but not rank one. It was previously unknown whether zero entropy was a sufficient (as well as necessary) condition for the AT property to hold. We resolve this by showing that a zero entropy transformation of the torus studied by Furstenberg [**8**] fails to have the AT property. If Ferenczi's conjecture [**6**] that the Morse system is not of funny rank one is true, it would of course demonstrate that the converse implication approximately transitive implies funny rank one is false.

In view of the positive result for the Morse system, it is natural to ask whether the AT property is preserved under finite extensions. In Theorem 5, we use a construction of Helson and Parry [**15**] to give an example of a system with the AT property with an ergodic two-point extension that fails to have the AT property.

A related question due to Thouvenot is whether funny rank one is equivalent to the condition of simple spectrum in L^p for all $p > 0$. No examples are known to have simple spectrum without the AT property or vice versa. We are unable to shed further light on this as the Morse system which is approximately transitive has simple (L^2) spectrum and the two examples that we give that are not approximately transitive have infinite multiplicity. In [**16**], questions of the existence of L^p simple spectrum are studied in the case of ergodic group automorphisms.

We would like to thank Toshihiro Hamachi, Ivo Klemeš and Jean-Paul Thouvenot for useful discussions during the preparation of this paper.

2. *Statements and Proofs*

Let $\Omega = \{0,1\}^{\mathbb{Z}^+}$ and define

$$\phi(\omega) = \begin{cases} -1 & \text{if } \omega \text{ has an even number of trailing 1s} \\ +1 & \text{otherwise} \end{cases}$$

We note that Ω is a group under the operation of addition with carry and we let 1 represent the element of ω whose zeroth term is 1 and whose other terms are 0. Let $X = \Omega \times \{\pm 1\}$ and define the Morse transformation on X by $T(\omega, t) = (\omega + 1, t \cdot \phi(\omega))$, where $\cdot$ denotes multiplication in $\{\pm 1\}$. We let μ denote the Haar measure on Ω and letting c be the normalized counting measure on $\{\pm 1\}$, $\mu \times c$ is known to be a uniquely ergodic invariant measure for the Morse transformation.

This defines the Morse system as an explicit two-point extension of the odometer.

THEOREM 1. *The Morse system has the AT property.*

Proof. We will demonstrate that for any n, the functions $g_{n,i}$ for $i = \pm 1$ may be arbitrarily well-approximated in L^1 by positive linear combinations of a single function f_n, where $g_{n,i}$ is the characteristic function of the set of (ω, t) such that $t = i$ and $\omega_{n-1} = \ldots = \omega_0 = 0$. To see the sufficiency of this, we note that if it is possible to approximate the functions $g_{n,i}$, then taking appropriate combinations of these functions, it is possible to approximate any positive function $f(\omega, t)$ depending only on t and the last n coordinates of ω. Since these functions are dense in the positive L^1 functions, the sufficiency follows.

To demonstrate that the functions $g_{n,i}$ may be well-approximated, we can in fact show that it is sufficient to approximate the functions $g_i = g_{0,i}$. To see this, let $\sigma(\omega, t) = (\sigma(\omega), t)$, where $\sigma \colon \Omega \to \Omega$ is the shift transformation and suppose that

$\|g_i - \sum_j a_{i,j} f \circ T^j\| < \epsilon$. Then using the relation $T \circ \sigma^n = \sigma^n \circ T^{2^n}$, we see

$$\|g_i \circ \sigma^n - \sum_j a_{i,j} f \circ \sigma^n \circ T^{2^n j}\| < \epsilon.$$

Letting h be the indicator function of the set of points where $\omega_{n-1} = \ldots = \omega_0 = 0$, we see that $h \circ T^{2^n} = h$. Multiplying the above inequality by h, we see that

$$\left\| g_{n,i} - \sum_j a_{i,j}((f \circ \sigma^n)h) \circ T^{2^n j} \right\| = \left\| h \left(g_i \circ \sigma^n - \sum_j a_{i,j}(f \circ \sigma^n) \circ T^{2^n j} \right) \right\|$$
$$= \left\| h \left(g_i - \sum_j a_{i,j} f \circ T^j \right) \circ \sigma^n \right\| = 2^{-n} \left\| g_i - \sum_j a_{i,j} f \circ T^j \right\| < 2^{-n}\epsilon,$$

showing that for arbitrary n, if it is possible to approximate the functions g_i, then it is possible to approximate the functions $g_{n,i}$.

Let $N = 2^{8M}$ be a large power of 2 and let A be the subset of X consisting of those points (ω, t) such that $t = 1$ and for each $0 \leq j < N$, the string $\omega_{8Mj+4M}\,\omega_{8Mj+4M-1} \ldots \omega_{8Mj}$ either ends with a 0 or is equal to $011\ldots1$.

We then define a function on A as follows:

$$f_M(\omega, 1) = 2^{N+1} 2^{-3Ms(\omega)} (1 - 2^{-7M})^{N-s(\omega)},$$

where $s(\omega)$ is the number of blocks in ω of the form $011\ldots1$ described above.

This function is chosen to have L^1 norm 1. We check this as follows. The measure of the set of $(\omega, 1)$ with s of the $011\ldots1$ blocks is given by

$$\frac{1}{2}\binom{N}{s}\left(\frac{1}{2^{4M+1}}\right)^s \left(\frac{1}{2}\right)^{N-s}$$
$$= \frac{1}{2^{N+1}}\binom{N}{s} 2^{-4Ms}.$$

The norm of the function is therefore given by

$$\|f_M\|_1 = \frac{1}{2^{N+1}} \sum_{s=0}^{N} \binom{N}{s} 2^{-4Ms} 2^{N+1} 2^{-3Ms} \left(1 - 2^{-7M}\right)^{N-s}$$
$$= \sum_{s=0}^{N} \binom{N}{s} \left(2^{-7M}\right)^s \left(1 - 2^{-7M}\right)^{N-s}$$
$$= 1.$$

We will show that the functions g_i can be arbitrarily closely approximated by suitable positive linear combinations of the function f_M for sufficiently large M.

The combinations that we consider are the following:

$$\Phi_i(\omega, t) = \frac{1}{2^N} \sum_{\mathbf{a} \in \{0,1\}^N :\, (-1)^{\sum \mathbf{a}} = i} f \circ T^{-n(\mathbf{a})}(\omega, t), \tag{1}$$

where $n(\mathbf{a})$ denotes $\sum_{j=0}^{N-1} a_j 2^{8Mj}$.

Since Φ_i clearly has norm 1/2 (being the sum of 2^{N-1} functions of norm 2^{-N}), in order to demonstrate that $\|\Phi_i - g_i\|$ is small, it is sufficient to demonstrate that $\Phi_i(\omega, i)$ is close to 1 for a set of $\omega \in \Omega$ of measure close to 1.

Fix $\omega \in \Omega$ with k_0 blocks of the form $011\dots1$ and k_1 blocks of the form $100\dots0$. For blocks not of the form $011\dots1$ or $100\dots0$, there is a unique choice of $\mathbf{a}_j$ that is forced in order to ensure that $T^{-n(\mathbf{a})}(\omega, t) \in A$, whereas for blocks of the form $011\dots1$ or $100\dots0$, a point of A is formed whether the corresponding $\mathbf{a}_j$ is taken to be 1 or 0.

We write p for $\omega_0 + \omega_{8M} + \dots + \omega_{(N-1)(8M)}$. We will consider the contribution to $\Phi_i(\omega, i)$ coming from terms where l_0 of the $\mathbf{a}_j$ corresponding to blocks of the form $011\dots1$ are 0 and l_1 of the $\mathbf{a}_j$ corresponding to blocks of the form $100\dots0$ are 1. The number of $1's$ in $\mathbf{a}$ will then be given by $(p - k_0) + l_1 + (k_0 - l_0)$ (where the $p - k_0$ is the number of 1's forced to be in $\mathbf{a}$ and $k_0 - l_0$ is the number of 1's that need to be in $\mathbf{a}$ in order to ensure that l_0 of the $011\dots1$ have the corresponding $\mathbf{a}_j = 0$). In order to ensure that $T^{-n(\mathbf{a})}(\omega, i) \in A$, we require $(-1)^{p+l_1-l_0} = i$. Clearly this is satisfied if and only if $l_0 + l_1 \sim q$, where $a \sim b$ means that a and b differ by a multiple of 2 and q is such that $(-1)^q = (-1)^p i$.

We therefore consider a fixed pair l_0, l_1 such that $l_0 + l_1$ has the required parity. We then look at the contribution to $\Phi_i(\omega, i)$ coming from the terms $\mathbf{a}$ with the correct forced terms and with l_0 of the $\mathbf{a}_j$ corresponding to blocks of the form $011\dots1$ equal to 0 and l_1 of the $\mathbf{a}_j$ corresponding to blocks of the form $100\dots0$ equal to 1. This contribution is given by

$$2\binom{k_0}{l_0}\binom{k_1}{l_1}2^{-3M(l_0+l_1)}\left(1-2^{-7M}\right)^{N-(l_0+l_1)}$$

Let K denote $k_0 + k_1$. Letting $L \sim q$ and summing over terms with $l_0 + l_1 = L$, we get

$$2\binom{K}{L}2^{-3ML}\left(1-2^{-7M}\right)^{N-L}$$

Summing this over the terms L of the appropriate parity, we get

$$2\sum_{\{L:\,0\le L\le K;\,L\sim q\}}\binom{K}{L}2^{-3ML}\left(1-2^{-7M}\right)^{N-L}.$$

We note that with probability exponentially close to 1, the quantity K (depending on ω) given by the number of blocks of the form $011\dots1$ plus the number of blocks of the form $100\dots0$ is between $2^{4M} - 2^{8M/3}$ and $2^{4M} + 2^{8M/3}$. For K in this range, the summand is uniformly exponentially small and unimodal, the sums over L odd and L even differ by an exponentially small quantity. Hence for K in the appropriate range, we observe that the sums above with either parity are (exponentially) close to

$$\sum_{L=0}^{K}\binom{K}{L}2^{-3ML}\left(1-2^{-7M}\right)^{N-L}$$
$$=(1-2^{-7M})^{N-K}\left(2^{-3M}+(1-2^{-7M})\right)^{K}$$

It is straightforward to verify that the quantity above differs only by a multiplicative constant exponentially close to 1 as K varies between the two limits. To finish the proof, it is sufficient to verify that when $K=2^{4M}$, this quantity is close to 1.

We are thus considering $(1-2^{-7M})^{2^{8M}-2^{4M}}(1+2^{-3M}-2^{-7M})^{2^{4M}}$.

Taking logarithms, we get

$$(2^{8M}-2^{4M})(-2^{-7M})+O(2^{-6M})+2^{4M}(2^{-3M}-2^{-7M})+O(2^{-2M})$$
$$=-2^{M}+2^{M}+O(2^{-2M})=O(2^{-2M}).$$

We therefore see that for points (ω,i) with ω belonging to a set of measure exponentially close to 1, $\Phi_i(\omega,i)$ is exponentially close to $g_i(\omega,i)$. This completes the proof.

□

We now give a necessary condition for a system to have the AT property. The corollary (Corollary 3) to the following theorem below was stated to us by J.-P. Thouvenot and appears in the (unpublished) thesis of David [**2**]. Although we have not seen a copy of the thesis, we understand that it contains a similar theorem to the one we prove below.

Let ν be a shift-invariant measure on $Y=\{1,\dots,l\}^{\mathbb{Z}}$. A *funny cylinder set* in Y is a set of the form $C_{x,\Lambda}=\{y\in Y\colon y_i=x_i \text{ for all } i\in\Lambda\}$, where $x\in Y$ and Λ is a finite set. The ϵ-ball about the funny cylinder $C_{x,\Lambda}$ is the set $B_{\epsilon,x,\Lambda}=\{y\in Y\colon y_i=x_i \text{ for all but at most } \epsilon|\Lambda| \text{ of the } i\in\Lambda\}$.

We note the similarity between the following condition and a necessary condition for rank one given by del Junco in [**4**].

Theorem 2. *Suppose that the system $(X,\mathcal{B},\mu,T)$ has the AT property. Then let $\mathcal{P}=\{A_1,\dots,A_l\}$ be an arbitrary finite measurable partition. Denote by π, the natural map induced by $\mathcal{P}$ from X to $\{1,\dots,l\}^{\mathbb{Z}}$ and by ν the measure $\mu\circ\pi^{-1}$ induced by μ on Y. Then for every $\delta>0$ and $\epsilon>0$, there exist arbitrarily large finite sets $\Lambda\subset\mathbb{Z}$ and points $x\in Y$ such that $\nu(B_{\epsilon,x,\Lambda})>(1-\delta)/|\Lambda|$.*

Proof. We check that the system Y has the AT property as follows: Let $g_1,\dots,g_n$ be a collection of positive functions on Y that we wish to approximate. Then let π be the natural factor map from X onto Y, let $\mathcal{F}$ be the σ-algebra on Y and let S denote the shift map on Y. Then there exists a function f on X, positive linear combinations of which may be used to approximate the functions $g_i\circ\pi$ so that $\|g_i\circ\pi-\sum_j a_{i,j}f\circ T^j\|<\epsilon$ for each i. Taking expectations with respect to $\pi^{-1}\mathcal{F}$ (which is a T-invariant σ-algebra), we see that $\|g_i\circ\pi-\sum_j a_{i,j}\mathbb{E}(f|\pi^{-1}\mathcal{F})\circ T^j\|<\epsilon$

for each i. We then observe that we can write $\mathbb{E}(f|\pi^{-1}\mathcal{F})$ as $F\circ\pi$ for a measurable function F on Y. It then follows that $\|g_i - \sum_j a_{i,j} F\circ S^j\| < \epsilon$ for each i as required.

Take an arbitrarily fine refinement $\mathcal{Q} = \{A_1, A_2, \ldots, A_l\}$ of the partition of Y into cylinder sets and define functions g_i to be the normalized indicator functions of the sets A_i in the partition. Since Y has the AT property, there exists a positive L^1 function f with $\|f\| = 1$ on Y and positive numbers $a_{i,j}$ summing to 1 such that $\|g_i - \sum_j a_{i,j} f\circ S^{-j}\| < \epsilon\delta^2/36$ (the negative powers of S being convenient later).

In particular, we see that for each i,

$$\int_{Y\setminus A_i} \sum_j a_{i,j} f\circ S^{-j} < \frac{\epsilon\delta^2}{36}.$$

Letting P_i be the set $\{j\colon \int_{Y\setminus A_i} f\circ S^{-j} \geq \epsilon\delta/6\}$, we check that $\sum_{j\in P_i} a_{i,j} < \delta/6$. By setting the $a_{i,j}$ to 0 for $j \in P_i$ and rescaling the remaining $a_{i,j}$ to give coefficients $b_{i,j}$ summing to 1, we are able to ensure that $\|g_i - \sum_j b_{i,j} f\circ S^{-j}\| < \delta/2$ with the additional condition that if $b_{i,j} > 0$, then

$$\int_{Y\setminus A_i} f\circ S^{-j} < \frac{\epsilon\delta}{6}.$$

Write Λ for $\{j\colon \text{there exists } i \text{ with } b_{i,j} > 0\}$. Then we see

$$\sum_i \sum_{\{j\colon b_{i,j}>0\}} \int_{Y\setminus A_i} f\circ S^{-j} < \frac{\epsilon\delta|\Lambda|}{6}.$$

This can be written as

$$\begin{aligned}\frac{\epsilon\delta|\Lambda|}{6} &> \sum_i \sum_{\{j\colon b_{i,j}>0\}} \int f\circ S^{-j} \mathbf{1}_{Y\setminus A_i}\, d\nu \\ &= \int f \sum_i \sum_{\{j\colon b_{i,j}>0\}} \mathbf{1}_{Y\setminus A_i}\circ S^j\, d\nu.\end{aligned}$$

Now, let

$$H = \frac{1}{|\Lambda|} \sum_i \sum_{\{j\colon b_{i,j}>0\}} \mathbf{1}_{Y\setminus A_i}\circ S^j.$$

For $j \in \Lambda$, there is by construction a unique cylinder set $[k]$ for which $\int_{[k]} f\circ S^{-j} > 1 - \epsilon\delta/6$. Define x_j to be this k. If i is such that $b_{i,j} > 0$ for this j, then $A_i \subset [k]$. We see that $H(y)$ is bounded below by the 'mean Hamming distance' between the point (y_j) and the partially defined sequence $(x_j)_{j\in\Lambda}$ where the Hamming distance is computed by looking only at differences occurring at positions $j \in \Lambda$. Since the above inequality demonstrates that $\int Hf\, d\nu < \epsilon\delta/6$, it follows that $\int_{H(y)>\epsilon} f(y) < \delta/6$. Accordingly, we define a new function $\tilde{f}$ as follows:

$$\tilde{f}(y) = \begin{cases} 0 & \text{if } H(y) > \epsilon; \\ f(y)/\int_{\{z\colon H(z)\leq\epsilon\}} f & \text{otherwise.}\end{cases}$$

Since $\|\tilde{f} - f\| < \delta/2$, we check that

$$\|g_i - \sum_j b_{i,j} \tilde{f} \circ S^{-j}\| < \delta.$$

Taking a weighted combination of the above, we see

$$\|1 - \sum_{i,j} \mu(A_i) b_{i,j} \tilde{f} \circ S^{-j}\| < \delta.$$

Since the support of the function $\tilde{f}$ is contained in $B_{\epsilon,x,\Lambda}$, the sum in question is supported on a set of measure at most $|\Lambda|\nu(B_{\epsilon,x,\Lambda})$, yielding the inequality

$$\nu(B_{\epsilon,x,\Lambda}) > (1-\delta)/|\Lambda|.$$

This completes the proof. □

Corollary 3. *(David [2]) If $(X, \mathcal{B}, \mu, T)$ has positive entropy, then it does not have the AT property.*

Proof. If $(X, \mathcal{B}, \mu, T)$ has positive entropy, then it has a Bernoulli factor $Y = \{0,1\}^{\mathbb{Z}}$ with probability p of 0's and q of 1's with $0 < p < q$. One can check that the measure of any set $B_{\epsilon,x,\Lambda}$ is bounded above by

$$\binom{n}{\lfloor \epsilon n \rfloor} q^{n - \lfloor \epsilon n \rfloor},$$

where $n = |\Lambda|$. It is not hard to show that for sufficiently small ϵ, this quantity is strictly less than $(1-\epsilon)/n$ for all large n. □

Corollary 4. *There exists a zero entropy system that does not have the AT property.*

Proof. Let α be an irrational number and consider the transformation of $\mathbb{T}^2$ given by $T(x, y) = (x+\alpha, y+2x+\alpha) \pmod 1$. This transformation, studied by Furstenberg [**8, 9**], is uniquely ergodic, preserving Haar measure on the torus. We will argue using a version of the weak law of large numbers that it is not approximately transitive.

Let $\mathcal{P}$ be the partition of $\mathbb{T}^2$ consisting of $A_0 = \mathbb{T} \times [0, \frac{1}{2})$ and $A_1 = \mathbb{T} \times [\frac{1}{2}, 1)$. Letting ν be the measure on $Y = \{0,1\}^{\mathbb{Z}}$ induced from Haar measure by this partition. If we denote by Y_n, the nth term of a sequence in Y, we will show that the terms are **pairwise** independent equidistributed on 0 and 1. To see this, we will calculate the probability that $Y_0 = i$ and $Y_n = j$. This is by definition equal to the Haar measure of the set of points (a, b) such that $b \in [i/2, (i+1)/2)$ and $\pi_2(T^n(a, b)) \in [j/2, (j+1)/2)$. This is the set of points (a, b) such that $b \in [i/2, (i+1)/2)$ and $\langle b + n^2\alpha + na \rangle \in [j/2, (j+1)/2)$, where $\langle t \rangle$ denotes the fractional part of t. It can then be seen that this set of points is of measure $1/4$ as required.

Let Λ be an arbitrary subset of $\mathbb{Z}$ of size n and fix an arbitrary point $x \in Y$. For $y \in Y$ and $j \in \Lambda$, we write

$$A_j(y) = \begin{cases} 1 & \text{if } y_j = x_j \\ -1 & \text{otherwise} \end{cases}$$

Setting $S = \sum_{j\in\Lambda} A_j$, we see that the mean of S is 0 and the variance of S is n. From Tchebychev's inequality, it follows that the probability that $|S| > 3n/4$ is at most $16/(9n)$. Since the transformation $R(x, y) = (x, y + 1/2)$ commutes with T and has the effect of reversing all of the values of the Y_i and hence all the values of the A_i, it follows that S has a distribution that is symmetric about the origin. We therefore see that the probability that $S > 3n/4$ is bounded above by $8/(9n)$. Since $S(y) > 3n/4$ is a necessary and sufficient condition for $y \in B_{1/8,x,\Lambda}$, we see that $\nu(B_{1/8,x,\Lambda}) < 8/(9n)$ and it follows from Theorem 2 that this system does not have the AT property. □

THEOREM 5. *There exists a rank 1 system with the AT property and an ergodic two-point extension of it that fails to have the AT property.*

Proof. We modify a construction of Helson and Parry [**15, 17**] and again apply a weak law of large numbers. Letting p_n denote the nth prime number, let X be the compact group (with the product topology) $\mathbb{Z}/(2^N\mathbb{Z}) \times \prod_{n=2}^{\infty} \mathbb{Z}/(p_n^2\mathbb{Z})$, where N is to be determined. Let 1 denote the element $(1, 1, 1 \ldots)$ of X_0, let T be the rotation given by $T(x) = x + 1$ and let μ be Haar measure on X.

Next on the nth factor in X for $n \geq 2$, define

$$\phi_n(i) = \begin{cases} -1 & \text{if } i = 0; \\ 1 & \text{otherwise.} \end{cases}$$

Then since ϕ_k is different from 1 only on a set of measure $1/p_k^2$, it follows that for almost every $(x_1, x_2, \ldots) \in X$, $\phi_k(x_k)$ is eventually equal to 1. Hence for almost every $(x_1, x_2, \ldots) \in X$, $\prod_{k=2}^{\infty} \phi_k(x_k)$ exists. We will define below a function $\phi_1 \colon \mathbb{Z}/(2^N\mathbb{Z}) \to \{\pm 1\}$ and then define a function $\phi \colon X \to \{\pm 1\}$ by

$$\phi(x_1, x_2, \ldots) = \prod_{k=1}^{\infty} \phi_k(x_k).$$

Let $\phi^{(n)}(x)$ denote $\phi(x)\phi(Tx)\ldots\phi(T^{n-1}x)$. We will give estimates due to Helson and Parry on $c(n) = |\int \phi^{(n)}(x)\, d\mu|$ that will be valid independent of the choices of N and ϕ_1.

Specifically since the measure μ is a product measure and $\phi^{(n)}$ is the product of the $\phi_k^{(n)}$, one has

$$c(n) = \prod_{k=1}^{\infty} \left| \int \phi_k^{(n)}\, d\mu_k \right|,$$

where μ_k denotes the Haar measure on the kth factor. Note that all factors in the product are bounded above by 1. Moreover, in any factor such that

$p_k^2 \geq 2n$, the quantity $\phi_k^{(n)}$ is equal to -1 on a set of measure n/p_k^2 and is equal to 1 elsewhere. Accordingly, the kth factor in the above product is equal to $1 - 2n/p_k^2 \leq \exp(-2n/p_k^2)$. Hence they derive the estimate

$$c(n) \leq \exp\left(-2n \sum_{p_k^2 \geq 2n} \frac{1}{p_k^2}\right),$$

from which one gets the crude inequality $c(n) \leq \exp(-2Kn^{1/4})$ for some $K > 0$. It then follows that the $c(n)$ are summable. Let M be such that $\sum_{n=M+1}^{\infty} c(n) < 1/8$. Now choose N such that $2^N > 64M^3$. We consider a function ϕ_1 chosen uniformly at random from the 2^{2^N} possible functions $\mathbb{Z}/(2^N\mathbb{Z}) \to \{\pm 1\}$. Write $\omega_i = \phi_1(i)$ for $0 \leq i < 2^N$. We define ± 1-valued random variables X_i^n for $1 \leq n \leq M$ and $0 \leq i < 2^N$ by $X_i^n = \omega_i \omega_{i+1} \cdots \omega_{i+k-1}$, where the quantities $i+1, i+2, \ldots$ are interpreted modulo 2^N. We see that $c(n) \leq |Y^n|$, where $Y^n = (X_0^n + X_1^n + \ldots + X_{2^N-1}^n)/2^N$. Since the variables X_i^n have expectation 0 and are pairwise (but not mutually) independent of variance 1, we see that Y^n has expectation 0 and variance $1/2^N$. Accordingly, the probability that $|Y^n|$ exceeds $1/(8M)$ is less than $64M^2/2^N$. Hence the probability that one of the $|Y^n|$ for $1 \leq n \leq M$ exceeds $1/(8M)$ is less than $64M^3/2^N < 1$. It follows that there exists a choice of ϕ_1 such that $|Y^n| \leq 1/(8M)$ for each $1 \leq n \leq M$. Fix this function on the first factor. We then see that the quantities $c(n)$ for $1 \leq n \leq M$ are bounded above by $1/(8M)$ so that $\sum_{n=1}^{\infty} c(n) < 1/4$.

We note that the difference between this construction and the original one due to Helson and Parry is that in their construction, they had the bound $\sum_{n=1}^{\infty} c(n) < \infty$. For our later use, the bound of $1/4$ is needed. The changes made are entirely in the first factor.

By the Chinese Remainder Theorem, every point of X has a dense orbit so T is uniquely ergodic. Let μ denote Haar measure on X. Since T is an ergodic rotation of a compact Abelian group, it has discrete spectrum. Hence by a result of del Junco [**3**], T is a rank 1 transformation.

We consider the extension $T_\phi \colon X \times \{\pm 1\} \to X \times \{\pm 1\}$ given by $T_\phi(x,t) = (T(x), t \cdot \phi(x))$ and introduce the partition of $X \times \{\pm 1\}$, $\mathcal{P} = \{X \times \{1\}, X \times \{-1\}\}$. As before, T_ϕ preserves the product measure $\mu \times c$. The partition induces a natural map from $X \times \{\pm 1\}$ to $Y = \{\pm 1\}^{\mathbb{Z}}$ as in Theorem 2. Let ν be the induced measure on Y.

We are now in a position to argue as in Corollary 4. Let Λ be any finite subset of $\mathbb{Z}$ and let $z \in Y$ be fixed. For $y \in Y$ and $i \in \Lambda$, let $A_i(y) = z_i y_i$. Let T be the sum of the A_i over $i \in \Lambda$. The A_i have expectation 0 (since ν is invariant under flipping the entire sequence). We have

$$\begin{aligned}\mathrm{Cov}(A_i, A_j) &= z_i z_j \int y_i y_j \, d\nu = z_i z_j \int y_0 y_{|i-j|} \, d\nu \\ &= z_i z_j \int \phi^{(|i-j|)}(x) \, d\mu(x),\end{aligned}$$

so that $|\mathrm{Cov}(A_i, A_j)| = |\mathrm{Cov}(A_0, A_{|i-j|})| = c(|i-j|)$. We estimate

$$\begin{aligned}\mathrm{Var}(T) &\le \sum_{i\in\Lambda}\sum_{j\in\Lambda} c(|i-j|)\\ &\le \sum_{i\in\Lambda}\sum_{j\in\mathbb{Z}} c(|i-j|) < 3|\Lambda|/2.\end{aligned}$$

Working exactly as before, we see that $\nu(B_{15/16,z,\Lambda}) < 48/49$. This contradicts the conclusion of Theorem 2 showing that T_ϕ fails to have the AT property. □

References

[1] A. Connes and E. Woods. Approximately transitive flows and ITPFI factors. *Ergodic Theory Dynam. Systems*, 5:203–236, 1985.

[2] M.-C. David. *Sur quelques problèmes de théorie ergodique non commutative.* PhD thesis, 1979.

[3] A. del Junco. Transformations with discrete spectrum are stacking transformations. *Canad J. Math.*, 28:836–839, 1976.

[4] A. del Junco. A transformation with simple spectrum which is not rank one. *Canad. J. Math.*, 29:655–663, 1977.

[5] S. Ferenczi. Systèmes de rang un gauche. *Ann. Inst. H. Poincaré Probab. Statist.*, 21:177–186, 1985.

[6] S. Ferenczi. Tiling and local properties of the Morse sequence. *Theoretical Computer Science*, 129:369–383, 1994.

[7] S. Ferenczi. Systems of finite rank. *Colloq. Math.*, 73:35–65, 1997.

[8] H. Furstenberg. Strict ergodicity and transformations of the torus. *Amer. J. Math*, 83:573–60, 1961.

[9] H. Furstenberg. *Recurrence in Ergodic Theory and Combinatorial Number Theory.* Princeton University Press, 1981.

[10] T. Giordano and G. Skandalis. Krieger factors isomorphic to their tensor square and pure point spectrum flows. *J. Funct. Anal.*, 64:209–226, 1985.

[11] T. Giordano and G. Skandalis. On infinite tensor products of factors of type I_2. *Ergodic Theory Dynam. Systems*, 5:565–586, 1985.

[12] T. Giordano, G. Skandalis, and E. Woods. On the computation of invariants for ITPFI factors. *J. Operator Theory*, 15:83–107, 1986.

[13] T. Hamachi. A measure theoretical proof of the Connes-Woods theorem on AT-flows. *Pacific J. Math.*, 154:67–85, 1992.

[14] J. Hawkins. Properties of ergodic flows associated to product odometers. *Pacific J. Math.*, 141:287–294, 1990.

[15] H. Helson and W. Parry. Cocycles and spectra. *Ark. Mat.*, 16:195–206, 1978.

[16] A. Iwanik. The problem of L^p simple spectrum for ergodic group automorphisms. *Bull. Soc. Math. France*, 119:91–96, 1991.

[17] W. Parry. *Topics in ergodic theory*, volume 75 of *Cambridge Tracts in Mathematics.* Cambridge University Press, Cambridge, 1981.