Measure-Theoretically Mixing Subshifts of Minimal Word Complexity

Darren Creutz*

10 October 2025

Abstract We resolve a long-standing open question on the relationship between measure-theoretic dynamical complexity and symbolic complexity by establishing the exact word complexity at which measure-theoretic strong mixing manifests:

For every superlinear $f : \mathbb{N} \to \mathbb{N}$, i.e. $f(q)/q \to \infty$, there exists a subshift admitting a (strongly) mixing of all orders probability measure with word complexity p such that $p(q)/f(q) \to 0$.

For a subshift with word complexity p which is non-superlinear, i.e. $\liminf p(q)/q < \infty$, every ergodic probability measure is partially rigid.

Introduction

Among measure-theoretic dynamical properties of measure-preserving transformations, strong mixing of all orders is the 'most complex': every finite collection of measurable sets tends asymptotically toward independence, necessarily implying a significant amount of randomness. Despite this, 'low complexity' mixing transformations exist–there are mixing transformation with zero entropy–raising the question of how deterministic a mixing transformation can be.

Word complexity, the number $p(q)$ of distinct words of length q appearing in the language of the subshift, provides a more fine-grained means of quantifying complexity in the zero entropy setting, leading to the question of how low the word complexity of a mixing transformation can be.

Ferenczi [Fer95] initially conjectured that mixing transformations' word complexity should be superpolynomial but quickly refuted this himself [Fer96] showing that the staircase transformation, proven mixing by Adams [Ada98], has quadratic word complexity. Recent joint work of the author and R. Pavlov and S. Rodock [CPR23] exhibited subshifts admitting mixing measures with word complexity functions which are subquadratic but superlinear by more than a logarithm. We exhibit subshifts admitting mixing measures with complexity arbitrarily close to linear:

Theorem A. For every $f : \mathbb{N} \to \mathbb{N}$ which is superlinear, $f(q)/q \to \infty$, there exists a subshift, admitting a strongly mixing probability measure, with word complexity p such that $p(q)/f(q) \to 0$.

Our examples, which we call quasi-staircase transformations, are mixing rank-one transformations hence mixing of all orders [Kal84], [Ryz93]. We establish their word complexity is optimal:

Theorem B. Every subshift of non-superlinear word complexity, $\liminf p(q)/q < \infty$, equipped with an ergodic probability measure is partially rigid hence not strongly mixing,

Non-superlinear complexity subshifts are conjugate to S-adic shifts (Donoso, Durand, Maass and Petite [DDMP21]). Named by Vershik and the subject of a well-known conjecture of Host, S-adic subshifts are quite structured (see e.g. [Ler12] for more information on S-adicity).

Our work may be viewed as saying there is a sharp divide in 'measure-theoretic complexity', precisely at superlinear word complexity, between highly structured and highly complicated: as soon as the word complexity is 'large enough' to escape the S-adic structure and partial rigidity, there is already 'enough room' for (strong) mixing of all orders.

---

[*] Vanderbilt University, 1326 Stevenson Center Lane, Nashville TN 37240 `darren.a.creutz@vanderbilt.edu`

2020 *Mathematics Subject Classification.* Primary: 37B10; Secondary 37A25
Key words and phrases. Symbolic dynamics, word complexity, strong mixing, rank-one transformations

Cyr and Kra established that superlinear complexity is the dividing line for a subshift admitting countably many ergodic measures: there exists subshifts with complexity arbitrarily close to linear which admit uncountably many ergodic measures [CK20b] and non-superlinear complexity implies at most countably many [CK19], [Bos85]. Our work implies that in the non-superlinear case, the at most countably many measures are all partially rigid (with a uniform rigidity constant). Their result, like ours, indicates that superlinear word complexity is the line at which complicated measure-theoretic phenomena can manifest.

Beyond the structure imposed by S-adicity, linear complexity subshifts are known to be structured in various ways (e.g. [CFPZ19], [CK20a], [DDMP16], [DOP21], [PS22], [PS23]). Our work indicates there is no hope for similar phenomena in any superlinear setting.

1. Definitions and preliminaries

1.1. Symbolic dynamics

Definition 1.1. A **subshift** on the finite set $\mathcal{A}$ is any subset $X \subset \mathcal{A}^{\mathbb{Z}}$ which is closed in the product topology and shift-invariant: for all $x = (x_n)_{n\in\mathbb{Z}} \in X$ and $k \in \mathbb{Z}$, the translate $(x_{n+k})_{n\in\mathbb{Z}}$ of x by k is also in X.

Definition 1.2. A **word** is any element of $\mathcal{A}^{\ell}$ for some ℓ, the **length** of w, written $\|w\|$. A word w is a **subword** of a word or bi-infinite sequence x if there exists k so that $w_i = x_{i+k}$ for all $1 \leq i \leq \|w\|$. A word u is a **prefix** of w when $u_i = w_i$ for $1 \leq i \leq \|u\|$ and a word v is a suffix of w when $v_i = w_{i+\|w\|-\|v\|}$ for $1 \leq i \leq \|v\|$.

For words v, w, we denote by vw their concatenation–the word obtained by following v immediately by w. We write such concatenations with product or exponential notation, e.g. $\prod_i w_i$ or 0^n.

Definition 1.3. The **language** of a subshift X is $\mathcal{L}(X) = \{w : w \text{ is a subword of some } x \in X\}$.

Definition 1.4. The **word complexity function** of a subshift X over $\mathcal{A}$ is the function $p_X : \mathbb{N} \to \mathbb{N}$ defined by $p_X(q) = |\mathcal{L}(X) \cap \mathcal{A}^q|$, the number of words of length q in the language of X.

When X is clear from context, we suppress the subscript and just write $p(n)$.

For subshifts on the alphabet $\{0, 1\}$, we consider:

Definition 1.5. The set of **right-special** words is $\mathcal{L}^{RS}(X) = \{w \in X : w0, w1 \in \mathcal{L}(X)\}$.

Cassaigne [Cas97] showed the well-known: $p(q) = p(m) + \sum_{\ell=m}^{q-1} |\{w \in \mathcal{L}^{RS} : \|w\| = \ell\}|$ for $m < q$.

1.2. Ergodic theory

Definition 1.6. A **transformation** T is a measurable map on a standard Borel or Lebesgue measure space $(Y, \mathcal{B}, \mu)$ that is measure-preserving: $\mu(T^{-1}B) = \mu(B)$ for all $B \in \mathcal{B}$.

Definition 1.7. Two transformations T on $(Y, \mathcal{B}, \mu)$ and T' on $(Y', \mathcal{B}', \mu')$ are **measure-theoretically isomorphic** when there exists a bijective map ϕ between full measure subsets $Y_0 \subset Y$ and $Y_0' \subset Y'$ where $\mu(\phi^{-1}A) = \mu'(A)$ for all measurable $A \subset Y_0'$ and $(\phi \circ T)(y) = (T' \circ \phi)(y)$ for all $y \in Y_0$.

Definition 1.8. A transformation T is **ergodic** when $A = T^{-1}A$ implies that $\mu(A) = 0$ or $\mu(A^c) = 0$.

Theorem 1.9 (Mean Ergodic Theorem). *If T is ergodic and on a finite measure space and $f \in L^2(Y)$,*

$$\lim_{n\to\infty} \int \Big|\frac{1}{n}\sum_{i=0}^{n-1} f \circ T^{-i} - \int f \, d\mu\Big| \, d\mu = 0$$

Definition 1.10. A transformation T is **mixing** when for all $A, B \in \mathcal{B}$, $\mu(T^nA \cap B) \to \mu(A)\mu(B)$.

1.3. Rank-one transformations

A **rank-one transformation** is a transformation T constructed by "cutting and stacking". Here Y represents a (possibly infinite) interval, $\mathcal{B}$ is the induced σ-algebra from $\mathbb{R}$, and μ is Lebesgue measure. We give a brief description, referring the reader to [FGH$^+$23] or [Sil08] for more details.

The transformation is defined inductively on larger and larger portions of the space through Rohlin towers or **columns**, denoted C_n. Each column C_n consists of **levels** $I_{n,j}$ where $0 \leq j < h_n$ is the height of the level within the column. All levels $I_{n,j}$ in C_n are intervals with the same length, $\mu(I_n)$, and the total number of levels in a column is the **height** of the column, denoted by h_n. The transformation T is defined on all levels $I_{n,j}$ except the top one I_{n,h_n-1} by sending each $I_{n,j}$ to $I_{n,j+1}$ using the unique order-preserving affine map.

Start with $C_1 = [0,1)$ with height $h_1 = 1$. To obtain C_{n+1} from C_n, we require a **cut sequence**, $\{r_n\}$ such that $r_n \geq 1$ for all n. Make r_n vertical cuts of C_n to create $r_n + 1$ **subcolumns** of equal width. Denote a **sublevel** of C_n by $I_{n,j}^{[i]}$ where $0 \leq a < h_n$ is the height of the level within that column, and i represents the position of the subcolumn, where $i = 0$ represents the leftmost subcolumn and $i = r_n$ is the rightmost subcolumn. After cutting C_n into subcolumns, add extra intervals called **spacers** on top of each subcolumn to function as levels of the next column. The **spacer sequence**, $\{s_{n,i}\}$ such that $0 \leq i \leq r_n$ and $s_{n,i} \geq 0$, specifies how many sublevels to add above each subcolumn. Spacers are the same width as the sublevels, act as new levels in the column C_{n+1}, and are taken to be the leftmost intervals in $[1,\infty)$ not in C_n. After the spacers are added, stack the subcolumns with their spacers right on top of left, i.e. so that $I_{n,0}^{[i+1]}$ is directly above $I_{n,h_n-1}^{[i]}$. This gives the next column, C_{n+1}.

Each column C_n defines T on $\bigcup_{j=0}^{h_n-2} I_{n,j}$ and the partially defined map T on C_{n+1} agrees with that of C_n, extending the definition of T to a portion of the top level of C_n where it was previously undefined. Continuing this process gives the sequence of columns $\{C_1, \ldots, C_n, C_{n+1}, \ldots\}$ and T is then the limit of the partially defined maps.

Though this construction could result in Y being an infinite interval with infinite Lebesgue measure, Y has finite measure if and only if $\sum_n \frac{1}{r_n h_n} \sum_{i=0}^{r_n} s_{n,i} < \infty$, see [CS10]. All rank-one transformations we define satisfy this condition, and for convenience we renormalize so that $Y = [0,1)$. Every rank-one transformation is ergodic and invertible.

The reader should be aware that we are making r_n cuts and obtaining $r_n + 1$ subcolumns (following Ferenczi [Fer96]), while other papers (e.g. [Cre21]) use r_n as the number of subcolumns.

1.4. Symbolic models of rank-one transformations

For a rank-one transformation defined as above, we define a subshift $X(T)$ on the alphabet $\{0,1\}$ which is measure-theoretically isomorphic to T:

Definition 1.11. The **symbolic model** $X(T)$ of a rank-one transformation T is given by the sequence of words: $B_1 = 0$ and

$$B_{n+1} = B_n 1^{s_{n,0}} B_n 1^{s_{n,1}} \cdots B_n 1^{s_{n,r_n}} = \prod_{i=0}^{r_n} B_n 1^{s_{n,i}}$$

and $X(T)$ is the set of all bi-infinite sequences such that every subword is a subword of some B_n.

The words B_n are a symbolic coding of the column C_n: 0 represents C_1 and 1 represents the spacers. There is a natural measure associated to $X(T)$:

Definition 1.12. The **empirical measure** for a symbolic model $X(T)$ of a rank-one transformation T is the measure ν defined by, for each word w,

$$\nu([w]) = \lim_{n\to\infty} \frac{|\{1 \leq j \leq \|B_n\| - \|w\| \ : \ B_{n[j,j+\|w\|)} = w\}|}{\|B_n\| - \|w\|}$$

Danilenko [Dan16] (combined with [dJ77] and [Kal84]) proved that the symbolic model $X(T)$ of a rank-one subshift, equipped with its empirical measure, is measure-theoretically isomorphic to the cut-and-stack construction (see [AFP17]; see [FGH$^+$23] for the full generality including odometers).

Due to this isomorphism, we move back and forth between rank-one and symbolic model terminology as needed and write $\mathcal{L}(T)$ for the language of $X(T)$.

2. Quasi-staircase transformations

Definition 2.1. Given nondecreasing sequences of integers $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$ tending to infinity such that $c_1 \geq 1$ and $c_{n+1} \geq c_n + b_n$, a **quasi-staircase transformation** is a rank-one transformation with cut sequence $r_n = a_n b_n$ and spacer sequence $s_{n,t} = c_n + \left\lfloor \frac{t}{a_n} \right\rfloor$ for $0 \leq t < r_n$ and $s_{n,r_n} = 0$.

The symbolic representation of a quasi-staircase is $B_1 = 0$ and

$$B_{n+1} = \Big(\prod_{i=0}^{b_n - 1} \big(B_n 1^{c_n+i}\big)^{a_n} \Big) B_n$$

The height sequence of a quasi-staircase is $h_1 = 1$ and $h_{n+1} = (a_n b_n + 1)h_n + a_n b_n c_n + \frac{1}{2} a_n b_n (b_n - 1)$.

2.1. Quasi-staircase right-special words

Lemma 2.2. *The following hold:*

(i) $1^\ell \in \mathcal{L}^{RS}(T)$ *for all* ℓ.

(ii) *If* w *is a suffix of* $1^{c_n}(B_n 1^{c_n})^{a_n}$ *then* $w \in \mathcal{L}^{RS}(T)$.

(iii) *If* w *is a suffix of* $1^{c_n+i-1}(B_n 1^{c_n+i})^{a_n}$ *for* $0 < i < b_n$ *then* $w \in \mathcal{L}^{RS}(T)$.

(iv) *If* w *is a suffix of* $1^{c_n+b_n-1} B_n 1^{c_n}$ *then* $w \in \mathcal{L}^{RS}(T)$.

Proof. Since suffixes of right-special words are right-special, it suffices to show the words w is claimed to be a suffix of are right-special.

(i): For n such that $\ell < c_n$, as the word $1^{c_n} B_n$ is a subword of B_{n+1}, so are $1^{\ell+1}$ and $1^\ell 0$ since $\ell < c_n$ and B_n starts with 0.

(ii): B_{n+2} has $1^{c_{n+1}} B_{n+1} = 1^{c_{n+1}-c_n} 1^{c_n} B_{n+1}$ as a subword which has $1^{c_n}(B_n 1^{c_n})^{a_n} B_n$ as a subword which gives $1^{c_n}(B_n 1^{c_n})^{a_n} 0 \in \mathcal{L}(T)$. B_{n+1} has $(B_n 1^{c_n})^{a_n} B_n 1^{c_n+1}$ as a prefix which has suffix $1^{c_n}(B_n 1^{c_n})^{a_n-1} B_n 1^{c_n+1}$ and that word is $1^{c_n}(B_n 1^{c_n})^{a_n} 1$ giving $1^{c_n-1}(B_n 1^{c_n})^{a_n} 1 \in \mathcal{L}(T)$.

(iii): B_{n+1} has $1^{c_n+i-1}(B_n 1^{c_n+i})^{a_n} B_n$ as a subword which gives $1^{c_n+i-1}(B_n 1^{c_n+i})^{a_n} 0 \in \mathcal{L}(T)$. When $i < b_n - 1$, B_{n+1} has $(1^{c_n+i} B_n)^{a_n} 1^{c_n+i+1}$ as a subword which gives $1^{c_n+i-1}(B_n 1^{c_n+i})^{a_n} 1 \in \mathcal{L}(T)$; when $i = b_n - 1$, B_{n+2} has the subword $(1^{c_n+b_n-1} B_n)^{a_n} 1^{c_{n+1}}$ so $1^{c_n+b_n-2}(B_n 1^{c_n+b_n-1})^{a_n} 1^{c_{n+1}-c_n-b_n+1} \in \mathcal{L}(T)$ so $1^{c_n+b_n-2}(B_n 1^{c_n+b_n-1})^{a_n} 1 \in \mathcal{L}(T)$ as $c_{n+1} \geq c_n + b_n$.

(iv): B_{n+2} has $B_{n+1} 1^{c_{n+1}} B_{n+1}$ as a subword which has $B_{n+1} 1^{c_{n+1}} B_n 1^{c_n} B_n$ as a prefix, and that word has $1^{c_n+b_n-1} B_n 1^{c_n} 0$ as a subword since $c_n + b_n - 1 < c_{n+1}$. Also B_{n+2} has $B_{n+1} 1^{c_{n+1}}$ as a subword which has $1^{c_n+b_n-1} B_n 1^{c_{n+1}}$ as a suffix which then has $1^{c_n+b_n-1} B_n 1^{c_n} 1$ as a subword. □

Lemma 2.3. *Let* $01^z0 \in \mathcal{L}(T)$. *Then there are unique* n *and* i *with* $0 \leq i < b_n$ *such that* $z = c_n + i$. $01^{c_n+i}0$ *is not a subword of* B_m *for* $m \leq n$ *and for every* $x \in X(T)$ *and every occurrence of* $01^{c_n+1}0$ *in* x, $01^{c_n+i}0$ *occurs as a suffix of* $1^{c_{n+1}}(\prod_{j=0}^{i-1}(B_n 1^{c_n+j})^{a_n})(B_n 1^{c_n+i})^q 0$ *for some* $1 \leq q \leq a_n$ *(adopting the convention that* $\prod_0^{-1}$ *is the empty word).*

Proof. As every B_n begins and ends with 0, the only such words are of the form $01^{c_n+i}0$. Since $c_{n+1} \geq c_n + b_n$, such n and i are unique. This also gives that 1^{c_n} is not a subword of B_n.

The word $01^{c_n+i}0$ only occurs inside B_{n+1} due to $c_{n+1} \geq c_n + b_n$, and only as part of the $(B_n 1^{c_n+i})^{a_n}$ in its construction, and B_{n+1} is always preceded by $1^{c_{n+1}}$ □

Proposition 2.4. *If $w \in \mathcal{L}^{RS}(T)$ then at least one of the following holds:*

(i) $w = 1^{\|w\|}$

(ii) *w is a suffix of $1^{c_n+i-1}\big(B_n 1^{c_n+i}\big)^{a_n}$ for some n and $0 \leq i < b_n$*

(iii) *w is a suffix of $1^{c_n+b_n-1} B_n 1^{c_n}$ for some n*

(iv) $w = 1^{c_n}(B_n 1^{c_n})^{a_n}$

Proof. Let $w \in \mathcal{L}^{RS}(T)$. Since $c_1 \geq 1$, the word $00 \notin \mathcal{L}(T)$ so w does not end in 0. If $w = 1^{\|w\|}$ then w is of form (i) so from here on, assume that w contains at least one 0.

Let $z \geq 1$ such that w has 01^z as a suffix. Then $w0$ has 01^z0 as a suffix so $z = c_n + i$ for some unique $n \geq 1$ and $0 \leq i < b_n$ by Lemma 2.3. As $w0$ has $01^{c_n+i}0$ as a suffix, $w0$ shares a suffix with the word $1^{c_{n+1}}(\prod_{j=0}^{i-1}(B_n 1^{c_n+j})^{a_n})(B_n 1^{c_n+i})^q 0$ for some $1 \leq q \leq a_n$.

First consider the case when $i > 0$. If w is a suffix of $1^{c_n+i-1}(B_n 1^{c_n+i})^{a_n}$ then it is of form (ii) so we need only consider w that have $01^{c_n+i-1}(B_n 1^{c_n+i})^q$ as a suffix. For such w, the word $w1$ has the suffix $01^{c_n+i-1}(B_n 1^{c_n+i})^{q-1} B_n 1^{c_n+i+1}$ but that word is only in $\mathcal{L}(T)$ if $q - 1 = a_n$ which is impossible.

Now consider the case when $i = 0$, i.e. $z = c_n$. If w is a suffix of $1^{c_n-1}(B_n 1^{c_n})^{a_n}$ then it is of form (ii) so we may assume that w has $1^{c_n-1}(B_n 1^{c_n})^q$ as a strict suffix for some $1 \leq q \leq a_n$. Since $B_n 1^{c_n}$ is always preceded by 1^{c_n} (possibly as part of some $1^{c_{n+1}+i}$ or 1^{c_n+i}), w cannot have $01^{c_n-1}B_n 1^{c_n}$ as a subword so w has $1^{c_n}(B_n 1^{c_n})^q$ as a suffix for some $1 \leq q \leq a_n$.

Take q maximal so that w has $1^{c_n}(B_n 1^{c_n})^q$ as a suffix.

Consider first when w has $1^{c_n}(B_n 1^{c_n})^{a_n}$ as a suffix, i.e. when $q = a_n$. If $w = 1^{c_n}(B_n 1^{c_n})^{a_n}$ then it is of form (iv). If w has $01^{c_n}(B_n 1^{c_n})^{a_n}$ as a suffix then $w0 \notin \mathcal{L}(T)$ as $0(1^{c_n}B_n)^{a_n}1^{c_n}0 \notin \mathcal{L}(T)$. If w has $11^{c_n}(B_n 1^{c_n})^{a_n}$ as a suffix then $w1$ has $1^{c_n+1}(B_n 1^{c_n})^{a_n-1}B_n 1^{c_n+1}$ as a suffix but that is not in $\mathcal{L}(T)$.

So we may assume $q < a_n$. Since $1^{c_n}(B_n 1^{c_n})^q$ is then of form (ii), we may assume $1^{c_n}(B_n 1^{c_n})^q$ is a strict suffix of w.

Consider when w has $01^{c_n}(B_n 1^{c_n})^q$ as a suffix. As $01^{c_n}(B_n 1^{c_n})^q$ only appears as a suffix of $B_n 1^{c_n}(B_n 1^{c_n})^q$ and that word is always preceded by 1^{c_n} (possibly as part of some $1^{c_{n+1}+i}$), w then shares a suffix with $1^{c_n}(B_n 1^{c_n})^{q+1}$. As q is maximal, then w is a suffix of $1^{c_n-1}(B_n 1^{c_n})^{q+1}$ and, as $q < a_n$, this means w is of form (ii).

We are left with the case when w has $1^{c_n+1}(B_n 1^{c_n})^q$ as a suffix for some $1 \leq q < a_n$. If $q \geq 2$ then $w1$ has $1^{c_n+1}(B_n 1^{c_n})^{q-1}B_n 1^{c_n+1}$ as a suffix but that is not in $\mathcal{L}(T)$ for $q - 1 \geq 1$. So we are left with the situation when w shares a suffix with $1^{c_n+1}B_n 1^{c_n}$. So $w0$ shares a suffix with $1^{c_n+1}B_n 1^{c_n}0$ which must share a suffix with $1^{c_{n+1}}B_n 1^{c_n}0$, meaning that w shares a suffix with $1^{c_{n+1}}B_n 1^{c_n}$. If w is a suffix of $1^{c_n+b_n-1}B_n 1^{c_n}$ then it is of form (iii). If not then w has the suffix $1^{c_n+b_n}B_n 1^{c_n}$ so $w1$ has suffix $1^{c_n+b_n}B_n 1^{c_n+1}$ which is not in $\mathcal{L}(T)$ since $B_n 1^{c_n+1}$ is always preceded by $B_n 1^{c_n}$ or $B_n 1^{c_n+1}$. □

2.2. The level-n complexity functions

Definition 2.5. For a word w, define the **tail length** $z(w)$ such that $w = u01^{z(w)}$ for some (possibly empty) word u with the conventions that $z(1^{\|w\|}) = \infty$ and $z(u0) = 0$.

Definition 2.6. For $1 \leq n < \infty$, the set of **level-n generating words** is

$$W_n = \{w \in \mathcal{L}^{RS}(T) : c_n \leq z(w) < c_{n+1}\}$$

Proposition 2.7. $\mathcal{L}^{RS}(T) = \{1^\ell : \ell \in \mathbb{N}\} \sqcup \bigsqcup_{n=1}^{\infty} W_n$.

Proof. $\{c_n\}$ is strictly increasing so the W_n are disjoint. Lemma 2.2 (i) says $1^\ell \in \mathcal{L}^{RS}(T)$ for all ℓ and as every word in W_n has 0 as a subword, these are disjoint from the W_n. If $z(w) < c_1$ then $w0 \notin \mathcal{L}(T)$ by Lemma 2.3 so all right-special words with 0 as a subword are in some W_n. $\square$

Definition 2.8. The **level-n complexity** is $p_n(q) = |\{w \in W_n : \|w\| < q\}|$.

By definition, $p_n(\ell+1) - p_n(\ell) = |\{w \in W_n : \|w\| = \ell\}|$.

Proposition 2.9. *The complexity function p satisfies $p(q) = 1 + q + \sum_{n=1}^{\infty} p_n(q)$.*

Proof. Using Proposition 2.7 and that $p(\ell+1) - p(\ell) = |\{w \in \mathcal{L}^{RS} : \|w\| = \ell\}|$,

$$\begin{aligned} p(q) - p(1) &= \sum_{\ell=1}^{q-1} (p(\ell+1) - p(\ell)) = \sum_{\ell=1}^{q-1} |\{w \in \mathcal{L}^{RS}(T) : \|w\| = \ell\}| \\ &= \sum_{\ell=1}^{q-1} \Big(\sum_{n=1}^{\infty} |\{w \in W_n : \|w\| = \ell\}| + |\{1^{\ell}\}| \Big) = \sum_{\ell=1}^{q-1} \Big(\sum_{n=1}^{\infty} (p_n(\ell+1) - p_n(\ell)) + 1 \Big) \\ &= \sum_{n=1}^{\infty} \Big(\sum_{\ell=1}^{q-1} (p_n(\ell+1) - p_n(\ell)) \Big) + q - 1 = \sum_{n=1}^{\infty} (p_n(q) - p_n(1)) + q - 1 \end{aligned}$$

All words in W_n have length at least $1 + c_n > 1$ so $p_n(1) = 0$. The claim follows as $p(1) = 2$. $\square$

2.3. Counting quasi-staircase words

Lemma 2.10. *If $w \in W_n$ then exactly one of the following holds:*

(i) w *is a suffix of* $1^{c_n+i-1}\big(B_n 1^{c_n+i}\big)^{a_n}$ *and* $\|w\| > c_n + i$ *for some* $0 \le i < b_n$*;*

(ii) w *is a suffix of* $1^{c_n+b_n-1} B_n 1^{c_n}$ *and* $\|w\| > h_n + 2c_n$*; or*

(iii) $w = 1^{c_n}(B_n 1^{c_n})^{a_n}$

Proof. The only words in Proposition 2.4 which have $c_n \le z(w) < c_{n+1}$ are of the stated forms; Lemma 2.2 (ii), (iii) and (iv) state that these words are in $\mathcal{L}^{RS}(T)$. The forms do not overlap due to the restriction on $\|w\|$ in form (*ii*). $\square$

Lemma 2.11. *Fix $0 \le i < b_n$. For $c_n + i < \ell < a_n h_n + (a_n+1)(c_n+i)$ there is exactly one word in W_n of form* (*i*) *for that value of i; for ℓ not in that range, there are no words in W_n of form* (*i*) *for that i.*

Proof. For $w \in W_n$ of form (*i*), $w = u1^{c_n+i}$ where u is a nonempty suffix of $1^{c_n+i-1}(B_n 1^{c_n+i})^{a_n-1} B_n$. The word u is unique if it exists which is exactly when $c_n + i = \|1^{c_n+i}\| < \|w\| \le \|1^{c_n+i-1}(B_n 1^{c_n+i})^{a_n}\| = a_n h_n + (a_n+1)(c_n+i) - 1$. $\square$

Lemma 2.12. *For $h_n + 2c_n < \ell < h_n + 2c_n + b_n$ there is exactly one word in W_n of form* (*ii*)*; for ℓ not in that range, there are no words in W_n of form* (*ii*).

Proof. To be of that form, $w = u1^{c_n}$ where u is a nonempty suffix of $1^{c_n+b_n-1}B_n$ that has 1^{c_n+1} as a prefix. The word u is unique if it exists and it exists exactly when $h_n + 2c_n + 1 = \|1^{c_n+1}B_n 1^{c_n}\| \le \|w\| \le \|1^{c_n+b_n-1}B_n 1^{c_n}\| = h_n + 2c_n + b_n - 1$. $\square$

Lemma 2.13. *If $\ell \le c_n$ then $p_n(\ell+1) - p_n(\ell) = 0$.*

Proof. Every $w \in W_n$ has subwords 1^{c_n} and 0 so $\|w\| \ge c_n + 1$ therefore $p_{n+1}(\ell) = p_n(\ell) = 0$. $\square$

Lemma 2.14. *If $c_n < \ell < c_n + b_n$ then $p_n(\ell+1) - p_n(\ell) = \ell - c_n$.*

Proof. Lemma 2.11 applies for $0 \le i < \ell - c_n$ but not for $\ell - c_n \le i < b_n$. Lemma 2.12 does not apply. $\square$

Lemma 2.15. *If* $c_n + b_n \leq \ell \leq h_n + 2c_n$ *then* $p_n(\ell+1) - p_n(\ell) = b_n$.

Proof. Lemma 2.11 applies for all $0 \leq i < b_n$ and Lemma 2.12 does not apply. ☐

Lemma 2.16. *If* $h_n + 2c_n < \ell < h_n + 2c_n + b_n$ *then* $p_n(\ell+1) - p_n(\ell) = b_n + 1$.

Proof. Lemma 2.11 applies for all $0 \leq i < b_n$ and Lemma 2.12 applies. ☐

Lemma 2.17. *If* $h_n + 2c_n + b_n \leq \ell < a_n h_n + (a_n+1)c_n$ *then* $p_n(\ell+1) - p_n(\ell) = b_n$.

Proof. Lemma 2.11 applies for all $0 \leq i < b_n$ and Lemma 2.12 does not apply. ☐

Lemma 2.18. $p_n(a_n h_n + (a_n+1)c_n + 1) - p_n(a_n h_n + (a_n+1)c_n) = b_n + 1$.

Proof. Lemma 2.11 applies for all $0 \leq i < b_n$ and Lemma 2.12 does not apply. Lemma 2.10 form (*iii*) gives one additional word in W_n. ☐

Lemma 2.19. *If* $a_n h_n + (a_n+1)c_n + 1 < \ell < a_n h_n + (a_n+1)(c_n + b_n - 1)$ *then* $p_n(\ell+1) - p_n(\ell) \leq b_n$.

Proof. Lemma 2.11 applies for some subset of $0 \leq i < b_n$ and Lemma 2.12 does not apply. ☐

Lemma 2.20. $p(a_n h_n + (a_n+1)c_n + (a_n+1)(b_n-1)) - p(a_n h_n + (a_n+1)c_n) = \frac{1}{2}(a_n+1)b_n(b_n-1) + 1$.

Proof. For each $0 \leq i < b_n$, Lemma 2.11 applies for $\ell = a_n h_n + (a_n+1)c_n + y$ exactly when $0 \leq y < (a_n+1)i$, therefore there are a total of $(a_n+1)\frac{1}{2}b_n(b_n-1)$ words in W_n of the enclosed lengths from Lemma 2.11. Lemma 2.12 does not apply and Lemma 2.10 form (*iii*) gives one additional word. ☐

Lemma 2.21. *If* $a_n h_n + (a_n+1)(c_n + b_n - 1) \leq \ell$ *then* $p_n(\ell+1) - p_n(\ell) = 0$.

Proof. Neither Lemma 2.11 nor 2.12 apply. ☐

2.4. Bounding the complexity of quasi-staircases

Since $p_n(\ell+1) - p_n(\ell) = 0$ for $\ell \geq a_n h_n + (a_n+1)(c_n + b_n - 1)$, we define:

Definition 2.22. The **post-productive sequence** is

$$m_n = a_n h_n + (a_n+1)(c_n + b_n - 1) \qquad m_0 = 0$$

Lemma 2.23. $p_n(m_n) = h_{n+1} - h_n$

Proof. By Lemma 2.13, $p_n(c_n) = \sum_{\ell=0}^{c_n-1}(p_n(\ell+1) - p_n(\ell)) = 0$.
By Lemma 2.14, $p_n(c_n + b_n) - p_n(c_n) = \sum_{\ell=c_n}^{c_n+b_n-1}(\ell - c_n) = \frac{1}{2}b_n(b_n-1)$.
By Lemma 2.15, $p_n(h_n + 2c_n + 1) - p_n(c_n + b_n) = (h_n + c_n + 1 - b_n)b_n$.
By Lemma 2.16, $p_n(h_n + 2c_n + b_n) - p_n(h_n + 2c_n + 1) = (b_n+1)(b_n-1)$.
By Lemma 2.17, $p_n(a_n h_n + (a_n+1)c_n) - p_n(h_n + 2c_n + b_n) = ((a_n-1)h_n + (a_n-1)c_n - b_n)b_n$.
By Lemma 2.20, $p_n(m_n) - p(a_n h_n + (a_n+1)c_n) = \frac{1}{2}(a_n+1)b_n(b_n-1) + 1$. Therefore

$$\begin{aligned} p_n(m_n) &= \frac{1}{2}b_n(b_n-1) + (h_n + c_n + 1 - b_n)b_n + (b_n+1)(b_n-1) \\ &\quad + ((a_n-1)h_n + (a_n-1)c_n - b_n)b_n + \frac{1}{2}(a_n+1)b_n(b_n-1) + 1 \\ &= a_n b_n h_n + a_n b_n c_n + \frac{1}{2}a_n b_n(b_n-1) + b_n(b_n-1) + b_n - b_n^2 + b_n^2 - 1 - b_n^2 + 1 = h_{n+1} - h_n \end{aligned}$$ ☐

Definition 2.24. For $q \in \mathbb{N}$ define

$$\rho(q) = \max\{n : m_n \leq q\} \qquad \text{and} \qquad \beta(q) = \min\{n : q < c_{n+1}\}$$

Lemma 2.25. $\rho(q) \leq \beta(q)$

Proof. If $\beta(q) \leq \rho(q) - 1$ then $m_{\rho(q)} \leq q < c_{\beta(q)+1} \leq c_{\rho(q)-1+1} = c_{\rho(q)} < m_{\rho(q)}$ is impossible. □

Lemma 2.26. *If $q < c_n$ then $p_n(q) = 0$. If $c_n \leq q < m_n$ then $p_n(q) \leq (q - c_n + 1)b_n$. If $m_n \leq q$ then $p_n(q) = h_{n+1} - h_n$.*

Proof. Lemma 2.13 gives $p_n(\ell+1) - p_n(\ell) = 0$ for $0 \leq \ell < c_n$. Lemmas 2.14, 2.15, 2.16, 2.17 and 2.19 all give $p_n(\ell+1) - p_n(\ell) \leq b_n$ for $c_n \leq \ell < m_n$ except for Lemma 2.16 which gives $p_n(\ell+1) - p_n(\ell) = b_n + 1$ for exactly $b_n - 1$ values of ℓ and Lemma 2.18 which gives one additional word. Then, for $c_n \leq q < m_n$,

$$p_n(q) = \sum_{\ell=0}^{q-1}(p_n(\ell+1) - p_n(\ell)) = \sum_{\ell=0}^{c_n-1} 0 + \sum_{\ell=c_n}^{q-1}(p_n(\ell+1) - p_n(\ell)) \leq (q - c_n)b_n + b_n$$

Lemma 2.21 says $p_n(\ell+1) - p_n(\ell) = 0$ for $\ell \geq m_n$ so when $q \geq m_n$, $p_n(q) = p_n(m_n)$ and Lemma 2.23 gives the final statement. □

Proposition 2.27. $p(q) \leq q\Big(2 + \sum_{n=\rho(q)}^{\beta(q)} b_n\Big)$ *for all q.*

Proof. For n such that $\beta(q) < n$, by Lemma 2.13, $p_n(q) = 0$. Proposition 2.9 and Lemma 2.26 give, using that $h_1 = 1$ so $1 + \sum_{n=1}^{\rho(q)}(h_{n+1} - h_n) = h_{\rho(q)+1}$,

$$\begin{aligned} p(q) &= q + 1 + \sum_{n=1}^{\rho(q)} p_n(q) + \sum_{n=\rho(q)+1}^{\beta(q)} p_n(q) + \sum_{n=\beta(q)+1}^{\infty} p_n(q) \\ &\leq q + 1 + \sum_{n=1}^{\rho(q)}(h_{n+1} - h_n) + \sum_{n=\rho(q)+1}^{\beta(q)}(q - c_n + 1)b_n + 0 \leq q + h_{\rho(q)+1} + \sum_{n=\rho(q)+1}^{\beta(q)} qb_n \end{aligned}$$

and

$$\begin{aligned} h_{\rho(q)+1} &= h_{\rho(q)} + b_{\rho(q)}(a_{\rho(q)}h_{\rho(q)} + a_{\rho(q)}c_{\rho(q)} + \frac{1}{2}a_{\rho(q)}(b_{\rho(q)} - 1)) \\ &\leq h_{\rho(q)} + b_{\rho(q)}m_{\rho(q)} \leq m_{\rho(q)}(1 + b_{\rho(q)}) \leq q(1 + b_{\rho(q)}) \end{aligned}$$

□

3. Quasi-staircase complexity arbitrarily close to linear

Lemma 3.1. *Let $\{d_n\}$ be a nondecreasing sequence of integers such that $d_n \to \infty$ and $d_1 = d_2 = 1$ and $d_{n+1} - d_n \in \{0, 1\}$ and $d_{n+1} - d_n$ does not take the value 1 for consecutive n.*

Let $\{b_n\}$ be a nondecreasing sequence of integers such that $b_n \to \infty$ and $b_1 = 3$ and $b_n \leq n + 2$.

Set $a_n = 2n^2 + 2$. Set $c_1 = 1$ and for $n > 1$,

$$c_n = \begin{cases} m_{n-d_n} & \text{when } d_n = d_{n-1} \\ c_{n-1} + b_{n-1} & \text{when } d_n = d_{n-1} + 1 \end{cases}$$

Then $\{a_n\}, \{b_n\}, \{c_n\}$ define a quasi-staircase such that $\sum \frac{a_n b_n^2 + a_{n+1}b_{n+1} + c_{n+1}}{h_n} < \infty$ and $\sum \frac{1}{a_n b_n} < \infty$.

Proof. Since $r_n = a_n b_n$, we have $6n^2 + 6 \leq r_n \leq (2n^2 + 2)(n + 2)$. Then $\prod_{j=1}^{n-1}(r_j + 1) \geq n!$ so $h_n \geq \prod_{j=1}^{n-1}(r_j + 1) \geq n!$ so $\sum \frac{a_n b_n^2 + a_{n+1}b_{n+1}}{h_n} \leq \sum \frac{(2n^2+2)(n+2)^2 + (2(n+1)^2+2)(n+3)}{n!} < \infty$.

Since $a_n + 1 < \frac{3}{2}a_n \leq \frac{1}{2}b_n a_n$, then $m_n < h_{n+1}$ for all n.

For n such that $c_{n+1} = c_n + b_n$, since $d_{n+1} = d_n + 1$, also $d_n = d_{n-1}$. So for sufficiently large such n,

$$\frac{c_{n+1}}{h_n} = \frac{c_n + b_n}{h_n} = \frac{m_{n-d_n} + b_n}{h_n} < \frac{2h_{n-d_n+1}}{h_n} < \frac{2}{\prod_{i=1}^{d_n-1} r_{n-d_n+i}} < \frac{2}{r_{n-1}} < \frac{2}{2(n-1)^2+2}$$

and for sufficiently large n such that $c_{n+1} = m_{n+1-d_n}$,

$$\frac{c_{n+1}}{h_n} = \frac{m_{n+1-d_n}}{h_n} < \frac{h_{n-d_n+2}}{h_n} < \frac{1}{r_{n-1}} < \frac{1}{2(n-1)^2+2}$$

Therefore, as $\sum \frac{2}{2(n-1)^2+2} < \infty$, it follows that $\sum \frac{c_{n+1}}{h_n} < \infty$ and the result follows. ☐

Lemma 3.2. *If $f : \mathbb{N} \to \mathbb{N}$ is any function such that $f(q) \to \infty$ then there exists $g : \mathbb{N} \to \mathbb{N}$ which is nondecreasing such that $g(1) = 1$ and $g(q) \leq f(q)$ and $g(q+2) - g(q) \leq 1$ for all q and $g(q) \to \infty$.*

Proof. Set $f^*(q) = \inf_{q' \geq q} f(q')$. Then $f^*(q) \to \infty$ and $f^*(q)$ is nondecreasing and $f^*(q) \leq f(q)$ for all q. Set $g(1) = 1 \leq f^*(1)$. For $n \geq 0$, set $g(2n+2) = g(2n+1)$ and for $n \geq 1$ set

$$g(2n+1) = g(2n) + \begin{cases} 1 & \text{when } f^*(2n+1) > f^*(2n-1) \\ 0 & \text{otherwise} \end{cases}$$

Then g is nondecreasing and $g(q+2) - g(q) \leq 1$ for all q. Since f^* is integer-valued, if $f^*(2n+1) - f^*(2n-1) \neq 0$ then $f^*(2n+1) - f^*(2n-1) \geq 1$. Then $g(2n+1) - g(2n-1) \leq f^*(2n+1) - f^*(2n-1)$ so for all n we have

$$g(2n+1) = g(1) + \sum_{m=1}^{n} (g(2m+1) - g(2m-1)) \leq f^*(1) + \sum_{m=1}^{n} (f^*(2m+1) - f^*(2m-1)) = f^*(2n+1)$$

so, as $g(2n+2) = g(2n+1) \leq f^*(2n+1) \leq f^*(2n+2)$, we have $g(q) \leq f^*(q) \leq f(q)$ for all q. If $g(q) \leq C$ for all q then $f^*(2n+1) = f^*(2n-1)$ eventually, contradicting that $f^*(q) \to \infty$. Therefore $g(q) \to \infty$. ☐

Theorem 3.3. *Let $f : \mathbb{N} \to \mathbb{N}$ be any function such that $f(q) \to \infty$. There exists a quasi-staircase transformation with $\sum \frac{a_n b_n^2 + a_{n+1} b_{n+1} + c_{n+1}}{h_n} < \infty$, $\sum \frac{1}{a_n b_n} < \infty$, $\frac{b_n}{a_n} \to 0$ and complexity satisfying $\frac{p(q)}{q f(q)} \to 0$.*

Proof. By Lemma 3.2, we may assume f is nondecreasing and that $f(n+2) - f(n) \leq 1$ for all n. Then $f(n+1) - f(n) \in \{0, 1\}$ and is never 1 for two consecutive values. We may also assume $f(1) = 1$.

Set $d_1 = d_2 = 1$ and $d_n = \left\lfloor \sqrt[3]{f(n)} \right\rfloor$ for $n > 2$. Then $d_n \to \infty$ is nondecreasing. Also $d_{n+1} - d_n \in \{0, 1\}$ and is never 1 for two consecutive values.

Set $b_n = 3$ for all n such that $\sqrt[3]{f(n)} < 3$ and $b_n = \left\lfloor \sqrt[3]{f(n)} \right\rfloor$ for n such that $\sqrt[3]{f(n)} \geq 3$. Then $b_n \to \infty$ is nondecreasing and $b_n \leq f(n) + 2 \leq n + 2$ as $f(n) \leq n$ since $f(1) = 1$ and $f(n+2) - f(n) \leq 1$ imply $f(n) \leq 1 + \frac{n}{2}$.

Take the quasi-staircase transformation from Lemma 3.1 with defining sequences $\{a_n\}$ and $\{c_n\}$. As $a_n = 2n^2 + 2$ and $b_n = \max(3, \sqrt[3]{f(n)}) \leq \sqrt[3]{n}$, we have $\frac{b_n}{a_n} \to 0$.

Since $0 \leq d_{n+1} - d_n \leq 1$, the sequence $n - d_n$ is nondecreasing and attains every value in $\mathbb{N}$. For each q, let n_q be the largest n such that $m_{n-d_n} \leq q$. Then $q < m_{n_q+1-d_{n_q+1}}$ so $n_q + 1 - d_{n_q+1} > n_q - d_{n_q}$ and so $1 > d_{n_q+1} - d_{n_q}$ meaning that $d_{n_q+1} = d_{n_q}$. Therefore $c_{n_q+1} = m_{n_q+1-d_{n_q+1}} = m_{n_q-d_{n_q}+1}$.

So $\rho(q) = n_q - d_{n_q}$ as $m_{n_q-d_{n_q}} \leq q < m_{n_q+1-d_{n_q+1}} = m_{n_q-d_{n_q}+1}$ and $\beta(q) \leq n_q$ since $q < m_{n_q-d_{n_q}+1} = c_{n_q+1}$. By Proposition 2.27, since $q \geq n_q$ and f is nondecreasing to infinity and $n_q \to \infty$,

$$\begin{aligned}
\frac{p(q)}{q f(q)} &\leq \frac{2 + \sum_{n=\rho(q)}^{\beta(q)} b_n}{f(q)} \leq \frac{2 + \sum_{n=n_q-d_{n_q}}^{n_q} b_n}{f(q)} \leq \frac{2 + (d_{n_q}+1) b_{n_q}}{f(n_q)} \\
&\leq \frac{2 + (\sqrt[3]{f(n_q)} + 1)\sqrt[3]{f(n_q)}}{f(n_q)} = \frac{2}{f(n_q)} + \frac{1}{\sqrt[3]{f(n_q)}} + \frac{1}{(\sqrt[3]{f(n_q)})^2} \to 0
\end{aligned}$$

☐

4. Mixing for quasi-staircase transformations

The goal of this section is to prove the following.

Theorem 4.1. *Let T be a quasi-staircase transformation such that $\sum \frac{a_n b_n + b_{n+1} + c_{n+1}}{h_n} < \infty$ and $\sum \frac{1}{a_n b_n} < \infty$ and $\frac{a_n b_n^2}{h_n} \to 0$ and $\frac{a_{n+1} b_{n+1}}{h_n} \to 0$ and $\frac{b_n}{a_n} \to 0$. Then T is mixing.*

Throughout this section, we assume that all transformations are on probability spaces. Recall that $b_n \to \infty$ by definition for quasi-staircase transformations.

We first introduce some notation.

Notation 4.2. *For measurable sets A and B, write*

$$\lambda_B(A) = \mu(A \cap B) - \mu(A)\mu(B)$$

So $\{t_n\}$ is mixing when $\lambda_B(T^{t_n}A) \to 0$ for all measurable A and B. The following is left to the reader:

Lemma 4.3. *If A and A' are disjoint then*

$$\lambda_B(A \sqcup A') = \lambda_B(A) + \lambda_B(A') \qquad \text{and} \qquad |\lambda_B(A)| \leq \mu(A)$$

and, writing $\chi_B(x) = \mathbb{1}_B(x) - \mu(B)$, for $n \in \mathbb{Z}$, $\lambda_B(T^n A) = \int_A \chi_B \circ T^n d\mu$.

For a rank-one transformation T, a sequence $\{t_n\}$ is **rank-one uniform mixing** when for every union of levels B, $\sum_{j=0}^{h_n-1} |\lambda_B(T^{t_n} I_{n,j})| \to 0$. Rank-one uniform mixing for a sequence implies mixing for that sequence [CS04] Proposition 5.6.

Notation 4.4. *For $h_n \leq j < h_n + c_n$, let $I_{n,j} = T^{j-h_n+1} I_{n,h_n-1}$ be the union of the $(j-h_n)^{th}$ stage of the c_n spacer levels added above every subcolumn. Write*

$$\tilde{h}_n = h_n + c_n$$

Lemma 4.5. *Let T be a rank-one transformation, B a union of levels in some column C_N and $n \geq N$. Then for any $0 \leq j < \tilde{h}_n$ and $0 \leq i \leq r_n$,*

$$\lambda_B(I_{n,j}^{[i]}) = \frac{1}{r_n + 1} \lambda_B(I_{n,j})$$

Proof. Since B is a union of levels in C_N, either $I_{n,j} \subseteq B$ or $I_{n,j} \cap B = \emptyset$. If $I_{n,j} \subseteq B$ then $\mu(I_{n,j}^{[i]} \cap B) = \mu(I_{n,j}^{[i]}) = \frac{1}{r_n+1}\mu(I_{n,j}) = \frac{1}{r_n+1}\mu(I_{n,j} \cap B)$ and if $I_{n,j} \cap B = \emptyset$ then $\mu(I_{n,j}^{[i]} \cap B) = 0 = \frac{1}{r_n+1}\mu(I_{n,j} \cap B)$. □

Proposition 4.6. *Let T be a quasi-staircase transformation given by $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$ with height sequence $\{h_n\}$. Then T is on a finite measure space if and only if $\sum \frac{c_n + b_n}{h_n} < \infty$.*

Proof. Writing S_n for the spacers added above the n^{th} column,

$$\mu(S_n) = (c_n r_n + \frac{1}{2} r_n (b_n - 1))\mu(I_{n+1}) = \left(c_n \frac{r_n}{r_n + 1} + \frac{1}{2} \frac{r_n (b_n - 1)}{r_n + 1} \right) \mu(I_n) \leq \frac{c_n + b_n}{h_n} \mu(C_n)$$

and therefore $\mu(C_{n+1}) = \mu(C_n) + \mu(S_n) \leq \left(1 + \frac{c_n + b_n}{h_n}\right)\mu(C_n)$. Then $\mu(C_{n+1}) \leq \prod_{j=1}^{n} \left(1 + \frac{c_j + b_j}{h_j}\right)\mu(C_1)$, so the claim follows from [Kno54] p.219 that $\prod_{j=1}^{\infty}(1 + \frac{c_j + b_j}{h_j}) < \infty$ if and only if $\sum_{j=1}^{\infty} \frac{c_j + b_j}{h_j} < \infty$. □

4.1. Mixing along most sequences

For clarity of exposition, we state the results which follow from now-standard techniques for proving mixing on staircases with explanations of how one could modify the corresponding proofs in the literature to our class of transformations. Detailed proofs, including all necessary modifications, are deferred to the appendix.

4.1.1. (Weak) power ergodicity

Proposition 4.7. *Let T be a quasi-staircase transformation and B a measurable set. Then*

$$\max_{1\leq k\leq n} \int \left| \frac{1}{n} \sum_{j=0}^{n-1} \chi_B \circ T^{-jk} \right| \, d\mu \to 0$$

The proof of Proposition 4.7 is essentially identical to the proof of (weak) power ergodicity for staircases except that one must replace the height sequence by $\{h_n + c_n\}$ and then observe that for a union of levels B in C_n, any fixed positive integer k, and a level I in C_n at least kb_n above the base,

$$\mu(T^{k(h_n+c_n)} I \cap B) \approx \frac{1}{b_n} \sum_{i=0}^{b_n-1} \mu(T^{-ik} I \cap B) \pm \frac{2k}{r_n} \mu(I)$$

which follows from the standard technique that $\mu(T^{-ik} I^{[i]} \cap B) = \frac{1}{r_n} \mu(T^{-ik} I \cap B)$ provided that I is at least ik levels above the base of the tower. From there, deducing the proposition is identical, modulo the obvious replacement of r_n by b_n throughout the proof, to the proof of (weak) power ergodicity for elevated staircases in [CPR23].

4.1.2. Mixing between $a_n\tilde{h}_n$ and $\tilde{h}_{n+1}$

Proposition 4.8. *Let T be a quasi-staircase transformation such that $\frac{a_n b_n^2}{h_n} \to 0$ and $\frac{b_n}{a_n} \to 0$ and B be a union of levels in some fixed C_N. For $n > N$, set*

$$M_{B,n} := \max_{a_n\tilde{h}_n \leq t < \tilde{h}_{n+1}} \sum_{j=0}^{h_n-1} |\lambda_B(T^t I_{n,j})|$$

Then $\lim_{n\to\infty} M_{B,n} = 0$.

The proof of Proposition 4.8 follows from the standard argument for proving mixing on (elevated) staircases, see e.g. [CPR23]. Consider $T^{a_n\tilde{h}_n}$ applied to a level $I_{n,j}$ in the n^{th} column. Since every sublevel is pushed through at least one spacer, $T^{a_n\tilde{h}_n} I_{n,j}$ consists of a_n sublevels in $T^{-i} I_{n,j}$ for each $1 \leq i < b_n$ plus a_n sublevels which are pushed through the top of the next column. Since $b_n \to \infty$, the convergence of the ergodic average $\frac{1}{b_n} \sum_{i=0}^{b_n-1} T^{-i}$ implies mixing along this sequence.

For the sequence $\{k_n a_n \tilde{h}_n\}$ where $1 \leq k_n < b_n$, the resulting average $\frac{1}{b_n} \sum_{i=0}^{b_n-1} T^{-ik_n}$ converges by (weak) power ergodicity so these sequences are likewise mixing. The general case of times between $a_n\tilde{h}_n$ and $\tilde{h}_{n+1}$ then follows from the standard interpolation argument and Blum-Hanson trick combined with the Block Lemma.

4.1.3. Mixing between $\tilde{h}_n$ and $b_n\tilde{h}_n$

Proposition 4.9. *Let T be a quasi-staircase transformation with $\frac{b_n^2}{h_n} \to 0$ and $\frac{b_n}{a_n} \to 0$ and B be a union of levels in some fixed C_N. For $n > N$, set*

$$\widehat{M}_{B,n} := \max_{\tilde{h}_n \leq t < b_n\tilde{h}_n} \sum_{j=0}^{h_n-1} |\lambda_B(T^t I_{n,j})|$$

Then $\lim_{n\to\infty} \widehat{M}_{B,n} = 0$.

The proof of Proposition 4.9 follows nearly immediately from the fact that the resulting ergodic average for $T^{k_n\tilde{h}_n}$ is already known to converge by weak power ergodicity.

4.2. Mixing between $b_n\tilde{h}_n$ and $a_n\tilde{h}_n$

The new techniques introduced here apply to the times not covered by the above results. For $b_n\tilde{h}_n \leq t < a_n\tilde{h}_n$, we write t uniquely as $k_n\tilde{h}_n + y_n$ for $b_n \leq k_n < a_n$ and $|y_n| \leq \frac{1}{2}\tilde{h}_n$ (taking y_n positive in the case when $|y_n| = \frac{1}{2}\tilde{h}_n$).

4.2.1. Mixing using the previous column (mixing when $|y_n| \geq a_{n-1}\tilde{h}_{n-1}$)

The first new idea we introduce is showing mixing by invoking known mixing times for the previous tower (y_n is already known to be a mixing time for the previous column in the case $|y_n| \geq a_{n-1}\tilde{h}_{n-1}$).

We first explain the argument, eliding many details, then present the detailed proofs. Since $k_n < a_n$, application of $T^{k_n\tilde{h}_n}$ to a level effectively gives b_n blocks of a_n sublevels, each block having passed through the same number of spacers. The dominant term in the number of spacers is ik_n where $1 \leq i < b_n$ where each block corresponds to a single i.

Since each block will then also have T^{y_n} applied to it, the mixing nature of y_n can be used to show that each block is mixed. The goal now is to prove the following.

Proposition 4.10. *Let T be a quasi-staircase transformation such that $\frac{a_{n+1}b_{n+1}+c_{n+1}+a_nb_n^2}{h_n} \to 0$. Let B be a union of levels in some column C_N. For $n > N$, set*

$$\widetilde{M}_{B,n} = \max_{b_n \leq k < a_n} \max_{a_{n-1}\tilde{h}_{n-1} \leq y \leq \tilde{h}_n - a_{n-1}\tilde{h}_{n-1}} \sum_{j=0}^{\tilde{h}_{n-1}-1} |\lambda_B(T^{k\tilde{h}_n+y}I_{n-1,j})|$$

Then $\lim_{n\to\infty} \widetilde{M}_{B,n} = 0$.

We remark that one could strengthen Proposition 4.10 to also include the times when $|y_n| < b_n$ but we will not need that here (and it is more natural to include those cases in a later argument).

We first establish that, for the sublevels not pushed through the top of the next column, we can replace $T^{k_n\tilde{h}_n+y_n}$ by an ergodic-like average preserving the value of y_n. Essentially the proof is the standard technique that sublevels pushed through the top of the next column are mixed (corresponding to the $\frac{k}{a_n}\epsilon$ term below) combined with a careful accounting of the sublevels for which that does not occur.

Lemma 4.11. *Let T be a quasi-staircase transformation, B a union of levels in some C_N, $n > N$, $b_n \leq k < a_n$ and $0 \leq y < \tilde{h}_n$. Let $\epsilon > 0$ such that $\sup_{t \geq b_n}\left(\int \left|\frac{1}{t}\sum_{i=0}^{t-1}\chi_B \circ T^{-i}\right| \, d\mu + \frac{2}{t}\right) < \epsilon$. Then*

$$\sum_{j=a_nb_n+b_{n+1}+c_{n+1}-c_n}^{\tilde{h}_n-y} \left|\lambda_B(T^{k\tilde{h}_n+y}I_{n,j}) - \frac{a_n-k}{r_n+1}\sum_{\ell=0}^{b_n-1}\lambda_B(T^{y-k\ell}I_{n,j})\right| < \frac{k}{a_n}\epsilon; \qquad \textit{and}$$

$$\sum_{j=a_nb_n+b_{n+1}+c_{n+1}-c_n+\tilde{h}_n-y}^{\tilde{h}_n} \left|\lambda_B(T^{k\tilde{h}_n+y}I_{n,j}) - \frac{a_n-k-1}{r_n+1}\sum_{\ell=0}^{b_n-1}\lambda_B(T^{y-\tilde{h}_n-(k+1)\ell}I_{n,j})\right| < \frac{k+1}{a_n}\epsilon$$

Proof. For $a_nb_n + b_{n+1} + c_{n+1} - c_n \leq j < \tilde{h}_n - y$, by Lemmas A.7 and A.8,

$$\begin{aligned}\lambda_B(T^{k\tilde{h}_n+y}I_{n,j}) &= \sum_{i=0}^{a_n-1}\sum_{\ell=0}^{b_n-1}\lambda_B(T^{k\tilde{h}_n+y}I_{n,j}^{[\ell a_n+i]}) + \lambda_B(T^{k\tilde{h}_n+y}I_{n,j}^{[r_n]}) \\ &= \sum_{i=0}^{a_n-k-1}\sum_{\ell=0}^{b_n-1}\lambda_B(T^{-k\ell}I_{n,j+y}^{[\ell a_n+i+k]}) + \sum_{i=a_n-k}^{a_n-1}\sum_{\ell=0}^{b_n-2}\lambda_B(T^{-k\ell-(i+k-a_n)}I_{n,j+y}^{[\ell a_n+i+k+1]})\end{aligned}$$

$$+\sum_{i=0}^{k} \lambda_B(T^{k\tilde{h}_n+y} I_{n,j}^{[r_n-i]})$$

and since $k\ell \leq a_n b_n$ and $j+y \geq j \geq a_n b_n$, using Lemma 4.5,

$$\sum_{i=0}^{a_n-k-1} \sum_{\ell=0}^{b_n-1} \lambda_B(T^{-k\ell} I_{n,j+y}^{[\ell a_n+i+k]}) = \frac{1}{r_n+1} \sum_{i=0}^{a_n-k-1} \sum_{\ell=0}^{b_n-1} \lambda_B(T^{-k\ell} I_{n,j+y}) = \frac{a_n-k}{r_n+1} \sum_{\ell=0}^{b_n-1} \lambda_B(T^{y-k\ell} I_{n,j})$$

Likewise, since $k\ell + (i+k-a_n) \leq a_n b_n$,

$$\begin{aligned}\left| \sum_{i=a_n-k}^{a_n-1} \sum_{\ell=0}^{b_n-2} \lambda_B(T^{-k\ell-(i+k-a_n)} I_{n,j+y}^{[\ell a_n+i+k+1]}) \right| &= \left| \frac{1}{r_n+1} \sum_{i=a_n-k}^{a_n-1} \sum_{\ell=0}^{b_n-2} \lambda_B(T^{-k\ell-(i+k-a_n)} I_{n,j+y}) \right| \\ = \left| \frac{1}{r_n+1} \sum_{i=0}^{k-1} \sum_{\ell=0}^{b_n-2} \lambda_B(T^{y-k\ell-i} I_{n,j}) \right| &\leq \frac{k}{r_n+1} \sum_{\ell=0}^{b_n-2} \int_{T^{y-k\ell} I_{n,j}} \left| \frac{1}{k} \sum_{i=0}^{k-1} \chi_B \circ T^{-i} \right| \, d\mu\end{aligned}$$

and therefore

$$\sum_{j=a_n b_n+b_{n+1}+c_{n+1}-c_n}^{\tilde{h}_n-y} \left| \sum_{i=a_n-k}^{a_n-1} \sum_{\ell=0}^{b_n-2} \lambda_B(T^{-k\ell-(i+k-a_n)} I_{n,j+y}^{[\ell a_n+i+k+1]}) \right| < \frac{k(b_n-1)}{r_n+1} \int \left| \frac{1}{k} \sum_{i=0}^{k-1} \chi_B \circ T^{-i} \right| \, d\mu$$

For $0 \leq i \leq k-1$, using that $j \geq c_{n+1} - c_n + b_{n+1} + a_n b_n$ and that $I_{n,j}^{[0]} = I_{n+1,j}$,

$$\begin{aligned}T^{k\tilde{h}_n+y} I_{n,j}^{[r_n-i]} &= T^{k\tilde{h}_n+y+h_{n+1}-h_n-i(\tilde{h}_n+b_n-1)} I_{n,j}^{[0]} \\ &= T^{\tilde{h}_{n+1}+(k-i-1)\tilde{h}_n+c_n-c_{n+1}-i(b_n-1)+y} I_{n,j}^{[0]} = T^{\tilde{h}_{n+1}} I_{n+1,j+(k-i-1)\tilde{h}_n+c_n-c_{n+1}-i(b_n-1)+y}\end{aligned}$$

therefore, since $|\lambda_B(T^{\tilde{h}_{n+1}} I_{n+1,j'})| = |\sum_{t=0}^{b_{n+1}-1} (\sum_{i=0}^{a_{n+1}-2} \lambda_B(T^{-t} I_{n+1,j'}^{[ta_{n+1}+i+1]}) + \lambda_B(T^{-t-1} I_{n+1,j'}^{[(t+1)a_{n+1}]})) + \lambda_B(T^{\tilde{h}_{n+1}} I_{n+1,j'}^{[r_{n+1}]})| \leq \frac{a_{n+1}}{r_{n+1}+1} |\sum_{t=0}^{b_{n+1}-1} \lambda_B(T^{-t} I_{n+1,j'})| + \frac{2\mu(I_{n+1,j'})}{r_{n+1}+1}$ whenever $j' \geq b_{n+1}$,

$$\begin{aligned}\left|\lambda_B(T^{k\tilde{h}_n+y} I_{n,j}^{[r_n-i]})\right| &\leq \left| \frac{a_{n+1}}{r_{n+1}+1} \sum_{t=0}^{b_{n+1}-1} \lambda_B\left(T^{-t} I_{n+1,j+(k-i-1)\tilde{h}_n+c_n-c_{n+1}-i(b_n-1)+y}\right) \right| + \frac{2\mu(I_{n+1})}{r_{n+1}+1} \\ &= \left| \frac{a_{n+1}}{r_{n+1}+1} \sum_{t=0}^{b_{n+1}-1} \lambda_B\left(T^{-t} I_{n,j+c_n-c_{n+1}-i(b_n-1)+y}^{[k-i-1]}\right) \right| + \frac{2\mu(I_{n+1})}{r_{n+1}+1} \\ &= \left| \frac{a_{n+1}}{r_{n+1}+1} \frac{1}{r_n+1} \sum_{t=0}^{b_{n+1}-1} \lambda_B\left(T^{-t} I_{n,j+c_n-c_{n+1}-i(b_n-1)+y}\right) \right| + \frac{2\mu(I_{n+1})}{r_{n+1}+1} \\ &\leq \frac{a_{n+1} b_{n+1}}{(r_{n+1}+1)(r_n+1)} \int_{T^{y+c_n-c_{n+1}-i(b_n-1)} I_{n,j}} \left| \frac{1}{b_{n+1}} \sum_{t=0}^{b_{n+1}-1} \chi_B \circ T^{-t} \right| \, d\mu + \frac{2\mu(I_{n+1,j})}{r_{n+1}+1}\end{aligned}$$

and so

$$\begin{aligned}&\sum_{j=a_n b_n+b_{n+1}+c_{n+1}-c_n}^{\tilde{h}_n-y} \sum_{i=0}^{k} \left|\lambda_B(T^{k\tilde{h}_n+y} I_{n,j}^{[r_n-i]})\right| \\ &\qquad \leq \frac{k}{r_n+1} \int \left| \frac{1}{b_{n+1}} \sum_{t=0}^{b_{n+1}-1} \chi_B \circ T^{-t} \right| \, d\mu + \frac{1}{r_n+1} + \frac{2}{(r_{n+1}+1)(r_n+1)}\end{aligned}$$

Therefore, since $\sup_{t \geq b_n} \left(\int \left| \frac{1}{t} \sum_{i=0}^{t-1} \chi_B \circ T^{-i} \right| \, d\mu + \frac{2}{t} \right) < \epsilon$,

$$\sum_{j=a_nb_n+b_{n+1}+c_{n+1}-c_n}^{\tilde{h}_n-y} \left| \lambda_B(T^{k\tilde{h}_n+y}I_{n,j}) - \frac{a_n-k}{r_n+1} \sum_{\ell=0}^{b_n-1} \lambda_B(T^{y-k\ell}I_{n,j}) \right| \leq \frac{kb_n}{r_n+1}\epsilon < \frac{k}{a_n}\epsilon$$

For $a_nb_n + c_{n+1} - c_n + \tilde{h}_n - y \leq j < \tilde{h}_n$,

$$T^{k\tilde{h}_n+y}I_{n,j} = T^{(k+1)\tilde{h}_n+0}I_{n,j-(\tilde{h}_n-y)}$$

and since $a_nb_n + b_{n+1} + c_{n+1} - c_n \leq j - (\tilde{h}_n - y) < \tilde{h}_n - 0$, the claim follows from the above replacing k by $k+1$, j by $j - (\tilde{h}_n - y)$ and y by 0. ☐

We are now able to prove Proposition 4.10 using the previous lemma and the mixing nature of $\{y_n\}$. Unfortunately, the proof cannot be written as directly as one might hope: one cannot just use that T^{y_n} is mixing on levels in C_{n-1} directly. Instead, the proof is essentially the same as the proof that T^{y_n} is mixing but applied to the blocks of sublevels. The techniques are standard for staircase mixing but care must be taken to keep track of the sublevels so there is a significant amount of bookkeeping.

Proof of Proposition 4.10. Let $\epsilon > 0$ such that $\sup_{t \geq b_n} \left(\int \left| \frac{1}{t} \sum_{i=0}^{t-1} \chi_B \circ T^{-i} \right| \, d\mu + \frac{2}{t} \right) < \epsilon$. Write $y = xa_{n-1}\tilde{h}_{n-1} + z\tilde{h}_{n-1} + w$ for $1 \leq x \leq b_n$ and $0 \leq z < a_{n-1}$ and $0 \leq w < \tilde{h}_{n-1}$. Observe that if $0 \leq i < (b_{n-1} - x)a_{n-1}$ then $I_{n-1,j}^{[i]}$ is a level in C_n below $I_{n,\tilde{h}_n-y}$ and that if $(b_{n-1} - x)a_{n-1} < i \leq r_{n-1}$ then $I_{n-1,j}^{[i]}$ is a level in C_n above $I_{n,\tilde{h}_n-y}$. Then by Lemma 4.11, as $\frac{2k+1}{a_n}\epsilon \leq \frac{3k}{a_n}\epsilon$,

$$\begin{aligned} \sum_{j=0}^{\tilde{h}_{n-1}-1} \Bigg| \lambda_B(T^{k\tilde{h}_n+y}I_{n-1,j}) - \frac{a_n-k}{r_n+1} \sum_{\ell=0}^{b_n-1} \sum_{i=0}^{(b_{n-1}-x)a_{n-1}-1} \lambda_B(T^{y-k\ell}I_{n-1,j}^{[i]}) \qquad\qquad (\dagger) \\ - \frac{a_n-k-1}{r_n+1} \sum_{\ell=0}^{b_n-1} \sum_{i=(b_{n-1}-x+1)a_{n-1}}^{r_{n-1}} \lambda_B(T^{y-\tilde{h}_n-(k+1)\ell}I_{n-1,j}^{[i]}) \Bigg| < \frac{3k}{a_n}\epsilon + \frac{a_n}{r_n+1} + \frac{4(a_nb_n+b_{n+1}+c_{n+1})}{\tilde{h}_n} \end{aligned}$$

Now observe that, via Lemma A.9, writing $k' = xa_{n-1} + z$,

$$\begin{aligned} \sum_{j=0}^{\tilde{h}_{n-1}-1} \left| \frac{1}{b_n} \sum_{\ell=0}^{b_n-1} \sum_{i=0}^{(b_{n-1}-x)a_{n-1}-1} \lambda_B(T^{y-k\ell}I_{n-1,j}^{[i]}) \right| &\leq \frac{1}{b_n} \sum_{\ell=0}^{b_n-1} \sum_{j=0}^{\tilde{h}_{n-1}-1} \left| \sum_{i=0}^{(b_{n-1}-x)a_{n-1}-1} \lambda_B(T^{y-k\ell}I_{n-1,j}^{[i]}) \right| \\ &\leq \frac{c_{n-1}}{\tilde{h}_{n-1}} + \sum_{j=0}^{\tilde{h}_{n-1}-1} \left(\left| \sum_{i=0}^{(b_{n-1}-x)a_{n-1}-1} \lambda_B(T^{k'\tilde{h}_{n-1}}I_{n-1,j}^{[i]}) \right| + \left| \sum_{i=0}^{(b_{n-1}-x)a_{n-1}-1} \lambda_B(T^{(k'+1)\tilde{h}_{n-1}}I_{n-1,j}^{[i]}) \right| \right) \end{aligned}$$

which are precisely the sums $(\star)$ in the proof Proposition 4.8 (since $x \geq 1$ so $k' \geq a_{n-1}$). Therefore

$$\sum_{j=0}^{\tilde{h}_{n-1}-1} \left| \frac{a_n-k}{r_n+1} \sum_{\ell=0}^{b_n-1} \sum_{i=0}^{(b_{n-1}-x)a_{n-1}-1} \lambda_B(T^{y-k\ell}I_{n-1,j}^{[i]}) \right| \to 0 \qquad (\ddagger)$$

Now observe that for $0 \leq i < a_{n-1}$ and $0 \leq q < b_{n-1}$,

$$I_{n-1,j}^{[qa_{n-1}+i]} = T^{qa_{n-1}\tilde{h}_{n-1}+\frac{1}{2}a_{n-1}q(q-1)+i\tilde{h}_{n-1}+iq}I_{n-1,j}^{[0]}$$

so for $0 \leq i < a_{n-1} - 1$, as $(b_{n-1} - x + q)(b_{n-1} - x + q - 1) - q(q-1) = (b_{n-1} - x)(b_{n-1} - x - 1 + 2q)$,

$$I_{n-1,j}^{[(b_{n-1}-x+q)a_{n-1}+i+1]} = T^{(b_{n-1}-x)a_{n-1}\tilde{h}_{n-1}+\frac{1}{2}a_{n-1}(b_{n-1}-x)(b_{n-1}-x-1+2q)+\tilde{h}_{n-1}+q+(i+1)(b_{n-1}-x)}I_{n-1,j}^{[qa_{n-1}+i]}$$

Set $Q = Q_q = -c_n + c_{n-1} - (k+1)\ell + \frac{1}{2}a_{n-1}(b_{n-1}-x)(b_{n-1}-x-1+2q) + q - \frac{1}{2}a_{n-1}b_{n-1}(b_{n-1}-1) + b_{n-1} - x$ and note that $|Q| \leq c_n + a_nb_n + 2a_{n-1}b_{n-1}^2$. Then, since $b_{n-1}a_{n-1}\tilde{h}_{n-1} + \tilde{h}_{n-1} - \tilde{h}_n = -c_n + c_{n-1} -$

$\frac{1}{2}a_{n-1}b_{n-1}(b_{n-1}-1)$,

$$T^{y-(k+1)\ell-\tilde{h}_n} I_{n-1,j}^{[(b_{n-1}-x+q)a_{n-1}+i]} = T^{z\tilde{h}_{n-1}+w+i(b_{n-1}-x)+Q} I_{n-1,j}^{[qa_{n-1}+i]}$$

Consider j such that $0 \leq j + Q - a_{n-1}b_{n-1} < \tilde{h}_{n-1} - w - a_{n-1}b_{n-1}$. If $z + i \geq a_{n-1}$,

$$\begin{aligned} T^{y-(k+1)\ell-\tilde{h}_n} I_{n-1,j}^{[(b_{n-1}-x+q)a_{n-1}+i]} &= T^{z\tilde{h}_{n-1}} I_{n-1,j+Q+w+i(b_{n-1}-x)}^{[qa_{n-1}+i]} = I_{n-1,j+Q+w+i(b_{n-1}-x)-zq-(z+i-a_{n-1})}^{[qa_{n-1}+i+z]} \\ &= T^{i(b_{n-1}-x-1)} I_{n-1,j+Q+w-zq-(z-a_{n-1})}^{[qa_{n-1}+i+z]} \end{aligned}$$

and therefore

$$\lambda_B(T^{y-(k+1)\ell-\tilde{h}_n} I_{n-1,j}^{[qa_{n-1}+i]}) = \frac{1}{r_{n-1}} \lambda_B \left(T^{i(b_{n-1}-x-1)} I_{n-1,j+Q+w-zq-(z-a_{n-1})} \right)$$

Similarly, if $z + i < a_{n-1}$,

$$\begin{aligned} T^{y-(k+1)\ell-\tilde{h}_n} I_{n-1,j}^{[(b_{n-1}-x+q)a_{n-1}+i]} &= T^{z\tilde{h}_{n-1}} I_{n-1,j+Q+w+i(b_{n-1}-x)}^{[qa_{n-1}+i]} = I_{n-1,j+Q+w+i(b_{n-1}-x)-zq}^{[qa_{n-1}+i+z]} \\ &= T^{i(b_{n-1}-x)} I_{n-1,j+Q+w-zq}^{[qa_{n-1}+i+z]} \end{aligned}$$

so

$$\lambda_B(T^{y-(k+1)\ell-\tilde{h}_n} I_{n-1,j}^{[qa_{n-1}+i]}) = \frac{1}{r_{n-1}} \lambda_B \left(T^{i(b_{n-1}-x)} I_{n-1,j+Q+w-zq-(z-a_{n-1})} \right)$$

Therefore, as $\frac{x}{r_{n-1}} \leq \frac{b_{n-1}}{r_{n-1}} < \frac{1}{a_{n-1}}$,

$$\begin{aligned} &\sum_{j=a_{n-1}b_{n-1}-Q}^{\tilde{h}_{n-1}-w-a_{n-1}b_{n-1}} \left| \sum_{i=(b_{n-1}-x+1)a_{n-1}}^{r_{n-1}} \lambda_B(T^{y-(k+1)\ell-\tilde{h}_n} I_{n-1,j}^{[i]}) \right| \\ &\quad \leq \sum_{j=a_{n-1}b_{n-1}-Q}^{\tilde{h}_{n-1}-w-a_{n-1}b_{n-1}} \left| \sum_{q=b_{n-1}-x+1}^{b_{n-1}-1} \sum_{i=0}^{a_{n-1}-2} \lambda_B(T^{y-(k+1)\ell-\tilde{h}_n} I_{n-1,j}^{[qa_{n-1}+i]}) \right| + \frac{x+1}{r_{n-1}} \\ &\quad \leq \frac{1}{r_{n-1}} \sum_{q=0}^{x-1} \int \left| \sum_{i=0}^{a_{n-1}-z-1} \chi_B \circ T^{i(b_{n-1}-x-1)} \right| d\mu + \frac{1}{r_{n-1}} \sum_{q=0}^{x-1} \int \left| \sum_{i=0}^{z-1} \chi_B \circ T^{i(b_{n-1}-x)} \right| d\mu + \frac{x+1}{r_{n-1}} \\ &\quad \leq \int \left| \frac{1}{a_{n-1}} \sum_{i=0}^{a_{n-1}-z-1} \chi_B \circ T^{i(b_{n-1}-x-1)} \right| d\mu + \int \left| \frac{1}{a_{n-1}} \sum_{i=0}^{z-1} \chi_B \circ T^{i(b_{n-1}-x)} \right| d\mu + \frac{x+1}{r_{n-1}} \end{aligned}$$

Now consider j such that $\tilde{h}_{n-1} - w + a_{n-1}b_{n-1} - Q \leq j < \tilde{h}_{n-1} - a_{n-1}b_{n-1}$. Then

$$T^{y-(k+1)\ell-\tilde{h}_n} I_{n-1,j}^{[qa_{n-1}+i]} = T^{(z+1)\tilde{h}_{n-1}} I_{n-1,j+Q+w+i(b_{n-1}-x)-\tilde{h}_{n-1}}$$

so similar reasoning as above shows that

$$\begin{aligned} &\sum_{j=\tilde{h}_{n-1}-w+a_{n-1}b_{n-1}-Q}^{\tilde{h}_{n-1}-a_{n-1}b_{n-1}} \left| \sum_{i=(b_{n-1}-x+1)a_{n-1}}^{r_{n-1}} \lambda_B(T^{y-(k+1)\ell-\tilde{h}_n} I_{n-1,j}^{[i]}) \right| \\ &\quad \leq \int \left| \frac{1}{a_{n-1}} \sum_{i=0}^{a_{n-1}-z-1} \chi_B \circ T^{i(b_{n-1}-x-1)} \right| d\mu + \int \left| \frac{1}{a_{n-1}} \sum_{i=0}^{z-1} \chi_B \circ T^{i(b_{n-1}-x)} \right| d\mu + \frac{x+1}{r_{n-1}} \end{aligned}$$

Note that $y < \tilde{h}_n - a_{n-1}\tilde{h}_{n-1} = (b_{n-1}-1)a_{n-1}\tilde{h}_{n-1} + h_{n-1} + \frac{1}{2}a_{n-1}b_{n-1}(b_{n-1}-1) + c_n < (b_{n-1}-1)a_{n-1}\tilde{h}_{n-1} + 2\tilde{h}_{n-1}$. Therefore $x \leq b_{n-1} - 1$ and if $x = b_{n-1} - 1$ then $z \leq 1$. When $x \leq b_{n-1} - 1$, both $b_{n-1} - x \geq 1$ and $b_{n-1} - x - 1 \geq 1$ so both integrals tend to zero by Proposition 4.7. When $x = b_{n-1} - 1$, the first integral tends to zero by Proposition 4.7 and the second is bounded by $\frac{z}{a_{n-1}} \to 0$.

Since $\frac{|Q|}{\tilde{h}_{n-1}} \leq \frac{c_n+a_nb_n+a_{n-1}b_{n-1}^2}{\tilde{h}_{n-1}} \to 0$, then

$$\sum_{j=0}^{\tilde{h}_{n-1}} \left| \sum_{i=(b_{n-1}-x+1)a_{n-1}}^{r_{n-1}} \lambda_B(T^{y-(k+1)\ell-\tilde{h}_n} I_{n-1,j}^{[i]}) \right| \to 0$$

Therefore equation (†) gives that $\sum_{j=0}^{\tilde{h}_{n-1}-1} |\lambda_B(T^{k\tilde{h}_n+y} I_{n-1,j})| \to 0$ as both the above quantity and that in (‡) tend to 0. Since this holds uniformly over k and y in the specified range, the claim follows. ☐

4.2.2. Mixing for the remaining times

The times not covered by the previous cases, namely those of the form $k_n\tilde{h}_n + y_n$ where $b_n \leq k_n < a_n$ and $|y_n| < a_{n-1}\tilde{h}_{n-1}$, require more care and the introduction of many new ideas; however, the majority of the argument is to split the sequence into cases where, in each case, a (generalization of a) standard mixing proof technique can be applied.

The final goal of this section is to prove the following proposition which ensures mixing on all remaining times.

Proposition 4.12. *Let T be a quasi-staircase transformation such that $\sum \frac{a_nb_n+b_{n+1}+c_{n+1}}{h_n} < \infty$ and $\sum \frac{1}{a_nb_n} < \infty$ and $\frac{a_nb_n^2}{h_n} \to 0$ and $\frac{a_{n+1}b_{n+1}}{h_n} \to 0$. Let B be a union of levels in some fixed C_N. Then*

$$\lim_{n\to\infty} \max_{b_n \leq k < a_n} \max_{|q| < a_{n-1}\tilde{h}_{n-1}} \left| \lambda_B(T^{k\tilde{h}_n+q}B) \right| = 0.$$

For ease of exposition, we denote the measure of the levels which may be 'safely' ignored as follows.

Notation 4.13. *Define $\tau_n = \frac{4(a_nb_n+b_{n+1}+c_{n+1})}{\tilde{h}_n}$.*

Our first lemma effectively generalizes the known technique of splitting the tower into the rightmost subcolumns, which will be mixed due to pushing through the top of the next tower, and those on the left which will need to be handled differently.

Lemma 4.14. *Let T be a quasi-staircase transformation, B a union of levels in some C_N, $\epsilon > 0$ such that $\sup_{t \geq b_N} \left(\int \left| \frac{1}{t} \sum_{i=0}^{t-1} \chi_B \circ T^{-i} \right| \, d\mu + \frac{2}{t} \right) < \frac{\epsilon}{3}$, $n > N$, $b_n \leq k < a_n$ and $0 \leq |y| < a_{n-1}\tilde{h}_{n-1}$. Then*

$$\left| \lambda_B(T^{k\tilde{h}_n+y}B) - \frac{a_n-k}{r_n+1} \sum_{\ell=0}^{b_n-1} \lambda_B(T^{y-k\ell}B) \right| \leq \frac{k}{a_n}\epsilon + \tau_n + \left(1 - \frac{k}{a_n}\right) \frac{1}{b_n} \sum_{\ell=0}^{b_n-1} \frac{2|y-k\ell|}{\tilde{h}_n}$$

Proof. Consider first when $y \geq 0$. Write $\beta = \{a_nb_n + b_{n+1} + c_{n+1} \leq j < \tilde{h}_n - y : I_{n,j} \subseteq B\}$ and $\beta' = \{a_nb_n + b_{n+1} + c_{n+1} + \tilde{h}_n - y \leq j < \tilde{h}_n : I_{n,j} \subseteq B\}$. By Lemma 4.11,

$$\left| \sum_{j \in \beta \cup \beta'} \lambda_B(T^{k\tilde{h}_n+y} I_{n,j}) - \sum_{\ell=0}^{b_n-1} \left(\frac{a_n-k}{r_n+1} \sum_{j\in\beta} \lambda_B(T^{y-k\ell} I_{n,j}) - \frac{a_n-k-1}{r_n+1} \sum_{j\in\beta'} \lambda_B(T^{y-\tilde{h}_n-(k+1)\ell} I_{n,j}) \right) \right|$$

is bounded by $\frac{k}{a_n}\frac{\epsilon}{3} + \frac{k+1}{a_n}\frac{\epsilon}{3} \leq \frac{k\epsilon}{a_n}$ and therefore

$$\begin{aligned} &\left| \lambda_B(T^{k\tilde{h}_n+y}B) - \frac{a_n-k}{r_n+1} \sum_{\ell=0}^{b_n-1} \sum_{j\in\beta\cup\beta'} \lambda_B(T^{y-k\ell} I_{n,j}) \right| \\ &\quad \leq \frac{k}{a_n}\epsilon + \frac{\tau_n}{2} + \frac{a_n-k}{r_n+1} \sum_{\ell=0}^{b_n-1} \sum_{j\in\beta'} \left| \lambda_B(T^{y-k\ell} I_{n,j}) - \frac{a_n-k-1}{a_n-k} \lambda_B(T^{y-\tilde{h}_n-(k+1)\ell} I_{n,j}) \right| \end{aligned}$$

$$\leq \frac{k}{a_n}\epsilon + \frac{\tau_n}{2} + \frac{a_n - k}{r_n + 1} b_n |\beta'| \mu(I_n) \frac{2a_n - 2k - 1}{a_n - k} < \frac{k}{a_n}\epsilon + \frac{\tau_n}{2} + \left(1 - \frac{k}{a_n}\right) \frac{2|\beta'|}{\tilde{h}_n}$$

so the claim follows for $y \geq 0$ as $|\beta'| = y - a_n b_n - c_{n+1} \leq |y - k\ell|$ for all $0 \leq \ell < b_n$ (and if $y < a_n b_n + b_{n+1} + c_{n+1}$ then $\beta' = \emptyset$) and since $|\lambda_B(T^{k\tilde{h}_n + y}B) - \sum_{j \in \beta \cup \beta'} \lambda_B(T^{k\tilde{h}_n + y} I_{n,j})| \leq \frac{\tau_n}{2}$.

Now consider when $y < 0$. Then $k\tilde{h}_n + y = (k-1)\tilde{h}_n + (\tilde{h}_n + y)$ so, following the same reasoning as above and swapping the roles of β' and β,

$$\left| \lambda_B(T^{k\tilde{h}_n + y}B) - \frac{a_n - (k-1) - 1}{r_n + 1} \sum_{\ell=0}^{b_n - 1} \sum_{j \in \beta \cup \beta'} \lambda_B(T^{(y + \tilde{h}_n) - (k-1+1)\ell} I_{n,j}) \right|$$
$$< \frac{k-1+1}{a_n}\epsilon + \frac{\tau_n}{2} + \left(1 - \frac{k-1+1}{a_n}\right) \frac{2|\beta|}{\tilde{h}_n} \leq \frac{k}{a_n}\epsilon + \tau_n + \left(1 - \frac{k}{a_n}\right) \frac{2|\beta|}{\tilde{h}_n}$$

Since in this case $|\beta| \leq |y - k\ell|$ for all $0 \leq \ell < b_n$, then $|\beta| \leq \frac{1}{b_n} \sum_{\ell=0}^{b_n - 1} |y - k\ell|$ so the claim follows. □

Our next lemma is a generalization of the now-standard Block Lemma [Ada98]. It is a 'weighted' version of that lemma, though we emphasize that it is not the case that the weights sum to 1 but rather that each term in the average may be weighted by a value between 0 and 1 (so e.g. the case when most of the weights are 0 follows trivially).

Lemma 4.15. *Let $\epsilon > 0$ and $q, k, p, Q, L \in \mathbb{N}$ and for all $0 \leq \ell < L$, let $0 \leq \delta_\ell \leq 1$. If $\frac{pQ}{L} < \epsilon$ and $\frac{1}{Q} < \epsilon$ and $|\lambda_B(T^{kpt}B)| < \epsilon$ for all $1 \leq t < Q$ then*

$$\left| \frac{1}{L} \sum_{\ell=0}^{L-1} \delta_\ell \lambda_B(T^{q - k\ell}B) \right| < (2\epsilon)^{1/2} + \epsilon$$

Proof. Using that T is measure-preserving and the Cauchy-Schwarz inequality,

$$\left| \frac{1}{L} \sum_{\ell=0}^{L-1} \delta_\ell \lambda_B(T^{q-k\ell}B) \right| = \left| \frac{1}{L} \sum_{\ell=0}^{L-1} \int_B \delta_\ell \chi_B \circ T^{q-k\ell} \, d\mu \right| \leq \int \left| \frac{1}{L} \sum_{\ell=0}^{L-1} \delta_\ell \chi_B \circ T^{q-k\ell} \right| \, d\mu$$

$$\leq \frac{pQ\lfloor \frac{L}{pQ} \rfloor}{L} \frac{1}{\lfloor \frac{L}{pQ} \rfloor} \sum_{j=0}^{\lfloor \frac{L}{pQ} \rfloor - 1} \frac{1}{p} \sum_{i=0}^{p-1} \int \left| \frac{1}{Q} \sum_{t=0}^{Q-1} \delta_{jpQ+i+pt} \chi_B \circ T^{-kpt} \right| \circ T^{q - kjpQ - ki} \, d\mu + \frac{pQ}{L}$$

$$< \frac{1}{\lfloor \frac{L}{pQ} \rfloor} \sum_{j=0}^{\lfloor \frac{L}{pQ} \rfloor - 1} \frac{1}{p} \sum_{i=0}^{p-1} \int \left| \frac{1}{Q} \sum_{t=0}^{Q-1} \delta_{jpQ+i+pt} \chi_B \circ T^{-kpt} \right| \, d\mu + \epsilon$$

$$\leq \frac{1}{\lfloor \frac{L}{pQ} \rfloor} \sum_{j=0}^{\lfloor \frac{L}{pQ} \rfloor - 1} \frac{1}{p} \sum_{i=0}^{p-1} \left(\int \left| \frac{1}{Q} \sum_{t=0}^{Q-1} \delta_{jpQ+i+pt} \chi_B \circ T^{-kpt} \right|^2 \, d\mu \right)^{1/2} + \epsilon$$

$$= \frac{1}{\lfloor \frac{L}{pQ} \rfloor} \sum_{j=0}^{\lfloor \frac{L}{pQ} \rfloor - 1} \frac{1}{p} \sum_{i=0}^{p-1} \left(\frac{1}{Q^2} \sum_{t,u=0}^{Q-1} \delta_{jpQ+i+pt} \delta_{jpQ+i+pu} \lambda_B(T^{kp(t-u)}B) \right)^{1/2} + \epsilon$$

$$= \frac{1}{\lfloor \frac{L}{pQ} \rfloor} \sum_{j=0}^{\lfloor \frac{L}{pQ} \rfloor - 1} \frac{1}{p} \sum_{i=0}^{p-1} \left(\frac{1}{Q^2} \sum_{t=0}^{Q-1} \delta_{jpQ+i+pt}^2 \lambda_B(B) + \frac{1}{Q^2} \sum_{t \neq u} \delta_{jpQ+i+pt} \delta_{jpQ+i+pu} \lambda_B(T^{kp(t-u)}B) \right)^{1/2} + \epsilon$$

$$< \frac{1}{\lfloor \frac{L}{pQ} \rfloor} \sum_{j=0}^{\lfloor \frac{L}{pQ} \rfloor - 1} \frac{1}{p} \sum_{i=0}^{p-1} \left(\frac{1}{Q} + \frac{1}{Q^2} \sum_{t \neq u} \delta_{jpQ+i+pt} \delta_{jpQ+i+pu} \epsilon \right)^{1/2} + \epsilon \leq \left(\frac{1}{Q} + \frac{1}{Q^2} \sum_{t \neq u} \epsilon \right)^{1/2} + \epsilon$$

□

Using the 'weighted' Block Lemma, we likewise generalize the Blum-Hanson trick (see e.g. [Ada98]) to a 'weighted' version.

Proposition 4.16. *Let T be a quasi-staircase transformation such that $\frac{b_n^2}{\tilde{h}_n} \to 0$, $\frac{a_n b_n}{\tilde{h}_n} \to 0$ and $\frac{b_n}{a_n} \to 0$. Let B be a union of levels in some column C_{N_0}. Then*

$$\lim_{N\to\infty} \max_{0\le\delta_\ell\le 1} \max_{1\le k\le N} \int \left| \frac{1}{N} \sum_{\ell=0}^{N-1} \delta_\ell \chi_B \circ T^{-\ell k} \right| \, d\mu = 0$$

Proof. Fix $\epsilon > 0$. Let m such that $b_m \ge 2\lceil\epsilon^{-1}\rceil$, $\frac{4(r_m+1)\lceil\epsilon^{-1}\rceil^2}{\tilde{h}_m} < \epsilon$ and $\sup_{n\ge m} \widehat{M}_{B,n} < \epsilon$ (using Proposition 4.9). Take any N such that $\frac{\tilde{h}_m\lceil\epsilon^{-1}\rceil}{N} < \epsilon$. Let k and δ_ℓ attain the maximum for N.

Consider first the case when $k \ge \tilde{h}_m$. Let $n \ge m$ such that $\tilde{h}_n \le k < \tilde{h}_{n+1}$. Let p such that $(p-1)k < \tilde{h}_{n+1} \le pk$ so that $pk < \tilde{h}_{n+1} + k < 2\tilde{h}_{n+1}$. Then for every $1 \le q < \lceil\epsilon^{-1}\rceil$, $\tilde{h}_{n+1} \le qpk < \lceil\epsilon^{-1}\rceil 2\tilde{h}_{n+1} \le b_n\tilde{h}_n$ meaning that $|\lambda_B(T^{qpk}B)| \le \widehat{M}_{B,n} < \epsilon$. Now

$$\frac{p\lceil\epsilon^{-1}\rceil}{N} = \frac{pk\lceil\epsilon^{-1}\rceil}{Nk} < \frac{2\tilde{h}_{n+1}\lceil\epsilon^{-1}\rceil}{N\tilde{h}_n} < \frac{4(r_n+1)\lceil\epsilon^{-1}\rceil}{N} \le \frac{4(r_n+1)\lceil\epsilon^{-1}\rceil}{k} \le \frac{4(r_n+1)\lceil\epsilon^{-1}\rceil}{\tilde{h}_n} < \epsilon$$

so Lemma 4.15 implies that $\int \left| \frac{1}{N} \sum_{\ell=0}^{N-1} \delta_\ell \chi_B \circ T^{-\ell k} \right| \, d\mu < (2\epsilon)^{1/2} + \epsilon$.

Consider now when $k < \tilde{h}_m$. Let p such that $(p-1)k < \tilde{h}_m \le pk$ so that $pk < 2\tilde{h}_m$ and $p \le \tilde{h}_m$. Then $\tilde{h}_m \le qpk < \lceil\epsilon^{-1}\rceil 2\tilde{h}_m \le b_m\tilde{h}_m$ for $1 \le q < \lceil\epsilon^{-1}\rceil$ so $|\lambda_B(T^{qpk}B)| \le \widehat{M}_{B,m} < \epsilon$. Since $\frac{p\lceil\epsilon^{-1}\rceil}{N} < \frac{\tilde{h}_m\lceil\epsilon^{-1}\rceil}{N} < \epsilon$, Lemma 4.15 again implies that $\int \left| \frac{1}{N} \sum_{\ell=0}^{N-1} \delta_\ell \chi_B \circ T^{-\ell k} \right| \, d\mu < (2\epsilon)^{1/2} + \epsilon$. ☐

Since we will have need to work with times not all ranging within the same interval $[h_n, h_{n+1})$ (which is all one needs to consider when working with staircases), we introduce the following notation to denote which column a time 'naturally wants to be thought of as applying to'.

Notation 4.17. *For $t \in \mathbb{Z}$, write $\alpha(t)$ for the unique positive integer such that $\tilde{h}_{\alpha(t)} \le |t| < \tilde{h}_{\alpha(t)+1}$.*

The next lemma is the main ingredient in the proof of mixing for our remaining times. Given a 'weighted' average of times which form an arithmetic progression (meaning the times are $q - \ell k$ for some fixed q and k as ℓ ranges from 0 to $L-1$), the lemma states that either the average is already mixed by the previously established facts or that the sequence of times which cannot be guaranteed to already be mixing times has a very specific structure closely resembling that of an arithmetic progression, along with control on the size of the gaps in the potentially 'bad' times $\mathcal{L}$. This structure, along with the control on the size of the gaps, will allow us to deduce mixing by dropping to previous towers in a suitable manner.

The proof breaks into cases. The bulk of them, where the average is shown to be mixed, follow from a combination of the already established (weak) power ergodicity (similar to how mixing is deduced for staircases) and the generalization of the Blum-Hanson trick combined with the generalized Block Lemma. Due to the length of the proof, we include expository text throughout, italicized to distinguish it.

Lemma 4.18. *Let $\epsilon > 0$ and set $\epsilon_0 = (2\lceil\epsilon^{-1}\rceil^{\lceil\epsilon^{-1}\rceil+1})^{-1}$. Let $L, k, q \in \mathbb{Z}$ with $L \ge \epsilon_0^{-1}$ and for each $0 \le \ell < L$, let $0 \le \delta_\ell \le 1$.*

Let $\alpha_0 = \max\{\alpha(q - \ell k) : 0 \le \ell < L\}$. Assume that $\max(M_{B,\alpha_0}, M_{B,\alpha_0-1}, \widehat{M}_{B,\alpha_0}, \widetilde{M}_{B,\alpha_0}) < \epsilon$ and $b_{\alpha_0-1} > 4\epsilon^{-1}\epsilon_0^{-1}$.

Write $k = z\tilde{h}_{\alpha_0} + y$ *for* $|y| \le \frac{1}{2}\tilde{h}_{\alpha_0}$ *and* $q = x\tilde{h}_{\alpha_0} + r$ *for* $|r| \le \frac{1}{2}\tilde{h}_{\alpha_0}$.

Let $k_\ell, y_\ell \in \mathbb{Z}$ such that $q - \ell k = k_\ell\tilde{h}_{\alpha_0} + y_\ell$ *with* $|y_\ell| \le \frac{1}{2}\tilde{h}_{\alpha_0}$ *and define*

$$\mathcal{L} = \left\{ 0 \le \ell < L : (k_\ell = 0 \text{ or } b_{\alpha_0} \le |k_\ell| < a_{\alpha_0}) \text{ and } |y_\ell| < a_{\alpha_0-1}\tilde{h}_{\alpha_0-1} \right\}$$

Then at least one of the following holds:

- $\left| \frac{1}{L} \sum_{\ell=0}^{L-1} \delta_\ell \lambda_B(T^{q-\ell k}B) \right| + \frac{1}{L} \sum_{\ell=0}^{L-1} (1-\delta_\ell)\epsilon < 4\epsilon^{1/2}$*; or*

- *there exists* $p \in \mathbb{Z}$*,* $t > 0$ *and* $0 \leq \ell_0 < L' \leq L$ *such that*

$$\mathcal{L} \subseteq \{\ell_0 + it : 0 \leq i < L'\} \qquad \textit{and} \qquad |ity - ip\tilde{h}_{\alpha_0}| < \frac{1}{3}\tilde{h}_{\alpha_0} \quad \textit{for all } 0 \leq i < L'$$

Proof. First observe that if $0 \leq \ell < L$ and $\ell \notin \mathcal{L}$ then one of the following must hold:

- $a_{\alpha_0}\tilde{h}_{\alpha_0} \leq |q - \ell k| < \tilde{h}_{\alpha_0+1}$ (using that α_0 is maximal);
- $b_{\alpha_0}\tilde{h}_{\alpha_0} \leq |q - \ell k| < a_{\alpha_0}\tilde{h}_{\alpha_0}$ and $|y_\ell| \geq a_{\alpha_0-1}\tilde{h}_{\alpha_0-1}$;
- $\tilde{h}_{\alpha_0} \leq |q - \ell k| < b_{\alpha_0}\tilde{h}_{\alpha_0}$; or
- $k = 0$ (so $y_\ell = q - \ell k = q$) and $a_{\alpha_0-1}\tilde{h}_{\alpha_0-1} \leq |q - \ell k| < \tilde{h}_{\alpha_0}$.

Each of the above cases falls under one of the hypotheses on times which are already known to be ϵ*-mixed, thus* $\mathcal{L}$ *represents the 'potentially bad times'. As such, the proof need only focus on those values in* $\mathcal{L}$ *since the rest are already mixed. We now make this concrete.*

Since $\lambda_B(T^{-t}B) = \lambda_B(T^tB)$,

- $a_{\alpha_0}\tilde{h}_{\alpha_0} \leq |q - \ell k| < \tilde{h}_{\alpha_0+1}$ implies $\left|\lambda_B(T^{q-\ell k}B)\right| \leq M_{B,\alpha_0}$;
- $b_{\alpha_0}\tilde{h}_{\alpha_0} \leq |q - \ell k| < a_{\alpha_0}\tilde{h}_{\alpha_0}$ and $|y_\ell| \geq a_{\alpha_0-1}\tilde{h}_{\alpha_0-1}$ implies $\left|\lambda_B(T^{q-\ell k}B)\right| \leq \widetilde{M}_{B,\alpha_0}$;
- $\tilde{h}_{\alpha_0} \leq |q - \ell k| < b_{\alpha_0}\tilde{h}_{\alpha_0}$ implies $\left|\lambda_B(T^{q-\ell k}B)\right| \leq \widehat{M}_{B,\alpha_0}$; and
- $a_{\alpha_0-1}\tilde{h}_{\alpha_0-1} \leq |q - \ell k| < \tilde{h}_{\alpha_0}$ implies $\left|\lambda_B(T^{q-\ell k}B)\right| \leq M_{B,\alpha_0-1}$.

Therefore, by hypothesis, for $\ell \notin \mathcal{L}$, it holds that $\left|\lambda_B(T^{q-\ell k}B)\right| < \epsilon$.

In particular, if $|\mathcal{L}| < \epsilon L$ then

$$\left|\frac{1}{L}\sum_{\ell=0}^{L-1}\delta_\ell\lambda_B(T^{q-\ell k}B)\right| + \frac{1}{L}\sum_{\ell=0}^{L-1}(1-\delta_\ell)\epsilon < \frac{1}{L}|\mathcal{L}| + \frac{1}{L}\left(L - |\mathcal{L}|\right)\epsilon + \epsilon < 3\epsilon$$

so the first possible conclusion will hold.

The main idea is to consider the fraction $\frac{y}{\tilde{h}_{\alpha_0}}$ *and break into cases. We begin by approximating* $\frac{y}{\tilde{h}_{\alpha_0}}$ *with the closest fraction* $\frac{p}{t}$ *that has denominator less than* $b_{\alpha_0-1}L$ *(and then do the same to approximate* $\frac{r}{\tilde{h}_{\alpha_0}}$*).*

Let $p, t \in \mathbb{Z}$ with $0 < t < b_{\alpha_0-1}L$ such that $\left|\frac{y}{\tilde{h}_{\alpha_0}} - \frac{p}{t}\right| < \frac{1}{Lb_{\alpha_0-1}}$ and either $(p = 0, t = 1)$ or p, t are relatively prime. Then let $u \in \mathbb{Z}$ such that $\left|\frac{r}{\tilde{h}_{\alpha_0}} - \frac{u}{t}\right| \leq \frac{1}{2t}$.

In the case when $\frac{p}{t}$ *has a large denominator and large numerator, we expect that* $T^{\ell y}$ *distributes sublevels somewhat evenly across many levels and so we can apply either (weak) power ergodicity or the (generalized) Blum-Hanson trick.*

In the case when $\frac{p}{t}$ *is very small,* $\ell|y| < \tilde{h}_{\alpha_0}$ *so the set of* ℓ *for which* $|y_\ell| \approx q - \ell y$ *is less than* $b_{\alpha_0-1}\tilde{h}_{\alpha_0-1}$ *should be a consecutive set of integers (as there is no 'wraparound' coming from* ℓy*).*

In the case when $\frac{p}{t}$ *has a small denominator, we expect* $T^{\ell y}$ *to send most of the sublevels into the same small set of levels. This the arithmetic progression like structure conclusion of the lemma. The idea (in a later lemma) for handling this case is to 'drop down' to an earlier column, similar to in Section 4.2.1.*

For $\ell \in \mathcal{L}$, since $q - \ell k = k_\ell\tilde{h}_{\alpha_0} + y_\ell$ for $|y_\ell| < a_{\alpha_0-1}\tilde{h}_{\alpha_0-1}$ and since

$$q - \ell k = (x\tilde{h}_{\alpha_0} + r - \ell(z\tilde{h}_{\alpha_)} + y) = (x - \ell z)\tilde{h}_{\alpha_0} + (r - \ell y)$$

it follows that $r - \ell y$ must be of the form $n_\ell\tilde{h}_{\alpha_)} + o_\ell$ for some $n_\ell, o_\ell \in \mathbb{Z}$ where $|o_\ell| < a_{\alpha_0-1}\tilde{h}_{\alpha_0-1}$. Then

$|r - \ell y - n_\ell \tilde{h}_{\alpha_0}| < a_{\alpha_0-1}\tilde{h}_{\alpha_0-1}$ so $\left|\frac{r-\ell y}{\tilde{h}_{\alpha_0}} - n_\ell\right| < \frac{1}{b_{\alpha_0-1}}$. Therefore

$$\left|\frac{u-\ell p}{t} - n_\ell\right| \leq \left|\frac{u-\ell p}{t} - \frac{r-\ell y}{\tilde{h}_{\alpha_0}}\right| + \left|\frac{r-\ell y}{\tilde{h}_{\alpha_0}} - n_\ell\right| < \frac{1}{2t} + \frac{\ell}{Lb_{\alpha_0-1}} + \frac{1}{b_{\alpha_0-1}} < \frac{1}{2t} + \frac{2}{b_{\alpha_0-1}}$$

The above calculation shows that $\frac{u-\ell p}{t}$ *is always very close to an integer when* $\ell \in \mathcal{L}$. *This will play a key role in establishing the claimed structure of* $\mathcal{L}$ *in the case when* $\frac{p}{t}$ *has small denominator.*

We now split into cases.

Case: $t \leq \dfrac{b_{\alpha_0-1}}{4}$ and $|p| > 0$

For $\ell \in \mathcal{L}$, $\left|\dfrac{u-\ell p}{t} - n_\ell\right| < \dfrac{2}{b_{\alpha_0-1}} + \dfrac{1}{2t} \leq \dfrac{1}{2t} + \dfrac{1}{2t} = \dfrac{1}{t}$ so $|u - \ell p - n_\ell t| < 1$. As those are integers, then $u - \ell p = n_\ell t$. Let ℓ_0 be the minimal element of $\mathcal{L}$. Then $(\ell - \ell_0)p = (u - \ell_0 p) - (u - \ell p) = (n_{\ell_0} - n_\ell)t$. As p and t are relatively prime, then t must divide $\ell - \ell_0$ so every $\ell \in \mathcal{L}$ is of the form $\ell = \ell_0 + ti$ for some $0 \leq i < L$. Also $|ty - p\tilde{h}_{\alpha_0}| < \dfrac{t}{Lb_{\alpha_0-1}}\tilde{h}_{\alpha_0} \leq \dfrac{1}{4L}\tilde{h}_{\alpha_0}$ meaning $|ity - ip\tilde{h}_{\alpha_0}| \leq \frac{1}{4}\tilde{h}_{\alpha_0}$ for $0 \leq i < L$ so $\mathcal{L}$ has the structure of the second conclusion.

Case: $\dfrac{b_{\alpha_0-1}}{4} < t \leq L$ and $|p| > 0$

In this case $y \approx \frac{p}{t}\tilde{h}_{\alpha_0} \gtrapprox \frac{p}{b_{\alpha_0}}\tilde{h}_{\alpha_0} \approx pa_{\alpha_0-1}\tilde{h}_{\alpha_0-1}$ *so we can deduce mixing from having already established that (most) such times are mixing (and that most times in the progression are of that form).*

For $\ell \in \mathcal{L}$, $\left|\dfrac{u-\ell p}{t} - n_\ell\right| < \dfrac{2}{b_{\alpha_0-1}} + \dfrac{1}{2t} < \dfrac{4}{b_{\alpha_0-1}}$ so $|u - \ell p - n_\ell t| < \frac{4}{b_{\alpha_0-1}}t$ which means that $u - \ell p \pmod{t} < \frac{4}{b_{\alpha_0-1}}t$ or $u - \ell p \pmod{t} > \left(1 - \frac{4}{b_{\alpha_0-1}}\right)t$ whenever $\ell \in \mathcal{L}$.
Since p and t are relatively prime, the map $\mathbb{Z} \to \mathbb{Z}/t\mathbb{Z}$ given by $z \mapsto zp \pmod{t}$ is cyclic and onto. So at most $\frac{8}{b_{\alpha_0-1}}t$ choices of $1 \leq \ell \leq t$ can be in $\mathcal{L}$. Likewise, for any range of t values of ℓ, at most $\frac{8}{b_{\alpha_0-1}}t$ choices of ℓ (out of the t possible) can be in $\mathcal{L}$. Therefore $|\mathcal{L}| \leq \frac{8}{b_{\alpha_0-1}}t\left\lceil\frac{L}{t}\right\rceil$. Then

$$|\mathcal{L}| \leq \frac{8}{b_{\alpha_0-1}}t\left\lceil\frac{L}{t}\right\rceil \leq \frac{8}{b_{\alpha_0-1}}(L+t) \leq \frac{8}{b_{\alpha_0-1}}(L+L) < 4\epsilon\epsilon_0 L$$

Therefore, as $|\lambda_B(T^{q-\ell k}B)| < \epsilon$ for $\ell \notin \mathcal{L}$, we have $\left|\frac{1}{L}\sum_{\ell=0}^{L-1}\delta_\ell\lambda_B(T^{q-\ell k}B)\right| + \frac{1}{L}\sum_{\ell=0}^{L-1}(1-\delta_\ell)\epsilon < \frac{1}{L}|\mathcal{L}| + \epsilon + \epsilon < 4\epsilon\epsilon_0 + 2\epsilon$.

Case: $|p| < \dfrac{2t}{b_{\alpha_0-1}\epsilon_0 L}$

In this case, p being very small means y is very small so $\{\ell : 0 \leq \ell < L\}$ *already satisfies the claimed structure (and* $\mathcal{L}$ *is a subset of that).*
For $0 \leq \ell < L$, then $|\ell p| < \dfrac{2t}{b_{\alpha_0-1}\epsilon_0} < \dfrac{2t}{4\epsilon^{-1}\epsilon_0^{-1}\epsilon_0} = 2\epsilon t$ so

$$\left|\frac{\ell y}{\tilde{h}_{\alpha_0}}\right| \leq \ell\left|\frac{y}{\tilde{h}_{\alpha_0}} - \frac{p}{t}\right| + \ell\left|\frac{p}{t}\right| \leq \frac{\ell}{b_{\alpha_0-1}L} + \left|\frac{\ell p}{t}\right| < \frac{1}{b_{\alpha_0-1}} + 2\epsilon < 3\epsilon$$

For all $0 \leq i < L$, then $|i \cdot 1 \cdot y - i \cdot 0 \cdot \tilde{h}_{\alpha_0}| = |iy| < 3\epsilon\tilde{h}_{\alpha_0} < \frac{1}{3}\tilde{h}_{\alpha_0}$ Therefore the second possible conclusion of the lemma holds with $L' = L$ and $\ell_0 = 0$ and $t = 1$ and $p = 0$.

From here on, we may then assume that $\quad L < t \quad$ and $\quad |p| \geq \frac{2t}{b_{\alpha_0-1}\epsilon_0 L}$

Our goal now is to apply the (generalized) Blum-Hanson trick combined with the Block Lemma. The complication is that, unlike for staircases, there is no single choice of the power to use in the Block Lemma to reach mixing times but there are at most $1/\epsilon$ possible values.

For $0 \leq j < \lceil \epsilon^{-1} \rceil$, define $\qquad p_0 = p \pmod{t} \qquad$ and $\qquad p_{j+1} = \left\lceil \frac{t}{p_j} \right\rceil p_j \pmod{t}$

Suppose that $\epsilon t \leq p_j$ and $p_{j+1} \leq p_j - \epsilon t$ for all $0 \leq j < \lceil \epsilon^{-1} \rceil$. Since $p_j \geq \epsilon t$, $\lceil \frac{t}{p_j} \rceil \leq \lceil \epsilon^{-1} \rceil$. Then $\lceil \frac{t}{p_0} \rceil \lceil \frac{t}{p_1} \rceil \cdots \lceil \frac{t}{p_{\lceil \epsilon^{-1} \rceil - 1}} \rceil \leq \lceil \epsilon^{-1} \rceil^{\lceil \epsilon^{-1} \rceil} < L < t$. Therefore $p_{j+1} = \lceil \frac{t}{p_0} \rceil \lceil \frac{t}{p_1} \rceil \cdots \lceil \frac{t}{p_j} \rceil p_0$ for $j < \lceil \epsilon^{-1} \rceil$.

Since $p_{j+1} \leq p_j - \epsilon t$, then $p_{j+1} \leq p_j - \epsilon t \leq p_{j-1} - 2\epsilon t \leq \ldots \leq p - j\epsilon t$. Then $p_{\lceil \epsilon^{-1} \rceil} \leq p - \lceil \epsilon^{-1} \rceil \epsilon t \leq p - t$. But then $\epsilon t \leq p_0 \leq p_{\lceil \epsilon^{-1} \rceil} \leq p - t < 0$ is a contradiction. So there exists $0 \leq m < \lceil \epsilon^{-1} \rceil$ such that $0 < p_m$ (since $p_0 \neq 0$ as p and t are relatively prime) and either $p_m < \epsilon t$ or $p_m > p_{m-1} - \epsilon t$.

We now establish that in the bulk of the cases remaining, we can apply the generalized Blum-Hanson trick using p_m to determine the power used in the Block Lemma.

Case: $\dfrac{2t}{b_{\alpha_0 - 1}} \leq p_m < \epsilon t$

Set $g = \lceil \frac{t}{p_{m-1}} \rceil \cdots \lceil \frac{t}{p_0} \rceil$. We may assume m is the minimal choice such that $p_m < \epsilon t$ so $p_j \geq \epsilon t$ for all $0 \leq j < m$. If $\frac{t}{p_j} > \epsilon^{-1}$ then $t > \epsilon^{-1} p_j \geq \epsilon^{-1} \epsilon t = t$ is a contradiction, so $\frac{t}{p_j} \leq \epsilon^{-1}$ for all $0 \leq j < m$. Then $g \leq \lceil \epsilon^{-1} \rceil^m \leq \lceil \epsilon^{-1} \rceil^{\lceil \epsilon^{-1} \rceil - 1}$.

For $1 \leq i < \lceil (2\epsilon)^{-1} \rceil$, then $\dfrac{2t}{b_{\alpha_0 - 1}} \leq i p_m < \lceil (2\epsilon)^{-1} \rceil \epsilon t < \dfrac{1}{2} t + \epsilon t$ so $igp \pmod{t} \in \Big[\frac{2}{b_{\alpha_0 - 1}} t, (\frac{1}{2} + \epsilon) t \Big)$.

Then there exists $n_i, w_i \in \mathbb{Z}$ so that $igp = nt + w$ and $w \in \Big[\frac{2}{b_{\alpha_0 - 1}} t, (\frac{1}{2} + \epsilon) t \Big)$.

Since $\left| \dfrac{igy}{\tilde{h}_{\alpha_0}} - \dfrac{igp}{t} \right| < \dfrac{ig}{L b_{\alpha_0 - 1}} < \dfrac{2 \lceil \epsilon^{-1} \rceil^{\lceil \epsilon^{-1} \rceil}}{L b_{\alpha_0 - 1}} < \dfrac{1}{b_{\alpha_0 - 1}}$, we have $\left| igy - n_i \tilde{h}_{\alpha_0} - \dfrac{w_i}{t} \tilde{h}_{\alpha_0} \right| < \dfrac{1}{b_{\alpha_0 - 1}}$ so

$$\left| igy - n_i \tilde{h}_{\alpha_0} \right| > \frac{w_i}{t} \tilde{h}_{\alpha_0} - \frac{1}{b_{\alpha_0 - 1}} \tilde{h}_{\alpha_0} \geq \frac{2}{b_{\alpha_0 - 1}} \tilde{h}_{\alpha_0} - \frac{1}{b_{\alpha_0 - 1}} \tilde{h}_{\alpha_0} = \frac{1}{b_{\alpha_0 - 1}} \tilde{h}_{\alpha_0} > a_{\alpha_0 - 1} \tilde{h}_{\alpha_0 - 1}$$

and likewise $|igy - n_i \tilde{h}_{\alpha_0}| \leq (\frac{1}{2} + \epsilon) \tilde{h}_{\alpha_0} + \frac{1}{b_{\alpha_0 - 1}} \tilde{h}_{\alpha_0} < \tilde{h}_{\alpha_0} - a_{\alpha_0 - 1} \tilde{h}_{\alpha_0 - 1}$.

Since $igk = igz \tilde{h}_{\alpha_0} + igy$, then igk is of the form $v_i \tilde{h}_{\alpha_0} + \tilde{y}_i$ for $v_i, \tilde{y}_i \in \mathbb{Z}$ where $a_{\alpha_0 - 1} \tilde{h}_{\alpha_0 - 1} \leq \tilde{y}_i < \tilde{h}_{\alpha_0} - a_{\alpha_0 - 1} \tilde{h}_{\alpha_0 - 1}$. Therefore $|\lambda_B(T^{igk} B)| \leq \widetilde{M}_{B, \alpha_0} < \epsilon$ (or instead $|\lambda_B(T^{igk} B)| < M_{B, \alpha_0 - 1} < \epsilon$ when $v_i \geq a_{\alpha_0}$).

Since $g \lceil (2\epsilon)^{-1} \rceil < 2 \lceil \epsilon^{-1} \rceil^{\lceil \epsilon^{-1} \rceil} < \lceil \epsilon^{-1} \rceil^{-1} L < \epsilon L$, by Lemma 4.15 (with $p = g$ and $Q = \lceil (2\epsilon)^{-1} \rceil$),

$$\left| \frac{1}{L} \sum_{\ell=0}^{L-1} \lambda_B(T^{q - \ell k} B) \right| + \frac{1}{L} \sum_{\ell=0}^{L-1} (1 - \delta_\ell) \epsilon < (2\epsilon)^{1/2} + \epsilon + \epsilon < 4 \epsilon^{1/2}$$

Case: $\dfrac{2t}{b_{\alpha_0 - 1} \epsilon_0 L} \leq p_m < \dfrac{2t}{b_{\alpha_0 - 1}}$

Let $g \in \mathbb{N}$ minimal so that $g p_m \geq \frac{2t}{b_{\alpha_0 - 1}}$. Then $g p_m < \frac{4t}{b_{\alpha_0 - 1}}$ (if not then $(g-1) p_m > \frac{4t}{b_{\alpha_0 - 1}} - \frac{2t}{b_{\alpha_0 - 1}}$ contradicts that g is minimal). For $1 \leq i < \lceil \epsilon^{-1} \rceil$,

$$\frac{2t}{b_{\alpha_0 - 1}} \leq i g p_m < 4 \lceil \epsilon^{-1} \rceil \frac{t}{b_{\alpha_0 - 1}} < \frac{1}{2} t \qquad \text{meaning} \qquad ig \left\lceil \frac{t}{p_{m-1}} \right\rceil \cdots \left\lceil \frac{t}{p_0} \right\rceil p \pmod{t} \in \left[\frac{2t}{b_{\alpha_0 - 1}}, \frac{t}{2} \right)$$

Set $g_0 = g \lceil \frac{t}{p_{m-1}} \rceil \cdots \lceil \frac{t}{p_0} \rceil$. By the same reasoning as in the $\frac{2t}{b_{\alpha_0 - 1}} \leq p_m < \epsilon t$ case, $\lceil \frac{t}{p_j} \rceil \leq \lceil \epsilon^{-1} \rceil$ for all $0 \leq j < m$. Then

$$g_0 = g \left\lceil \frac{t}{p_{m-1}} \right\rceil \cdots \left\lceil \frac{t}{p_0} \right\rceil < \frac{4t}{b_{\alpha_0}} \frac{1}{p_m} \lceil \epsilon^{-1} \rceil^m \leq \frac{4t}{b_{\alpha_0}} \frac{b_{\alpha_0} \epsilon_0 L}{2t} \lceil \epsilon^{-1} \rceil^{\lceil \epsilon^{-1} \rceil} = 2 \epsilon_0 L \lceil \epsilon^{-1} \rceil^{\lceil \epsilon^{-1} \rceil} < \epsilon L$$

Since $i g_0 p \pmod{t} \in \Big[\frac{2t}{b_{\alpha_0 - 1}}, \frac{t}{2} \Big)$, Lemma 4.15 then gives the first conclusion.

Case: $0 < p_m < \dfrac{2t}{b_{\alpha_0-1}\epsilon_0 L}$ and $t < \epsilon_0 b_{\alpha_0-1} L$

Set $g = \lceil \frac{2t}{b_{\alpha_0-1}p_m} \rceil$ so $g < \frac{2t}{b_{\alpha_0-1}} + 1 < \frac{2\epsilon_0 b_{\alpha_0-1}L}{b_{\alpha_0-1}} + 1 = 2\epsilon_0 L + 1$. For $1 \leq i < \lceil \epsilon^{-1} \rceil$, then $\frac{2t}{b_{\alpha_0-1}} \leq igp_m < \lceil \epsilon^{-1} \rceil (2\epsilon_0 L + 1) \frac{2t}{b_{\alpha_0-1}\epsilon_0 L} < \frac{t}{2}$. Since $g \lceil \frac{t}{p_{m-1}} \rceil \cdots \lceil \frac{t}{p_0} \rceil < (2\epsilon_0 L + 1) \lceil \epsilon^{-1} \rceil^{\lceil \epsilon^{-1} \rceil - 1} < \lceil \epsilon^{-1} \rceil^{-2} L$, then, using the same reasoning as in the above cases, Lemma 4.15 gives the first conclusion.

Unfortunately, there is one remaining case for which no choice of p will allow us to apply the Blum-Hanson trick. However, in this case, we can establish that the potentially bad times $\mathcal{L}$ themselves have arithmetic-like structure with control over the gaps.

Case: $0 < p_m < \frac{2t}{b_{\alpha_0-1}\epsilon_0 L}$ and $t \geq \epsilon_0 b_{\alpha_0-1} L$

Since $|p| \geq \frac{2t}{b_{\alpha_0-1}\epsilon_0 L}$ and $p_m < \frac{2t}{b_{\alpha_0-1}\epsilon_0 L}$, it follows that $m > 0$. Set $g = \lceil \frac{t}{p_{m-1}} \rceil \cdots \lceil \frac{t}{p_0} \rceil < \lceil \epsilon^{-1} \rceil^{\lceil \epsilon^{-1} \rceil - 1} = \frac{1}{2}\epsilon_0^{-1} \lceil \epsilon^{-1} \rceil^{-2} < \epsilon^2 \epsilon_0^{-1}$. Since $gp \pmod{t} = p_m$, then there exists an integer $v > 0$ (since $m > 0$) such that $gp = vt + p_m$. Then $|gp - vt| = p_m < \frac{2}{b_{\alpha_0-1}\epsilon_0 L}$.
For $\ell \in \mathcal{L}$, recall that there exists $n_\ell \in \mathbb{Z}$ such that $|u - \ell p - n_\ell t| < \frac{1}{2} + \frac{2t}{b_{\alpha_0-1}}$. Let ℓ_0 minimal such that $\ell_0 \in \mathcal{L}$. Since $|p| \geq \frac{2t}{b_{\alpha_0-1}\epsilon_0 L}$, then $\frac{1}{b_{\alpha_0}} \leq \frac{1}{2}\epsilon_0 \frac{|p|L}{t}$ so

$$|n_\ell - n_{\ell_0}| \leq \left| \frac{u - \ell p}{t} - \frac{u - \ell_0 p}{t} \right| + \frac{1}{t} + \frac{4}{b_{\alpha_0-1}} < \frac{(\ell - \ell_0)|p|}{t} + \frac{5}{b_{\alpha_0-1}} < \frac{L|p|}{t} + \frac{5}{2}\epsilon_0 \frac{|p|L}{t} = \left(1 + \frac{5\epsilon_0}{2}\right) \frac{|p|L}{t}$$

Then, since $b_{\alpha_0-1}\epsilon_0 > 4\epsilon^{-1}$,

$$\begin{aligned} |(n_\ell - n_{\ell_0})vt - (n_\ell - n_{\ell_0})gp| &< |n_\ell - n_{\ell_0}| \frac{2t}{b_{\alpha_0-1}\epsilon_0 L} < \left(1 + \frac{5\epsilon_0}{2}\right) \frac{2}{b_{\alpha_0-1}\epsilon_0} |p| \\ &= \frac{2}{b_{\alpha_0-1}\epsilon_0}|p| + \frac{5}{b_{\alpha_0-1}}|p| < \frac{\epsilon}{2}|p| + \frac{5\epsilon\epsilon_0}{4}|p| \end{aligned}$$

and

$$\begin{aligned} |(n_\ell - n_{\ell_0})vt - v(\ell - \ell_0)p| &= |v|\,|(n_\ell - (u - \ell p)) - (n_{\ell_0} - (u - \ell_0 p))| < |v| \left(1 + \frac{4t}{b_{\alpha_0-1}}\right) \\ &< \frac{5|v|t}{b_{\alpha_0-1}} < \frac{5g|p|}{b_{\alpha_0-1}} < \frac{5\lceil \epsilon^{-1} \rceil^{\lceil \epsilon^{-1} \rceil - 1}}{b_{\alpha_0-1}}|p| < \frac{5\epsilon^2\epsilon_0^{-1}}{b_{\alpha_0-1}}|p| < \frac{5\epsilon^3}{4}|p| \end{aligned}$$

Therefore $|(n_\ell - n_{\ell_0})gp - v(\ell_0 - \ell)p| < \left(\frac{\epsilon}{2} + \frac{5\epsilon\epsilon_0}{4} + \frac{5\epsilon^3}{4}\right)|p|$ so $|(n_\ell - n_{\ell_0})g - v(\ell_0 - \ell)| < 1$. As these are integers, then $(n_\ell - n_{\ell_0})g = v(\ell_0 - \ell)$ for all $\ell \in \mathcal{L}$. Set $g^\star = \frac{g}{gcd(g,v)}$ and $v^\star = \frac{v}{gcd(g,v)}$. Then $g^\star$ divides $\ell_0 - \ell$ for all $\ell \in \mathcal{L}$ so every $\ell \in \mathcal{L}$ is of the form $\ell = \ell_0 + ig^\star$. Also for $0 \leq i < L$,

$$\begin{aligned} |ig^\star y - iv^\star \tilde{h}_{\alpha_0}| &= \frac{i\tilde{h}_{\alpha_0}}{gcd(g,v)} \left| \frac{gy}{\tilde{h}_{\alpha_0}} - v \right| < \frac{i\tilde{h}_{\alpha_0}}{gcd(g,v)} \left(\left| \frac{gp}{t} - v \right| + \frac{g}{b_{\alpha_0-1}L} \right) \\ &< L\tilde{h}_{\alpha_0} \left(\frac{2}{b_{\alpha_0-1}\epsilon_0 L} + \frac{\epsilon^2\epsilon_0^{-1}}{b_{\alpha_0-1}L} \right) = \frac{2 + \epsilon^2}{\epsilon_0 b_{\alpha_0-1}} \tilde{h}_{\alpha_0} < \frac{2 + \epsilon^2}{\epsilon_0 4\epsilon_0^{-1}\epsilon^{-1}} \tilde{h}_{\alpha_0} < (2\epsilon + \epsilon^3)\tilde{h}_{\alpha_0} \end{aligned}$$

so the second possible conclusion holds with $g^\star$ for t and $v^\star$ for p.

The above cases cover all possibilities where $p_m < \epsilon t$. Therefore we may assume from here on that

$$p_m > p_{m-1} - \epsilon t$$

Set $p^\star = \left\lfloor \dfrac{t}{p_{m-1}} \right\rfloor p_{m-1} = \left\lceil \dfrac{t}{p_{m-1}} \right\rceil p_{m-1} - p_{m-1} = p_m + t - p_{m-1} > t - \epsilon t$.

The remaining cases follow along similar lines as the cases where $p_n < \epsilon t$.

Note that $p_j \geq \epsilon t$ for all $j < m$ by the minimality of m so we again have that $\frac{t}{p_j} < \epsilon^{-1}$ for all $j < m$.

Case: $t - \epsilon t < p^\star \leq t - \dfrac{2t}{b_{\alpha_0 - 1}}$

For $1 \leq i < \lceil (2\epsilon)^{-1} \rceil$, $\frac{1}{2}t - \epsilon t < ip^\star \pmod{t} \leq t - \frac{2t}{b_{\alpha_0-1}}$. Then $i \lfloor \frac{t}{p_{m-1}} \rfloor \lceil \frac{t}{p_{m-2}} \rceil \cdots \lceil \frac{t}{p_0} \rceil p \pmod{t}$ is nonzero and at least $\frac{2t}{b_{\alpha_0-1}}$ away from every multiple of t. Since $\lceil \epsilon^{-1} \rceil \lfloor \frac{t}{p_{m-1}} \rfloor \cdots \lceil \frac{t}{p_0} \rceil < \lceil \epsilon^{-1} \rceil^{\lceil \epsilon^{-1} \rceil} \epsilon_0 L < \epsilon L$, as above Lemma 4.15 gives the claim.

Case: $t - \dfrac{2t}{b_{\alpha_0-1}} < p^\star \leq t - \dfrac{2t}{b_{\alpha_0-1}\epsilon_0 L}$

Let $g \in \mathbb{N}$ minimal such that $gp^\star \pmod{t} \leq t - \frac{2t}{b_{\alpha_0-1}}$. As in the case where $\frac{2t}{b_{\alpha_0-1}\epsilon_0 L} \leq p_m < \frac{2t}{b_{\alpha_0-1}}$, $g < 2\epsilon_0 L$ and then similar reasoning as there using Lemma 4.15 gives the claim.

Case: $t - \dfrac{2t}{b_{\alpha_0-1}\epsilon_0 L} < p^\star < t$ and $t \leq \epsilon_0 b_{\alpha_0} L$

Set $g = \lceil \frac{2t}{b_{\alpha_0-1}(t-p^\star)} \rceil < 2\epsilon_0 L + 1$. Then $igp^\star \pmod{t} < t - \frac{2t}{b_{\alpha_0-1}(t-p^\star)}(t - p^\star) = t - \frac{2t}{b_{\alpha_0-1}}$ and $igp^\star \pmod{t} > t - \lceil \epsilon^{-1} \rceil (2\epsilon_0 L + 1) \frac{2}{b_{\alpha_0-1}\epsilon_0 L}$ so again similar reasoning gives the claim.

Case: $t - \dfrac{2t}{b_{\alpha_0-1}\epsilon_0 L} < p^\star < t$ and $t \geq \epsilon_0 b_{\alpha_0-1} L$ and $|p| \geq \frac{2t}{b_{\alpha_0-1}\epsilon_0 L}$

Set $g = \lfloor \frac{t}{p_{m-1}} \rfloor \lceil \frac{t}{p_{m-2}} \rceil \cdots \lceil \frac{t}{p_0} \rceil < \lceil \epsilon^{-1} \rceil^{\lceil \epsilon^{-1} \rceil - 1}$. Then $p^\star = gp \pmod{t}$. Here, as above, $m \neq 0$ so there exists $v \neq 0$ such that $vt - \frac{2t}{b_{\alpha_0-1}\epsilon_0 L} < gp < vt$ and the same argument as in the $0 < p_m < \frac{2t}{b_{\alpha_0-1}\epsilon_0}$ case shows that the second possible conclusion holds.

Therefore the claim is proved as all cases have been covered. ☐

For clarity of exposition, we introduce notation for the amount of the leftmost subcolumns which will not be pushed through the top of the next tower upon application of $T^{q-\ell k}$. However, we will declare it to be 0 if $q - \ell k$ is already known to be a mixing time.

Notation 4.19. *For $\ell, q, k \in \mathbb{Z}$ and $n > 0$ such that $|q - \ell k| < \tilde{h}_{n+1}$, let d be the unique integer such that $|(q - \ell k) - d\tilde{h}_n| \leq \frac{1}{2}\tilde{h}_n$ and define*

$$\gamma_\ell^{n,q,k} = \begin{cases} \frac{a_n - |d|}{a_n} & \text{if } (b_n \leq |d| < a_n \text{ or } d = 0) \text{ and } |(q - \ell k) - d\tilde{h}_n| < a_{n-1}\tilde{h}_{n-1} \\ 0 & \text{otherwise} \end{cases}$$

Our next lemma is an application of Lemma 4.14 to precisely the set of potentially bad times $\mathcal{L}$ in the previous lemma. The critical point here is that when we split $T^{q-k(\ell_0+ti)}B$ into a $\frac{1}{b_{\alpha_0}} \sum_{j=0}^{b_{\alpha_0}}$ average, the resulting new q and k for each j is the same over all i. This relies in a crucial way on the bounded gaps and arithmetic structure established in the preceding lemma. Since Lemma 4.14 does produce an error term, the γ terms are necessary bookkeeping devices to ensure the total error is bounded.

Lemma 4.20. *Let $\epsilon > 0$. Let $q, k, t, p, \ell_0, L' \in \mathbb{Z}$ with $\ell_0, t, L' > 0$. Set $\alpha_0 = \max\{\alpha(q - k(\ell_0 + it)) : 0 \leq i < L'\}$.*

Assume that $b_{\alpha_0-1} > 2\epsilon^{-1}$ and $\sup_{m \geq b_{\alpha_0-1}} \left(\int \left| \frac{1}{m} \sum_{i=0}^{m-1} \chi_B \circ T^{-i} \right| d\mu + \frac{2}{m} \right) < \frac{\epsilon}{3}$.

Let $y, z \in \mathbb{Z}$ such that $k = z\tilde{h}_{\alpha_0} + y$ with $|y| \leq \frac{1}{2}\tilde{h}_{\alpha_0}$.

Assume, for $0 \leq i < L'$, that $\left| ity - ip\tilde{h}_{\alpha_0} \right| < \frac{1}{3}\tilde{h}_{\alpha_0}$.

For each i, let $q - k(\ell_0 + ti) = k_i\tilde{h}_{\alpha_0} + y_i$ with $|y_i| \leq \frac{1}{2}\tilde{h}_{\alpha_0}$. Assume that $|k_0| < a_{\alpha_0}$ and $|y_0| < a_{\alpha_0-1}\tilde{h}_{\alpha_0-1}$.

Then for each $0 \le j < b_{\alpha_0}$ *such that* $|k_i| < a_{\alpha_0}$ *and* $|y_i| < a_{\alpha_0-1}\tilde{h}_{\alpha_0-1}$*, there exists* $q_j, h_j \in \mathbb{Z}$ *such that* $\alpha_j = \max\{\alpha(q_j - h_j i) : 0 \le i < L'\} < \alpha_0$ *and for every* $0 \le i < L'$*,*

$$\left| \lambda_B(T^{q-k(\ell_0+it)}B) - \frac{r_{\alpha_0}}{r_{\alpha_0}+1}\gamma_{\ell_0+it}^{\alpha_0,q,k}\frac{1}{b_{\alpha_0}}\sum_{j=0}^{b_{\alpha_0}-1}\lambda_B(T^{q_j-h_j i}B) \right|$$

$$\le \left(1-\gamma_{\ell_0+it}^{\alpha_0,q,k}\right)\epsilon + \gamma_{\ell_0+it}^{\alpha_0,q,k}\frac{1}{b_{\alpha_0}}\sum_{j=0}^{b_{\alpha_0}-1}\left(1-\gamma_i^{\alpha_j,q_j,h_j}\right)\epsilon + \tau_{\alpha_0} + \frac{1}{a_{\alpha_0-1}b_{\alpha_0-1}}$$

Proof. We first use the bounded gaps property to write a concrete expression for k_i *and* y_i*, specifcally to write* $q - k(\ell_0 + ti)$ *as a multiple of* $\tilde{h}_{\alpha_0}$ *plus a remainder term where the remainder term and multiple term are linear functions of* i*.*

Since $|y_0 - ity + ip\tilde{h}_{\alpha_0}| < a_{\alpha_0-1}\tilde{h}_{\alpha_0-1} + \frac{1}{3}\tilde{h}_{\alpha_0} < \frac{1}{2}\tilde{h}_{\alpha_0}$ and since

$$\begin{aligned} q - (\ell_0+it)k &= q - \ell_0 k - it(z\tilde{h}_{\alpha_0} + y) = k_0\tilde{h}_{\alpha_0} + y_0 - itz\tilde{h}_{\alpha_0} - ity \\ &= (k_0 - itz)\tilde{h}_{\alpha_0} + y_0 - ity = (k_0 - itz - ip)\tilde{h}_{\alpha_0} + y_0 - ity + ip\tilde{h}_{\alpha_0} \end{aligned}$$

then $\quad k_i = k_0 - itz - ip \quad$ and $\quad y_i = y_0 - ity + ip\tilde{h}_{\alpha_0}$.

From the linear expressions, we can define what the new powers q_j *and* h_j *are and show that they are independent of the choice of* i*.*

For $0 \le j < b_{\alpha_0}$, set $\quad q_j = y_0 - jk_0 \quad$ and $\quad h_j = ty - p\tilde{h}_{\alpha_0} - jt - jp \quad$ and then

$$y_i - k_i j = y_0 - ity + ip\tilde{h}_{\alpha_0} - (k_0 - itz - ip)j = (y_0 - k_0 j) - (ty - p\tilde{h}_{\alpha_0} - jtz - jp)i = q_j - h_j i$$

Since $|k_i| < a_{\alpha_0}$ and $j < b_{\alpha_0}$, then $|q_j - h_j i| = |y_i - k_i j| < a_{\alpha_0-1}\tilde{h}_{\alpha_0-1} + a_{\alpha_0}b_{\alpha_0} < \tilde{h}_{\alpha_0}$. Therefore $\alpha_j \le \alpha_0 - 1$.

Consider first i such that $k_i = 0$. Then for any j, $q - (\ell_0 + it)k = y_i = y_i - k_i j = q_j - h_j i$ so

$$\lambda_B(T^{q-(\ell_0+it)}B) = \frac{1}{b_{\alpha_0}}\sum_{j=0}^{b_{\alpha_0}-1}\lambda_B(T^{q_j-h_j i}B)$$

Since $\gamma_{\ell_0+it}^{\alpha_0,q,k} = 1$ (as $k_i = 0$),

$$\left| \lambda_B(T^{q-k(\ell_0+it)}B) - \frac{r_{\alpha_0}}{r_{\alpha_0}+1}\gamma_{\ell_0+it}^{\alpha_0,q,k}\frac{1}{b_{\alpha_0}}\sum_{j=0}^{b_{\alpha_0}-1}\lambda_B(T^{q_j-h_j i}B) \right| = 1 - \frac{r_{\alpha_0}}{r_{\alpha_0}+1} < \frac{1}{a_{\alpha_0}b_{\alpha_0}} < \frac{1}{a_{\alpha_0-1}b_{\alpha_0-1}}$$

so the claim holds for i such that $k_i = 0$.

From here on, assume $k_i \neq 0$.

Next we bound the error term that will come from Lemma 4.14. For j *such that* $\alpha(q_j - h_j i) \le \alpha_0 - 2$*, this is straightforwardly bounded by* $\frac{1}{a_{\alpha_0-1}b_{\alpha_0-1}}$ *which we will later insist be summable. The* γ *terms appear exactly for* j *such that* $\alpha(q_j - h_j i) = \alpha_0 - 1$*.*

For i, j such that $\alpha(q_j - h_j i) < \alpha_0 - 1$, we have $\dfrac{|q_j - h_j i|}{\tilde{h}_{\alpha_0}} < \dfrac{\tilde{h}_{\alpha_0-1}}{\tilde{h}_{\alpha_0}} < \dfrac{1}{a_{\alpha_0-1}b_{\alpha_0-1}}$.

For i, j such that $\alpha(q_j - h_j i) = \alpha_0 - 1$, let $d_{i,j}$ be the unique integer such that $|q_j - h_j i - d_{i,j}\tilde{h}_{\alpha_0-1}| \le \frac{1}{2}\tilde{h}_{\alpha_0-1}$. Then $|d_{i,j}| \le a_{\alpha_0-1}$ since $|q_j - h_j i| < a_{\alpha_0-1}\tilde{h}_{\alpha_0-1} + a_{\alpha_0}b_{\alpha_0}$ so $\gamma_i^{\alpha_0-1,q_j,h_j} = \frac{a_{\alpha_0-1} - |d_{i,j}|}{a_{\alpha_0-1}}$. Then

$$\frac{|q_j - h_j i|}{\tilde{h}_{\alpha_0}} < \frac{(|d_{i,j}|+1)\tilde{h}_{\alpha_0-1}}{\tilde{h}_{\alpha_0}} < \frac{|d_{i,j}|+1}{a_{\alpha_0-1}b_{\alpha_0-1}} < \frac{2|d_{i,j}|}{a_{\alpha_0-1}b_{\alpha_0-1}} < \frac{|d_{i,j}|}{a_{\alpha_0-1}}\frac{\epsilon}{2} = \left(1 - \gamma_i^{\alpha_0-1,q_j,h_j}\right)\frac{\epsilon}{2}$$

Therefore, for every i with $|k_i| < a_{\alpha_0}$ and $k_i \neq 0$,

$$\frac{1}{b_{\alpha_0}} \sum_{j=0}^{b_{\alpha_0}-1} \frac{2|y_i - k_i j|}{\tilde{h}_{\alpha_0}} \leq \frac{1}{b_{\alpha_0}} \sum_{j=0}^{b_{\alpha_0}-1} \left(1 - \gamma_i^{\alpha_j,q_j,h_j}\right)\epsilon + \frac{1}{a_{\alpha_0-1}b_{\alpha_0-1}}$$

By Lemma 4.14,

$$\left|\lambda_B(T^{q-(\ell_0+ti)k}B) - \frac{a_{\alpha_0} - |k_{ii}|}{r_{\alpha_0}+1} \sum_{j=0}^{b_{\alpha_0}-1} \lambda_B(T^{y_i - k_i j}B)\right|$$
$$\leq \frac{a_{\alpha_0} - |k_i|}{a_{\alpha_0}} \frac{1}{b_{\alpha_0}} \sum_{j=0}^{b_{\alpha_0}-1} \frac{2|y_i - k_i j|}{\tilde{h}_{\alpha_0}} + \frac{|k|}{a_{\alpha_0}}\epsilon + \tau_{\alpha_0}$$

Since $\gamma_{\ell_0+it}^{\alpha_0,q,k} = \frac{a_{\alpha_0} - |k_i|}{a_{\alpha_0}}$, then

$$\left|\lambda_B(T^{q-(\ell_0+ti)k}B) - \frac{r_{\alpha_0}}{r_{\alpha_0}+1}\gamma_{\ell_0+it}^{\alpha_0,q,k} \sum_{j=0}^{b_{\alpha_0}-1} \lambda_B(T^{y_i - k_i j}B)\right|$$
$$\leq \gamma_{\ell_0+it}^{\alpha_0,q,k}\left(\frac{1}{b_{\alpha_0}} \sum_{j=0}^{b_{\alpha_0}-1} \left(1 - \gamma_i^{\alpha_j,q_j,h_j}\right)\epsilon + \frac{1}{a_{\alpha_0}b_{\alpha_0}}\right) + \left(1 - \gamma_{\ell_0+it}^{\alpha_0,q,k}\right)\epsilon + \tau_{\alpha_0} \qquad \square$$

Our final lemma addresses the situation when the potentially bad times themselves form a controlled arithmetic progression. In essence, it says that such a progression of potentially bad times can themselves be shown to lead to a 'weighted' average with (much) smaller gaps. The reason for the weights becomes apparent in the proof: $1-\gamma$ of the tower is already known to be mixing and γ of the tower might potentially not be. The proof is really just combining our previous lemmas without new ideas but performing necessary bookkeeping and sum rearranging.

Lemma 4.21. *Let $\epsilon > 0$ and set $\epsilon_0 = (2\lceil\epsilon^{-1}\rceil^{\lceil\epsilon^{-1}\rceil+1})^{-1}$. Let $L, k, q \in \mathbb{Z}$ with $L \geq \epsilon_0^{-1}$ and for each $0 \leq \ell < L$, let $0 \leq \delta_\ell \leq 1$.*

Let $\alpha_0 = \max\{\alpha(q - \ell k) : 0 \leq \ell < L\}$. Assume that $\max(M_{B,\alpha_0}, M_{B,\alpha_0-1}, \widehat{M}_{B,\alpha_0}, \widetilde{M}_{B,\alpha_0}) < \epsilon$ and $b_{\alpha_0-1} > 4\epsilon^{-1}\epsilon_0^{-1}$ and $\sup_{m \geq b_{\alpha_0-1}} \left(\int \left|\frac{1}{m}\sum_{i=0}^{m-1} \chi_B \circ T^{-i}\right| d\mu + \frac{2}{m}\right) < \frac{\epsilon}{3}$.

Then either

$$\left|\frac{1}{L}\sum_{\ell=0}^{L-1} \delta_\ell \lambda_B(T^{q-\ell k}B)\right| + \frac{1}{L}\sum_{\ell=0}^{L-1}(1-\delta_\ell)\epsilon < 4\epsilon^{1/2}$$

or there exist integers $0 < t, L' < L$ and, for each $0 \leq j < b_{\alpha_0}$, there exist $h_j, q_j \in \mathbb{Z}$ such that $\alpha_j = \max\{\alpha(q_j - h_j\ell) : 0 \leq \ell < L'\} < \alpha_0$ and

$$\left|\frac{1}{L}\sum_{\ell=0}^{L-1} \delta_\ell \lambda_B(T^{q-\ell k}B)\right| + \frac{1}{L}\sum_{\ell=0}^{L-1}(1-\delta_\ell)\epsilon$$
$$< \frac{1}{b_{\alpha_0}}\sum_{j=0}^{b_{\alpha_0}-1}\left(\left|\frac{1}{L}\sum_{\ell=0}^{L-1} \mathbb{1}_{\ell<L'}\delta_{\ell_0+\ell t}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\lambda_B(T^{q_j-h_j\ell}B)\right| + \frac{1}{L}\sum_{\ell=0}^{L-1}\left(1 - \mathbb{1}_{\ell<L'}\delta_{\ell_0+\ell t}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\right)\epsilon\right)$$
$$+ \frac{1}{b_{\alpha_0}}\sum_{j=0}^{b_{\alpha_0}-1}\frac{1}{L}\sum_{\ell=0}^{L-1}\mathbb{1}_{\ell<L'}\delta_{\ell_0+t\ell}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\left(1 - \gamma_\ell^{\alpha_j,q_j,h_j}\right)\epsilon + \tau_{\alpha_0} + \frac{1}{a_{\alpha_0-1}b_{\alpha_0-1}}$$

Proof. Write $q - \ell k = k_\ell \tilde{h}_{\alpha_0} + y_\ell$ with $|y_\ell| \leq \frac{1}{2}\tilde{h}_{\alpha_0}$. Define

$$\mathcal{L} = \left\{0 \leq \ell < L : (k_\ell = 0 \text{ or } b_{\alpha_0} \leq |k_\ell| < a_{\alpha_0}) \text{ and } |y_\ell| < a_{\alpha_0-1}\tilde{h}_{\alpha_0-1}\right\}$$

By Lemma 4.18, we may assume from here on that there exists $p \in \mathbb{Z}$, $t > 0$ and $0 \leq \ell_0 < L' \leq L$ such that $\mathcal{L} \subseteq \{\ell_0 + it : 0 \leq i < L'\}$ and $|ity - ip\tilde{h}_{\alpha_0}| < \frac{1}{3}\tilde{h}_{\alpha_0}$ for all $0 \leq i < L'$ as otherwise the claim follows from Lemma 4.18.

By Lemma 4.20 applied to each $\ell_0 + ti$,

$$\begin{aligned}
\left| \sum_{\ell_0+ti\in\mathcal{L}} \delta_{\ell_0+ti} \lambda_B(T^{q-(\ell_0+ti)k}B) \right| &< \left| \sum_{\ell_0+ti\in\mathcal{L}} \delta_{\ell_0+ti} \frac{r_{\alpha_0}}{r_{\alpha_0}+1} \gamma^{\alpha_0,q,k}_{\ell_0+it} \frac{1}{b_{\alpha_0}} \sum_{j=0}^{b_{\alpha_0}-1} \lambda_B(T^{q_j - h_j i}B) \right| \\
&\quad + \sum_{\ell_0+ti\in\mathcal{L}} \delta_{\ell_0+ti} \left(\left(1 - \gamma^{\alpha_0,q,k}_{\ell_0+it}\right)\epsilon + \gamma^{\alpha_0,q,k}_{\ell_0+it} \frac{1}{b_{\alpha_0}} \sum_{j=0}^{b_{\alpha_0}-1} \left(1 - \gamma_i^{\alpha_j,q_j,h_j}\right)\epsilon + \tau_{\alpha_0} + \frac{1}{a_{\alpha_0-1}b_{\alpha_0-1}} \right) \\
&< \left| \sum_{\ell_0+ti\in\mathcal{L}} \delta_{\ell_0+ti} \gamma^{\alpha_0,q,k}_{\ell_0+it} \frac{1}{b_{\alpha_0}} \sum_{j=0}^{b_{\alpha_0}-1} \lambda_B(T^{q_j-h_j i}B) \right| + \sum_{\ell_0+ti\in\mathcal{L}} \delta_{\ell_0+ti} \left(1 - \gamma^{\alpha_0,q,k}_{\ell_0+it}\right)\epsilon \\
&\quad + \sum_{\ell_0+ti\in\mathcal{L}} \delta_{\ell_0+ti} \gamma^{\alpha_0,q,k}_{\ell_0+it} \frac{1}{b_{\alpha_0}} \sum_{j=0}^{b_{\alpha_0}-1} \left(1 - \gamma_i^{\alpha_j,q_j,h_j}\right)\epsilon + \sum_{\ell_0+ti\in\mathcal{L}} \delta_{\ell_0+ti} \left(\tau_{\alpha_0} + \frac{1}{a_{\alpha_0-1}b_{\alpha_0-1}}\right)
\end{aligned}$$

Then, since $(1-\delta) + \delta(1-\gamma) = 1 - \delta\gamma$,

$$\begin{aligned}
&\left| \sum_{\ell_0+ti\in\mathcal{L}} \delta_{\ell_0+ti} \lambda_B(T^{q-(\ell_0+ti)k}B) \right| + \sum_{\ell_0+ti\in\mathcal{L}} (1 - \delta_{\ell_0+ti})\epsilon \\
&\quad < \frac{1}{b_{\alpha_0}} \sum_{j=0}^{b_{\alpha_0}-1} \left| \sum_{\ell_0+ti\in\mathcal{L}} \delta_{\ell_0+ti} \gamma^{\alpha_0,q,k}_{\ell_0+ti} \lambda_B(T^{q_j-h_j i}B) \right| + \sum_{\ell_0+ti\in\mathcal{L}} \left(1 - \delta_{\ell_0+ti}\gamma^{\alpha_0,q,k}_{\ell_0+ti}\right)\epsilon \\
&\qquad + \sum_{\ell_0+ti\in\mathcal{L}} \delta_{\ell_0+ti} \gamma^{\alpha_0,q,k}_{\ell_0+it} \frac{1}{b_{\alpha_0}} \sum_{j=0}^{b_{\alpha_0}-1} \left(1 - \gamma_i^{\alpha_j,q_j,h_j}\right)\epsilon + \sum_{\ell_0+ti\in\mathcal{L}} \delta_{\ell_0+ti} \left(\tau_{\alpha_0} + \frac{1}{a_{\alpha_0-1}b_{\alpha_0-1}}\right)
\end{aligned}$$

Since $\lambda_B(T^{-t}B) = \lambda_B(T^{t}B)$, then for $\ell \notin \mathcal{L}$, $|\lambda_B(T^{q-\ell k}B)| < \epsilon$ as it is bounded by one of M_{B,α_0}, M_{B,α_0-1}, $\widehat{M}_{B,\alpha_0}$ or $\widetilde{M}_{B,\alpha_0}$. For $\ell_0 + ti \notin \mathcal{L}$, since $\gamma^{\alpha_0,q,k}_{\ell_0+ti} = 0$ and $|\lambda_B(T^{q-(\ell_0+ti)}B)| < \epsilon$, then

$$\left|\delta_{\ell_0+ti}\lambda_B(T^{q-(\ell_0+ti)k}B)\right| + (1 - \delta_{\ell_0+ti})\epsilon < \epsilon = \left(1 - \delta_{\ell_0+ti}\gamma^{\alpha_0,q,k}_{\ell_0+ti}\right)\epsilon$$

Therefore, again using that $\gamma^{\alpha_0,q,k}_{\ell_0+ti} = 0$ for $\ell_0 + ti \notin \mathcal{L}$,

$$\begin{aligned}
&\left| \sum_{i=0}^{L'-1} \delta_{\ell_0+ti} \lambda_B(T^{q-(\ell_0+ti)k}B) \right| + \sum_{i=0}^{L'-1} (1 - \delta_{\ell_0+ti})\epsilon \\
&\quad < \frac{1}{b_{\alpha_0}} \sum_{j=0}^{b_{\alpha_0}-1} \left| \sum_{i=0}^{L'-1} \delta_{\ell_0+ti} \gamma^{\alpha_0,q,k}_{\ell_0+ti} \lambda_B(T^{q_j-h_j i}B) \right| + \sum_{i=0}^{L'-1} \left(1 - \delta_{\ell_0+ti}\gamma^{\alpha_0,q,k}_{\ell_0+ti}\right)\epsilon \\
&\qquad + \sum_{i=0}^{L'-1} \delta_{\ell_0+ti} \gamma^{\alpha_0,q,k}_{\ell_0+ti} \frac{1}{b_{\alpha_0}} \sum_{j=0}^{b_{\alpha_0}-1} \left(1 - \gamma_i^{\alpha_j,q_j,h_j}\right)\epsilon + \sum_{i=0}^{L'-1} \delta_{\ell_0+ti} \left(\tau_{\alpha_0} + \frac{1}{a_{\alpha_0-1}b_{\alpha_0-1}}\right)
\end{aligned}$$

Rewriting the sums over $0 \leq i < L'$ as over $0 \leq \ell < L$ with the indicator function $\mathbb{1}_{\ell<L'}$,

$$\begin{aligned}
&\left| \frac{1}{L} \sum_{i=0}^{L'-1} \delta_{\ell_0+ti} \lambda_B(T^{q-(\ell_0+ti)k}B) \right| + \frac{1}{L} \sum_{i=0}^{L'-1} (1 - \delta_{\ell_0+ti})\epsilon \\
&\quad < \frac{1}{b_{\alpha_0}} \sum_{j=0}^{b_{\alpha_0}-1} \left| \frac{1}{L} \sum_{\ell=0}^{L-1} \mathbb{1}_{\ell<L'} \delta_{\ell_0+t\ell} \gamma^{\alpha_0,q,k}_{\ell_0+t\ell} \lambda_B(T^{q_j-h_j\ell}B) \right| + \frac{1}{L} \sum_{\ell=0}^{L-1} \mathbb{1}_{\ell<L'} \left(1 - \delta_{\ell_0+t\ell}\gamma^{\alpha_0,q,k}_{\ell_0+t\ell}\right)\epsilon
\end{aligned}$$

$$+\frac{1}{L}\sum_{\ell=0}^{L-1}\mathbb{1}_{\ell<L'}\delta_{\ell_0+t\ell}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\frac{1}{b_{\alpha_0}}\sum_{j=0}^{b_{\alpha_0}-1}\left(1-\gamma_{\ell}^{\alpha_j,q_j,h_j}\right)\epsilon+\tau_{\alpha_0}+\frac{1}{a_{\alpha_0-1}b_{\alpha_0-1}}$$

Since $|\lambda_B(T^{\ell}B)|<\epsilon$ for $\ell\neq\ell_0+ti$ as then $\ell\notin\mathcal{L}$,

$$\left|\frac{1}{L}\sum_{\ell\neq\ell_0+ti}\delta_{\ell_0+ti}\lambda_B(T^{q-(\ell_0+ti)k}B)\right|+\frac{1}{L}\sum_{\ell\neq\ell_0+ti}(1-\delta_{\ell_0+ti})\epsilon<\frac{1}{L}\sum_{\ell\neq\ell_0+ti}\epsilon$$

$$=\frac{L-L'}{L}\epsilon=\frac{1}{L}\sum_{\ell=L'}^{L-1}\epsilon=\frac{1}{L}\sum_{\ell=0}^{L-1}(1-\mathbb{1}_{\ell<L'})\,\epsilon=\frac{1}{L}\sum_{\ell=0}^{L-1}\left(1-\mathbb{1}_{\ell<L'}\delta_{\ell_0+t\ell}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\right)\epsilon$$

Therefore the claim follows from the triangle inequality by splitting $0\leq\ell<L$ into $\{\ell_0+ti:0\leq i<L'\}$ and $\{0\leq\ell<L:\ell\neq\ell_0+ti\}$. □

We are now ready to prove mixing for the remaining times. The method of proof is to apply Lemma 4.21 iteratively in the sense that that lemma either gives that the average over the k powers is already mixed or that it can be reduced to a convex combination of 'weighted' averages of the same form Lemma 4.21 applies to. This iterative process must either terminate, in which case mixing is shown, or reach a point where the powers appearing the average are smaller than L, in which case mixing follows from the 'weighted' (weak) power ergodicity already established (it is crucial here that L not decrease).

Proof of Proposition 4.12. Fix $\epsilon>0$ and set $\epsilon_0=(2\lceil\epsilon^{-1}\rceil^{\lceil\epsilon^{-1}\rceil+1})^{-1}$. Using Propositions 4.8, 4.9, 4.10 and 4.16 and that $\sum_n\tau_n<\infty$ and $\sum_n\frac{1}{a_nb_n}<\infty$, there exists N such that

- $b_N>4\epsilon^{-1}\epsilon_0^{-1}$;
- $\sup_{m\geq N-1}M_{B,m}<\epsilon$;
- $\sup_{m\geq N}\widehat{M}_{B,m}<\epsilon$;
- $\sup_{m\geq N}\widetilde{M}_{B,m}<\epsilon$;
- $\sum_{n=N}^{\infty}\tau_n<\epsilon$;
- $\sum_{n=N-1}^{\infty}\frac{1}{a_nb_n}<\epsilon$; and
- $\sup_{m\geq b_N-1}\sup_{k\leq m}\left(\int\left|\frac{1}{m}\sum_{j=0}^{m-1}\chi_B\circ T^{-jk}\right|\,d\mu+\frac{2}{m}\right)<\frac{\epsilon}{3}$.

Take any n such that $b_n>\tilde{h}_{N+1}$. For $b_n\leq k<a_n$ and $|q|<a_{n-1}\tilde{h}_{n-1}$, by Lemma 4.14,

$$\left|\lambda_B(T^{k\tilde{h}_n+q}B)\right|<\frac{a_n-k}{a_n}\left|\frac{1}{b_n}\sum_{\ell=0}^{b_n-1}\lambda_B(T^{q-k\ell}B)\right|+\frac{k}{a_n}\epsilon+\tau_n<\left|\frac{1}{b_n}\sum_{\ell=0}^{b_n-1}\lambda_B(T^{q-k\ell}B)\right|+2\epsilon$$

Set $L=b_n$. By Lemma 4.21, $\left|\frac{1}{L}\sum_{\ell=0}^{L-1}\lambda_B(T^{q-k\ell}B)\right|<4\epsilon^{1/2}$ or there exists $q_{\ell'},k'_{\ell'},L',\ell_0,t$ such that

$$\left|\frac{1}{L}\sum_{\ell=0}^{L-1}\lambda_B(T^{q-k\ell}B)\right|<\frac{1}{b_{\alpha_0}}\sum_{\ell'=0}^{b_{\alpha_0}-1}\left(\left|\frac{1}{L}\sum_{\ell=0}^{L-1}\mathbb{1}_{\ell<L'}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\lambda_B(T^{q_{\ell'}-k'_{\ell'}\ell}B)\right|+\frac{1}{L}\sum_{\ell=0}^{L-1}\left(1-\mathbb{1}_{\ell<L'}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\right)\epsilon\right)$$

$$+\frac{1}{b_{\alpha_0}}\sum_{\ell'=0}^{b_{\alpha_0}-1}\frac{1}{L}\sum_{\ell=0}^{L-1}\mathbb{1}_{\ell<L'}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\left(1-\gamma_{\ell}^{\alpha_{\ell'},q_{\ell'},k'_{\ell'}}\right)\epsilon+\tau_{\alpha_0}+\frac{1}{a_{\alpha_0-1}b_{\alpha_0-1}}$$

The first average above consists of terms which are exactly of the form that Lemma 4.21 applies to. By Lemma 4.21, some of those terms are bounded by $4\epsilon^{1/2}$ and the rest likewise split into a convex combination of terms of that same form so our aim is to iterate this process. At each stage of this process, the 'error bound' increases by τ_α and $\frac{1}{a_\alpha b_\alpha}$ (for strictly decreasing α) and by a term similar to the 'extra ϵ' term above involving two γ's.

Let $\mathcal{L}' = \{0 \leq \ell' < b_{\alpha_0} : \alpha_{\ell'} > N$ and Lemma 4.21 does not bound the ℓ' weighted average by $4\epsilon^{1/2}\}$. Then

$$\frac{1}{b_{\alpha_0}}\sum_{\ell'=0}^{b_{\alpha_0}-1}\left(\left|\frac{1}{L}\sum_{\ell=0}^{L-1}\mathbb{1}_{\ell<L'}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\lambda_B(T^{q_{\ell'}-k'_{\ell'}\ell}B)\right| + \frac{1}{L}\sum_{\ell=0}^{L-1}\left(1-\mathbb{1}_{\ell<L'}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\right)\epsilon\right) < \left(1-\frac{|\mathcal{L}'|}{b_{\alpha_0}}\right)4\epsilon^{1/2}$$
$$+\frac{|\mathcal{L}'|}{b_{\alpha_0}}\frac{1}{|\mathcal{L}'|}\sum_{\ell'\in\mathcal{L}'}\left(\left|\frac{1}{L}\sum_{\ell=0}^{L-1}\mathbb{1}_{\ell<L'}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\lambda_B(T^{q_{\ell'}-k'_{\ell'}\ell}B)\right| + \frac{1}{L}\sum_{\ell=0}^{L-1}\left(1-\mathbb{1}_{\ell<L'}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\right)\epsilon\right)$$

Therefore, applying Lemma 4.21 to each ℓ' weighted average, since $\alpha_{\ell'} \leq \alpha_0 - 1$ (and suppressing the explicit dependence on ℓ' of L'', ℓ'_0, t' for clarity),

$$\left|\frac{1}{L}\sum_{\ell=0}^{L-1}\lambda_B(T^{q-k\ell}B)\right| < \left(1-\frac{|\mathcal{L}'|}{b_{\alpha_0}}\right)4\epsilon^{1/2} + \tau_{\alpha_0} + \tau_{\alpha_0-1} + \frac{1}{a_{\alpha_0-1}b_{\alpha_0-1}} + \frac{1}{a_{\alpha_0-1}b_{\alpha_0-1}}$$
$$+\frac{|\mathcal{L}'|}{b_{\alpha_0}}\frac{1}{|\mathcal{L}'|}\sum_{\ell'\in\mathcal{L}'}\frac{1}{b_{\alpha_{\ell'}}}\sum_{\ell''=0}^{b_{\alpha_{\ell'}}-1}\left(\left|\frac{1}{L}\sum_{\ell=0}^{L-1}\mathbb{1}_{\ell<L''}\mathbb{1}_{\ell'_0+t'\ell<L'}\gamma_{\ell_0+t(\ell'_0+t'\ell)}^{\alpha_0,q,k}\gamma_{\ell'_0+t'\ell}^{\alpha_{\ell'},q_{\ell'},k'_{\ell'}}\lambda_B(T^{q_{\ell',\ell''}-k_{\ell',\ell''}\ell}B)\right|\right.$$
$$\left.+\frac{1}{L}\sum_{\ell=0}^{L-1}\left(1-\mathbb{1}_{\ell<L''}\mathbb{1}_{\ell'_0+t'\ell<L'}\gamma_{\ell_0+t(\ell'_0+t'\ell)}^{\alpha_0,q,k}\gamma_{\ell'_0+t'\ell}^{\alpha_{\ell'},q_{\ell'},k'_{\ell'}}\right)\epsilon\right)$$
$$+\frac{|\mathcal{L}'|}{b_{\alpha_0}}\frac{1}{|\mathcal{L}'|}\sum_{\ell'\in\mathcal{L}'}\frac{1}{b_{\alpha_{\ell'}}}\sum_{\ell''=0}^{b_{\alpha_{\ell'}}-1}\frac{1}{L}\sum_{\ell=0}^{L-1}\left(\mathbb{1}_{\ell<L'}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\left(1-\gamma_{\ell}^{\alpha_{\ell'},q_{\ell'},k'_{\ell'}}\right)\right.$$
$$\left.+\mathbb{1}_{\ell<L''}\mathbb{1}_{\ell'_0+t'\ell<L'}\gamma_{\ell_0+t(\ell'_0+t'\ell)}^{\alpha_0,q,k}\gamma_{\ell'_0+t'\ell}^{\alpha_{\ell'},q_{\ell'},k'_{\ell'}}\left(1-\gamma_{\ell}^{\alpha_{\ell''},q_{\ell',\ell''},k_{\ell',\ell''}}\right)\right)\epsilon$$

We now consider the 'extra ϵ' terms resulting from this process. At this stage, there are two such terms.

Now observe that

$$\frac{1}{L}\sum_{\ell=0}^{L-1}\left(\mathbb{1}_{\ell<L'}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\left(1-\gamma_{\ell}^{\alpha_{\ell'},q_{\ell'},k'_{\ell'}}\right)+\mathbb{1}_{\ell<L''}\mathbb{1}_{\ell'_0+t'\ell<L'}\gamma_{\ell_0+t(\ell'_0+t'\ell)}^{\alpha_0,q,k}\gamma_{\ell'_0+t'\ell}^{\alpha_{\ell'},q_{\ell'},k'_{\ell'}}\left(1-\gamma_{\ell}^{\alpha_{\ell''},q_{\ell',\ell''},k_{\ell',\ell''}}\right)\right)$$
$$=\frac{1}{L}\sum_{\substack{\ell\neq\ell'_0+t'i\\ \ell<L'}}\gamma_{\ell_0+t\ell}^{\alpha_0,q,k}\left(1-\gamma_{\ell}^{\alpha_{\ell'},q_{\ell'},k'_{\ell'}}\right)+\frac{1}{L}\sum_{\ell=0}^{L''-1}\mathbb{1}_{\ell'_0+t'\ell<L'}\gamma_{\ell_0+t(\ell'_0+t'\ell)}^{\alpha_0,q,k}\left(1-\gamma_{\ell'_0+t'\ell}^{\alpha_{\ell'},q_{\ell'},k'_{\ell'}}\gamma_{\ell}^{\alpha_{\ell''},q_{\ell',\ell''},k_{\ell',\ell''}}\right)$$

and that the sets of the original $0 \leq \ell < L$ the two sums range over are disjoint. The left sum in the above is precisely those ℓ for which we do not need to continue iteratively applying Lemma 4.21 (as they are exactly those which that lemma will bound by $4\epsilon^{1/2}$). The right sum in the above are those ℓ for which the iterative process will need to continue.

The key point is that the terms on the right have the form $\gamma(1-\gamma'\gamma'')$ rather than just $\gamma(1-\gamma')$.

Continue iteratively applying Lemma 4.21 until all terms are bounded by $4\epsilon^{1/2}$ or have $k_{\ell',\ell'',\dots} < L$, which must occur as α decrements at each application of the lemma and if $k_{\ell,\ell',\dots} \geq L = b_n > \tilde{h}_{N+1}$ then $\alpha_{\ell',\ell'',\dots} \geq N+1$ (and the hypotheses of the lemma hold as long as $\alpha_{\ell',\ell'',\dots} > N$).

If the process terminates with one or more terms having k values less than L then the already established weighted weak power ergodicity (Proposition 4.16) tells us those terms are already bounded by ϵ.

If $k_{\ell',\ell'',\dots} < L$ then, by the seventh requirement on the choice of N, as $L = b_n > \tilde{h}_{N+1} > b_N$,

$$\left|\frac{1}{L}\sum_{\ell=0}^{L-1}\lambda_B(T^{q_{\ell',\ell'',\dots}-\ell k_{\ell',\ell'',\dots}}B)\right| \leq \int\left|\frac{1}{L}\sum_{\ell=0}^{L-1}\chi_B\circ T^{-\ell k_{\ell',\ell'',\dots}}\right|\,d\mu < \frac{\epsilon}{3}$$

The 'extra ϵ' terms weighted by γ terms will be of the form

$$\gamma(1-\gamma')+\gamma\gamma'(1-\gamma'')+\cdots+\gamma\gamma'\cdots\gamma^{(m-1)}(1-\gamma^{(m)})=\gamma(1-\gamma\gamma'\cdots\gamma^{(m-1)}\gamma^{(m)})$$

where m is the number of applications of Lemma 4.21 needed for that ℓ (this is written carefully in the case of two applications above).

Then $\left|\frac{1}{L}\sum_{\ell=0}^{L-1}\lambda_B(T^{q-k\ell}B)\right|$ is bounded by a convex combination of terms less than $4\epsilon^{1/2}$ plus a sum of τ's bounded by $\sum_{n=N}^{\infty}\tau_n<\epsilon$ plus a sum of $\frac{1}{ab}$ terms bounded by $\sum_{n=N}^{\infty}\frac{1}{a_nb_n}<\epsilon$ plus an average over $0\leq\ell<L$ of terms of the form

$$\gamma^{\alpha_0}\,(1-\gamma^{\alpha_{\ell'}}\gamma^{\alpha_{\ell''}}\cdots\gamma^{\alpha_{\ell^{(m)}}})\,\epsilon$$

which are all bounded by ϵ as $0\leq\gamma\leq 1$. Therefore

$$\left|\frac{1}{L}\sum_{\ell=0}^{L-1}\lambda_B(T^{q-k\ell}B)\right|<4\epsilon^{1/2}+\epsilon+\epsilon+\epsilon\qquad\text{meaning that}\qquad\left|\lambda_B(T^{k\tilde{h}_n+q}B)\right|<4\epsilon^{1/2}+5\epsilon$$ ☐

4.3. Proof of mixing

Proof of Theorem 4.1. By Propositions 4.8, 4.9, 4.10 and 4.12, for any B which is a union of levels in some C_N, $\lim_{n\to\infty}\max_{\tilde{h}_n\leq t<\tilde{h}_{n+1}}|\lambda_B(T^tB)|=0$. As unions of levels generate the measure algebra, T is Renyi mixing hence mixing. ☐

5. Non-superlinear word complexity implies partial rigidity

Theorem 5.1. *Let X be a subshift with word complexity p such that $\liminf\frac{p(q)}{q}<\infty$. Then there exists a constant $\delta_X>0$ such that every ergodic probability measure μ on X is at least δ_X-partially rigid.*

5.1. Word combinatorics

Notation 5.2. *For x a finite or infinite word and $-\infty\leq i<j\leq\infty$,*

$$x_{[i,j)}=\text{ the subword of } x \text{ from position } i \text{ through position } j-1$$

Notation 5.3. $[w]=\{x\in X: x_{[0,\|w\|)}=w\}$ *for finite words w.*

Notation 5.4. *For a word v and $0\leq q<\|v\|$, let $v^{q/\|v\|}$ be the suffix of v of length q.*
Let $v^{n+q/\|v\|}=v^{q/\|v\|}v^n$ for $n\in\mathbb{N}$.

Definition 5.5. Let $w\in\mathcal{L}(X)$ be a word in the language of a subshift. A word $v\in\mathcal{L}(X)$ is a **root** of w if $wv\in\mathcal{L}(X)$ and $\|v\|\leq\|w\|$ and w is a suffix of v^∞, i.e. there exists $q={p}/{\|v\|}$ with $p\geq\|v\|$ such that $w=v^q$. The **minimal root** of w is the shortest v which is a root of w.

Every word has a unique minimal root as it is a root of itself.

Lemma 5.6. *If $uw=wv$ and $\|v\|\leq\|w\|$ then v is a root of w.*

Proof. As w has v as a suffix, $w=w'v$. Then $uw'v=uw=wv=w'vv$ so $uw'=w'v$. If $\|w'\|\geq\|v\|$, repeat this process until it terminates at $w=w''v^n$ with $\|w''\|<\|v\|$. Then $uw''=w''v$ so w'' is a suffix of v. ☐

Lemma 5.7. *If $uv=vu$ then $u=v_0^t$ and $v=v_0^s$ for some word v_0 and $t,s\in\mathbb{N}$.*

Proof. If $\|u\|=\|v\|$ then $uv=vu$ immediately implies $u=v$. Let

$$V=\{(u,v):uv=vu,\|v\|<\|u\|,\text{ there is no word } v_0 \text{ with } u=v_0^t \text{ and } v=v_0^s \text{ for } s,t\in\mathbb{N}\}$$

and suppose $V \neq \emptyset$. Let $(u,v) \in V$ such that $\|u\|$ is minimal. As $\|u\| > \|v\|$, $uv = vu$ implies $u = vu' = u''v$ for some nonempty words u', u''. Then $vu'v = uv = vu = vu''v$ so $u' = u''$ and $vu' = u'v$. If $\|u'\| = \|v\|$ then $u' = v$ so $u = v^2$ contradicting that $(u,v) \in V$.

Consider when $\|u'\| < \|v\|$. Since $\|u'\| < \|u\|$ and $\|v\| < \|u\|$, the minimality of $\|u\|$ implies that $(v,u') \notin V$. Then $v = v_0^n$ and $u' = v_0^m$ for some word v_0 and $n, m \in \mathbb{N}$. So $u = v_0^{n+m}$ meaning $(u,v) \notin V$. When $\|v\| < \|u'\|$, we have $(u',v) \notin V$ so $u' = v_0^n$ and $u = v_0^{n+m}$. So $V = \emptyset$. □

Lemma 5.8. *If u and v are both roots of a word w and uu is a suffix of w and $\|v\| < \|u\|$ then there exists a suffix v_0 of v such that $u = v_0^n$ and $v = v_0^m$ for some $n, m \in \mathbb{N}$.*

In particular, if v is the minimal root of w and u is a root of w and uu is a suffix of w then u is a multiple of v, i.e. there exists $n \in \mathbb{N}$ such that $u = v^n$.

Proof. Writing u' and v' for the appropriate suffixes of u and v, we have $w = u'u^t = v'v^q$ for some $t, q \in \mathbb{N}$. Then $u = u_0v^a$ for some proper (possibly empty) suffix u_0 of v and $1 \leq a \leq q$. So $u'(u_0v^a)^t = v'v^q$ meaning that $u'(u_0v^a)^{t-1}u_0 = v'v^{q-a}$. As $t \geq 2$, $\|v'v^{q-a}\| = \|u'(u_0v^a)^{t-1}u_0\| \geq \|u_0v^au_0\| \geq \|vu_0\|$ so, as u_0 is a suffix of v, then $v'v^{q-a}$ has u_0v as a suffix. This means $vu_0 = u_0v$ so Lemma 5.7 gives v_0 such that $v = v_0^n$ and $u_0 = v_0^m$ so $u = v_0^{m+an}$. If v is the minimal root then $v = v_0$ since v_0 is a root of w. □

Lemma 5.9. *Let w be a word with minimal root v. If $0 \leq i \leq \frac{1}{2}\|w\|$ and $T^i[w] \cap [w] \neq \emptyset$ then i is a multiple of $\|v\|$.*

Proof. Let u be the prefix of B of length i and v_0 be the suffix of B of length i. For $x \in T^i[w] \cap [w]$, then $x_{[-i,\|w\|)} = uw = wv_0$. By Lemma 5.6, then v_0 is a root of w. As $\|v_0\| = i \leq \frac{1}{2}\|w\|$, w has v_0v_0 as a suffix. By Lemma 5.8, since v is the minimal root then v_0 is a multiple of v. □

5.2. Language analysis

Proposition 5.10. *There exists $C, k > 0$, depending only on X, and $\ell_n \to \infty$ and, for each n, at most C words $B_{n,j}$ so that $X_0 = \{x \in X :$ every finite subword of x is a subword of a concatenation of the $B_{n,j}\}$ has measure one.*

Let $h_{n,j} = \|B_{n,j}\|$. Then $\max_j h_{n,j} \leq k\ell_n$ and $\min_j h_{n,j} \to \infty$. Let

$$W_{B_{n,j}} = W_{n,j} = \{x \in X_0 : x \text{ can be written as a concatenation such that } x_{[0,h_{n,j})} = B_{n,j}\} \subseteq [B_{n,j}]$$

There exists $c_{n,j} \leq k\ell_n$ such that the sets $T^iW_{n,j}$ are disjoint over $0 \leq i < c_{n,j}$.

For j such that $h_{n,j} > \frac{1}{2}\ell_n$, $c_{n,j} \geq \frac{1}{2}\ell_n$.

For j such that $h_{n,j} \leq \frac{1}{2}\ell_n$, $c_{n,j} = h_{n,j}$. For such j, also $W_{n,j} = T^{\ell_n}[B_{n,j}^{\ell_n/h_{n,j}}B_{n,j}]$ and $B_{n,j}$ is the minimal root of $B_{n,j}^{\ell_n/h_{n,j}}B_{n,j}$.

If $x \in T^{h_{n,j}}W_{n,j} \cap W_{n,j'}$ for $j \neq j'$ and $h_{n,j'} \leq \frac{1}{2}\ell_n$ then $x_{(-\infty,0)}$ has $B_{n,j'}^{\ell_n/h_{n,j'}}$ as a suffix and does not have $B_{n,j'}^{\ell_n/h_{n,j'}}B_{n,j'}$ as a suffix.

Proof. Since $\liminf \frac{p(q)}{q} < \infty$, [Bos85] Theorem 2.2 gives a constant k and $\ell_n \to \infty$ such that $p(\ell_n + 1) - p(\ell_n) \leq k$ and $p(\ell_n) \leq k\ell_n$. We perform an analysis similar to Ferenczi [Fer96] Proposition 4.

Let G_q be the Rauzy graphs: the vertices are the words of length q in $\mathcal{L}(X)$ and the directed edges are from words w to w' such that $wa = bw' \in \mathcal{L}(X)$ for some letters a and b and we label the edge with the letter a. As μ is ergodic, exactly one strongly connected component has measure one and the rest have measure zero so we may assume G_q is strongly connected.

Let V_q^{RS} be the set of all vertices with more than one outgoing edge, i.e. the right-special vertices. Let $\mathcal{B}_q$ be the set of all paths from some $v \in V_q^{RS}$ to some $v' \in V_q^{RS}$ that do not pass through any $v'' \in V_q^{RS}$. Then every $v \in V_q$ is necessarily along such a path. Given any word w in $\mathcal{L}(X)$, there exists $x \in X$ such that $x_{[0,\|w\|)} = w$ so w is the label of the path from the vertex corresponding to $x_{[-q,0)}$ to the vertex corresponding to $x_{[\|w\|-q,\|w\|)}$ hence is a subword of some concatenation of labels of paths in $\mathcal{B}_q$.

The labels of the paths between right-special vertices are nested: $\mathcal{B}_{q+1}$ is necessarily a concatenation of paths in $\mathcal{B}_q$ since words corresponding to elements of V_{q+1}^{RS} necessarily have right-special suffixes. There are therefore recursion formulas defining $\mathcal{B}_{q+1}$ in terms of $\mathcal{B}_q$ though we do not make use of this fact.

Writing $\operatorname{outdeg}(v)$ for the number of outgoing edges of a vertex, $\sum_{v \in V_{\ell_n}^{RS}}(\operatorname{outdeg}(v) - 1) = p(\ell_n + 1) - p(\ell_n) \leq k$ meaning that $|V_{\ell_n}^{RS}| \leq k$ and therefore $\sum_{v \in V_{\ell_n}^{RS}} \operatorname{outdeg}(v) \leq 2k$. Therefore $|\mathcal{B}_{\ell_n}| \leq 2k$. No path in $\mathcal{B}_{\ell_n}$ properly contains a cycle so $\|B\| \leq p(\ell_n) \leq k\ell_n$ for any label B of a path in $\mathcal{B}_{\ell_n}$.

Let $\mathcal{B}_n^g$ be the set of all concatenations of paths in $\mathcal{B}_{\ell_n}$ of total length at least $\frac{3}{2}\ell_n$ and at most $k\ell_n$ not properly containing any cycles. As such a path contains no cycle properly, it has at most $|\mathcal{B}_{\ell_n}| \leq 2k$ segments from some vertex in $V_{\ell_n}^{RS}$ to another, so there are at most $K = \sum_{t=1}^{2k}(2k)^t$ such paths.

Let $\mathcal{B}_n^c$ be the set of all concatenations of paths in $\mathcal{B}_{\ell_n}$ of total length less than $\frac{3}{2}\ell_n$ which are simple cycles. Then $|\mathcal{B}_n^c| \leq K$ as each path has at most $2k$ segments and at most $2k$ choices for each segment. Every bi-infinite concatenation of paths in $\mathcal{B}_{\ell_n}$ is necessarily a concatenation of paths in $\mathcal{B}_n^g \cup \mathcal{B}_n^c$.

Let B be the label of a path in $\mathcal{B}_n^g$ and let v be its minimal root. Suppose that $\|v\| < \frac{1}{2}\ell_n$. Then the vertex at which the path corresponding to B ends is the word $v^{\ell_n/\|v\|}$ as it must be a suffix of B. Let B' such that $B = B'v$. Then $\|B'\| = \|B\| - \|v\| \geq \frac{3}{2}\ell_n - \|v\| > \ell_n$. Then the path corresponding to B reaches its final vertex twice as B' has suffix $v^{\ell_n/\|v\|}$ corresponding to that vertex. This means the path properly contains a cycle which is a contradiction. So all labels of paths in $\mathcal{B}_n^g$ have minimal root of length at least $\frac{1}{2}\ell_n$. By Lemma 5.9, then $T^i W_{n,j} \cap W_{n,j} \neq \emptyset$ for $0 < i \leq \frac{1}{2}\|B\|$ only when i is a multiple of $\|v\|$. Set $c_{n,j} = \min(\|v\|, \frac{1}{2}\|B\|) \geq \frac{1}{2}\ell_n$ and then $T^i W_{n,j}$ are disjoint over $0 \leq i < c_{n,j}$.

Let B be the label of a simple cycle beginning and ending at the word w. Since B is the label of a path beginning at w, every appearance of B as a label in $x \in X$ is preceded by w, i.e. $W_B \subseteq T^{\ell_n}[wB]$. Since B either has w as a suffix or B is a root of w by Lemma 5.6, B is a root of wB. Let v be the minimal root of wB and write $B = B'v$. Then wB' has v as a root and $\|wB'\| = \ell_n + \|B'\|$ so wB' has suffix $v^{\ell_n/\|v\|}$. If B' is nonempty then the path corresponding to B passes through its final vertex before the path ends, contradicting that it is a simple cycle. So $B = v$ is the minimal root of wB.

Then Lemma 5.9 implies that $T^i W_B \cap W_B \neq \emptyset$ for $0 < i \leq \frac{1}{2}\|wB\|$ only when i is a multiple of $\|B\|$. So if $\|B\| > \frac{1}{2}\ell_n$ then set $c_{n,j} = \min(\|B\|, \frac{1}{2}\|wB\|) > \frac{1}{2}\ell_n$. If $\|B\| \leq \frac{1}{2}\ell_n$, set $c_{n,j} = \|B\|$. For such B, since $W_B \subseteq T^{\ell_n}[wB]$, we have that every occurrence of B as a label of a path is preceded by $w = B^{\ell_n/\|B\|}$. Moreover, if $x_{[-\ell_n,\|B\|)} = wB$ then $x_{[0,\|B\|)}$ is the label of a path beginning at the vertex w and ending at w so $x \in W_B$.

For $x \in W_B$, if $x_{(-\infty,0)}$ has $B^{\ell_n/\|B\|}B$ as a suffix then the path reaches w prior to the final B in that suffix. As no word B' appearing in the concatenation is the label of a path properly containing a cycle, this means the word preceding $x_{[0,\|B\|)} = B$ in x must be B, i.e. $x \in T^{\ell_n+\|B\|}[B^{\ell_n/\|B\|}B]$ so $x \in T^{\|B\|}W_B \cap W_B$ and $x \notin T^{\|B'\|}W_{B'} \cap W_B$ for every $B' \neq B$ as the path for B' does not properly contain a cycle.

Let $\mathcal{B}_n^* = \mathcal{B}_n^g \cup \mathcal{B}_n^c$. Then $|\mathcal{B}_n^*| \leq 2K = C$ for all n and every word in $\mathcal{L}(X)$ is a subword of some concatenation of labels of paths in $\mathcal{B}_n^*$. Let $\mathcal{R}_n$ be the set of all labels of paths in $\mathcal{B}_n^*$.

Let $\mathcal{D}_M = \{B : \|B\| \leq M \text{ and } B \in \mathcal{R}_n \text{ infinitely often}\}$. Then $|D_M| < \infty$ as there only finitely many words of length at most M (as non-superlinear complexity implies finite alphabet rank [DDMP21]). Let X_M be the set of $x \in X$ such that for infinitely many n, x cannot be written as a concatenation of labels in $\mathcal{B}_n^*$ without using at least one label in $\mathcal{D}_M$.

For $x \in X_M$, there exist infinitely many t such that x has $B_t^{r_t}$ as a subword for some $B_t \in \mathcal{D}_M$ and $r_t \to \infty$ (since the label B_t is preceded by the word $B_t^{\lfloor \ell_n/(\|B_t\|) \rfloor}$). As $|\mathcal{D}_M| < \infty$, there exists B such that $B_t = B$ infinitely often. Then B^{r_t} is a subword of x for $r_t \to \infty$ meaning x is periodic. Therefore $\bigcup_M X_M \subseteq \{\text{periodic words}\}$ so $\mu(\bigcup_M X_M) = 0$ as μ is ergodic hence nonatomic and a periodic word of positive measure would be an atom (there are at most countably many periodic words).

Define $\{B_{n,j}\}$ to be the set of all labels of paths in $\mathcal{B}_n^*$ which are in $\mathcal{R}_n \setminus \bigcup_M \mathcal{D}_M$. If $\liminf_n \min_j \|B_{n,j}\| <$

∞ then $B_{n,j} = B$ for some fixed B infinitely often (as there are only finitely many words of up to some fixed length). But then $B \in \mathcal{D}_{\|B\|}$, a contradiction, so $\lim_n \min_j \|B_{n,j}\| = \infty$. As $X_0 = X \setminus \bigcup_M X_M$, we have $\mu(X_0) = 1$. □

5.3. Measure-theoretic analysis

Definition 5.11. Let $C_{n,j} = \bigcup_{i=0}^{h_{n,j}-1} T^i W_{n,j}$.

Definition 5.12. For j such that $\|B_{n,j}\| \leq \frac{1}{2}\ell_n$, let

$$\begin{aligned} Z_{n,j} &= [B_{n,j}^{\ell_n/h_{n,j}} B_{n,j}] \setminus T^{h_{n,j}}[B_{n,j}^{\ell_n/h_{n,j}} B_{n,j}] \\ &= \{x \in X : x_{[0,\ell_n+h_{n,j})} = B_{n,j}^{\ell_n/h_{n,j}} B_{n,j} \text{ and } x_{[-h_{n,j},\ell_n)} \neq B_{n,j}^{\ell_n/h_{n,j}} B_{n,j}\} \end{aligned}$$

Proposition 5.13. *For j such that $\|B_{n,j}\| \leq \frac{1}{2}\ell_n$, the sets $T^{ah_{n,j}} Z_{n,j}$ are disjoint over $0 \leq a \leq \left\lfloor \frac{\ell_n}{h_{n,j}} \right\rfloor$.*

Proof. For $0 \leq a < b \leq \left\lfloor \frac{\ell_n}{h_{n,j}} \right\rfloor$ and $x \in T^{ah_{n,j}} Z_{n,j} \cap T^{bh_{n,j}} Z_{n,j}$, writing $z = \ell_n - \left\lfloor \frac{\ell_n}{h_{n,j}} \right\rfloor h_{n,j}$, we would have $x_{[z-(a+1)h_{n,j}, z-ah_{n,j})} \neq B_{n,j}$ but $x_{[z-bh_{n,j}, z)} = B_{n,j}^b$ which is impossible. □

Proposition 5.14. *For j such that $\|B_{n,j}\| \leq \frac{1}{2}\ell_n$, the sets $T^i Z_{n,j}$ are disjoint over $0 \leq i < c_{n,j}$.*

Proof. Lemma 5.9 as $B_{n,j}$ is the minimal root of $B_{n,j}^{\ell_n/h_{n,j}} B_{n,j}$ and $c_{n,j} \leq \frac{1}{2}\ell_n < \frac{1}{2}\|B_{n,j}^{\ell_n/h_{n,j}} B_{n,j}\|$. □

Definition 5.15. For j such that $\|B_{n,j}\| > \frac{1}{2}\ell_n$, let, for $0 \leq i < c_{n,j}$,

$$I_{n,j,i} = T^i W_{n,j}$$

and for j such that $\|B_{n,j}\| \leq \frac{1}{2}\ell_n$, let, for $0 \leq i < c_{n,j}$,

$$I_{n,j,i} = T^i \Big(\bigsqcup_{a=0}^{\lfloor \frac{\ell_n}{h_{n,j}} \rfloor} T^{ah_{n,j}} Z_{n,j} \Big)$$

As T is measure-preserving, $\mu(I_{n,j,i}) = \mu(I_{n,j,0})$ for all n, j and $0 \leq i < c_{n,j}$.

Definition 5.16. Let $\tilde{C}_{n,j} = \bigsqcup_{i=0}^{c_{n,j}-1} I_{n,j,i}$. For j such that $\|B_{n,j}\| \leq \frac{1}{2}\ell_n$, let $\widehat{C}_{n,j} = \bigsqcup_{i=0}^{h_{n,j}-1} T^i W_{n,j}$.

Proposition 5.17. *For j such that $\|B_{n,j}\| > \frac{1}{2}\ell_n$, we have $\mu(\tilde{C}_{n,j}) \geq \frac{1}{2k}\mu(C_{n,j})$.*

Proof. $\mu(C_{n,j}) \leq h_{n,j}\mu(W_{n,j}) = h_{n,j}\mu(I_{n,j,0}) = \dfrac{h_{n,j}}{c_{n,j}}\mu(\tilde{C}_{n,j}) \leq \dfrac{k\ell_n}{\frac{1}{2}\ell_n}\mu(\tilde{C}_{n,j}) = 2k\mu(\tilde{C}_{n,j})$. □

Proposition 5.18. $\lim_n \max_j\{\mu(I_{n,j,0})\} = 0$.

Proof. For j such that $\|B_{n,j}\| > \frac{1}{2}\ell_n$, we have $1 \geq \mu(\tilde{C}_{n,j}) = c_{n,j}\mu(I_{n,j,0}) \geq \frac{1}{2}\ell_n\mu(I_{n,j,0})$ and $\ell_n \to \infty$. For j such that $\|B_{n,j}\| \leq \frac{1}{2}\ell_n$, we have $1 \geq \mu(\tilde{C}_{n,j}) = h_{n,j}\mu(I_{n,j,0})$ and $\min_j h_{n,j} \to \infty$. □

Proposition 5.19. *$T^{h_{n,j}} W_{n,j} \subseteq \bigcup_{j'} W_{n,j'}$ and $X_0 = \bigcup_j C_{n,j}$.*

Proof. Every $x \in X_0$ is a concatenation of words of the form $B_{n,j}$ so every occurrence of $B_{n,j}$ is followed immediately by some $B_{n,j'}$ and $x_{[0,\infty)} = uB_1B_2\cdots$ for some u a suffix of some $B_{n,j}$ and $B_\ell \in \{B_{n,j}\}$. □

Proposition 5.20. *Let $E \subseteq W_{n,j}$. Then there exists j' such that $\mu(T^{h_{n,j}}E \cap W_{n,j'}) \geq \frac{1}{C}\mu(E)$.*

Proof. $T^{h_n}E = T^{h_n}E \cap T^{h_{n,j}}W_{n,j} \subseteq T^{h_n}E \cap \bigcup_{j'} W_{n,j'}$ and there are at most C choices of j'. □

Lemma 5.21. $\mu(W_{n,j}) \geq \frac{1}{k\ell_n}\mu(\tilde{C}_{n,j})$.

Proof. For j such that $\|B_{n,j}\| \leq \frac{1}{2}\ell_n$, by Proposition 5.10, $T^{-\ell_n}W_{n,j} = [B_{n,j}^{\ell_n/h_{n,j}}B_{n,j}] \supseteq Z_{n,j}$ so

$$\mu(W_{n,j}) \geq \mu(Z_{n,j}) = \frac{1}{\lfloor \frac{\ell_n}{h_{n,j}} \rfloor + 1}\mu(I_{n,j,0}) \geq \frac{1}{\frac{\ell_n}{h_{n,j}}}\frac{1}{h_{n,j}}\mu(\tilde{C}_{n,j}) = \frac{1}{\ell_n}\mu(\tilde{C}_{n,j})$$

and for j such that $\|B_{n,j}\| > \frac{1}{2}\ell_n$, we have $\mu(W_{n,j}) = \frac{1}{c_{n,j}}\mu(\tilde{C}_{n,j}) \geq \frac{1}{k\ell_n}\mu(\tilde{C}_{n,j})$ since $c_{n,j} \leq k\ell_n$. □

Proposition 5.22. *If $\mu(T^{h_{n,j}}W_{n,j} \cap W_{n,j'}) \geq \delta\mu(W_{n,j''})$ for $j \neq j'$ then $\mu(\tilde{C}_{n,j'}) \geq \frac{1}{2k}\delta\mu(\tilde{C}_{n,j''})$.*

Proof. For j' such that $h_{n,j'} < \frac{1}{2}\ell_n$, Proposition 5.10 states that, as $j \neq j'$, for $x \in T^{h_{n,j}}W_{n,j} \cap W_{n,j'}$, the word $x_{(-\infty,0)}$ has $B_{n,j'}^{\ell_n/h_{n,j'}}$ as a suffix but does not have $B_{n,j'}^{\ell_n/h_{n,j'}}B_{n,j'}$ as a suffix. Therefore $T^{-\ell_n}(T^{h_{n,j}}W_{n,j} \cap W_{n,j'}) \subseteq [B_{n,j'}^{\ell_n/h_{n,j'}}B_{n,j'}] \setminus T^{h_{n,j'}}[B_{n,j'}^{\ell_n/h_{n,j'}}B_{n,j'}] = Z_{n,j'}$. This means that $\mu(Z_{n,j'}) \geq \mu(T^{h_{n,j}}W_{n,j} \cap W_{n,j'}) \geq \delta\mu(W_{n,j''})$ so

$$\begin{aligned}\mu(\tilde{C}_{n,j'}) &= h_{n,j'}\mu(I_{n,j',0}) = h_{n,j'}\Big(\Big\lfloor \frac{\ell_n}{h_{n,j'}} \Big\rfloor + 1\Big)\mu(Z_{n,j'}) \geq h_{n,j'}\frac{\ell_n}{h_{n,j'}}\delta\mu(W_{n,j''}) \\ &\geq \ell_n\delta\frac{1}{c_{n,j''}}\mu(\tilde{C}_{n,j''}) \geq \ell_n\delta\frac{1}{k\ell_n}\mu(\tilde{C}_{n,j''}) = \delta\frac{1}{k}\mu(\tilde{C}_{n,j''})\end{aligned}$$

For j' such that $h_{n,j'} > \frac{1}{2}\ell_n$, using Lemma 5.21 and that $\mu(W_{n,j'}) \geq \delta\mu(W_{n,j''})$,

$$\mu(\tilde{C}_{n,j'}) = c_{n,j'}\mu(W_{n,j'}) \geq c_{n,j'}\delta\mu(W_{n,j''}) \geq c_{n,j'}\delta\frac{1}{k\ell_n}\mu(\tilde{C}_{n,j''}) \geq \frac{\ell_n}{2}\delta\frac{1}{k\ell_n}\mu(\tilde{C}_{n,j''}) = \frac{1}{2k}\delta\mu(\tilde{C}_{n,j''})$$ □

Proposition 5.23. *For j such that $\|B_{n,j}\| \leq \frac{1}{2}\ell_n$, we have $\mu(T^{h_{n,j}}I_{n,j,0} \cap I_{n,j,0}) \geq \dfrac{1}{2}\mu(I_{n,j,0})$.*

Proof.

$$\mu(T^{h_{n,j}}I_{n,j,0} \cap I_{n,j,0}) \geq \mu(\bigsqcup_{a=1}^{\lfloor \frac{\ell_n}{h_{n,j}} \rfloor} T^{ah_{n,j}}Z_{n,j}) = \Big\lfloor \frac{\ell_n}{h_{n,j}} \Big\rfloor\mu(Z_{n,j}) = \frac{\Big\lfloor \frac{\ell_n}{h_{n,j}} \Big\rfloor}{\Big\lfloor \frac{\ell_n}{h_{n,j}} \Big\rfloor + 1}\mu(I_{n,j,0}) \geq \frac{1}{2}\mu(I_{n,j,0})$$ □

5.4. Partial rigidity

We employ ideas similar to Danilenko's [Dan16] proof that exact finite rank implies partial rigidity:

Proposition 5.24. *If there exists $\delta > 0$ and j_n and $t_n \to \infty$ with $\mu(\tilde{C}_{n,j_n}) \geq \delta$ (or $\mu(\widehat{C}_{n,j_n}) \geq \delta$ when applicable) and $\mu(T^{t_n}I_{n,j_n} \cap I_{n,j_n}) \geq \delta\mu(I_{n,j_n})$ then (X,μ) is $\frac{1}{2}\delta^2$-partially rigid.*

Proof. Let $A = W_{N,J}$ for some fixed N and J. Define $\alpha_n = \{0 \leq i < c_{n,j_n} - h_{N,J} : I_{n,j_n,i} \subseteq A\}$.

For j_n such that $h_{n,j_n} > \frac{1}{2}\ell_n$, if $x \in I_{n,j_n,i} \cap W_{N,J}$ then $x_{[-i,-i+h_{n,j_n})} = B_{n,j_n}$ and $x_{[0,h_{N,J})} = B_{N,J}$ meaning that $(B_{n,j_n})_{[i,i+h_{N,J})} = B_{N,J}$. This implies that $T^iW_{n,j_n} \subseteq W_{N,J}$ provided $i < h_{n,j_n} - h_{N,J}$.

For j_n such that $h_{n,j_n} \leq \frac{1}{2}\ell_n$, if $x \in I_{n,j_n,i} \cap W_{N,J}$ then $x_{[-i,-i+\ell_n/h_{n,j_n})} = B_{n,j_n}^{\ell_n/h_{n,j_n}}$ and $x_{[0,h_{N,J})} = B_{N,J}$ so $(B_{n,j_n}^{\ell_n/h_{n,j_n}})_{[i,i+h_{N,J})} = B_{N,J}$ which implies $I_{n,j_n,i} \subseteq W_{N,J}$ provided $i < h_{n,j_n} - h_{N,J}$.

Therefore $(|\alpha_n| + h_{N,J})\mu(I_{n,j_n,0}) \geq \mu(A \cap \tilde{C}_{n,j_n}) \geq |\alpha_n|\mu(I_{n,j_n,0})$. Likewise, if $\|B_{n,j_n}\| \leq \frac{1}{2}\ell_n$ then $(|\alpha_n| + h_{N,J})\mu(W_{n,j_n}) \geq \mu(A \cap \widehat{C}_{n,j_n}) \geq |\alpha_n|\mu(W_{n,j_n})$ using $\alpha_n = \{0 \leq i < h_{n,j_n} - h_{N,J} : T^i W_{n,j_n} \subseteq A\}$. For $m < c_{n,j_n}$, $\mu(T^m \tilde{C}_{n,j_n} \triangle \tilde{C}_{n,j_n}) \leq 2m\mu(I_{n,j_n,0})$, (and likewise $\mu(T^m \widehat{C}_{n,j} \triangle \widehat{C}_{n,j}) \leq 2m\mu(W_{n,j})$ when applicable) therefore

$$\int \big|\mathbb{1}_{\tilde{C}_{n,j_n}} \circ T^{-m} - \mathbb{1}_{\tilde{C}_{n,j_n}}\big|^2 \, d\mu = 2\mu(\tilde{C}_{n,j_n}) - 2\mu(T^m \tilde{C}_{n,j_n} \cap \tilde{C}_{n,j_n}) \leq 2m\mu(I_{n,j_n,0})$$

Therefore for $M < c_{n,j_n}$,

$$\begin{aligned}
\Big|\frac{1}{M}\sum_{m=1}^{M} \mu(T^{-m}A \cap \tilde{C}_{n,j_n}) - \mu(A \cap \tilde{C}_{n,j_n})\Big| &= \Big|\frac{1}{M}\sum_{m=1}^{M} \mu(A \cap T^m \tilde{C}_{n,j_n}) - \mu(A \cap \tilde{C}_{n,j_n})\Big| \\
&\leq \frac{1}{M}\sum_{m=1}^{M} \big|\mu(A \cap T^m \tilde{C}_{n,j_n}) - \mu(A \cap \tilde{C}_{n,j_n})\big| \leq \frac{1}{M}\sum_{m=1}^{M} \int_A \big|\mathbb{1}_{\tilde{C}_{n,j_n}} \circ T^{-m} - \mathbb{1}_{\tilde{C}_{n,j_n}}\big| \, d\mu \\
&\leq \frac{1}{M}\sum_{m=1}^{M} \Big(\int \big|\mathbb{1}_{\tilde{C}_{n,j_n}} \circ T^{-m} - \mathbb{1}_{\tilde{C}_{n,j_n}}\big|^2 \, d\mu\Big)^{1/2} \leq \frac{1}{M}\sum_{m=1}^{M} \sqrt{2m\mu(I_{n,j_n,0})} \leq \sqrt{2M\mu(I_{n,j_n,0})}
\end{aligned}$$

The mean ergodic theorem gives M such that $\int \big|\frac{1}{M}\sum_{m=1}^{M} \mathbb{1}_A \circ T^m - \mu(A)\big|^2 d\mu < (\frac{1}{4}\delta\mu(A))^2$ so

$$\begin{aligned}
\Big|\frac{1}{M}\sum_{m=1}^{M} \mu(T^{-m}A \cap \tilde{C}_{n,j_n}) - \mu(A)\mu(\tilde{C}_{n,j_n})\Big| &= \Big|\int_{\tilde{C}_{n,j_n}} \frac{1}{M}\sum_{m=1}^{M} \mathbb{1}_A \circ T^m - \mu(A) \, d\mu\Big| \\
&\leq \int_{\tilde{C}_{n,j_n}} \Big|\frac{1}{M}\sum_{m=1}^{M} \mathbb{1}_A \circ T^m - \mu(A)\Big| \, d\mu \leq \Big(\int \Big|\frac{1}{M}\sum_{m=1}^{M} \mathbb{1}_A \circ T^m - \mu(A)\Big|^2 \, d\mu\Big)^{1/2} < \frac{1}{4}\delta\mu(A)
\end{aligned}$$

For n large enough that $c_{n,j_n} > M$ and $\sqrt{2M\mu(I_{n,j_n,0})} < \frac{1}{4}\delta\mu(A)$ (Proposition 5.18 states $\mu(I_{n,j_n,0}) \to 0$) then $|\mu(A \cap \tilde{C}_{n,j_n}) - \mu(A)\mu(\tilde{C}_{n,j_n})| < \frac{1}{2}\delta\mu(A)$. Then

$$\begin{aligned}
\mu(T^{t_n}A \cap A) &\geq \mu(T^{t_n}(A \cap \tilde{C}_{n,j_n}) \cap (A \cap \tilde{C}_{n,j_n})) \geq \sum_{i \in \alpha_n} \mu(T^{t_n}T^i I_{n,j_n,0} \cap T^i I_{n,j_n,0}) \\
&= |\alpha_n|\mu(T^{t_n} I_{n,j_n,0} \cap I_{n,j_n,0}) \geq |\alpha_n|\delta\mu(I_{n,j_n,0}) \geq \delta(\mu(A \cap \tilde{C}_{n,j_n}) - h_{N,J}\mu(I_{n,j_n,0})) \\
&> \delta\Big(\mu(A)\mu(\tilde{C}_{n,j_n}) - \frac{1}{2}\delta\mu(A)\Big) - \delta h_{N,J}\mu(I_{n,j_n,0}) \\
&\geq \delta\Big(\mu(A)\delta - \frac{1}{2}\delta\mu(A)\Big) - \delta h_{N,J}\mu(I_{n,j_n,0}) = \frac{1}{2}\delta^2\mu(A) - \delta h_{N,J}\mu(I_{n,j_n,0})
\end{aligned}$$

with the same applying to $\hat{C}_{n,j_n}$ when applicable. Therefore for fixed N and J and $0 \leq i < h_{N,J}$,

$$\liminf \mu(T^{t_n}T^i W_{N,J} \cap T^i W_{N,J}) = \liminf \mu(T^{t_n} W_{N,J} \cap W_{N,J}) \geq \frac{1}{2}\delta^2\mu(W_{N,J}) = \frac{1}{2}\delta^2\mu(T^i W_{N,J})$$

and since the sets $T^i W_{N,J}$ generate the Borel algebra, μ is $\frac{1}{2}\delta^2$-partially rigid. □

Proof of Theorem 5.1. We aim to apply Proposition 5.24. Set $\delta = \frac{1}{4k^2C^{C+1}}$ which depends only on X. There exists a_0 such that $\mu(C_{n,a_0}) \geq \frac{1}{C}$ since $X_0 = \bigcup_j C_{n,j}$. If $\|B_{n,a_0}\| \leq \frac{1}{2}\ell_n$ then $\mu(\widehat{C}_{n,a_0}) = \mu(C_{n,a_0}) \geq \frac{1}{C}$ and Proposition 5.23 implies $\mu(T^{h_{n,a_0}} I_{n,a_0,0} \cap I_{n,a_0,0}) \geq \frac{1}{2}\mu(I_{n,a_0,0})$ so take $t_n = h_{n,a_0}$ and $j_n = a_0$. Now consider when $\|B_{n,a_0}\| > \frac{1}{2}\ell_n$ so Proposition 5.17 implies $\mu(\tilde{C}_{n,a_0}) \geq \frac{1}{2k}\mu(C_{n,a_0}) \geq \frac{1}{2kC}$. By Proposition 5.20, there exists a_1 such that $\mu(T^{h_{n,a_0}} W_{n,a_0} \cap W_{n,a_1}) \geq \frac{1}{C}\mu(W_{n,a_0})$. If $a_1 = a_0$ then $\mu(\tilde{C}_{n,a_1}) = \mu(\tilde{C}_{n,a_0}) \geq \frac{1}{2kC}$ and if $a_1 \neq a_0$ then Proposition 5.22 implies $\mu(\tilde{C}_{n,a_1}) \geq \frac{1}{2k}\mu(\tilde{C}_{n,a_0}) \geq \frac{1}{4k^2C}$. Proposition 5.20 then says there exists a_2 such that

$$\mu(T^{h_{n,a_1}}(T^{h_{n,a_0}} W_{n,a_0} \cap W_{n,a_1}) \cap W_{n,a_2}) \geq \frac{1}{C}\mu(T^{h_{n,a_0}} W_{n,a_0} \cap W_{n,a_1}) \geq \frac{1}{C^2}\mu(W_{n,a_0})$$

and then Proposition 5.22 gives $\mu(\tilde{C}_{n,a_2}) \geq \frac{1}{C^2}\frac{1}{2k}\mu(\tilde{C}_{n,a_0}) \geq \frac{1}{4k^2C^3}$.

Repeating this process, we obtain a_ℓ for $0 \leq \ell \leq C$ such that $\mu(\tilde{C}_{n,a_\ell}) \geq \frac{1}{4k^2C^{\ell+1}} \geq \frac{1}{4k^2C^{C+1}}$ and

$$\mu(W_{n,a_C} \cap \bigcap_{\ell=0}^{C-1} T^{\sum_{z=\ell}^{C-1} h_{n,a_z}} W_{n,a_\ell}) \geq \frac{1}{C^C}\mu(W_{n,a_0})$$

If any of the a_ℓ are such that $h_{n,a_\ell} \leq \frac{1}{2}\ell_n$ then Proposition 5.23 implies $\mu(T^{h_{n,a_\ell}} I_{n,a_\ell,0} \cap I_{n,a_\ell,0}) \geq \frac{1}{2}\mu(I_{n,a_\ell,0})$ so take $t_n = h_{n,a_\ell}$ and $j_n = a_\ell$.

If $h_{n,a_\ell} > \frac{1}{2}\ell_n$ for all $0 \leq \ell \leq C$ then, since there are at most C choices of j, for some $q < s$ we must have $a_q = a_s$ so setting $j_n = a_q$ and $t_n = \sum_{z=q}^{s-1} h_{n,a_z}$,

$$\mu(T^{t_n} I_{n,j_n,0} \cap I_{n,j_n,0}) = \mu(T^{\sum_{z=q}^{s-1} h_{n,a_z}} W_{n,a_q} \cap W_{n,a_s}) \geq \mu(W_{n,a_C} \cap \bigcap_{\ell=0}^{C-1} T^{\sum_{z=\ell}^{C-1} h_{n,a_z}} W_{n,a_\ell}) \geq \frac{1}{C^C}\mu(W_{n,a_0})$$

As
$$\begin{aligned}\mu(W_{n,a_0}) = \mu(I_{n,a_0,0}) &= \frac{1}{c_{n,a_0}}\mu(\tilde{C}_{n,a_0}) \geq \frac{1}{h_{n,a_0}}\frac{1}{2kC} \geq \frac{1}{k\ell_n}\frac{1}{2kC} \geq \frac{1}{k\ell_n}\frac{1}{2kC}\mu(\tilde{C}_{n,j_n}) \\ &= \frac{1}{k\ell_n}\frac{1}{2kC}c_{n,j_n}\mu(I_{n,j_n,0}) \geq \frac{1}{k\ell_n}\frac{1}{2kC}\frac{\ell_n}{2}\mu(I_{n,j_n,0}) = \frac{1}{4k^2C}\mu(I_{n,j_n,0})\end{aligned}$$

we then have $\mu(T^{t_n} I_{n,j_n,0} \cap I_{n,j_n,0}) \geq \frac{1}{4k^2C^{C+1}}\mu(I_{n,j_n,0})$.

In all cases, by Proposition 5.24, we have that (X, μ, T) is $\frac{1}{2}\delta^2$-partially rigid. ☐

Acknowledgments The author thanks Ronnie Pavlov for introducing him to this question and for numerous enjoyable discussions. The author thanks the referees for many helpful suggestions for improving the clarity of exposition, in particular for the suggestion to move the parts of the proof of mixing which follow from standard arguments to the appendix in order to focus on the novel techniques and for suggesting adding significant exposition in and around the use of said techniques.

A. Proofs of mixing properties

We conclude with detailed proofs of mixing for sequences for which the mixing proof is essentially standard arguments for staircase adapted to our notation.

A.1. Mixing between $a_n\tilde{h}_n$ and $\tilde{h}_{n+1}$

Lemma A.1. *Let T be a quasi-staircase transformation. Then for any n and $0 \leq \ell < b_n$ and $k, i \geq 0$ such that $i + k \leq a_n$ and any $j \geq k\ell$,*

$$T^{k\tilde{h}_n} I_{n,j}^{[\ell a_n + i]} = I_{n,j-k\ell}^{[\ell a_n + i + k]}$$

Proof. There are $c_n + \left\lfloor \frac{\ell a_n + i}{a_n} \right\rfloor = c_n + \ell$ spacers above $I_{n,j}^{[\ell a_n+i]}$ so $T^{\tilde{h}_n} I_{n,j}^{[\ell a_n+i]} = I_{n,j-\ell}^{[\ell a_n+i+1]}$. Since $i + k \leq a_n$, there are also $c_n + \ell$ spacers above each $I_{n,j-v\ell}^{[\ell a_n+i+v]}$ for $1 \leq v < k$ so applying $T^{h_n+c_n}$ repeated k times, the claim follows. ☐

Lemma A.2. *Let T be a quasi-staircase transformation, $k \in \mathbb{N}$, B a union of levels in some C_N and $n \geq N$. If $k < a_n$ and $kb_n < h_n$ then*

$$\sum_{j=0}^{h_n-1} |\lambda_B(T^{k\tilde{h}_n} I_{n,j})| \leq \int \left| \frac{1}{b_n} \sum_{\ell=0}^{b_n-1} \chi_B \circ T^{-k\ell} \right| d\mu + \frac{k+1}{a_n} + \frac{kb_n}{h_n}$$

Proof. By Lemma A.1 and then Lemma 4.5, for $kb_n \leq j < h_n$,

$$\begin{aligned}
|\lambda_B(T^{k\tilde{h}_n} I_{n,j})| &= \left| \sum_{\ell=0}^{b_n-1} \sum_{i=0}^{a_n-1} \lambda_B(T^{k\tilde{h}_n} I_{n,j}^{[\ell a_n + i]}) + \lambda_B(T^{k\tilde{h}_n} I_{n,j}^{[r_n]}) \right| \\
&\leq \left| \sum_{\ell=0}^{b_n-1} \sum_{i=0}^{a_n-k-1} \lambda_B(T^{k\tilde{h}_n} I_{n,j}^{[\ell a_n + i]}) \right| + (b_n k + 1)\mu(I_{n+1}) \\
&= \left| \sum_{\ell=0}^{b_n-1} \sum_{i=0}^{a_n-k-1} \lambda_B(I_{n,j-k\ell}^{[\ell a_n + i + k]}) \right| + (b_n k + 1)\mu(I_{n+1}) \\
&= \left| \sum_{\ell=0}^{b_n-1} \sum_{i=0}^{a_n-k-1} \frac{1}{r_n+1} \lambda_B(I_{n,j-k\ell}) \right| + \frac{b_n k + 1}{r_n + 1}\mu(I_n) \\
&= \left| \frac{1}{r_n+1} \sum_{\ell=0}^{b_n-1} \sum_{i=0}^{a_n-k-1} \lambda_B(T^{-k\ell} I_{n,j}) \right| + \frac{b_n k + 1}{r_n + 1}\mu(I_n) \\
&= \frac{a_n - k}{r_n + 1} \left| \sum_{\ell=0}^{b_n-1} \lambda_B(T^{-k\ell} I_{n,j}) \right| + \frac{b_n k + 1}{r_n + 1}\mu(I_n) \leq \frac{1}{b_n} \left| \sum_{\ell=0}^{b_n-1} \lambda_B(T^{-k\ell} I_{n,j}) \right| + \frac{k+1}{a_n}\mu(I_n) \\
&= \left| \frac{1}{b_n} \sum_{\ell=0}^{b_n-1} \int_{I_{n,j}} \chi_B \circ T^{-k\ell}\, d\mu \right| + \frac{k+1}{a_n}\mu(I_n) \leq \int_{I_{n,j}} \left| \frac{1}{b_n} \sum_{\ell=0}^{b_n-1} \chi_B \circ T^{-k\ell} \right| d\mu + \frac{k+1}{a_n}\mu(I_n)
\end{aligned}$$

Therefore

$$\begin{aligned}
\sum_{j=0}^{h_n-1} |\lambda_B(T^{k\tilde{h}_n} I_{n,j})| &\leq \sum_{j=kb_n}^{h_n-1} |\lambda_B(T^{k\tilde{h}_n} I_{n,j})| + kb_n\mu(I_n) \\
&\leq \sum_{j=kb_n}^{h_n-1} \left(\int_{I_{n,j}} \left| \frac{1}{b_n} \sum_{\ell=0}^{b_n-1} \chi_B \circ T^{-k\ell} \right| d\mu + \frac{k+1}{a_n}\mu(I_{n,j}) \right) + kb_n\mu(I_n) \\
&\leq \int \left| \frac{1}{b_n} \sum_{\ell=0}^{b_n-1} \chi_B \circ T^{-k\ell} \right| d\mu + \frac{k+1}{a_n} + \frac{kb_n}{h_n}
\end{aligned}$$

□

Proposition A.3. *Let T be a quasi-staircase transformation and $k \in \mathbb{N}$. If T^k is ergodic then $\{k\tilde{h}_n\}$ and $\{kh_n\}$ are rank-one uniform mixing.*

Proof. Since $\frac{b_n}{h_n} \to 0$ and $a_n \to \infty$ there exists N such that for all $n \geq N$ we have $k < a_n$ and $kb_n < h_n$. That $\{k\tilde{h}_n\}$ is rank-one uniform mixing follows from Lemma A.2 since T^k is ergodic, $b_n \to \infty$, $a_n \to \infty$ and $\frac{b_n}{h_n} \to 0$. Then

$$\sum_{j=0}^{h_n-1} |\lambda_B(T^{kh_n} I_{n,j})| \leq \sum_{j=kc_n}^{h_n} |\lambda_B(T^{kh_n} I_{n,j})| + \frac{kc_n}{h_n} = \sum_{j=0}^{h_n-kc_n} |\lambda_B(T^{k\tilde{h}_n} I_{n,j})| + \frac{kc_n}{h_n} \to 0$$

as $\frac{c_n}{h_n} \to 0$, k is fixed and $\{k\tilde{h}_n\}$ is rank-one uniform mixing. □

Lemma A.4 ([CPR23] Proposition A.13)**.** *Let T be a rank-one transformation and $\{c_n\}$ a sequence such that $\frac{c_n}{h_n} \to 0$. If $k \in \mathbb{N}$ and $\{q(h_n + c_n)\}$ is rank-one uniform mixing for each $q \leq k+1$ and $\{t_n\}$ is a sequence such that $h_n + c_n \leq t_n < (q+1)(h_n + c_n)$ for all n then $\{t_n\}$ is mixing.*

Lemma A.5 ([CPR23] Proposition A.16)**.** *Let T be a rank-one transformation and $\{c_n\}$ a sequence such that $\frac{c_n}{h_n} \to 0$. If $\{q(h_n + c_n)\}$ is rank-one uniform mixing for each fixed q and $k_n \to \infty$ is such that $\frac{k_n}{n} \leq 1$ then for any measurable set B, $\int |\frac{1}{n} \sum_{j=0}^{n-1} \chi_B \circ T^{-jk_n}|\, d\mu \to 0$.*

Proof of Proposition 4.7. As T is ergodic, Proposition A.3 with $k = 1$ gives that $\{\tilde{h}_n\}$ is rank-one uniform mixing, hence mixing, so T is totally ergodic. Then Proposition A.3 gives that for each fixed k the sequence $\{k\tilde{h}_n\}$ is rank-one uniform mixing so Lemma A.5 gives the claim. ☐

Proposition A.6. *Let T be a quasi-staircase transformation, B a measurable set and $Q > 0$. Then*

$$\max_{h_n+c_n \leq t < Q\tilde{h}_n} |\lambda_B(T^t B)| \to 0$$

Proof. As in the proof of Proposition 4.7, for each fixed k the sequence $\{k\tilde{h}_n\}$ is rank-one uniform mixing so Lemma A.4 gives the claim. ☐

Lemma A.7. *Let T be a quasi-staircase transformation. Let $n > 0$ and $0 \leq x < b_n$ and $0 \leq q < a_n$. If $0 \leq \ell < b_n - x$ and $0 \leq i < a_n - q$ and $j \geq \frac{1}{2}a_n x(x-1) + qx + ix + \ell(xa_n + q)$ then*

$$T^{(xa_n+q)\tilde{h}_n} I_{n,j}^{[\ell a_n+i]} = I_{n,j-\frac{1}{2}a_n x(x-1)-qx-ix-\ell(xa_n+q)}^{[(\ell+x)a_n+i+q]}$$

Proof. If $x = 0$ then Lemma A.1 applied with q in place of k gives the claim. So we can write

$$xa_n + q = (a_n - i) + (x-1)a_n + (q+i)$$

and assume all three terms on the right are nonnegative.

Using Lemma A.1,

$$T^{(a_n-i)\tilde{h}_n} I_{n,j}^{[\ell a_n+i]} = I_{n,j-(a_n-i)\ell}^{[\ell a_n+i+a_n-i]} = I_{n,j-(a_n-i)\ell}^{[(\ell+1)a_n]}$$

Now observe that, by Lemma A.1 with 0 as i and a_n as k, for any $0 \leq v < x$ and any $a_n v \leq z < h_n$,

$$T^{a_n\tilde{h}_n} I_{n,z}^{[va_n]} = I_{n,z-a_n v}^{[(v+1)a_n]}$$

so applying that $x-1$ times for $v = \ell+1, \ell+2, \ldots, \ell+x-1$,

$$T^{(x-1)a_n\tilde{h}_n} I_{n,j-(a_n-i)\ell}^{[(\ell+1)a_n]} = I_{n,j-(a_n-i)\ell-(x-1)\ell a_n-\frac{1}{2}x(x-1)a_n}^{[(\ell+x)a_n]}$$

since $\sum_{v=\ell+1}^{\ell+x-1} v = \frac{1}{2}(\ell+x)(\ell+x-1) - \frac{1}{2}\ell(\ell+1) = (x-1)\ell + \frac{1}{2}x(x-1)$. Then applying Lemma A.1 one final time with $q+i$ in place of k,

$$\begin{aligned} T^{(q+i)\tilde{h}_n} I_{nj-(a_n-i)\ell-(x-1)\ell a_n-\frac{1}{2}x(x-1)a_n}^{[(\ell+x)a_n]} &= I_{n,j-(a_n-i)\ell-(x-1)\ell a_n-\frac{1}{2}x(x-1)a_n-(x+\ell)(q+i)}^{[(\ell+x)a_n+q+i]} \\ &= I_{n,j-x\ell a_n-\frac{1}{2}x(x-1)a_n-xi-xq-\ell q}^{[(\ell+x)a_n+q+i]} \end{aligned}$$

☐

Lemma A.8. *Let T be a quasi-staircase transformation. Let $n > 0$ and $0 \leq x < b_n$ and $0 \leq q < a_n$. If $0 \leq \ell < b_n - x - 1$ and $a_n - q \leq i < a_n$ and $j \geq \frac{1}{2}a_n x(x+1) + q(x+1) + i(x+1) + \ell(xa_n + 1)$ then*

$$T^{(xa_n+q)\tilde{h}_n} I_{n,j}^{[\ell a_n+i]} = I_{n,j-\frac{1}{2}a_n x(x+1)-(q+i-a_n)(x+1)-\ell(xa_n+q)}^{[(\ell+x)a_n+i+q]}$$

Proof. The same proof as Lemma A.7 except we write $xa_n + q = (a_n - i) + xa_n + (q + i - x)$. ☐

Lemma A.9. *Let T be a quasi-staircase transformation. Let B be a union of levels C_N. For $n \geq N$ and $k_n\tilde{h}_n \leq t_n < (k_n+1)\tilde{h}_n$,*

$$\sum_{j=0}^{h_n-1} |\lambda_B(T^{t_n} I_{n,j})| \leq \sum_{x=0}^{h_n-1} |\lambda_B(T^{k_n\tilde{h}_n} I_{n,x})| + c_n\mu(I_n) + \sum_{x=0}^{h_n-1} |\lambda_B(T^{(k_n+1)\tilde{h}_n} I_{n,x})|$$

Proof. Write $t_n = k_n\tilde{h}_n + z_n$ for $0 \leq z_n < \tilde{h}_n$. Then

$$\begin{aligned}
\sum_{j=0}^{h_n-1} |\lambda_B(T^{t_n} I_{n,j})| &\leq \sum_{j=0}^{h_n-z_n-1} |\lambda_B(T^{t_n} I_{n,j})| + c_n\mu(I_n) + \sum_{j=h_n-z_n+c_n}^{h_n-1} |\lambda_B(T^{t_n} I_{n,j})| \\
&\leq \sum_{j=0}^{h_n-z_n-1} |\lambda_B(T^{k_n\tilde{h}_n} I_{n,j+z_n})| + c_n\mu(I_n) + \sum_{j=\tilde{h}_n-z_n}^{h_n-1} |\lambda_B(T^{(k_n+1)\tilde{h}_n} I_{n,j+z_n-\tilde{h}_n})| \\
&\leq \sum_{x=0}^{h_n-1} |\lambda_B(T^{k_n\tilde{h}_n} I_{n,x})| + c_n\mu(I_n) + \sum_{x=0}^{h_n-1} |\lambda_B(T^{(k_n+1)\tilde{h}_n} I_{n,x})| \qquad \square
\end{aligned}$$

Proof of Proposition 4.8. Let t_n attain the maximum in $M_{B,n}$. If $t_n \geq (r_n-1)\tilde{h}_n$ then $h_{n+1}+c_{n+1}-t_n \leq c_{n+1} + 2h_n + c_n + \frac{1}{2}a_nb_n(b_n-1)$ so

$$\begin{aligned}
\sum_{j=0}^{h_{n+1}-1} |\lambda_B(T^{t_n} I_{n+1,j})| &\leq \sum_{j=h_{n+1}+c_{n+1}-t_n}^{h_{n+1}-1} |\lambda_B(T^{t_n} I_{n+1,j})| + (h_{n+1}+c_{n+1}-t_n)\mu(I_{n+1}) \\
&\leq \sum_{j=0}^{t_n-c_{n+1}-1} |\lambda_B(T^{\tilde{h}_{n+1}} I_{n+1,j})| + \frac{c_{n+1}+2h_n+c_n\frac{1}{2}a_nb_n(b_n-1)}{h_{n+1}} \to 0
\end{aligned}$$

since $\{\tilde{h}_{n+1}\}$ is rank-one uniform mixing.

So we may assume $t_n < (r_n-1)\tilde{h}_n$ and therefore write $t_n = k_n\tilde{h}_n + z_n$ for $a_n \leq k_n < r_n - 1$ and $0 \leq z_n < \tilde{h}_n$. By Lemma A.9,

$$\sum_{j=0}^{h_n-1} |\lambda_B(T^{t_n} I_{n,j})| \leq \sum_{x=0}^{h_n-1} |\lambda_B(T^{k_n\tilde{h}_n} I_{n,x})| + c_n\mu(I_n) + \sum_{x=0}^{h_n-1} |\lambda_B(T^{(k_n+1)\tilde{h}_n} I_{n,x})|$$

We will show the sum on the left tends to zero; the same argument with k_n+1 in place of k_n gives the same for the right sum. As $c_n\mu(I_n) \to 0$, this will complete the proof.

Write $k_n = x_na_n + q_n$ for $0 \leq q_n < a_n$ and $1 \leq x_n < b_n$. Observe that

$$\begin{aligned}
\sum_{j=0}^{h_n-1} |\lambda_B(T^{k_n\tilde{h}_n} I_{n,j})| \leq \sum_{j=0}^{h_n-1} \left| \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=0}^{a_n-1} \lambda_B(T^{k_n\tilde{h}_n} I_{n,j}^{[\ell a_n+i]}) \right| + 2a_nh_n\mu(I_{n+1}) \qquad (\star) \\
+ \sum_{j=0}^{h_n-1} \left| \sum_{\ell=b_n-x_n+1}^{b_n-1} \sum_{i=0}^{a_n-1} \lambda_B(T^{k_n\tilde{h}_n} I_{n,j}^{[\ell a_n+i]}) \right| + \frac{1}{r_n+1} \qquad (\star\star)
\end{aligned}$$

We handle the sum $(\star\star)$ first and return to the sum in $(\star)$ shortly.

For $0 \leq \ell < b_n$ and $0 \leq i < a_n$, we have that

$$I_{n,0}^{[\ell a_n+i]} = T^{(\ell a_n+i)\tilde{h}_n} I_{n,\frac{1}{2}\ell(\ell-1)a_n+i\ell}^{[0]}$$

since $\frac{1}{2}\ell(\ell-1)a_n + i\ell \leq a_nb_n^2 + a_nb_n < h_n$ (as $\frac{a_nb_n^2}{h_n} \to 0$).

For $b_n - x_n + 1 \leq \ell < b_n$ and $0 \leq i < a_n$, since $x + \ell \geq b_n + 1$,

$$\begin{aligned}
k_n\tilde{h}_n + (\ell a_n + i)\tilde{h}_n &= (x_na_n + q_n + \ell a_n + i)(h_n + c_n) \\
&\geq (b_na_n + a_n)\tilde{h}_n \\
&= (b_na_n + 1)h_n + b_na_nc_n + (a_n-1)h_n + a_nc_n \geq h_{n+1}
\end{aligned}$$

since $\frac{1}{2}a_nb_n(b_n-1) \leq h_n$. Also,

$$\begin{aligned} k_n\tilde{h}_n + (\ell a_n + i)\tilde{h}_n + \frac{1}{2}\ell(\ell-1)a_n + i\ell &= ((x_n+\ell)a_n + q_n + i)(h_n + c_n) + \frac{1}{2}\ell(\ell-1)a_n + i\ell \\ &\leq 2b_na_n(h_n+c_n) + \frac{1}{2}b_n(b_n-1)a_n + a_nb_n < 2h_{n+1} \end{aligned}$$

Since a sublevel in I_n is a level in I_{n+1} and $\{h_{n+1}\}$ is rank-one uniform mixing (Proposition A.3),

$$\sum_{j=0}^{h_n-1} \sum_{\ell=b_n-x_n+1}^{b_n-1} \sum_{i=0}^{a_n-1} |\lambda_B(T^{k_n\tilde{h}_n} I_{n,j}^{[\ell a_n+i]})| \leq \sum_{y=0}^{h_{n+1}-1} |\lambda_B(T^{h_{n+1}} I_{n+1,y})| \to 0$$

As $2a_nh_n\mu(I_{n+1}) \leq \frac{2a_nh_n}{h_{n+1}} \leq \frac{2}{b_n} \to 0$ and $r_n \to \infty$, it remains only to show that the sum in $(\star)$ tends to zero. Observe that

$$\begin{aligned} \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=0}^{a_n-1} \lambda_B(T^{k_n\tilde{h}_n} I_{n,j}^{[\ell a_n+i]}) &= \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=0}^{a_n-q_n-1} \lambda_B(T^{k_n\tilde{h}_n} I_{n,j}^{[\ell a_n+i]}) && (\dagger) \\ &\quad + \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=a_n-q_n}^{a_n-1} \lambda_B(T^{k_n\tilde{h}_n} I_{n,j}^{[\ell a_n+i]}) && (\ddagger) \end{aligned}$$

First, we address (†): set $y_n = \frac{1}{2}a_nx_n(x_n-1) + q_nx_n$. For $i < a_n - q_n$ and $\ell < b_n - x_n - 1$, we have $y_n + ix_n + \ell k_n \leq 3a_nb_n^2$ so for $j \geq 3a_nb_n^2$, by Lemma A.7 and Lemma 4.5,

$$\begin{aligned} \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=0}^{a_n-q_n-1} \lambda_B(T^{k_n\tilde{h}_n} I_{n,j}^{[\ell a_n+i]}) &= \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=0}^{a_n-q_n-1} \lambda_B(I_{n,j-y_n-ix_n-\ell k_n}^{[(\ell+x_n)a_n+i+q_n]}) \\ = \frac{1}{r_n+1} \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=0}^{a_n-q_n-1} \lambda_B(I_{n,j-y_n-ix_n-\ell k_n}) &= \frac{1}{r_n+1} \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=0}^{a_n-q_n-1} \lambda_B(T^{-\ell k_n - ix_n - y_n} I_{n,j}) \end{aligned}$$

Then, summing over all $3a_nb_n^2 \leq j < h_n$,

$$\begin{aligned} &\sum_{j=3a_nb_n^2}^{h_n-1} \left| \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=0}^{a_n-q_n-1} \lambda_B(T^{k_n\tilde{h}_n} I_{n,j}^{[\ell a_n+i]}) \right| \\ &= \sum_{j=3a_nb_n^2}^{h_n-1} \left| \frac{1}{r_n+1} \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=0}^{a_n-q_n-1} \lambda_B(T^{-\ell k_n - ix_n - y_n} I_{n,j}) \right| \\ &\leq \frac{1}{r_n+1} \sum_{j=0}^{h_n-1} \sum_{\ell=0}^{b_n-x_n-2} \left| \sum_{i=0}^{a_n-q_n-1} \lambda_B(T^{-\ell k_n - ix_n - y_n} I_{n,j}) \right| \\ &\leq \frac{1}{r_n+1} \sum_{\ell=0}^{b_n-x_n-2} \int \left| \sum_{i=0}^{a_n-q_n-1} \chi_B \circ T^{-\ell k_n - ix_n - y_n} \right| \, d\mu \\ &= \frac{(b_n-x_n-2)(a_n-q_n)}{r_n+1} \int \left| \frac{1}{a_n-q_n} \sum_{i=0}^{a_n-q_n-1} \chi_B \circ T^{-ix_n} \right| \, d\mu \\ &\leq \min\left(\frac{a_n-q_n}{a_n}, \int \left| \frac{1}{a_n-q_n} \sum_{i=0}^{a_n-q_n-1} \chi_B \circ T^{-ix_n} \right| \, d\mu \right) \end{aligned}$$

since $\frac{(b_n-2)}{r_n+1} < \frac{1}{a_n}$ and $\int |\chi_B| \, d\mu \leq 1$. For a subsequence along which $x_n \leq a_n - q_n$, Proposition 4.7 implies the integral tends to zero. For n such that $a_n - q_n < x_n < b_n$, the quantity on the left is bounded by $\frac{b_n}{a_n} \to 0$.

For (‡): set $y'_n = \frac{1}{2}a_n x_n(x_n+1) + (q_n - a_n)(x_n+1)$. By Lemma A.8 and Lemma 4.5, for $j \geq 3a_n b_n^2$,

$$\sum_{\ell=0}^{b_n-x_n-2} \sum_{i=a_n-q_n}^{a_n-1} \lambda_B(T^{k_n\tilde{h}_n} I_{n,j}^{[\ell a_n+i]}) = \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=a_n-q_n}^{a_n-1} \lambda_B(I_{n,j-y'_n-i(x_n+1)-\ell k_n}^{[(\ell+x_n)a_n+i+q_n]})$$

$$= \frac{1}{r_n+1} \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=a_n-q_n}^{a_n-1} \lambda_B(I_{n,j-y'_n-i(x_n+1)-\ell k_n}) = \frac{1}{r_n+1} \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=a_n-q_n}^{a_n-1} \lambda_B(T^{-\ell k_n - i(x_n+1) - y'_n} I_{n,j})$$

Similar to the sum (†), then

$$\begin{aligned}
&\sum_{j=3a_n b_n^2}^{h_n-1} \left| \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=a_n-q_n}^{a_n-1} \lambda_B(T^{k_n\tilde{h}_n} I_{n,j}^{[\ell a_n+i]}) \right| \\
&= \sum_{j=3a_n b_n^2}^{h_n-1} \left| \frac{1}{r_n+1} \sum_{\ell=0}^{b_n-x_n-2} \sum_{i=a_n-q_n}^{a_n-1} \lambda_B(T^{-\ell k_n - i(x_n+1) - y'_n} I_{n,j}) \right| \\
&\leq \frac{(b_n - x_n - 2)q_n}{r_n+1} \int \left| \frac{1}{q_n} \sum_{i=a_n-q_n}^{a_n-1} \chi_B \circ T^{-i(x_n+1)} \right| \, d\mu \\
&= \frac{(b_n - x_n - 2)q_n}{r_n+1} \int \left| \frac{1}{q_n} \sum_{i'=0}^{q_n-1} \chi_B \circ T^{-i'(x_n+1)} \right| \, d\mu \leq \min\Big(\frac{q_n}{a_n}, \int \left| \frac{1}{q_n} \sum_{i'=0}^{q_n-1} \chi_B \circ T^{-i'(x_n+1)} \right| \, d\mu\Big)
\end{aligned}$$

and along any subsequence where $x_n + 1 \leq q_n$, this tends to zero by Proposition 4.7, and for $q_n \leq x_n + 1 < b_n + 1$, the quantity on the left is bounded by $\frac{b_n}{a_n} \to 0$, completing the proof. ☐

A.2. Mixing between $\tilde{h}_n$ and $b_n\tilde{h}_n$

Proof of Proposition 4.9. Let t_n attain the maximum in $\widehat{M}_{B,n}$. By Lemma A.9, writing $t_n = k_n\tilde{h}_n + z_n$ for $1 \leq k_n < b_n$ and $0 \leq z_n < \tilde{h}_n$,

$$\sum_{j=0}^{h_n-1} |\lambda_B(T^{t_n} I_{n,j})| \leq \sum_{x=0}^{h_n-1} |\lambda_B(T^{k_n\tilde{h}_n} I_{n,x})| + c_n\mu(I_n) + \sum_{x=0}^{h_n-1} |\lambda_B(T^{(k_n+1)\tilde{h}_n} I_{n,x})|$$

By Lemma A.2,

$$\sum_{j=0}^{h_n-1} |\lambda_B(T^{k_n\tilde{h}_n} I_{n,j})| \leq \int \left| \frac{1}{b_n} \sum_{\ell=0}^{b_n-1} \chi_B \circ T^{-k_n\ell} \right| \, d\mu + \frac{k_n+1}{a_n} + \frac{k_n b_n}{h_n} \to 0$$

since $k_n < b_n$ so Proposition 4.7 implies the integral tends to zero. Similar reasoning for $k_n + 1 \leq b_n$ then completes the proof. ☐

References

[Ada98] Terrence M. Adams, *Smorodinsky's conjecture on rank-one mixing*, Proc. Amer. Math. Soc. **126** (1998), no. 3, 739–744. MR 1443143

[AFP17] Terrence Adams, Sébastien Ferenczi, and Karl Petersen, *Constructive symbolic presentations of rank one measure-preserving systems*, Colloq. Math. **150** (2017), no. 2, 243–255. MR 3719459

[Bos85] Michael Boshernitzan, *A unique ergodicity of minimal symbolic flows with linear block growth*, J. Analyse Math. **44** (1984/85), 77–96. MR 801288

[Cas97] Julien Cassaigne, *Complexité et facteurs spéciaux*, Journées Montoises (Mons, 1994), no. 1, vol. 4, Bull. Belg. Math. Soc. Simon Stevin, 1997, pp. 67–88. MR 1440670

[CFPZ19] Julien Cassaigne, Anna E. Frid, Svetlana Puzynina, and Luca Q. Zamboni, *A characterization of words of linear complexity*, Proc. Amer. Math. Soc. **147** (2019), no. 7, 3103–3115. MR 3973910

[CK19] Van Cyr and Bryna Kra, *Counting generic measures for a subshift of linear growth*, J. Eur. Math. Soc. (JEMS) **21** (2019), no. 2, 355–380. MR 3896204

[CK20a] ———, *The automorphism group of a shift of slow growth is amenable*, Ergodic Theory Dynam. Systems **40** (2020), no. 7, 1788–1804. MR 4108905

[CK20b] ———, *Realizing ergodic properties in zero entropy subshifts*, Israel J. Math. **240** (2020), no. 1, 119–148. MR 4193129

[CPR23] Darren Creutz, Ronnie Pavlov, and Shaun Rodock, *Measure-theoretically mixing subshifts with low complexity*, Ergodic Theory and Dynamical Systems **43** (2023), no. 7, 2293–2316.

[Cre21] Darren Creutz, *Mixing on stochastic staircase tansformations*, Studia Math. **257** (2021), no. 2, 121–153. MR 4194950

[CS04] Darren Creutz and Cesar E. Silva, *Mixing on a class of rank-one transformations*, Ergodic Theory Dynam. Systems **24** (2004), no. 2, 407–440. MR 2054050

[CS10] ———, *Mixing on rank-one transformations*, Studia Math. **199** (2010), no. 1, 43–72. MR 2652597

[Dan16] Alexandre I. Danilenko, *Actions of finite rank: weak rational ergodicity and partial rigidity*, Ergodic Theory Dynam. Systems **36** (2016), no. 7, 2138–2171. MR 3568975

[DDMP16] Sebastián Donoso, Fabien Durand, Alejandro Maass, and Samuel Petite, *On automorphism groups of low complexity subshifts*, Ergodic Theory Dynam. Systems **36** (2016), no. 1, 64–95. MR 3436754

[DDMP21] ———, *Interplay between finite topological rank minimal Cantor systems, $\mathcal{S}$-adic subshifts and their complexity*, Trans. Amer. Math. Soc. **374** (2021), no. 5, 3453–3489. MR 4237953

[dJ77] Andrés del Junco, *A transformation with simple spectrum which is not rank one*, Canadian J. Math. **29** (1977), no. 3, 655–663. MR 466489

[DOP21] Andrew Dykstra, Nicholas Ormes, and Ronnie Pavlov, *Subsystems of transitive subshifts with linear complexity*, Ergodic Theory and Dynamical Systems (2021), 1–27.

[Fer95] Sébastien Ferenczi, *Les transformations de Chacon : combinatoire, structure géométrique, lien avec les systèmes de complexité* $2n+1$, Bulletin de la Société Mathématique de France **123** (1995), no. 2, 271–292 (fr).

[Fer96] ———, *Rank and symbolic complexity*, Ergodic Theory Dynam. Systems **16** (1996), no. 4, 663–682. MR 1406427

[FGH$^+$23] M. Foreman, S. Gao, A. Hill, C.E. Silva, and B. Weiss, *Rank one transformations, odometers and finite factors*, Israel J. Math **255** (2023), 231–249.

[Kal84] Steven Arthur Kalikow, *Twofold mixing implies threefold mixing for rank one transformations*, Ergodic Theory Dynam. Systems **4** (1984), no. 2, 237–259. MR 766104

[Kno54] Konrad Knopp, *Theory and application of infinite series*, Blackie and Son Limited, 1954.

[Ler12] Julien Leroy, *Some improvements of the S-adic conjecture*, Adv. in Appl. Math. **48** (2012), no. 1, 79–98. MR 2845508

[PS22] Ronnie Pavlov and Scott Schmeiding, *Local finiteness and automorphism groups of low complexity subshifts*, Ergodic Theory and Dynamical Systems **43** (2022), 1980–2001.

[PS23] ———, *On the structure of generic subshifts*, Nonlinearity **36** (2023), no. 9, 4904–4953.

[Ryz93] V. V. Ryzhikov, *Joinings and multiple mixing of the actions of finite rank*, Funktsional. Anal. i Prilozhen. **27** (1993), no. 2, 63–78, 96. MR 1251168

[Sil08] C. E. Silva, *Invitation to ergodic theory*, Student Mathematical Library, vol. 42, American Mathematical Society, Providence, RI, 2008. MR 2371216