On the multifractal spectrum of weighted Birkhoff averages

Balázs Bárány, Micha Rams, Ruxi Shi

HAL Id: hal-03534271

https://hal.sorbonne-universite.fr/hal-03534271v1

Submitted on 19 Jan 2022

ON THE MULTIFRACTAL SPECTRUM OF WEIGHTED
BIRKHOFF AVERAGES

Balázs Bárány*

Budapest University of Technology and Economics
Department of Stochastics, MTA-BME Stochastics Research Group
P.O.Box 91, 1521 Budapest, Hungary

Michał Rams

Institute of Mathematics, Polish Academy of Sciences
ul. Śniadeckich 8, 00-656 Warszawa, Poland

Ruxi Shi

Sorbonne Université, LPSM
75005 Paris, France

(Communicated by Jairo Bochi)

Abstract. In this paper, we study the topological spectrum of weighted Birk–
hoff averages over aperiodic and irreducible subshifts of finite type. We show
that for a uniformly continuous family of potentials, the spectrum is continuous
and concave over its domain. In case of typical weights with respect to some
ergodic quasi-Bernoulli measure, we determine the spectrum. Moreover, in
case of full shift and under the assumption that the potentials depend only on
the first coordinate, we show that our result is applicable for regular weights,
like Möbius sequence.

1. **Introduction.** Let $T\colon X \mapsto X$ be a measure preserving transformation of the standard Borel probability space $(X, \mathcal{B}, \nu)$. The well-known Theorem of Birkhoff states that for any $f \in L^1(X, \mathcal{B}, \nu)$, the limit

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) \text{ exists for } \nu\text{-almost every } x.$$

Moreover, if ν is an ergodic measure with respect to T then

$$\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x) = \int f d\nu \text{ for } \nu\text{-almost every } x.$$

The "time" averages $\frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)$ are called the Birkhoff averages. If T is uniquely ergodic then for a continuous potential $f \in C(X)$, the limit $\lim_{n\to\infty} \frac{1}{n} \sum_{k=0}^{n-1} f(T^k x)$ exists for all $x \in X$ and converges to a constant. That is, there is no multifractal behaviour. However, if the system (X, T) has a large

---

family of ergodic measures (for example, a full shift), one may expect that the limit of the Birkhoff averages can take a wide variety of values. It is a natural question to ask, how large (for example, topological entropy, Hausdorff or packing dimension) is the set of points in X for which the Birkhoff average converges to a prescribed value α? It leads to the multifractal analysis and there has been a considerable amount of works on this.

As far as we know, the first work is due to Besicovitch [4] where he studied the Hausdorff dimension of sets given by the frequency of digits in dyadic expansions. Then it was subsequently extended by Eggleston [7]. For further results on digit frequencies, see Barreira, Saussol and Schmeling [2]. For multifractal analysis of Birkhoff averages, we refer to [12, 14, 24, 28, 18, 16, 13, 3] and references therein.

Let X be a compact metric space, let $T \colon X \mapsto X$ be a continuous transformation, and let $\varphi \colon X \mapsto \mathbb{R}$ be a continuous potential. Takens and Verbitskiy [29] showed under the condition that the system (X, T) has specification that for an $\alpha \in \mathbb{R}$, the topological entropy of the set

$$E(\alpha) = \left\{ x \in X : \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \varphi(T^k x) = \alpha \right\}$$

equals to the Legendre transform of the topological pressure, which is equal to the supremum of the entropy of all invariant and ergodic measures for which the "space" average (i.e. the integral of φ) equals to α.

In this paper, we are interested in the generalisation of the problem above for weighted Birkhoff averages. Let $\mathbf{w} = \{w_k\}_{k \in \mathbb{N}}$ be a sequence of bounded reals and let $\varphi \colon X \mapsto \mathbb{R}^d$ be a continuous potential and let $\alpha \in \mathbb{R}^d$. Is it possible to determine

$$h_{\text{top}} \left(\left\{ x \in X : \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} w_k \varphi(T^k x) = \alpha \right\} \right) = ?$$

Weighted Birkhoff averages were studied since 1940', since the celebrated Theorem of Wiener and Wintner. Since then several generalisations of the weighted ergodic theorems appeared, see for example [8, Chapter 21]. The first study of weighted Birkhoff averages in the context of spectrum was by Fan [9]. Lately, there were an additional motivation for this problem by Sarnak's conjecture [27]. Let us recall the definition of the Möbius sequence, $\boldsymbol{\mu} \colon \mathbb{N} \mapsto \{-1, 0, 1\}$,

$$\boldsymbol{\mu}(n) = \begin{cases} (-1)^k & \text{if } n \text{ is a product of } k \text{ distinct primes,} \\ 0 & \text{if there exists } a \geq 2 \text{ such that } a^2 | n. \end{cases}$$

Sarnak's conjecture [27] claims that if $T \colon X \mapsto X$ is continuous over the compact metric space X with topological zero entropy then for every $x \in X$ and every continuous potential $\varphi \colon X \mapsto \mathbb{R}$

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \boldsymbol{\mu}(k) \varphi(T^k x) = 0.$$

Even though Sarnak's conjecture has been verified for various special dynamical systems (e.g. rotations on the circle, automorphism of the torus with entropy zero etc.), it is still widely open in general. We refer to [19] for a survey of many recent results on Sarnak conjecture. El Abdalaoui, Kułaga-Przymus, Lemańczyk and de la Rue [1] showed Birkhoff's type ergodic theorem with Möbius weight.

Theorem ([1]). *Let T be an automorphism of a standard Borel probability space $(X, \mathcal{B}, \nu)$ and let $f \in L^1(X, \mathcal{B}, \nu)$. Then, for ν-almost every $x \in X$, we have*

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \boldsymbol{\mu}(k) f(T^k(x)) = 0.$$

Fan [10] proved a similar result for a more general family of sequences, like Davenport's type. Hence, the usual method of calculating the spectrum for weighted Birkhoff averages, that is, to show that it is equal to the supremum of the entropy of invariant measures, is not applicable. This paper is devoted to present a method, which allows us to calculate the spectrum. Recently, Fan [11] studied the same question, but with strictly different methods. We will point out the main differences between our and his results.

2. **Results.** Let $\mathcal{A} = \{1, \ldots, K\}$ be a finite alphabet, and let $\Sigma = \mathcal{A}^{\mathbb{N}}$. Let us denote the left-shift operator on Σ by σ. Denote Σ_n the set of n-length finite word. Moreover, denote Σ_* the set of all finite prefixes of the infinite words in Σ. For an $\mathbf{i} = (i_0, i_1, \ldots) \in \Sigma$ and $m > n \geq 0$ let $\mathbf{i}|_n^m = (i_n, \ldots, i_m)$ be the subword of $\mathbf{i}$ between the positions n and m, and for short denote by $\mathbf{i}|_n$ the first n element of $\mathbf{i}$, i.e. $\mathbf{i}|_n = \mathbf{i}|_0^{n-1}$. For an $\mathbf{i} \in \Sigma_*$, denote $|\mathbf{i}|$ the length of $\mathbf{i}$ and let $[\mathbf{i}]$ denote the corresponding cylinder set, that is, $[\mathbf{i}] := \{\mathbf{j} \in \Sigma : \mathbf{j}|_{|\mathbf{i}|} = \mathbf{i}\}$. We use $l(\cdot)$ to denote the level of cylinder. Moreover, The space Σ is clearly metrisable with metric

$$d(\mathbf{i}, \mathbf{j}) = e^{-\min\{n \geq 0 : i_n \neq j_n\}}. \tag{2.1}$$

In some cases, we extend our interest to a special family of σ-invariant compact sets. Let $\mathbf{A}$ be a $K \times K$ matrix with entries $0, 1$, and we say that the set $\Sigma_{\mathbf{A}} \subseteq \Sigma$ is *subshift of finite type* if

$$\Sigma_{\mathbf{A}} = \{\mathbf{i} = (i_0, i_1, \ldots) \in \mathcal{A}^{\mathbb{N}} : \mathbf{A}_{i_k, i_{k+1}} = 1 \text{ for every } k = 0, 1, \ldots\}.$$

We call the matrix $\mathbf{A}$ the adjacency matrix. Let us denote the set of admissible words with length n (i.e. n-length subwords of some element in $\Sigma_{\mathbf{A}}$) by $\Sigma_{\mathbf{A},n}$ and denote $\Sigma_{\mathbf{A},*}$ the set of all admissible words. Without loss of generality, we may assume that $\Sigma_{\mathbf{A},1} = \mathcal{A}$. Moreover, we say that $\Sigma_{\mathbf{A}}$ is *aperiodic and irreducible* if there exists $r \geq 1$ such that every entry of $\mathbf{A}^r$ is strictly positive.

2.1. **Topological entropy.** Before we turn to our main results, let us recall here the definition of topological entropy on the shift space. Let $\Sigma = \mathcal{A}^{\mathbb{N}}$ be the symbolic space. Let $E \subset \Sigma$. Define

$$\mathcal{H}_r^s(E) := \inf_{\alpha} \sum_{C \in \alpha} e^{-sl(C)}$$

where α is taken over all covers consisting of cylinders of levels large than r. Clearly, $\mathcal{H}_r^s(E)$ is increasing as a function of r. We define

$$\mathcal{H}^s(E) := \lim_{r \to \infty} \mathcal{H}_r^s(E) \in [0, +\infty].$$

The *topological entropy* of E is the value where the above limit jumps from $+\infty$ to 0, that is,

$$h_{\text{top}}(E) := \inf\{s \geq 0 : \mathcal{H}^s(E) < +\infty\}.$$

An upper bound of $h_{\text{top}}(E)$ is given by

$$h_{\text{top}}(E) \leq \liminf_{n \to \infty} \frac{1}{n} \log \#\{\mathbf{i} \in \Sigma_n : [\mathbf{i}] \cap E \neq \emptyset\}. \tag{2.2}$$

In fact, the reason is that we can always take a cover with cylinders of level n when estimating $\mathcal{H}_n^s(E)$. If E is a closed σ-invariant set, then the equality holds (see for example [22, Theorem 2.6]). However, the equality does not necessarily hold in general, because there might exist a better cover (in the sense that we could get smaller value of $\sum_{C \in \alpha} e^{-sl(C)}$) than covers consisting of cylinders of level n.

To get the lower bound, one has a version of Frostman Lemma as follows.

Lemma 2.1. *Let $E \subset \Sigma$. Suppose that there exists a probabilistic measure μ on E satisfying that there is a constant c such that for every cylinder C, we have $\mu(C \cap E) \leq ce^{-sl(C)}$. Then $h_{\mathrm{top}}(E) \geq s$.*

2.2. **Continuity of the entropy.** The first aspect of the study is the continuity of the entropy in a more general setting than weighted Birkhoff averages. That is, let $\Sigma_{\mathbf{A}}$ be an aperiodic and irreducible subshift of finite type and let $\phi_i : \Sigma_{\mathbf{A}} \to \mathbb{R}$ be a sequence of continuous potentials. We say that the sequence of potentials $\{\phi_i\}$ are *uniformly equicontinuous* if

$$\rho_n^{(1)} := \sup_i \mathrm{var}_n(\phi_i), \tag{2.3}$$

is finite for every n and converges to 0 as n tends to ∞, where

$$\mathrm{var}_n(\phi) := \sup_{\mathbf{i} \in \Sigma_n} \sup_{\mathbf{j}, \mathbf{k} \in [\mathbf{i}]} \{|\phi(\mathbf{j}) - \phi(\mathbf{k})|\}.$$

For $\mathbf{i} \in \Sigma_{\mathbf{A}}$, let

$$\overline{A}(\mathbf{i}) := \limsup_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi_i(\sigma^i \mathbf{i}),$$

$$\underline{A}(\mathbf{i}) := \liminf_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi_i(\sigma^i \mathbf{i}).$$

Moreover, if the limit exists let

$$A(\mathbf{i}) := \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi_i(\sigma^i \mathbf{i}).$$

Given $\alpha \leq \beta \in \mathbb{R}$, let

$$L_{\mathbf{A}}(\alpha, \beta) = \{\mathbf{i} \in \Sigma_{\mathbf{A}} : \underline{A}(\mathbf{i}) = \alpha \text{ and } \overline{A}(\mathbf{i}) = \beta\}.$$

For short, let $L_{\mathbf{A}}(\alpha) := L_{\mathbf{A}}(\alpha, \alpha)$. Now we state our first main result.

Theorem 2.2. *Let $\Sigma_{\mathbf{A}} \subseteq \Sigma$ be an aperiodic and irreducible subshift of finite type. For every sequence $\phi_i \colon \Sigma_{\mathbf{A}} \mapsto \mathbb{R}$ of uniformly equicontinuous potentials, the function $\alpha \mapsto h_{\mathrm{top}}(L_{\mathbf{A}}(\alpha))$ is continuous and concave over its domain, which is a (possibly empty) closed interval.*

In his paper, Fan [11] gave upper and lower bounds for $h_{\mathrm{top}}(L_{\mathbf{A}}(\alpha))$ in case of full shift by using a generalized topological pressure generated by the sequence ϕ_i. If the pressure is sufficiently smooth then these bounds agree.

It is a natural question how large is the set of irregular points, that is, let

$$D := \left\{ \mathbf{i} \in \Sigma_{\mathbf{A}} : \lim_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi_i(\sigma^i \mathbf{i}) \text{ does not exists} \right\}.$$

Theorem 2.3. *Let $\Sigma_{\mathbf{A}} \subseteq \Sigma$ be an aperiodic and irreducible subshift of finite type. Let $\phi_i \colon \Sigma_{\mathbf{A}} \mapsto \mathbb{R}$ be a sequence of uniformly equicontinuous potentials. Assume that $A(\mathbf{i})$ takes at least two possible values, that is, the domain of the function $\alpha \mapsto h_{\mathrm{top}}(L_{\mathbf{A}}(\alpha))$ is a non-trivial interval. Then*

$$h_{\mathrm{top}}(D) = h_{\mathrm{top}}(\Sigma_{\mathbf{A}}).$$

2.3. **Random weights.** Let us now extend our symbolic space $\Sigma = \mathcal{A}^{\mathbb{N}}$. Namely, Let $\Lambda = \{1, \ldots, N\}$ be another finite alphabet, and let $\Omega = \Lambda^{\mathbb{N}}$ be compact left-shift invariant subsets. Let us define the extended symbolic space $\Gamma := \Omega \times \Sigma$. As an abuse of notation, we denote the left-shift operator on Ω, and Γ by σ too. Adapting the notations for Ω and Γ, let Ω_n and Γ_n be the set of n-length finite words, and denote Ω_* and Γ_* the set of all finite words. The spaces Ω, Σ and Γ are clearly metrisable with the same metric defined in (2.1). For short, denote $\mathbf{i} \wedge \mathbf{j} = \min\{n \geq 0 : i_n \neq j_n\}$.

For an aperiodic and irreducible subshift of finite type $\Sigma_{\mathbf{A}} \subseteq \Sigma$, the set $\Gamma_{\mathbf{A}} = \Omega \times \Sigma_{\mathbf{A}}$ is an aperiodic and irreducible subshift of finite type as well. Denote the set of finite admissible words by $\Gamma_{\mathbf{A},*}$ and that of words of length n by $\Gamma_{\mathbf{A},n}$. Let $f \colon \Gamma_{\mathbf{A}} \mapsto \mathbb{R}^d$ be a continuous potential. For a given sequence $\mathbf{w} \in \Omega$ and $\alpha \in \mathbb{R}^d$ let

$$E_{\mathbf{w}}(\alpha) := \left\{ \mathbf{i} \in \Sigma_{\mathbf{A}} : \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\sigma^k \mathbf{w}, \sigma^k \mathbf{i}) = \alpha \right\}.$$

Our goal is to determine the topological entropy of $E_{\mathbf{w}}(\alpha)$, at least for the case of typical $\mathbf{w} \in \Omega$. In order to do so, we need to introduce further regularity properties on f and on the choice of $\mathbf{w}$.

We say that the potential $f \colon \Omega \times \Sigma_{\mathbf{A}} \mapsto \mathbb{R}^d$ has *summable variation* if

$$\sum_{k=0}^{\infty} \max_{\substack{(\mathbf{w},\mathbf{i}),(\mathbf{z},\mathbf{j}) \in \Gamma_{\mathbf{A},*}: \\ (\mathbf{w},\mathbf{i}) \wedge (\mathbf{z},\mathbf{j}) = k}} |f(\mathbf{w}, \mathbf{i}) - f(\mathbf{z}, \mathbf{j})| < \infty.$$

Let ν be a σ-invariant ergodic measure on Ω. We say that ν is *quasi-Bernoulli* if there exists $C > 0$ such that for every $\mathbf{w}, \mathbf{z} \in \Omega_*$ with $\mathbf{w}\mathbf{z} \in \Omega_*$

$$C^{-1}\nu([\mathbf{w}])\nu([\mathbf{z}]) \leq \nu([\mathbf{w}\mathbf{z}]) \leq C\nu([\mathbf{w}])\nu([\mathbf{z}]).$$

Denote by Π the natural projection $\Pi \colon \Omega \times \Sigma \mapsto \Omega$, that is, $\Pi(\mathbf{w}, \mathbf{i}) = \mathbf{w}$. Denote by $\mathcal{E}_{\nu}(\Gamma), \mathcal{E}_{\nu}(\Gamma_{\mathbf{A}})$ the set of ergodic σ-invariant measures on Γ and $\Gamma_{\mathbf{A}}$ respectively, whose marginal is ν, i.e., $\Pi_* \mu = \nu$. Denote by $\mathcal{M}_{\nu}(\Gamma)$ and $\mathcal{M}_{\nu}(\Gamma_{\mathbf{A}})$ the set of σ-invariant measures on Γ and $\Gamma_{\mathbf{A}}$ with marginal ν. Let

$$\mathcal{P}_{\mathbf{A}} = \{\alpha \in \mathbb{R}^d : \text{ there exists } \mu \in \mathcal{M}_{\nu}(\Gamma_{\mathbf{A}}) \text{ such that } \int f d\mu = \alpha\}. \qquad (2.4)$$

Denote the relative interior of $\mathcal{P}_{\mathbf{A}}$ by $\mathcal{P}_{\mathbf{A}}^o$.

Moreover, let us define the conditional pressure of a potential $f \colon \Gamma_{\mathbf{A}} \mapsto \mathbb{R}$ by

$$P_{\nu}(f) = \lim_{n \to \infty} \frac{1}{n} \int \log \sum_{\mathbf{i} \in \Sigma_n} \sup_{\mathbf{j} \in [\mathbf{i}]} e^{S_n f(\mathbf{w}, \mathbf{j})} d\nu(\mathbf{w}), \qquad (2.5)$$

where $S_n f = f + f \circ \sigma + \cdots + f \circ \sigma^{n-1}$ and log is taken in the base e. Throughout the paper we will use the convention that $0 \cdot \log 0 = 0$. Moreover, we note that we define the supremum over an empty set as $-\infty$ and the topological entropy of an empty set as $-\infty$. Now, we can formalise our second theorem.

Theorem 2.4. *Let $\Sigma_{\mathbf{A}} \subseteq \Sigma$ be an aperiodic and irreducible subshift of finite type, and let ν be a quasi-Bernoulli σ-invariant ergodic measure on Ω. Moreover, let $f \colon \Omega \times \Sigma_{\mathbf{A}} \mapsto \mathbb{R}^d$ be a continuous map with summable variation. Then for every $\alpha \in \mathcal{P}_{\mathbf{A}}^o$ and for ν-almost every $\mathbf{w} \in \Omega$,*

$$h_{\mathrm{top}}(E_{\mathbf{w}}(\alpha)) = \sup\{h_\mu : \mu \in \mathcal{E}_\nu(\Gamma_{\mathbf{A}}) \ and \ \int f d\mu = \alpha\} - h_\nu$$

$$= \sup\{h_\mu : \mu \in \mathcal{M}_\nu(\Gamma_{\mathbf{A}}) \ and \ \int f d\mu = \alpha\} - h_\nu$$

$$= \inf_{\underline{p} \in \mathbb{R}^d} P_\nu(\langle \underline{p}, f - \alpha\rangle).$$

Furthermore, there exists $\alpha_0 \in \mathbb{R}^d$ such that for ν-almost every $\mathbf{w}$,

$$h_{\mathrm{top}}(E_{\mathbf{w}}(\alpha_0)) = h_{\mathrm{top}}(\Sigma_{\mathbf{A}}). \tag{2.6}$$

Combining Theorem 2.2 and Theorem 2.4 we get the following stronger result for real valued potentials, which shows that for a typical sequence of weights it is possible to calculate the whole spectrum.

Theorem 2.5. *Let $\Sigma_{\mathbf{A}} \subseteq \Sigma$ be an aperiodic and irreducible subshift of finite type, and let ν be a quasi-Bernoulli σ-invariant ergodic measure on Ω. Moreover, let $f \colon \Omega \times \Sigma_{\mathbf{A}} \mapsto \mathbb{R}$ be a continuous map with summable variation. Then for ν-almost every $\mathbf{w} \in \Omega$,*

$$h_{\mathrm{top}}(E_{\mathbf{w}}(\alpha)) = \sup\{h_\mu : \mu \in \mathcal{E}_\nu(\Gamma_{\mathbf{A}}) \ and \ \int f d\mu = \alpha\} - h_\nu$$

$$= \sup\{h_\mu : \mu \in \mathcal{M}_\nu(\Gamma_{\mathbf{A}}) \ and \ \int f d\mu = \alpha\} - h_\nu$$

$$= \inf_{p \in \mathbb{R}} (P_\nu(p \cdot f) - \alpha \cdot p) \ for \ every \ \alpha \in \mathbb{R}.$$

Moreover, for ν-almost every $\mathbf{w}$, the map $\alpha \mapsto h_{\mathrm{top}}(E_{\mathbf{w}}(\alpha))$ is continuous and concave over its domain $\mathcal{P}_{\mathbf{A}}$.

Fan [11] proved some similar results. Namely, he showed a version of Theorem 2.5 for full shifts with the choice $f(\sigma^k \mathbf{w}, \sigma^k \mathbf{i}) = w_k \varphi(\sigma^k \mathbf{i})$, where $(w_k)_k$ is an ergodic sequence of real random variables or deduced from a uniquely ergodic dynamical system, and φ depends only on a finite number of coordinates. In this cases, he shows analyticity of the conditional topological pressure, while our result only gives continuity.

2.4. **Potentials depending on the first coordinate.** Let us assume that $f \colon \Omega \times \Sigma \mapsto \mathbb{R}$ depends only on the first symbol, that is, $f(\mathbf{w}, \mathbf{i}) = f_{w_0, i_0}$. Then for a $\mathbf{w} \in \Omega$,

$$E_{\mathbf{w}}(\alpha) := \left\{ \mathbf{i} \in \Sigma : \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} f_{w_k, i_k} = \alpha \right\}.$$

Let $\underline{q} = (q_1, \ldots, q_N) \in \mathcal{S}_N$ be a probability vector, where $\mathcal{S}_N$ denotes the $(N-1)$-dimensional simplex. We say that $\mathbf{w} \in \Omega$ is $\underline{q}$-*frequency regular*, if

$$\lim_{n \to \infty} \frac{\#\{k \in [0, n] \cap \mathbb{Z} : \omega_k = i\}}{n} = q_i \ \text{for every } i = 1, \ldots, N. \tag{2.7}$$

Notice that there is a bijection between $\mathcal{S}_N$ and the probability Bernoulli measure on Ω. In this case, we choose $\nu = \nu_{\underline{q}}$ to be the Bernoulli measure on Ω. Then for

the potential $f \colon \Gamma \mapsto \mathbb{R}$, the conditional pressure has the form

$$P_{\underline{q}}(\langle \underline{p}, f - \alpha \rangle) = \sum_{j=1}^{N} q_j \log \sum_{i=1}^{K} e^{\langle \underline{p}, f_{j,i} - \alpha \rangle}. \tag{2.8}$$

Denote by $\mathcal{B}_{\underline{q}}(\Gamma)$ the set of all Bernoulli measures on Γ with marginal ν. That is, let $(p_{j,i})_{j=1,i=1}^{N,K} \in \mathcal{B}_{\underline{q}}(\Gamma) \subset \mathcal{S}_{NK}$ such that $\sum_{i=1}^{K} p_{j,i} = q_j$. Our third main result is as follows.

Theorem 2.6. *Let* $\mathbf{w} \in \{1, \ldots, N\}^{\mathbb{N}}$ *be a $\underline{q}$-frequency regular sequence with frequencies* $(q_1, \ldots, q_N)$. *Then for every* $\alpha \in \mathbb{R}$.

$$h_{\text{top}}(E_{\mathbf{w}}(\alpha)) = \sup_{(p_{j,i}) \in \mathcal{B}_{\underline{q}}(\Gamma)} \left\{ -\sum_{i,j} p_{j,i} \log p_{ji} : \sum_{i,j} p_{j,i} f_{j,i} = \alpha \right\} + \sum_{i=1}^{N} q_i \log q_i$$

$$= \inf_{p \in \mathbb{R}} \left\{ P_{\underline{q}}(p \cdot f) - p\alpha \right\}.$$

Comparing Theorem 2.4 with Theorem 2.5, in the general setup of Theorem 2.4 we are only able to show that for any possible value of α one can find a full measure set Ω_α, which might depend on α, while in Theorem 2.5 in the one dimensional case, we manage to show that there exists a universal full measure set Ω, for which any $\mathbf{w} \in \Omega$ the spectrum $\alpha \mapsto h_{\text{top}}(E_{\mathbf{w}}(\alpha))$ can be determined. Comparing Theorem 2.6 with Theorem 2.5, we can construct a Bernoulli measure with probabilities $\underline{q} = (q_i)_i$, for which the $\underline{q}$-frequency regular sequences will be a set of full measure. That is, we can explicitly construct the set Ω of full measure for which the spectrum can be determined. In this sense Theorem 2.6 is a strengthening of Theorem 2.5 for this particular class of systems.

Fan [11] also gave a similar result in his recent paper. Namely, Fan shows Theorem 2.6 under a weaker condition that φ depends on finitely many coordinates but under the stronger assumption that it takes only values $-1, 1$.

Now we state the corresponding version of Theorem 2.3 for the frequency regular case. Similarly, let

$$D_{\mathbf{w}} = \left\{ \mathbf{i} \in \Sigma : \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} f_{w_k, i_k} \text{ does not exists} \right\}.$$

Theorem 2.7. *Let* $\mathbf{w} \in \{1, \ldots, N\}^{\mathbb{N}}$ *be a $\underline{q}$-frequency regular sequence with frequencies* $(q_1, \ldots, q_N)$. *Suppose that* $g_i = \sum_{j=1}^{N} q_j f_{j,i}$ *is not constant as function of* i. *Then*

$$h_{\text{top}}(D_{\mathbf{w}}) = \log K.$$

2.5. Examples: weighted Birkhoff averages with frequency regular weights.

Now, we show examples and demonstrate our result on the spectrum of real valued potentials depending on the first coordinate and frequency regular weights. Here we assume again that our potentials are supported on the whole spaces $\Sigma = \mathcal{A}^{\mathbb{N}}$, $\Omega = \Lambda^{\mathbb{N}}$ and the potential $\varphi \colon \Sigma \mapsto \mathbb{R}$ and the weight $\lambda \colon \Omega \mapsto \mathbb{R}$ depend only on the first symbol, that is, $\varphi(\mathbf{i}) = \varphi_{i_0}$ and $\lambda(\mathbf{w}) = \lambda_{w_0}$. Then for a $\mathbf{w} \in \Omega$, let

$$E_{\mathbf{w}}(\alpha) := \left\{ \mathbf{i} \in \Sigma : \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \lambda_{w_k} \varphi_{i_k} = \alpha \right\}.$$

Let $\underline{q} = (q_1, \ldots, q_N) \in \mathcal{S}_N$ be a probability vector, and let $\mathbf{w} \in \Omega$ be an arbitrary $\underline{q}$-frequency regular sequence. Denote by $\varphi_{\max} = \max\{\varphi_i : 1 \leq i \leq K\}$ and $\varphi_{\min} = \min\{\varphi_i : 1 \leq i \leq K\}$. To avoid the trivial case, we assume $\varphi_{\max} \neq \varphi_{\min}$. Finally, let

$$I = \left[\varphi_{\min} \sum_{\lambda_j > 0} q_j \lambda_j + \varphi_{\max} \sum_{\lambda_j < 0} q_j \lambda_j, \varphi_{\max} \sum_{\lambda_j > 0} q_j \lambda_j + \varphi_{\min} \sum_{\lambda_j < 0} q_j \lambda_j \right]. \quad (2.9)$$

Now we show a compatible form of $h_{\text{top}}(E_{\mathbf{w}}(\alpha))$ in order to compute some examples.

Example 2.8. *Let $\mathbf{w} \in \{1, \ldots, N\}^{\mathbb{N}}$ be a $\underline{q}$-frequency regular sequence with frequencies $(q_1, \ldots, q_N)$. Then for every $\alpha \in I$*

$$h_{\text{top}}(E_{\mathbf{w}}(\alpha)) = \sum_{j=1}^{N} q_j \log \sum_{i=1}^{K} e^{p(\lambda_j \varphi_i - \alpha)},$$

where p is the unique solution of the equation

$$\sum_{j=1}^{N} q_j \lambda_j \frac{\sum_{i=1}^{N} \varphi_i e^{p\lambda_j \varphi_i}}{\sum_{i=1}^{N} e^{p\lambda_j \varphi_i}} = \alpha. \quad (2.10)$$

Moreover, if $\alpha \notin I$, $\inf_p P_\nu(f_p) = -\infty$, that is, there is no $\underline{p}^* \in \mathbb{R}^d$ such that $\inf_p P_\nu(f_p) = P_\nu(f_{p^*})$.

Proof. Let $\alpha \in I$. For sake of simplicity, denote $P(p) = P_{\underline{q}}(p(\lambda\varphi - \alpha))$. It is easy to check by (2.8) that

$$P(p) = \sum_{j=1}^{N} q_j \log \sum_{i=1}^{N} e^{p\lambda_j \varphi_i} - p\alpha.$$

It follows that

$$P'(p) = \sum_{j=1}^{N} q_j \lambda_j \frac{\sum_{i=1}^{N} \varphi_i e^{p\lambda_j \varphi_i}}{\sum_{i=1}^{N} e^{p\lambda_j \varphi_i}} - \alpha,$$

and

$$P''(p) = \sum_{j=1}^{N} q_j \lambda_j^2 \frac{(\sum_{i=1}^{N} \varphi_i^2 e^{p\lambda_j \varphi_i})(\sum_{i=1}^{N} e^{p\lambda_j \varphi_i}) - (\sum_{i=1}^{N} \varphi_i e^{p\lambda_j \varphi_i})^2}{(\sum_{i=1}^{N} e^{p\lambda_j \varphi_i})^2}.$$

Since $\varphi_{\max} \neq \varphi_{\min}$, by Cauchy-Schwarz inequality, we see that $P''(p) > 0$ for all $p \in \mathbb{R}$. A simple computation shows that

$$P'(-\infty) = \varphi_{\min} \sum_{\lambda_j > 0} q_j \lambda_j + \varphi_{\max} \sum_{\lambda_j < 0} q_j \lambda_j - \alpha < 0,$$

and

$$P'(+\infty) = \varphi_{\max} \sum_{\lambda_j > 0} q_j \lambda_j + \varphi_{\min} \sum_{\lambda_j < 0} q_j \lambda_j - \alpha > 0.$$

Thus $P'(p) = 0$ has a unique solution at which P achieves minima.

Now let $\alpha \notin I$. It is easy to calculate that

$$P(-\infty) = \lim_{p \to -\infty} p(\varphi_{\min} \sum_{\lambda_j > 0} q_j \lambda_j + \varphi_{\max} \sum_{\lambda_j < 0} q_j \lambda_j - \alpha),$$

and

$$P(+\infty) = \lim_{p \to -\infty} p(\varphi_{\max} \sum_{\lambda_j > 0} q_j \lambda_j + \varphi_{\min} \sum_{\lambda_j < 0} q_j \lambda_j) - \alpha).$$

Thus $\inf_p P(p) = -\infty$. □

Example 2.9. *Let us consider again the Möbius sequence with the potential $\varphi(\mathbf{i}) = i_0$ for $\mathbf{i} \in \Sigma = \{0, \dots, N-1\}^{\mathbb{N}}$. The Möbius function is frequency regular with*

$$\lim_{n \to \infty} \frac{\#\{0 \le i \le n-1 : \boldsymbol{\mu}(i) = \pm 1\}}{n} = \frac{3}{\pi^2} \text{ and}$$

$$\lim_{n \to \infty} \frac{\#\{0 \le i \le n-1 : \boldsymbol{\mu}(i) = 0\}}{n} = 1 - \frac{6}{\pi^2},$$

see for example [6]. *As a special case of Example* 2.8 *for $\varphi \colon \{0, \dots, N-1\}^{\mathbb{N}} \mapsto \mathbb{R}$ with $\varphi(\mathbf{i}) = i_0$, we get*

$$h_{\text{top}}(E_{\boldsymbol{\mu}}(\alpha)) = \left(1 - \frac{6}{\pi^2}\right) \log(N) + \frac{6}{\pi^2} \log\left(\frac{e^{pN} - 1}{e^p - 1}\right) - \left((N-1)\frac{3}{\pi^2} + \alpha\right) p,$$

where $H(\underline{p}) = -\sum_i p_i \log p_i$. and p is the unique solution of

$$\frac{(e^{(N+1)p} - 1)(N-1) - (N+1)(e^{Np} - e^p)}{(e^{Np} - 1)(e^p - 1)} = \frac{\pi^2 \alpha}{3}, \text{ for } \alpha \in \left[\frac{-(N-1)3}{\pi^2}, \frac{(N-1)3}{\pi^2}\right].$$

A corollary of the above results is that non-degenerate weights and potentials give us non-degenerate weighted spectrum.

Corollary 2.10. *Let $\mathbf{w} \in \Omega$ be a frequency regular sequence with frequencies $(q_1, \dots, q_N)$ with non-degenerate weights, i.e. $\sum_{j=1}^{N} q_j |\lambda_j| > 0$. Let $\varphi \colon \Sigma \mapsto \mathbb{R}$ be a potential depending only on the first coordinate. Then there exists $\alpha_0 \in I$ such that $h_{\text{top}}(E_{\mathbf{w}}(\alpha_0)) = \log K$. Moreover, the domain I is a non-degenerate closed interval unless the potential $\varphi(\mathbf{i}) = \varphi_{i_0}$ is constant. In particular, either the limit of the weighted Birkhoff average at every point exists and equals α_0 or the set of points at which the limit of the weighted Birkhoff average does not exist has full topological entropy.*

Proof. The first assertion follows by Theorem 2.6 for $f_{j,i} = \lambda_j \varphi_i$ with the choice $p_{j,i} = \frac{q_j}{K}$ and $\alpha_0 = \left(\sum_{i=1}^{K} \frac{\varphi_i}{K}\right)\left(\sum_{j=1}^{N} q_j \lambda_j\right)$. Moreover, (2.9), Theorem 2.8 and the continuity of the spectrum give the second claim by some algebraic manipulation. The proof can be finished by applying Theorem 2.7. □

The difference between the usual Birkhoff averages and weighted Birkhoff averages is shown by the following example:

Example 2.11. *On the full shift system $\Sigma = \{0,1\}^{\mathbb{N}}$ there exist a potential $\varphi \colon \Sigma \mapsto \mathbb{R}$ depending only on the first symbol and a bounded sequence of weights $\mathbf{w} = (w_i)_i$ (which is not frequency regular) such that*

- *The $\alpha = 0$ is the only possible value of the limit of a weighted Birkhoff average,*
- *$0 < h_{\text{top}}(E_{\mathbf{w}}(0)) < \log 2$. At all the points in $\Sigma \setminus E_{\mathbf{w}}(0)$ the limit of the weighted Birkhoff average does not exist.*

In particular, to have non-degenerate weighted spectrum, the frequency regularity of the weights is somewhat necessary. The proof of the example will be given in the last section.

Structure of paper. In Section 2.1, we recall the definition and basic properties of topological entropy. In Section 3, we prove Theorem 2.2 and Theorem 2.3. In Section 4, we prove Theorem 2.4 and Theorem 2.5. In Section 5, as an application of Theorem 2.2 and Theorem 2.4, we show Theorem 2.6 and Theorem 2.7. We remark that the proof of Theorem 2.2 and that of Theorem 2.4 are independent. Thus the readers who are interested in Theorem 2.4 may read directly Section 4.

3. **Continuity and concavity of the spectrum.** Let us recall the conditions and notations of Theorem 2.2. That is, we assume that $\Sigma_{\mathbf{A}} \subseteq \Sigma = \mathcal{A}^{\mathbb{N}}$ is an aperiodic and irreducible subshift of finite type. Moreover, let $\phi_i : \Sigma_{\mathbf{A}} \to \mathbb{R}$ be a sequence of uniformly equicontinuous potentials. For $\mathbf{i} \in \Sigma_{\mathbf{A}}$, let

$$\overline{A}(\mathbf{i}) := \limsup_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi_i(\sigma^i \mathbf{i}) \text{ and } \underline{A}(\mathbf{i}) := \liminf_{n \to \infty} \frac{1}{n} \sum_{i=0}^{n-1} \phi_i(\sigma^i \mathbf{i}).$$

Given $\alpha \leq \beta \in \mathbb{R}$, let

$$L_{\mathbf{A}}(\alpha, \beta) = \{\mathbf{i} \in \Sigma : \underline{A}(\mathbf{i}) = \alpha \text{ and } \overline{A}(\mathbf{i}) = \beta\}.$$

For short, let $L_{\mathbf{A}}(\alpha) := L_{\mathbf{A}}(\alpha, \alpha)$. Define

$$B_m^n(\mathbf{i}) := \sum_{i=m}^{n-1} \phi_i(\sigma^i \mathbf{i})$$

and $A_m^n(\mathbf{i}) = \frac{1}{n-m} B_m^n(\mathbf{i})$. Let

$$\rho_n^{(2)} := \sup_{\mathbf{i} \in \Sigma_{\mathbf{A},n}} \sup_{\mathbf{j},\mathbf{k} \in [\mathbf{i}]} |A_0^n(\mathbf{j}) - A_0^n(\mathbf{k})|,$$

for $m, n \in \mathbb{N}$ with $n > m$. It is clear that

$$\rho_n^{(2)} \leq \frac{1}{n} \sum_{i=1}^n \rho_i^{(1)}. \tag{3.1}$$

Since $\rho_n^{(1)}$ converges to 0 as n tends to ∞, so does $\rho_n^{(2)}$.

Lemma 3.1. *Let $\varepsilon > 0$ and $N \in \mathbb{N}$. Suppose that $|A_0^n(\mathbf{i}) - \alpha| < \varepsilon$ for all $n > N$. Then for $m, n > N$ we have*

$$|A_m^n(\mathbf{i}) - \alpha| \leq \varepsilon \frac{n+m}{n-m}.$$

Proof. The statement follows simply from $(n - m)A_m^n(\mathbf{i}) = nA_0^n(\mathbf{i}) - mA_0^m(\mathbf{i})$. □

We remind that for an aperiodic and irreducible subshift of finite type $\Sigma_{\mathbf{A}}$ there exists a constant r such that for any two admissible words $\mathbf{i}, \mathbf{j} \in \Sigma_{\mathbf{A},*}$ there exists a word $\mathbf{k}$ of length r such that the concatenation $\mathbf{ikj}$ is admissible, moreover one can choose $\mathbf{k}$ depending only on the last symbol of $\mathbf{i}$ and the first symbol of $\mathbf{j}$. We fix r for the rest of the section.

We will need the following technical lemma. Note that although the sequence ϕ_i is defined only on $\Sigma_{\mathbf{A}} \subseteq \Sigma$, it can be naturally extended to Σ in such a way that the sequence remains uniformly equicontinuous. For instance, for every $\mathbf{i} \in \Sigma$ let $n(\mathbf{i}) = \inf\{n \geq 0 : \mathbf{i}|_0^n \in \Sigma_{\mathbf{A},*}\}$, that is, $\mathbf{i}|_0^{n(\mathbf{i})}$ is the longest admissible prefix of $\mathbf{i}$ and let $\phi_i(\mathbf{i}) := \max_{\mathbf{j} \in [\mathbf{i}|_0^{n(\mathbf{i})}]} \phi_i(\mathbf{j})$. We consider a map π in the following lemma, which illustrates that the concatenation of a sequence of admissible words can be changed into an admissible infinite sequence without changing the weighted Birkhoff average.

Lemma 3.2. *Let $(q_j)_{j=1}^{\infty}$ be an increasing sequence of integers satisfying $q_j/j \to \infty$ and $q_{j+1} - q_j > 2r$. Let $\pi \colon \Sigma \to \Sigma$ be a map satisfying the following properties for every $n \in \mathbb{N}$:*

i) if $\mathbf{i}|_0^n = \mathbf{j}|_0^n$ then $(\pi\mathbf{i})|_0^{q_j} = (\pi\mathbf{j})|_0^{q_j}$ for j such that $q_j < n \leq q_{j+1}$,
ii) if $\mathbf{i}|_n^n \neq (\pi\mathbf{i})|_n^n$ then $n \in \{q_j + 1, \ldots, q_j + r\}$ for some j.

Then there exists a sequence $\rho_n^{(3)} \searrow 0$ such that for every $\mathbf{i} \in \Sigma$ and for every n

$$|A_0^n(\pi\mathbf{i}) - A_0^n(\mathbf{i})| < \rho_n^{(3)}.$$

Moreover, for every $X \subset \Sigma$

$$h_{\mathrm{top}}(\pi(X)) = h_{\mathrm{top}}(X).$$

Proof. Taking j such that $q_j < n \leq q_{j+1}$ we get

$$|A_0^n(\pi\mathbf{i}) - A_0^n(\mathbf{i})| \leq \frac{(j+1)r}{n} \max_{i \geq 0, \mathbf{i} \in \mathcal{A}^{\mathbb{N}}} |\phi_i(\mathbf{i})| + \frac{1}{n} \sum_{i=1}^{j} \sum_{\ell=0}^{q_i - q_{i-1} - r} \rho_\ell^{(1)} + \frac{1}{n} \sum_{\ell=\max\{q_{j+1} - q_j - r - n, 0\}}^{q_{j+1} - q_j - r} \rho_\ell^{(1)},$$

$$\leq \frac{(j+1)r}{n} \max_{i \geq 0, \mathbf{i} \in \mathcal{A}^{\mathbb{N}}} |\phi_i(\mathbf{i})| + \frac{1}{n} \sum_{i=1}^{j} (q_i - q_{i-1} - r)\rho_{q_i - q_{i-1} - r}^{(2)} + \rho_n^{(2)}.$$

Observe that $\frac{1}{n} \sum_{i=1}^{j} (q_i - q_{i-1} - r)\rho_{q_i - q_{i-1} - r}^{(2)} \to 0$ as $n \to \infty$. Indeed, since $q_j - q_{j-1} - r \to \infty$ as $j \to \infty$, for every $\varepsilon > 0$ there exists $J > 0$ so that for every $i \geq J$ $\rho_{q_i - q_{i-1} - r}^{(2)} < \varepsilon$ and thus, $\frac{1}{n} \sum_{i=1}^{j} (q_i - q_{i-1} - r)\rho_{q_i - q_{i-1} - r}^{(2)} \leq \frac{q_j - q_{J-1}}{n}\varepsilon + \frac{q_J \rho_1^{(2)}}{n}$. This proves the first assertion.

To prove the second assertion, we need a lower and an upper bound. For the upper bound we notice that the image under π of a cylinder whose level is not of form $\{q_j + 1, \ldots, q_j + 2r\}$ is contained in a cylinder of the same level. As for any set X we can construct a family of covers realizing the topological entropy using only cylinders of levels not of form $\{q_j + 1, \ldots, q_j + 2r\}$, the images of those cylinders will give us a family of covers of $\pi(X)$ realizing the same topological entropy.

For the lower bound, let μ be a measure supported on X such that for every cylinder C of level $\ell(C) = n$ we have

$$\mu(C \cap X) \leq e^{(h_{\mathrm{top}}(X) + \varepsilon)\ell(C)}.$$

Then if n is not of form $\{q_j + 1, \ldots, q_j + 2r\}$ but $q_j < n \leq q_{j+1}$ then for every cylinder C' of level n we have

$$\pi_*(\mu)(C') \leq K^{(j+1)r} e^{(h_{\mathrm{top}}(X) + \varepsilon)\ell(C)}.$$

Intuitively speaking but not quite precisely, the map π acting on initial words of length $\leq q_{j+1}$ is at most $K^{(j+1)r}$-to-1. As $j = o(q_j)$, the factor $K^{r(j+1)}$ is subexponential in q_j and thus we get the lower bound from Lemma 2.1. $\qquad\square$

The proof of Theorem 2.2 relies on the following technical proposition. It is a weighted version of the w-measure construction in Gelfert and Rams [20, Section 5.2]. In simple but very vague words, we have some collections of sequences with given weighted Birkhoff averages α_i and we concatenate proper parts of them to construct 'Frankenstein' sequences with weighted Birkhoff average $\lim \alpha_i$. Important part is that if our starting collections were large (of large topological entropy), we can do it in a way that the constructed set of sequences also has large topological entropy.

Proposition 3.3. *Let $\varepsilon_n > 0$ and α_n be sequences of reals such that $\limsup_{n\to\infty} \alpha_n = \alpha_{\max}$, $\liminf_{n\to\infty} \alpha_n = \alpha_{\min}$ and $\lim_{n\to\infty} \varepsilon_n = 0$. Moreover, assume that for every $n \geq 1$ there exists a set $M_n \subset \Sigma_{\mathbf{A}}$ and a positive integer $T_n > 0$ such that for every $\mathbf{i} \in M_n$ and $m \geq T_n$*

$$\left| \frac{1}{m} \sum_{k=0}^{m-1} \phi_k(\sigma^k \mathbf{i}) - \alpha_n \right| < \varepsilon_n.$$

Then $h_{\text{top}}(L_{\mathbf{A}}(\alpha_{\min}, \alpha_{\max})) \geq \liminf_{n\to\infty} h_{\text{top}}(M_n)$.

Moreover, in case $\lim_{n\to\infty} \alpha_n = \alpha$ then there exists a set $M \subset L_{\mathbf{A}}(\alpha)$ such that the convergence $A_0^n(\mathbf{i}) \to \alpha$ is uniform on M and $h_{\text{top}}(M) \geq \limsup_{n\to\infty} h_{\text{top}}(M_n)$.

For a subset $M \subset \Sigma$, we denote by $M[a,b] = \{\mathbf{i} \in \mathcal{A}^{b-a+1} : \exists \mathbf{j} \in M, \mathbf{j}|_a^b = \mathbf{i}\}$. That is, the collection of $(b-a+1)$-words occurring in certain element of M starting at place a and ending at place b. Moreover, we use the notation $Z_a^b(M) = \#M[a, b-1]$ for convenience. It is clear that for $a < b < c$, we have $Z_a^c(M) \leq Z_a^b(M) \cdot Z_b^c(M)$.

Lemma 3.4. *Let M be a set with $h_{\text{top}}(M) > 0$. Then for every $h < h_{\text{top}}(M)$ there exists a sequence $(z_i)_{i\in\mathbb{N}}$ of $\mathbb{N}$ such that for every z_i and for every $n > z_i$ we have*

$$\log Z_{z_i}^n(M) > (n - z_i)h.$$

Proof. Indeed, if it fails then we would be able to find an increasing subsequence $(n_i)_{i\in\mathbb{N}}$ of $\mathbb{N}$ such that $\log Z_{n_i}^{n_{i+1}}(M) \leq (n_{i+1} - n_i)h$, and by summing them up this would imply $\log Z_0^{n_i}(M) \leq (n_i - n_0)h + \log Z_0^{n_0}(M)$, hence $h_{\text{top}}(M) \leq h$, which is a contradiction. $\square$

Proof of Proposition 3.3. Let M_k be the sequence of subsets and T_k as in the assumption. Moreover, let $\inf_k h_{\text{top}}(M_k) > \delta > 0$ be arbitrary but fixed. Then by Lemma 3.4, for every $k \in \mathbb{N}$ there exists a sequence $(z_i^k)_{i\in\mathbb{N}}$ such that

$$\log Z_{z_i^k}^n(M_k) > (n - z_i^k)(h_{\text{top}}(M_k) - \delta) \text{ for every } n \geq z_i^k. \tag{3.2}$$

We choose a subsequence $(N_k)_{k\in\mathbb{N}}$ of $\mathbb{N}$ satisfying the following properties:

- $N_0 = 0$, $N_{k-1} > T_k$;
- $N_k \in (z_i^{k+1})_{i\in\mathbb{N}}$;
- $\lim_{k\to\infty} \frac{N_{k+1}}{\sum_{j=1}^k N_j} = \infty$;
- $\log Z_0^n(M_1) \geq n(h_{\text{top}}(M_1) - \delta)$ for all $n \geq N_1$.

Now, let us define sequences $2 \geq r_k > 1$ and $m(k) \in \mathbb{N}$ such that

$$(r_k)^{m(k)} = \frac{N_k}{N_{k-1}}, \quad \lim_{k\to\infty} r_k = 1 \text{ and } \lim_{k\to\infty} (r_k - 1)\varepsilon_k^{-1} = \infty.$$

Define a sequence $(t_i^k)_{i=0}^{m(k)}$ by $t_i^k = \lfloor (r_k)^i N_{k-1} \rfloor$ for $i = 0, \ldots, m(k)$. By definition, $t_{m(k)}^k = t_0^{k+1}$. It is easy to check that

$$r_k - \frac{1}{N_{k-1}} \leq \frac{t_{i+1}^k}{t_i^k} \leq r_k + \frac{2}{N_{k-1}} \text{ for } 1 \leq i \leq m(k) - 1.$$

Finally, let

$$\widetilde{M} = \{\mathbf{i} \in \mathcal{A}^{\mathbb{N}} : \mathbf{i}|_0^{N_1-1} \in M_1[0, N_1 - 1],$$
$$\mathbf{i}|_{t_i^k}^{t_{i+1}^k-1} \in M_k[t_i^k, t_{i+1}^k - 1], \forall 0 \leq i \leq m(k) - 1, \forall k \geq 2\}$$
$$= M_1[0, N_1 - 1] \times \prod_{k=2}^{\infty} \prod_{i=0}^{m(k)-1} M_k[t_i^k, t_{i+1}^k - 1].$$

In other words, on positions $0, \ldots, N_1 - 1$ we can put any sequence that appears in M_1. For $k > 1$, on positions in each $[t_i^k, t_{i+1}^k - 1]$ we can put any sequence that can appear (on those positions) in M_k. Note that $\widetilde{M}$ is not necessarily a subset of $\Sigma_{\mathbf{A}}$, since it might happen that these concatenations are forbidden. We will use this set to construct one with the properties claimed in the statement, but first we show that $\widetilde{M}$ is a prototype of our goal set. Namely, we will first show that the set $\widetilde{M} \subseteq \Sigma$ satisfies

(i) $\alpha_{\min} \leq \liminf_{n \to \infty} A_0^n(\mathbf{i}) \leq \limsup_{n \to \infty} A_0^n(\mathbf{i}) \leq \alpha_{\max}$ for every $\mathbf{i} \in \widetilde{M}$,
(ii) $h_{\text{top}}(\widetilde{M}) \geq \liminf_{n \to \infty} h_{\text{top}}(M_n)$.

Consider $\mathbf{i} \in \widetilde{M}$ and $n \in \mathbb{N}$. Take $k \in \mathbb{N}$ with $N_k \leq n < N_{k+1}$. Let m be the largest number such that $n - t_m^{k+1} > 0$. Remark that

$$B_0^n(\mathbf{i}) = B_0^{N_1}(\mathbf{i}) + \sum_{j=2}^{k} \sum_{\ell=0}^{m(j)-1} B_{t_\ell^j}^{t_{\ell+1}^j}(\sigma^{t_\ell^j}\mathbf{i}) + \sum_{\ell=0}^{m-1} B_{t_\ell^{k+1}}^{t_{\ell+1}^{k+1}}(\sigma^{t_\ell^{k+1}}\mathbf{i}) + B_{t_m^{k+1}}^{n}(\sigma^{t_m^{k+1}}\mathbf{i}).$$

Observe that for every t_ℓ^j there exists a $\mathbf{j} \in M_j$ such that for every $t_\ell^j \leq i < t_{\ell+1}^j$, $|\phi_i(\sigma^i\mathbf{i}) - \phi_i(\sigma^i\mathbf{j})| \leq \text{var}_{t_{\ell+1}^j - i}(\phi_i)$, and thus

$$B_{t_\ell^j}^{t_{\ell+1}^j}(\sigma^{t_\ell^j}\mathbf{i}) = \sum_{i=t_\ell^j}^{t_{\ell+1}^j-1} \phi_i(\sigma^i\mathbf{i}) \leq \sum_{i=t_\ell^j}^{t_{\ell+1}^j-1} \phi_i(\sigma^i\mathbf{j}) + \sum_{i=t_\ell^j}^{t_{\ell+1}^j-1} \text{var}_{t_{\ell+1}^j-i}(\phi_i).$$

Hence, by Lemma 3.1

$$B_0^n(\mathbf{i}) \leq \alpha_1 N_1 + \sum_{j=2}^{k} \alpha_j(N_j - N_{j-1}) + (n - N_k)\alpha_{k+1} + \varepsilon_1 N_1$$
$$+ \sum_{j=2}^{k} \sum_{\ell=0}^{m(j)-1} \varepsilon_j(t_{\ell+1}^j + t_\ell^j) + \sum_{\ell=0}^{m-1} \varepsilon_{k+1}(t_{\ell+1}^{k+1} + t_\ell^{k+1}) + (n + t_m^{k+1})\varepsilon_{k+1}$$
$$+ \sum_{j=2}^{k} \sum_{\ell=0}^{m(j)-1} \sum_{i=t_\ell^j}^{t_{\ell+1}^j-1} \text{var}_{t_{\ell+1}^j-i}(\phi_i) + \sum_{\ell=0}^{m-1} \sum_{i=t_\ell^{k+1}}^{t_{\ell+1}^{k+1}-1} \text{var}_{t_{\ell+1}^{k+1}-i}(\phi_i) + \sum_{i=t_m^{k+1}}^{n-1} \text{var}_{t_{m+1}^{k+1}-i}(\phi_i).$$

Observe that

$$\sum_{\ell=0}^{m(j)-1} \varepsilon_j(t_{\ell+1}^j + t_\ell^j) \leq \sum_{\ell=0}^{m(j)-1} \varepsilon_j r_j^\ell(r_j + 1)N_{j-1} \leq \frac{3\varepsilon_j N_{j-1}(r_j^{m(j)} - 1)}{r_j - 1} \leq \frac{3\varepsilon_j N_j}{r_j - 1}.$$

Hence,

$$\frac{1}{n}\left(\sum_{j=2}^{k}\sum_{\ell=0}^{m(j)-1}\varepsilon_j(t_{\ell+1}^j+t_\ell^j)+\sum_{\ell=0}^{m-1}\varepsilon_{k+1}(t_{\ell+1}^{k+1}+t_\ell^{k+1})+(n+t_m^{k+1})\varepsilon_{k+1}\right)$$

$$\leq\frac{1}{n}\left(\sum_{j=2}^{k}\frac{3\varepsilon_j}{r_j-1}N_j+\frac{3\varepsilon_{k+1}N_k(r_{k+1}^m-1)}{r_{k+1}-1}+(n+t_m^{k+1})\varepsilon_{k+1}\right)$$

$$\leq O(\frac{\sum_{j=2}^{k-1}N_j}{N_k})+\frac{3\varepsilon_k}{r_k-1}+\frac{3\varepsilon_{k+1}(t_m^{k+1}+1)}{n(r_{k+1}-1)}+\frac{(n+t_m^{k+1})\varepsilon_{k+1}}{n}$$

$$\leq O(\frac{\sum_{j=2}^{k-1}N_j}{N_k})+\frac{3\varepsilon_k}{r_k-1}+\frac{6\varepsilon_{k+1}}{r_{k+1}-1}+2\varepsilon_{k+1}=o(1).$$

On the other hand, since $\rho_i^{(1)}\to 0$ as $i\to\infty$, where $\rho_i^{(1)}$ is defined in (2.3), we get $\frac{1}{i}\sum_{j=1}^{i}\rho_j^{(1)}\to 0$ as $i\to\infty$ and hence,

$$\sum_{j=2}^{k}\sum_{\ell=0}^{m(j)-1}\sum_{i=t_\ell^j}^{t_{\ell+1}^j-1}\mathrm{var}_{t_{\ell+1}^j-i}(\phi_i)+\sum_{\ell=0}^{m-1}\sum_{i=t_\ell^{k+1}}^{t_{\ell+1}^{k+1}-1}\mathrm{var}_{t_{\ell+1}^{k+1}-i}(\phi_i)+\sum_{i=t_m^{k+1}}^{n-1}\mathrm{var}_{t_{m+1}^{k+1}-i}(\phi_i)$$

$$\leq\sum_{j=2}^{k}\sum_{\ell=0}^{m(j)-1}\sum_{i=t_\ell^j}^{t_{\ell+1}^j-1}\rho_{t_{\ell+1}^j-i}^{(1)}+\sum_{\ell=0}^{m-1}\sum_{i=t_\ell^{k+1}}^{t_{\ell+1}^{k+1}-1}\rho_{t_{\ell+1}^{k+1}-i}^{(1)}+\sum_{i=t_m^{k+1}}^{n-1}\rho_{t_{m+1}^{k+1}-i}^{(1)}$$

$$\leq\sum_{j=2}^{k}\sum_{\ell=0}^{m(j)-1}(t_{\ell+1}^j-t_\ell^j)o(1)+\sum_{\ell=0}^{m-1}(t_{\ell+1}^{k+1}-t_\ell^{k+1})o(1)+(n-t_m^{k+1}-1)o(1)$$

$$=n\cdot o(1).$$

The lower bound is similar, thus we get for every $N_k\leq n<N_{k+1}$

$$A_0^n(\mathbf{i})=\frac{\alpha_k N_k+(n-N_k)\alpha_{k+1}}{n}+o(1).$$

This together with $N_{k+1}/N_k\to\infty$ implies (i). Moreover, if $\alpha_k\to\alpha$, this shows that the convergence $A_0^n\to\alpha$ is uniform on $\widetilde{M}$. So it only remains to show (ii).

We pick arbitrarily $t_\ell^k\leq n<t_{\ell+1}^k$ for some $k\in\mathbb{N}$ and $0\leq\ell\leq m(k)-1$. By definition of $\widetilde{M}$ and (3.2), we have

$$Z_{N_{k-1}}^{t_\ell^k}(\widetilde{M})\geq Z_{N_{k-1}}^{t_\ell^k}(M_k)\geq\exp((t_\ell^k-N_{k-1})(h_{\mathrm{top}}(M_k)-\delta))$$

The last inequality is due to the fact that $N_{k-1}\in(z_i^k)_{i\in\mathbb{N}}$. Similarly, we see that

$$Z_{N_{k-1}}^{N_k}(\widetilde{M})\geq\exp((N_k-N_{k-1})(h_{\mathrm{top}}(M_k)-\delta)).$$

Since $\widetilde{M}[N_{i-1}, N_i - 1]\, \widetilde{M}[N_{j-1}, N_j - 1]$ are independent for $i \neq j$, we have

$$Z_0^{t_\ell^k}(\widetilde{M}) = Z_{N_{k-1}}^{t_\ell^k}(\widetilde{M}) \cdot \prod_{i=1}^{k-1} Z_{N_{i-1}}^{N_i}(\widetilde{M})$$

$$\geq \exp\left((t_\ell^k - N_{k-1})(h_{\text{top}}(M_k) - \delta) + \sum_{i=1}^{k-1}(h_{\text{top}}(M_i) - \delta)(N_i - N_{i-1}) \right)$$

$$\geq \exp\left(t_\ell^k (\liminf_{i\to\infty} h_{\text{top}}(M_i) - o(1) - \delta) \right). \quad (3.3)$$

We define a probability measure μ as follows. For any $\mathbf{i} \in \Sigma_n$, let $k \in \mathbb{N}$ and $0 \leq \ell \leq m(k) - 1$ be the unique integer such that $t_\ell^k < n \leq t_{\ell+1}^k$, and let

$$\mu([\mathbf{i}]) = \frac{\#\{A \in M[0, t_{\ell+1}^k - 1] : [\mathbf{i}] \supset [A]\}}{Z_0^{t_{\ell+1}^k}(\widetilde{M})}$$

It is easy to see that μ is a well defined measure supported on $\widetilde{M}$. Indeed, if $|\mathbf{i}| < t_{\ell+1}^k$ then

$$\sum_{j \in \mathcal{A}} \mu[\mathbf{i}j] = \sum_{j \in \mathcal{A}} \frac{\#\{A \in \widetilde{M}[0, t_{\ell+1}^k - 1] : [\mathbf{i}j] \supset [A]\}}{Z_0^{t_{\ell+1}^k}(\widetilde{M})}$$

$$= \frac{\#\{A \in \widetilde{M}[0, t_{\ell+1}^k - 1] : \text{ there exists } j \in \mathcal{A} \text{ such that } [\mathbf{i}j] \supset [A]\}}{Z_0^{t_{\ell+1}^k}(\widetilde{M})}$$

$$= \frac{\#\{A \in \widetilde{M}[0, t_{\ell+1}^k - 1] : [\mathbf{i}] \supset [A]\}}{Z_0^{t_{\ell+1}^k}(\widetilde{M})},$$

and if $|\mathbf{i}| = t_{\ell+1}^k$ then

$$\sum_{j \in \mathcal{A}} \mu([\mathbf{i}j]) = \sum_{j \in \mathcal{A}} \frac{\#\{A \in \widetilde{M}[0, t_{\ell+2}^k - 1] : [\mathbf{i}j] \supset [A]\}}{Z_0^{t_{\ell+2}^k}(\widetilde{M})}$$

$$= \frac{\#\{A \in \widetilde{M}[0, t_{\ell+2}^k - 1] : \text{ there exists } j \in \mathcal{A} \text{ such that } [\mathbf{i}j] \supset [A]\}}{Z_0^{t_{\ell+1}^k}(\widetilde{M}) Z_{t_{\ell+1}^k}^{t_{\ell+2}^k}(\widetilde{M})}$$

$$= \frac{Z_{t_{\ell+1}^k}^{t_{\ell+2}^k}(\widetilde{M}) \delta_{\mathbf{i} \in \widetilde{M}[0, t_{\ell+1}^k - 1]}}{Z_0^{t_{\ell+1}^k}(\widetilde{M}) Z_{t_{\ell+1}^k}^{t_{\ell+2}^k}(\widetilde{M})},$$

where with a slight abuse of notation we used the $t_{m(k)+1}^k := t_1^{k+1}$.

By (3.3), we have that for every $\mathbf{i} \in \widetilde{M}$

$$\liminf_{n\to\infty} \frac{-\log \mu([\mathbf{i}|_0^n])}{n} \geq \liminf_{n\to\infty} \frac{t_\ell^k}{n}\left(\liminf_{i\to\infty} h_{\text{top}}(M_i) - o(1) - \delta \right)$$

$$\geq \liminf_{k\to\infty} r_k \left(\liminf_{i\to\infty} h_{\text{top}}(M_i) - o(1) - \delta \right)$$

$$= \liminf_{i\to\infty} h_{\text{top}}(M_i) - \delta.$$

By Lemma 2.1, we get (ii).

We are now almost done. We have constructed the set $\widetilde{M}$ which has almost all the demanded properties, the only one that is still missing is that $\widetilde{M}$ is not necessarily a subset of $\Sigma_{\mathbf{A}}$. The last step is to find a map π satisfying the assumptions of Lemma 3.2 and such that $\pi(\widetilde{M}) \subset \Sigma_{\mathbf{A}}$. Observe that the assertion of Lemma 3.2 will guarantee that the set $M = \pi(\widetilde{M})$ will satisfy the assertion of Proposition 3.3.

It is easy enough to do. Let us put the points $(t_i^k)_{i,k}$ in the increasing order and denote this sequence by (q_j) (ignoring the initial finitely many terms we can freely assume that $q_{j+1} - q_j > 2r$). Observe that each sequence $\mathbf{i}|_{q_j+1}^{q_{j+1}}$ is an admissible word in $\Sigma_{\mathbf{A},*}$. We can thus modify $\mathbf{i}$ only on positions $q_j + 1, \ldots, q_j + r; j = 1, 2, \ldots$ so that we obtain a sequence in $\Sigma_{\mathbf{A}}$. Each modification on positions $q_j + 1, \ldots, q_j + r$ can be chosen depending only on $\mathbf{i}|_{q_{j-1}+1}^{q_j}$ and $\mathbf{i}|_{q_j+r+1}^{q_{j+1}}$, that is, there exists $\mathbf{j} \in \Sigma_{\mathbf{A},r}$ such that $\mathbf{i}|_{q_{j-1}+1}^{q_j} \mathbf{j} \mathbf{i}|_{q_j+r+1}^{q_{j+1}} \in \Sigma_{\mathbf{A},*}$. Thus, choosing those modifications in a consistent way we can construct a map $\pi : \Sigma \to \Sigma$ satisfying the assumptions of Lemma 3.2 and such that $\pi(\widetilde{M}) \subset \Sigma_{\mathbf{A}}$.

Finally, to obtain the second part of the assertion let us consider the case when $\lim_{n\to\infty} \alpha_n = \alpha$. By taking a subsequence n_k such that $\limsup_{n\to\infty} h_{\text{top}}(M_n) = \lim_{k\to\infty} h_{\text{top}}(M_{n_k})$, and applying the previous argument for the sequences $\{\alpha_{n_k}\}_k$ and $\{\varepsilon_{n_k}\}_k$ and $\{M_{n_k}\}_k$ we get the claimed statement. $\qquad\square$

Corollary 3.5. *If $L_{\mathbf{A}}(\alpha) \neq \emptyset$ then for every $\delta > 0$ there exists $\emptyset \neq M \subset L_{\mathbf{A}}(\alpha)$ such that $h_{\text{top}}(M) > h_{\text{top}}(L_{\mathbf{A}}(\alpha)) - \delta$ and the convergence of $A_0^n(\mathbf{i}) \to \alpha$ on M is uniform.*

Proof. Take a sequence $\varepsilon_n \to 0$ be arbitrary but fixed. Since $A_0^n(\mathbf{i}) \to \alpha$ as $n \to \infty$ for every $\mathbf{i} \in L(\alpha)$, there exists $N_n(\mathbf{i})$ such that for every $m \geq N_n(\mathbf{i})$, $|A_0^m(\mathbf{i}) - \alpha| < \varepsilon_n$. For every $n \geq 1$ and $T \geq 1$, let

$$M_{n,T} = \{\mathbf{i} \in L_{\mathbf{A}}(\alpha) : N_n(\mathbf{i}) \leq T\}.$$

Since $L_{\mathbf{A}}(\alpha) = \bigcup_{T=1}^{\infty} M_{n,T}$ we get that there exists a T_n such that $h_{\text{top}}(M_{n,T_n}) > h_{\text{top}}(L(\alpha)) - \delta$.

By applying Proposition 3.3 for the sequence $\alpha_n \equiv \alpha$, ε_n and $M_n := M_{n,T_n}$, we get that there exists a set $M \subset L_{\mathbf{A}}(\alpha)$ such that $h_{\text{top}}(M) \geq \limsup_{n\to\infty} h_{\text{top}}(M_{n,T_n}) \geq h_{\text{top}}(L_{\mathbf{A}}(\alpha)) - \delta$, and the convergence is uniform on M. $\qquad\square$

Corollary 3.6. *The map $\alpha \mapsto h_{\text{top}}(L_{\mathbf{A}}(\alpha))$ is upper semi-continuous.*

Proof. Let $\alpha_n \to \alpha$ be such that $L_{\mathbf{A}}(\alpha_n) \neq \emptyset$. Then we can use Corollary 3.5 to find in each $L_{\mathbf{A}}(\alpha_n)$ a large entropy subset M_n with uniform convergence of the Birkhoff averages, then we apply Proposition 3.3 to get the assertion. $\qquad\square$

Lemma 3.7. *The domain of $\alpha \mapsto h_{\text{top}}(L_{\mathbf{A}}(\alpha))$ is compact.*

Proof. Suppose $A(\mathbf{i}_k) = \alpha_k$ for $\mathbf{i}_k \in \Sigma_{\mathbf{A}}, k \in \mathbb{N}$ satisfying that $\alpha_k \to \alpha$ as $k \to \infty$. We will show $\alpha \in \mathcal{D}(L_{\mathbf{A}})$. Fix $\epsilon > 0$. Then there exists $(N_k)_{k\in\mathbb{N}}$ of positive integers such that for any $n \geq N_k$, $|A_0^n(\mathbf{i}_k) - \alpha_k| < \epsilon_k$ with $\epsilon_k \to 0$. We pick two sequences $(m_k)_{k\in\mathbb{N}}$ and $(n_k)_{k\in\mathbb{N}}$ satisfying the following conditions:

- $m_0 = 0$, $m_k < n_k = m_{k+1} - r$ for $k \in \mathbb{N}$.
- $n_k \geq N_{k+1}$ for $k \in \mathbb{N}$.
- $\frac{n_{k-1}}{n_k} \to 0$ as $k \to \infty$.
- The set $\{n \in \mathbb{N} : n \notin [m_k, n_k], \forall k\}$ has density 0.

Then we take an $\mathbf{i} \in \Sigma_{\mathbf{A}}$ such that $\mathbf{i}|_{m_k}^{n_k} = \mathbf{i}_k|_{m_k}^{n_k}$ for $k \in \mathbb{N}$ (such $\mathbf{i}$ exists but may not be unique due to that $\Sigma_{\mathbf{A}}$ is irreducible). It follows from Lemma 3.1 that for $m_k \leq n \leq n_k$, we have

$$|A_0^n(\mathbf{i}) - \alpha| \leq \frac{n_{k-1}}{n}|A_0^{n_{k-1}}(\mathbf{i}) - \alpha| + \frac{n - m_k}{n}|\alpha_k - \alpha| + \frac{r}{n} + 2\epsilon_k.$$

Since $\frac{n_{k-1}}{n_k} \to 0$, $\alpha_k \to \alpha$ and r is a constant, we conclude that $A_0^n(\mathbf{i}) \to \alpha$ as $n \to \infty$. $\qquad\square$

The following proposition is in a sense similar to Proposition 3.3. Like there, we have some given collections of sequences with prescribed weighted Birkhoff averages and use them to construct the large set of their 'Frankenstein' offspring. However, the technical process of constructing the concatenated sequences is noticeably different.

Proposition 3.8. *The domain of $f \colon \alpha \mapsto h_{\text{top}}(L_{\mathbf{A}}(\alpha))$ is a (possibly empty) closed convex set and f is a concave function.*

Proof. Let α, α' be in the domain of f. Assuming that $L_{\mathbf{A}}(\alpha)$ and $L_{\mathbf{A}}(\alpha')$ are non-empty, we want to prove that $L_{\mathbf{A}}(p\alpha + (1-p)\alpha')$ is non-empty and that $f(p\alpha + (1-p)\alpha') \geq pf(\alpha) + (1-p)f(\alpha')$ for all $p \in (0,1)$. Pick arbitrarily $\epsilon > 0$. By Corollary 3.5, for every $\varepsilon > 0$ there exist subsets $M(\alpha) \subset L_{\mathbf{A}}(\alpha)$ and $M(\alpha') \subset L_{\mathbf{A}}(\alpha')$ such that

- $h_{\text{top}}(M(\alpha)) > f(\alpha) - \epsilon$ and $h_{\text{top}}(M(\alpha')) > f(\alpha') - \epsilon$;
- there exists an increasing sequence $(N_k)_{k\in\mathbb{N}}$ such that for every $\mathbf{i} \in M(\alpha)$ and every $\mathbf{i}' \in M(\alpha')$, for every k for every $n > N_k$ we have $|A_0^n(\mathbf{i}) - \alpha| \leq 1/k$ and $|A_0^n(\mathbf{i}') - \alpha'| \leq 1/k$.

We choose two sequences $(t_i)_{i\in\mathbb{N}}, (s_i)_{i\in\mathbb{N}}$ satisfying the following conditions.

(i) $t_0 = 0$, $t_i \nearrow \infty$ and $t_{i+1}/t_i \searrow 1$.

(ii) $s_i \to \infty$.

(iii) $(t_{i+1} - t_i)$ is divisible by s_i and $\frac{t_{i+1} - t_i}{s_i} \nearrow \infty$.

(iv) $\frac{2s_i t_{i+1}}{n(t_{i+1} - t_i)} \to 0$ where n is the largest number such that $N_n < t_i$.

For example, we can choose $t_{i+1}/t_i \sim 1 + n^{-1/2}$ and $s_i \sim n^{1/3}$ where n is the largest number such that $N_n < t_i$. We divide each interval $[t_i, t_{i+1} - 1]$ into s_i equal subintervals, with endpoints $z_0^i = t_i, z_1^i = t_i + (t_{i+1} - t_i)/s_i, \ldots, z_{s_i}^i = t_{i+1}$. We will construct a set $\widetilde{M} \subset \Sigma$ step by step as follows.

Step 0. At positions $0, \ldots, t_1 - 1$ we can put anything.

Step i ($i \geq 1$). We put the s_i numbers

$$W_k^i := \log Z_{z_k^i}^{z_{k+1}^i}(\alpha) - \log Z_{z_k^i}^{z_{k+1}^i}(\alpha'); k = 0, 1, \ldots s_i - 1$$

in an increasing order and we choose $\lfloor ps_i \rfloor$ largest ones. At those chosen intervals the sequences in $\widetilde{M}$ will be taken from $M_{z_k^i}^{z_{k+1}^i}(\alpha)$, at the not chosen intervals from $M_{z_k^i}^{z_{k+1}^i}(\alpha')$.

It is enough to show that $\widetilde{M} \subset \Sigma$ has the following properties:

Claim 1: for $\mathbf{i} \in \widetilde{M}$ we have

$$A(\mathbf{i}) = p\alpha + (1-p)\alpha'.$$

Claim 2: $h_{\text{top}}(\widetilde{M}) \geq pf(\alpha) + (1-p)f(\alpha')$.

Indeed, just like in the proof of Proposition 3.3, we will prove that there exists a map $\pi\colon \Sigma \mapsto \Sigma$ such that $\pi(\widetilde{M}) \subseteq \Sigma_{\mathbf{A}}$ and the assumptions of Lemma 3.2 hold.

Proof of Claim 1. As $t_{i+1}/t_i \to 1$, it is enough to check that $A_0^{t_i}(\mathbf{i}) \to p\alpha + (1-p)\alpha'$ as i tends to ∞. Pick i and n such that $N_n < t_i \leq N_{n+1}$. By Lemma 3.1, we have

$$|A_{t_i}^{t_{i+1}}(\mathbf{i}) - (p\alpha + (1-p)\alpha')|$$

$$= |\sum_{k=0}^{s_i-1} \frac{1}{s_i} A_{z_k^i}^{z_{k+1}^i}(\mathbf{i}) - (p\alpha + (1-p)\alpha')|$$

$$\leq I_1^i + I_2^i + I_3^i,$$

where

$$I_1^i = \sum_{k=0}^{s_i-1} \frac{1}{s_i} \rho_{\frac{t_{i+1}-t_i}{s_i}}^{(2)} = \rho_{\frac{t_{i+1}-t_i}{s_i}}^{(2)},$$

$$I_2^i = \sum_{k=0}^{s_i-1} \frac{1}{s_i} \cdot \frac{z_k^i + z_{k+1}^i}{n\left(\frac{t_{i+1}-t_i}{s_i}\right)} \leq \frac{2s_i t_{i+1}}{n(t_{i+1}-t_i)},$$

and

$$I_3^i = \left| \frac{1}{s_i} \left(\lfloor ps_i \rfloor \alpha + (s_i - \lfloor ps_i \rfloor)\alpha \right) - (p\alpha + (1-p)\alpha') \right|.$$

By (ii), it is easy to see that I_3^i converges 0 as i tends to ∞. By (iii) and the fact that $\rho_\ell^{(2)} \to 0$ as $\ell \to \infty$, we see that I_1^i converges 0 as i tends to ∞. By (iv), I_2^i converges 0 as i tends to ∞. Thus we obtain that $A_{t_i}^{t_{i+1}}(\mathbf{i})$ converges $p\alpha + (1-p)\alpha'$ as i tends to ∞. Since $A_0^{t_i}(\mathbf{i}) = \frac{1}{t_i}\sum_{j=0}^{i-1}(t_{j+1}-t_j)A_{t_j}^{t_{j+1}}(\mathbf{i})$, we complete the proof. $\square$

Proof of Claim 2. Observe that the constructions for different j are completely independent from each other: whatever the initial t_j symbols of $\mathbf{i} \in \widetilde{M}$, we allow any admissible $t_{j+1} - t_j$ symbols to follow. Thus we have

$$Z_0^{t_i}(\widetilde{M}) = Z_0^{t_1}(\widetilde{M}) \cdot \prod_{k=1}^{i-1} Z_{t_k}^{t_{k+1}}(\widetilde{M}) \tag{3.4}$$

and

$$Z_{t_k}^{t_{k+1}}(\widetilde{M}) \geq \left(\prod_{\ell=0}^{s_k-1} Z_{z_\ell^k}^{z_{\ell+1}^k}(M(\alpha)) \right)^{\lfloor ps_k \rfloor/s_k} \cdot \left(\prod_{\ell=0}^{s_k-1} Z_{z_\ell^k}^{z_{\ell+1}^k}(M(\alpha')) \right)^{1-\lfloor ps_k \rfloor/s_k}, \tag{3.5}$$

for $1 \leq k \leq i-1$. Moreover, we have

$$\prod_{\ell=0}^{s_k-1} Z_{z_\ell^k}^{z_{\ell+1}^k}(M(\alpha)) \geq Z_{t_k}^{t_{k+1}}(M(\alpha)) \tag{3.6}$$

and

$$\prod_{k=1}^{i-1} Z_{t_k}^{t_{k+1}}(M(\alpha))| \geq Z_{t_1}^{t_i}(M(\alpha)). \tag{3.7}$$

The same holds for α'. We define the probability measure μ such that for an $\mathbf{i} \in \Sigma_n$ let $t_{i-1} < n \leq t_i$ and

$$\mu([\mathbf{i}]) = \frac{\#\{A \in \widetilde{M}[0, t_i - 1] : [\mathbf{i}] \supset A\}}{Z_0^{t_i}(\widetilde{M})}.$$

Similarly to the proof of Proposition 3.3, μ is a well defined probability measure supported on $\widetilde{M}$. By (3.4), (3.5), (3.6) and (3.7), as $t_{i+1}/t_i \to 1$, we have that

$$\liminf_{n \to \infty} \frac{-\log \mu(C_n \cap \widetilde{M})}{n}$$
$$\geq \liminf_{i \to \infty} \frac{1}{t_i} \left(p \log Z_0^{t_i}(M(\alpha)) + (1-p) \log Z_0^{t_i}(M(\alpha')) \right)$$
$$\geq pf(\alpha) + (1-p)f(\alpha'),$$

for any decreasing sequence $(C_n)_{n \in \mathbb{N}}$ of cylinders with $C_n \cap \widetilde{M} \neq \emptyset$. By Lemma 2.1, this completes the proof. $\square$

As in the proof of Proposition 3.3, we have now obtained a set $\widetilde{M}$ satisfying all the necessary properties except for one: it does not have to be contained in $\Sigma_{\mathbf{A}}$. Again, we have the same solution to this problem: we will find a map π satisfying the assumptions of Lemma 3.2 such that $\pi(\widetilde{M}) \subset \Sigma_{\mathbf{A}}$. It is done in almost the same manner: we define $(q_j)_j = (z_k^i)_{k,i}$ and then we modify each sequence $\mathbf{i} \in \widetilde{M}$ on the initial r positions of every interval $(q_j, q_{j+1}]$.

Therefore, we complete the proof. $\square$

Proof of Theorem 2.2. Since any concave function is clearly lower semi-continuous, Lemma 3.7 and Proposition 3.8 together with Corollary 3.6 implies the claim. $\square$

Proof of Theorem 2.3. There are two cases. Consider first the simple case: $h_{\text{top}}(L_{\mathbf{A}}(\alpha_0)) = h_{\text{top}}(\Sigma_{\mathbf{A}})$.

Fix some $\varepsilon > 0$. We assume that the spectrum domain is larger than one point, hence by Theorem 2.2 we can find a value α_1 such that $h_{\text{top}}(L_{\mathbf{A}}(\alpha_1)) > h_{\text{top}}(L_{\mathbf{A}}(\alpha_0)) - \varepsilon$. By Corollary 3.5 we can find a set M_0 such that $h_{\text{top}}(M_0) > h_{\text{top}}(L_{\mathbf{A}}(\alpha_0)) - \varepsilon$ and that the convergence $A_0^n(\mathbf{i}) \to \alpha_0$ is uniform in M_0 and we can find a set M_1 such that $h_{\text{top}}(M_1) > h_{\text{top}}(L_{\mathbf{A}}(\alpha_1)) - \varepsilon$ and that the convergence $A_0^n(\mathbf{i}) \to \alpha_1$ is uniform in M_1. We then apply the Proposition 3.3 to the sequence of sets $M_1, M_0, M_1, M_0, \ldots$, with α_n being α_0 or α_1 depending on n being even or odd. We get

$$h_{\text{top}}(L_{\mathbf{A}}(\alpha_0, \alpha_1)) \geq h_{\text{top}}(L_{\mathbf{A}}(\alpha_0)) - 2\varepsilon.$$

Naturally, $L_{\mathbf{A}}(\alpha_0, \alpha_1) \subset D$, hence passing with ε to zero ends the proof.

The complicated case is when $h_{\text{top}}(L_{\mathbf{A}}(\alpha_0)) < h_{\text{top}}(\Sigma_{\mathbf{A}})$. Note that we can still freely assume that $h_{\text{top}}(\Sigma_{\mathbf{A}} \setminus D) = h_{\text{top}}(\Sigma_{\mathbf{A}})$, otherwise we would have $h_{\text{top}}(D) = h_{\text{top}}(\Sigma_{\mathbf{A}})$ immediately. We start with a simple observation.

Lemma 3.9. *There exists β_0 such that the sets $\{\mathbf{i} \in \Sigma_{\mathbf{A}}; A(\mathbf{i}) < \beta_0\}$ and $\{\mathbf{i} \in \Sigma_{\mathbf{A}}; A(\mathbf{i}) > \beta_0\}$ are both of full entropy $h_{\text{top}}(\Sigma_{\mathbf{A}})$.*

Proof. The function $\beta \to h_{\text{top}}(\bigcup_{\alpha < \beta} L_{\mathbf{A}}(\alpha))$ is non-decreasing and left continuous, hence the set $\{\beta : h_{\text{top}}(\bigcup_{\alpha < \beta} L_{\mathbf{A}}(\alpha)) = h_{\text{top}}(\Sigma_{\mathbf{A}})\}$ is closed. So is the set $\{\beta : h_{\text{top}}(\bigcup_{\alpha > \beta} L_{\mathbf{A}}(\alpha)) = h_{\text{top}}(\Sigma_{\mathbf{A}})\}$, for analogous reason. Hence, the two sets must intersect – otherwise we would have some β which would belong to neither, and this is impossible because

$$\Sigma_{\mathbf{A}} \setminus D = \bigcup_{\alpha < \beta} L_{\mathbf{A}}(\alpha) \cup \bigcup_{\alpha > \beta} L_{\mathbf{A}}(\alpha) \cup L_{\mathbf{A}}(\beta)$$

and all three sets on the right would have entropy strictly smaller than the one on the left. $\square$

We fix $\varepsilon > 0$. Using again the left-continuity of the function $\beta \to h_{\text{top}}(\bigcup_{\alpha<\beta} L_{\mathbf{A}}(\alpha))$ we can find some $\beta_1 < \beta_0$ such that $h_{\text{top}}(\bigcup_{\alpha<\beta_1} L_{\mathbf{A}}(\alpha)) > h_{\text{top}}(\Sigma_{\mathbf{A}}) - \varepsilon$. Let $M_+ = \bigcup_{\alpha>\beta_0} L_{\mathbf{A}}(\alpha)$ and $M_- = \bigcup_{\alpha<\beta_1} L_{\mathbf{A}}(\alpha)$.

We now need a one-sided version of Proposition 3.3.

Proposition 3.10. *Let $\varepsilon_n > 0$, $\varepsilon_n \to 0$. Let α_n be a sequence such that $\alpha_{2k} \to \beta_0$ and $\alpha_{2k+1} \to \beta_1$. Moreover, assume that for every $n \geq 1$ there exists a set $M_n \subset \Sigma_{\mathbf{A}}$ and a positive integer $T_n > 0$ such that for every $\mathbf{i} \in M_n$ and $m \geq T_n$ we have*

$$\frac{1}{m} \sum_{k=0}^{m-1} \phi_k(\sigma^k \mathbf{i}) > \alpha_n - \varepsilon_n$$

(if n is even) or

$$\frac{1}{m} \sum_{k=0}^{m-1} \phi_k(\sigma^k \mathbf{i}) < \alpha_n + \varepsilon_n$$

(if n is odd). Then we can find a set $M \subset \Sigma_{\mathbf{A}}$ such that for $\mathbf{i} \in M$ we have $\underline{A}(\mathbf{i}) \leq \beta_1$ and $\overline{A}(\mathbf{i}) \geq \beta_0$ and that $h_{\text{top}}M \geq \liminf h_{\text{top}}M_i$.

Proof. The proof is virtually identical with the proof of Proposition 3.3. The construction and the calculation of entropy is the same, the only difference is that when the sets M_i give only one-sided bounds on the behaviour of the Birkhoff sums, we can only get a weaker statement about $\underline{A}$ and $\overline{A}$. We skip the details. $\square$

We can now fix any sequence $\varepsilon_n \to 0$ and use the sets M_- and M_+ defined above to construct the sets M_n satisfying the assumptions of Proposition 3.10, in such a way that $h_{\text{top}}M_{2k} > h_{\text{top}}M_+ - \varepsilon$ and $h_{\text{top}}M_{2k+1} > h_{\text{top}}M_- - \varepsilon$ (by choosing T_{2k}, resp. T_{2k+1}, large enough). Using now Proposition 3.10 with those sets M_n we construct a set M which is by construction contained in D, moreover $h_{\text{top}}M > h_{\text{top}}(\Sigma_{\mathbf{A}}) - 2\varepsilon$. Passing with ε to 0 we end the proof of this case. $\square$

4. **Typical weights.** First, we need to introduce some notations. Let $\Sigma_{\mathbf{A}}$ be an aperiodic and irreducible subshift of finite type, $\Omega = \Lambda^{\mathbb{N}}$ and $\Gamma_{\mathbf{A}} = \Omega \times \Sigma_{\mathbf{A}}$. Let $f: \Gamma_{\mathbf{A}} \mapsto \mathbb{R}$ be a continuous potential. Let us recall that $S_n f$ denotes the nth Birkhoff sum of f, that is, $S_n f = f + f \circ \sigma + \cdots + f \circ \sigma^{n-1}$. For every $\mathbf{w} \in \Omega$ let

$$Z_n(f, \mathbf{w}) = \sum_{\mathbf{i} \in \Sigma_{\mathbf{A},n}} \sup_{\mathbf{j} \in [\mathbf{i}]} e^{S_n f(\mathbf{w},\mathbf{j})},$$

and define the conditional pressure of f on $\xi(\mathbf{w})$ by

$$P(f, \mathbf{w}) = \limsup_{n \to \infty} \frac{1}{n} \log Z_n(f, \mathbf{w}). \tag{4.1}$$

The pressure defined in (4.1) corresponds to the definition of the pressure given in Fan [11, page 3] in case of $f_k(\mathbf{i}) := f(\sigma^k \mathbf{w}, \mathbf{i})$ without the extra requirement that it exists as a limit. Later, we will show that for typical $\mathbf{w}$ with respect to an ergodic quasi-Bernoulli measures ν it equals to the pressure defined in (2.5).

The following theorem was shown by Ledrappier and Walters [23]. They proved a more general statement but we state here only the form which corresponds to our main setup.

Theorem 4.1 (Ledrappier, Walters). *Let ν be a σ-invariant measure on Ω and let $f\colon \Gamma_{\mathbf{A}} \mapsto \mathbb{R}$ be a continuous potential. Then*

$$\sup\{h_\mu^\xi + \int f d\mu : \mu \in \mathcal{M}_\nu(\Gamma_{\mathbf{A}})\} = \int P(f, \mathbf{w}) d\nu(\mathbf{w}).$$

Unfortunately, this theorem itself does not provide enough regularity conditions in order to do multifractal analysis on weighted Birkhoff averages. So we adapt the idea of Ledrappier and Walters [23] combining with the methods of Takens and Verbitskiy [29], Feng [17] and Heurteaux [21].

4.1. **Pinsker's formula.** Let us recall that Π is the natural projection $\Pi\colon \Omega \times \Sigma \mapsto \Omega$, that is, $\Pi(\mathbf{w}, \mathbf{i}) = \mathbf{w}$. Let μ be an ergodic σ-invariant measure on Γ. Clearly, if μ is σ-invariant and ergodic then $\Pi_*\mu$ is σ-invariant and ergodic on Ω too. By Shannon-McMillan-Breiman's Theorem,

$$\begin{aligned}
h_\mu &= \lim_{n\to\infty} \frac{-1}{n} \log \mu[(\mathbf{w}, \mathbf{i})|_n] \text{ for } \mu\text{-almost every } (\mathbf{w}, \mathbf{i}), \\
h_{\Pi_*\mu} &= \lim_{n\to\infty} \frac{-1}{n} \log \Pi_*\mu[\mathbf{w}|_n] \text{ for } \Pi_*\mu\text{-almost every } \mathbf{w}.
\end{aligned} \tag{4.2}$$

Denote ξ the partition generated by the inverse branches $\Pi^{-1}(\mathbf{w}) = \{\mathbf{w}\} \times \Sigma = \xi(\mathbf{w})$. By Rohlin's Disintegration Theorem, there exists a family of probability measures $\{\mu_{\mathbf{w}}^\xi\}$ such that

1. $\mu_{\mathbf{w}}^\xi$ is supported on $\xi(\mathbf{w})$;
2. for every $A \in \mathcal{B}_\Gamma$, the map $\mathbf{w} \mapsto \mu_{\mathbf{w}}^\xi(A)$ is $\mathcal{B}_\Omega$-measurable;
3. $\mu = \int \mu_{\mathbf{w}}^\xi d\Pi_*\mu(\mathbf{w})$.

The family $\{\mu_{\mathbf{w}}^\xi\}$ of measures is unique up to a zero $\Pi_*\mu$-measure set. Let us define the conditional entropy of $\mu_{\mathbf{w}}^\xi$ by

$$h_\mu^\xi := \int -\log \mu_{\Pi(\mathbf{w}, \mathbf{i})}^\xi([i_0]) d\mu(\mathbf{w}, \mathbf{i}).$$

The following theorem is the corresponding version of Pinsker's formula [26], which we need to establish relation between the conditional entropy, and the entropy of the projection.

Theorem 4.2 (Pinsker's formula). *If μ is an ergodic σ-invariant measure then for $\Pi_*\mu$-almost every $\mathbf{w}$, we have*

$$\lim_{n\to\infty} \frac{-1}{n} \log \mu_{\mathbf{w}}^\xi([\mathbf{i}|_n]) = h_\mu^\xi \text{ for } \mu_{\mathbf{w}}^\xi\text{-a.e. } \mathbf{i}. \tag{4.3}$$

Moreover,

$$h_\mu = h_{\Pi_*\mu} + h_\mu^\xi.$$

For completeness, we give a proof here. Observe that the map $(\mathbf{w}, \mathbf{i}) \mapsto -\log \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}([i_0])$ is in $L^1(\Gamma, \mu)$. Indeed,

$$
\begin{aligned}
h_\mu^\xi &= \int -\log \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}([i_0]) d\mu(\mathbf{w}, \mathbf{i}) \\
&= \int_0^\infty \mu(\{(\mathbf{w}, \mathbf{i}) : -\log \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}([i_0]) > x\}) dx \\
&= \int_0^\infty \int \mathbb{1}_{\{-\log \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}([i_0]) > x\}}(\mathbf{w}, \mathbf{i}) d\mu(\mathbf{w}, \mathbf{i}) dx \\
&= \sum_{k \in \mathcal{A}} \int_0^\infty \int \mathbb{1}_{\{-\log \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}([i_0]) > x\}}(\mathbf{w}, \mathbf{i}) \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}([k]) d\mu(\mathbf{w}, \mathbf{i}) dx \\
&\leq \sum_{k \in \mathcal{A}} \int_0^\infty \int e^{-x} d\mu(\mathbf{w}, \mathbf{i}) dx = K.
\end{aligned}
$$

Let us denote the partition with respect to the cylinders on Γ by $\mathfrak{P}$. Then clearly,

$$
\sigma_* \left(\mu_{\mathbf{w}}^{\xi \vee \mathfrak{P}}\right) = \mu_{\sigma\mathbf{w}}^{\xi}, \text{ for } \Pi_* \mu - a.e. \ \mathbf{w} \tag{4.4}
$$

Indeed, $\sigma_* \mu_{\mathbf{w}}^{\xi \vee \mathfrak{P}}$ is supported on $\Pi^{-1}(\sigma\mathbf{w})$, and by the definition of conditional measures,

$$
\begin{aligned}
\int \sigma_* \left(\mu_{\mathbf{w}}^{\xi \vee \mathfrak{P}}\right) d\Pi_* \mu(\mathbf{w}) = \sigma_* \int \left(\mu_{\mathbf{w}}^{\xi \vee \mathfrak{P}}\right) d\Pi_* \mu(\mathbf{w}) = \sigma_* \mu = \mu \\
= \int \mu_{\mathbf{w}}^{\xi} d\Pi_* \mu(\mathbf{w}) = \int \mu_{\sigma\mathbf{w}}^{\xi} d\Pi_* \mu(\mathbf{w}).
\end{aligned}
$$

Thus, (4.4) follows by the uniqueness of the conditional measures.

Proof of Theorem 4.2. Let us first show the first assertion of the theorem. By (4.4), we have

$$
\begin{aligned}
\mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}([\mathbf{i}|_n]) &= \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}([\mathbf{i}|_1]) \prod_{k=2}^n \frac{\mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}([\mathbf{i}|_k])}{\mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}([\mathbf{i}|_{k-1}])} \\
&= \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}([\mathbf{i}|_1]) \prod_{k=2}^n \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi \vee \mathcal{P}_{k-1}}(\sigma^{-(k-1)}[\sigma^{k-1}\mathbf{i}|_1]) \\
&= \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}([\mathbf{i}|_1]) \prod_{k=2}^n \mu_{\Pi \circ \sigma^{k-1}(\mathbf{w}, \mathbf{i})}^{\xi}([\sigma^{k-1}\mathbf{i}|_1]).
\end{aligned}
$$

Taking logarithm and applying Birkhoff's Ergodic Theorem, we get $\frac{-1}{n}\log \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}([\mathbf{i}|_n]) = \frac{1}{n} \sum_{k=0}^{n-1} -\log \mu_{\Pi \circ \sigma^k(\mathbf{w}, \mathbf{i})}^{\xi}([\sigma^k\mathbf{i}|_1]) \to h_\mu^\xi$ for μ-almost every $(\mathbf{w}, \mathbf{i})$. Thus, (4.3) follows by Fubini's Theorem.

Now, we show that $h_\mu = h_{\Pi_* \mu} + h_\mu^\xi$. By Egorov's Theorem, for every $\varepsilon > 0$ there exists $J_1 \subset \Gamma$ such that $\mu(J_1) > 1 - \varepsilon$ and the convergences (4.2) and (4.3) are uniform. That is, there exists $C > 0$ such that for every $n \geq 1$ and every $(\mathbf{w}, \mathbf{i}) \in J_1$

$$
C^{-1} e^{-h_{\Pi_* \mu} n} \leq \Pi_* \mu([\mathbf{w}|_n]) \leq C e^{-h_{\Pi_* \mu} n} \text{ and } C^{-1} e^{-nh_\mu^\xi} \leq \mu_{\mathbf{w}}^{\xi}([\mathbf{i}|_n]) \leq C e^{-nh_\mu^\xi}.
$$

By Lebesgue's density Theorem and Egorov's Theorem, there exists $J_2 \subset J_1$ such that $\mu(J_2) > 1 - 2\varepsilon$ and there exists $N \geq 1$ such that for every $(\mathbf{w}, \mathbf{i}) \in J_2$ and $n \geq N$

$$
\mu(J_1 \cap [(\mathbf{w}, \mathbf{i})|_n]) \geq \frac{1}{2}\mu([(\mathbf{w}, \mathbf{i})|_n]) \text{ and } \mu_{\mathbf{w}}^{\xi}(J_1 \cap [(\mathbf{w}, \mathbf{i})|_n]) \geq \frac{1}{2}\mu_{\mathbf{w}}^{\xi}([(\mathbf{w}, \mathbf{i})|_n]).
$$

Thus, for every $(\mathbf{w}, \mathbf{i}) \in J_2$ and every $n \geq N$

$$
\begin{aligned}
\mu([(\mathbf{w}, \mathbf{i})|_n]) &\leq 2\mu(J_1 \cap [(\mathbf{w}, \mathbf{i})|_n]) \\
&= 2 \int \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}(J_1 \cap [(\mathbf{w}, \mathbf{i})|_n]) d\mu(\mathbf{w}, \mathbf{i}) \\
&= 2 \int_{\Pi^{-1}[\mathbf{w}|_n]} \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}(J_1 \cap [(\mathbf{w}, \mathbf{i})|_n]) d\mu(\mathbf{w}, \mathbf{i}) \\
&\leq 2\Pi_* \mu([\mathbf{w}|_n]) C e^{-nh_\mu^\xi} \leq 2C^2 e^{-n(h_{\Pi_* \mu} + h_\mu^\xi)}.
\end{aligned}
$$

On the other hand, for every $(\mathbf{w}, \mathbf{i}) \in J_2$

$$
\begin{aligned}
\mu([(\mathbf{w}, \mathbf{i})|_n]) &\geq \mu(J_1 \cap [(\mathbf{w}, \mathbf{i})|_n]) \\
&= \int \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}(J_1 \cap [(\mathbf{w}, \mathbf{i})|_n]) d\mu(\mathbf{w}, \mathbf{i}) \\
&= \int_{\Pi^{-1}[\mathbf{w}|_n]} \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}(J_1 \cap [(\mathbf{w}, \mathbf{i})|_n]) d\mu(\mathbf{w}, \mathbf{i}) \\
&\geq \frac{1}{2} \int_{\Pi^{-1}[\mathbf{w}|_n]} \mu_{\Pi(\mathbf{w}, \mathbf{i})}^{\xi}([(\mathbf{w}, \mathbf{i})|_n]) d\mu(\mathbf{w}, \mathbf{i}) \\
&\geq \frac{1}{2} \Pi_* \mu([\mathbf{w}|_n]) C^{-1} e^{-nh_\mu^\xi} \geq \frac{1}{2} C^{-2} e^{-n(h_{\Pi_* \mu} + h_\mu^\xi)}.
\end{aligned}
$$

Thus, the statement follows by Shannon-McMillan-Breiman Theorem. $\qquad \square$

4.2. **Regularity of conditional pressure.** In this part of the section, we study the regularity properties of the conditional pressure $P(f, \mathbf{w})$ under stronger assumptions than the setup of Ledrappier and Walters. Namely, we assume that f has summable variation, that is,

$$
\sum_{k=0}^{\infty} \max_{(\mathbf{x}, \mathbf{k}) \in \Gamma_{\mathbf{A}, k}} \sup_{(\mathbf{w}, \mathbf{i}), (\mathbf{z}, \mathbf{j}) \in [(\mathbf{x}, \mathbf{k})]} |f(\mathbf{w}, \mathbf{i}) - f(\mathbf{z}, \mathbf{j})| < \infty.
$$

Moreover, we assume that the measure ν is quasi-Bernoulli. Note that for a quasi-Bernoulli measure ν, the transformation σ^m is ergodic for every $m \geq 1$.

The following lemma is an easy calculation.

Lemma 4.3. *For every* $\mathbf{w} \in \Omega$,

$$
P(f, \mathbf{w}) = P(f, \sigma\mathbf{w}).
$$

Moreover, if $f \to g$ *uniformly then* $P(f, \mathbf{w}) \to P(g, \mathbf{w})$.

Proof. Since $f \colon \Gamma_{\mathbf{A}} \mapsto \mathbb{R}$ is continuous over a compact set, we get that $|f|$ is bounded by C. Hence,

$$
\begin{aligned}
\sum_{\mathbf{i} \in \Sigma_{\mathbf{A}, n+1}} \sup_{\mathbf{j} \in [\mathbf{i}]} e^{S_{n+1} f(\mathbf{w}, \mathbf{j})} &= \sum_{\mathbf{i} \in \Sigma_{\mathbf{A}, n+1}} \sup_{\mathbf{j} \in [\mathbf{i}]} e^{S_n f(\sigma\mathbf{w}, \sigma\mathbf{j})} e^{f(\mathbf{w}, \mathbf{j})} \\
&\leq \sum_{\mathbf{i} \in \Sigma_{\mathbf{A}, n+1}} \sup_{\mathbf{j} \in [\mathbf{i}]} e^{S_n f(\sigma\mathbf{w}, \sigma\mathbf{j})} e^{C} \\
&\leq K e^{C} \sum_{\mathbf{i} \in \Sigma_{\mathbf{A}, n}} \sup_{\mathbf{j} \in [\mathbf{i}]} e^{S_n f(\sigma\mathbf{w}, \mathbf{j})}.
\end{aligned}
$$

The direction $\sum_{\mathbf{i} \in \Sigma_{\mathbf{A}, n+1}} \sup_{\mathbf{j} \in [\mathbf{i}]} e^{S_{n+1} f(\mathbf{w}, \mathbf{j})} \geq e^{-C} \sum_{\mathbf{i} \in \Sigma_{\mathbf{A}, n}} \sup_{\mathbf{j} \in [\mathbf{i}]} e^{S_n f(\sigma\mathbf{w}, \mathbf{j})}$ is similar.

The second observation follows by the fact that if $\sup_{(\mathbf{w},\mathbf{i})\in\Gamma_{\mathbf{A}}}|f(\mathbf{w},\mathbf{i})-g(\mathbf{w},\mathbf{i})|<\varepsilon$ then $|S_nf-S_ng|\leq\varepsilon n$. $\qquad\square$

Since ν is ergodic, a simple corollary of Lemma 4.3 is that we can define the conditional pressure with respect to ν

$$P_\nu(f) := \int P(f,\mathbf{w})d\nu(\mathbf{w}) = P(f,\mathbf{w}) \text{ for } \nu\text{-almost every } \mathbf{w}. \qquad (4.5)$$

Here, we abused a notation slightly, since $P_\nu(f)$ of (4.5) does not necessarily equal to the defined conditional pressure in (2.5). However, we will show in equation (4.12) that it is indeed equal to the pressure defined in (2.5).

For short, for $\mathbf{w}\in\Omega$ and $\mathbf{i}\in\Sigma_{\mathbf{A},*}$ let

$$V(f,\mathbf{w},\mathbf{i}) := \sup_{\mathbf{j}\in[\mathbf{i}]} e^{S_{|\mathbf{i}|}f(\mathbf{w},\mathbf{j})},$$

and for an $\mathbf{w}\in\Omega_*$ let

$$Y(f,\mathbf{w},\mathbf{i}) := \sup_{\mathbf{z}\in[\mathbf{w}]} V(f,\mathbf{z},\mathbf{i}) \text{ and } W(f,\mathbf{w}) := \sup_{\mathbf{z}\in[\mathbf{w}]} Z_{|\mathbf{w}|}(f,\mathbf{z}).$$

We also use the convention that $Z_m(f,\mathbf{w})=1$ for $m\leq 0$.

Since f has summable variation, there exists constant $C>0$ such that for every $n\geq 1$ and every $(\mathbf{w},\mathbf{i}),(\mathbf{z},\mathbf{j})\in\Gamma_{\mathbf{A}}$ with $|(\mathbf{w},\mathbf{i})\wedge(\mathbf{z},\mathbf{j})|=n$

$$|S_nf(\mathbf{w},\mathbf{i}) - S_nf(\mathbf{z},\mathbf{j})| < C. \qquad (4.6)$$

Thus, for every $\mathbf{w}\in\Omega$ and every, $\mathbf{i},\mathbf{j}\in\Sigma_{\mathbf{A},*}$ with $\mathbf{ij}\in\Sigma_{\mathbf{A},*}$

$$V(f,\mathbf{w},\mathbf{ij}) \leq V(f,\mathbf{w},\mathbf{i})V(f,\sigma^{|\mathbf{i}|}\mathbf{w},\mathbf{j}) \leq e^C\cdot V(f,\mathbf{w},\mathbf{ij}). \qquad (4.7)$$

So clearly, for every $\mathbf{w}\in\Omega$

$$Z_{n+m}(f,\mathbf{w}) \leq Z_n(f,\mathbf{w})Z_m(f,\sigma^n\mathbf{w}). \qquad (4.8)$$

On the other hand,

$$\begin{aligned} Z_n(f,\mathbf{w})Z_m(f,\sigma^n\mathbf{w}) &\leq K^r e^{r|f|}Z_n(f,\mathbf{w})Z_{m-r}(f,\sigma^{n+r}\mathbf{w}) \\ &\leq K^r e^{2r|f|+2C}Z_{n+m}(f,\mathbf{w}), \end{aligned} \qquad (4.9)$$

where $r\geq 1$ is such that $\mathbf{A}^r$ is strictly positive.

Applying the bounded distortion again, we get for every $(\mathbf{w},\mathbf{i})\in\Gamma_{\mathbf{A},*}$, and every $\mathbf{z}\in[\mathbf{w}]$ that

$$V(f,\mathbf{z},\mathbf{i}) \leq Y(f,\mathbf{w},\mathbf{i}) \leq e^C V(f,\mathbf{z},\mathbf{i}) \qquad (4.10)$$

and therefore

$$Z_{|\mathbf{w}|}(f,\mathbf{z}) \leq W(f,\mathbf{w}) \leq e^C Z_{|\mathbf{w}|}(f,\mathbf{z}). \qquad (4.11)$$

By (4.8) and Kingman's subadditive ergodic theorem, we have that for ν-almost every $\mathbf{w}\in\Omega$ the limit

$$\lim_{n\to\infty} \frac{1}{n}\log Z_n(f,\mathbf{w}) = P(f,\mathbf{w}) = P_\nu(f)$$

exists and

$$\begin{aligned} P_\nu(f) &= \lim_{n\to\infty} \frac{1}{n}\int \log Z_n(f,\mathbf{w})d\nu(\mathbf{w}) \\ &= \lim_{n\to\infty} \frac{1}{n}\sum_{\mathbf{w}\in\Omega_n} \nu([\mathbf{w}])\log W(f,\mathbf{w}), \end{aligned} \qquad (4.12)$$

where in the last equation we used (4.11) too.

The next theorem is adapting the result and method of Feng [17, Section 4] for the situation of subshift of finite type and to the condition on the marginal measures.

Theorem 4.4. *Let ν be an ergodic σ-invariant quasi-Bernoulli measure on Ω and let $f: \Gamma_{\mathbf{A}} \mapsto \mathbb{R}$ be a continuous potential with summable variation. Then there exists a unique ergodic σ-invariant measure μ such that there exists a constant $C > 0$ such that for every $(\mathbf{w}, \mathbf{i}) \in \Gamma_{\mathbf{A}, *}$*

$$C^{-1} \frac{Y(f, \mathbf{w}, \mathbf{i})}{W_{|\mathbf{w}|}(f, \mathbf{w})} \nu([\mathbf{w}]) \leq \mu([\mathbf{w}, \mathbf{i}]) \leq C \frac{Y(f, \mathbf{w}, \mathbf{i})}{W_{|\mathbf{w}|}(f, \mathbf{w})} \nu([\mathbf{w}]). \tag{4.13}$$

In particular, $\Pi_* \mu = \nu$ and

$$h_\mu^\xi + \int f d\mu = P_\nu(f).$$

Proof. Let $\mathbf{z}$ be a generic point such that $\frac{1}{n} \sum_{k=0}^{n-1} \delta_{\sigma^k \mathbf{z}} \to \nu$ as $n \to \infty$. Then let

$$\eta_m = Z_m(f, \mathbf{z})^{-1} \sum_{\mathbf{i} \in \Sigma_{\mathbf{A}, m}} V(f, \mathbf{z}, \mathbf{i}) \delta_{(\mathbf{z}, \mathbf{ij})},$$

where $\mathbf{j} \in \Sigma_{\mathbf{A}}$ is arbitrary but fixed. Moreover, let

$$\nu_n = \frac{1}{n} \sum_{k=0}^{n-1} \eta_{2n} \circ \sigma^{-k}.$$

Let $\{n_j\}$ be a subsequence such that $\lim_{j \to \infty} \frac{1}{n_j} \log Z_{n_j}(f, \mathbf{w}) = P_\nu^f$ and $\nu_{n_j} \to \mu$. Clearly, μ is a σ-invariant measure on $\Gamma_{\mathbf{A}}$.

Fix $(\mathbf{w}, \mathbf{i}) \in \Gamma_{\mathbf{A}, *}$ with $|\mathbf{w}| = |\mathbf{i}|$. Choose n sufficiently large such that $n > |\mathbf{w}| = |\mathbf{i}|$. Then by (4.7) and (4.9) there exists $C' > 0$ such that

$$\nu_n([\mathbf{w}, \mathbf{i}]) = \frac{1}{n} \sum_{k=0}^{n-1} \sum_{\substack{(\alpha, \beta) \in \Gamma_{\mathbf{A}, k}, (\gamma, \tau) \in \Gamma_{\mathbf{A}, 2n - |\mathbf{i}| - k}: \\ (\alpha \mathbf{w} \gamma, \beta \mathbf{i} \tau) \in \Gamma_{\mathbf{A}, 2n}}} \eta_{2n}([(\alpha \mathbf{w} \gamma, \beta \mathbf{i} \tau)])$$

$$= \frac{1}{n} \sum_{k=0}^{n-1} \sum_{\substack{\beta \in \Sigma_{\mathbf{A}, k}, \tau \in \Sigma_{\mathbf{A}, 2n - |\mathbf{i}| - k}: \\ \beta \mathbf{i} \tau \in \Sigma_{\mathbf{A}, 2n}}} \frac{V(f, \mathbf{z}, \beta \mathbf{i} \tau))}{Z_{2n}(\mathbf{z})} \mathbb{1}_{[\mathbf{w}]}(\sigma^k \mathbf{z})$$

$$\leq \frac{C'}{n} \sum_{k=0}^{n-1} \sum_{\substack{\beta \in \Sigma_{\mathbf{A}, k}, \\ \tau \in \Sigma_{\mathbf{A}, 2n - k - |\mathbf{i}|}}} \frac{V(f, \mathbf{z}, \beta) V(f, \sigma^k \mathbf{z}, \mathbf{i}) V(f, \sigma^{|\mathbf{w}| + k} \mathbf{z}, \tau)}{Z_k(\mathbf{z}) Z_{|\mathbf{i}|}(\sigma^k \mathbf{z}) Z_{2n - k - |\mathbf{i}|}(\sigma^{|\mathbf{i}| + k} \mathbf{z})} \mathbb{1}_{[\mathbf{w}]}(\sigma^k \mathbf{z})$$

$$= \frac{C'}{n} \sum_{k=0}^{n-1} \frac{V(f, \sigma^k \mathbf{z}, \mathbf{i})}{Z_{|\mathbf{i}|}(\sigma^k \mathbf{z})} \mathbb{1}_{[\mathbf{w}]}(\sigma^k \mathbf{z}).$$

Thus, by Birkhoff's ergodic theorem

$$\mu([\mathbf{w}, \mathbf{i}]) = \lim_{j \to \infty} \nu_{n_j}([\mathbf{w}, \mathbf{j}])$$

$$\leq C' \int \frac{V(f, \mathbf{z}, \mathbf{i})}{Z_{|\mathbf{i}|}(\mathbf{z})} \mathbb{1}_{[\mathbf{w}]}(\mathbf{z}) d\nu(\mathbf{z})$$

$$\leq C'' \frac{Y(f, \mathbf{w}, \mathbf{i})}{W_{|\mathbf{w}|}(f, \mathbf{w})} \nu([\mathbf{w}]),$$

where we used (4.10) and (4.11).

Now, we show the other inequality. Similarly by using (4.7), (4.8), we have

$$
\nu_n([\mathbf{w}, \mathbf{i}])
$$

$$
\geq \frac{1}{n} \sum_{k=0}^{n-1} \sum_{\substack{\beta \in \Sigma_{\mathbf{A},k}, \tau \in \Sigma_{\mathbf{A}, 2n-|\mathbf{i}|-k}: \\ \beta \mathbf{i} \tau \in \Sigma_{\mathbf{A}, 2n}}} \frac{V(f, \mathbf{z}, \beta) V(f, \sigma^k \mathbf{z}, \mathbf{i}) V(f, \sigma^{|\mathbf{w}|+k} \mathbf{z}, \tau)}{Z_k(\mathbf{z}) Z_{|\mathbf{i}|}(\sigma^k \mathbf{z}) Z_{2n-k-|\mathbf{i}|}(\sigma^{|\mathbf{i}|+k} \mathbf{z})} \mathbb{1}_{[\mathbf{w}]}(\sigma^k \mathbf{z})
$$

$$
\geq \frac{e^{-2|f|r}}{n} \sum_{k=0}^{n-1} \sum_{\substack{\beta' \in \Sigma_{\mathbf{A},k-r} \\ \tau' \in \Sigma_{\mathbf{A}, 2n-|\mathbf{i}|-k-r}}} \frac{V(f, \mathbf{z}, \beta') V(f, \sigma^k \mathbf{z}, \mathbf{i}) V(f, \sigma^{|\mathbf{w}|+k+r} \mathbf{z}, \tau')}{Z_k(\mathbf{z}) Z_{|\mathbf{i}|}(\sigma^k \mathbf{z}) Z_{2n-k-|\mathbf{i}|}(\sigma^{|\mathbf{i}|+k} \mathbf{z})} \mathbb{1}_{[\mathbf{w}]}(\sigma^k \mathbf{z})
$$

$$
\geq \frac{e^{-2|f|r}}{n} \sum_{k=0}^{n-1} \frac{V(f, \sigma^k \mathbf{z}, \mathbf{i})}{Z_{|\mathbf{i}|}(\sigma^k \mathbf{z}) Z_r(\mathbf{z}) Z_r(\sigma^{|\mathbf{i}|+k} \mathbf{z})} \mathbb{1}_{[\mathbf{w}]}(\sigma^k \mathbf{z})
$$

$$
\geq \frac{e^{-4|f|r} K^{-22r}}{n} \sum_{k=0}^{n-1} \frac{V(f, \sigma^k \mathbf{z}, \mathbf{i})}{Z_{|\mathbf{i}|}(\sigma^k \mathbf{z})} \mathbb{1}_{[\mathbf{w}]}(\sigma^k \mathbf{z})
$$

and thus, taking the subsequence n_j and using (4.10) and (4.11), we have

$$
\mu([\mathbf{w}, \mathbf{i}]) \geq C'^{-1} \frac{Y(f, \mathbf{w}, \mathbf{i})}{W_{|\mathbf{w}|}(f, \mathbf{w})} \nu([\mathbf{w}]).
$$

Now, since ν is quasi-Bernoulli, by (4.7)-(4.8) and (4.10)-(4.11) we have

$$
\mu([(\mathbf{wx}, \mathbf{ij})]) \geq C'^{-1} \frac{Y(f, \mathbf{wx}, \mathbf{ij})}{W_{|\mathbf{wx}|}(f, \mathbf{wx})} \nu([\mathbf{wx}])
$$

$$
\geq C'^{-2} \frac{Y(f, \mathbf{w}, \mathbf{i})}{W_{|\mathbf{w}|}(f, \mathbf{w})} \nu([\mathbf{w}]) \frac{Y(f, \mathbf{x}, \mathbf{j})}{W_{|\mathbf{x}|}(f, \mathbf{x})} \nu([\mathbf{x}])
$$

$$
\geq C'^{-4} \mu([(\mathbf{w}, \mathbf{i})]) \mu([(\mathbf{x}, \mathbf{j})]).
$$

This implies that μ is ergodic. Since μ was an arbitrary accumulation point and two equivalent ergodic measures are equal, we get that μ is unique.

For every $\mathbf{w} \in \Omega_*$, and every $\mathbf{z} \in [\mathbf{w}]$

$$
\Pi_* \mu([\mathbf{w}]) = \sum_{\mathbf{i} \in \Sigma_{\mathbf{A}, |\mathbf{w}|}} \mu([\mathbf{w}, \mathbf{i}])
$$

$$
\leq C \sum_{\mathbf{i} \in \Sigma_{\mathbf{A}, |\mathbf{w}|}} \frac{Y(f, \mathbf{w}, \mathbf{i})}{W_{|\mathbf{w}|}(f, \mathbf{w})} \nu([\mathbf{w}])
$$

$$
\leq C^3 \sum_{\mathbf{i} \in \Sigma_{\mathbf{A}, |\mathbf{w}|}} \frac{V(f, \mathbf{z}, \mathbf{i})}{Z_{|\mathbf{w}|}(f, \mathbf{z})} \nu([\mathbf{w}])
$$

$$
= C^3 \nu([\mathbf{w}]).
$$

The other inequality $\Pi_* \mu([\mathbf{w}]) \geq C^{-3} \nu([\mathbf{w}])$ is similar. Since $\Pi_* \mu$ and ν are both ergodic, we have $\Pi_* \mu = \nu$.

Finally, by (4.12)

$$h_\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{(\mathbf{w},\mathbf{i})\in\Gamma_{\mathbf{A},n}} \mu([\mathbf{w},\mathbf{i}]) \log\mu([\mathbf{w},\mathbf{i}])$$

$$= \lim_{n\to\infty} \frac{-1}{n} \sum_{(\mathbf{w},\mathbf{i})\in\Gamma_{\mathbf{A},n}} \mu([\mathbf{w},\mathbf{i}]) \log\left(\frac{Y(f,\mathbf{w},\mathbf{i})}{W_{|\mathbf{w}|}(f,\mathbf{w})}\nu([\mathbf{w}])\right)$$

$$= h_\nu - \int f d\mu + P_\nu(f).$$

By Theorem 4.2, $h_\mu^\xi = h_\mu - h_\nu$, which proves the statement. $\qquad\square$

The next theorem is a modification of the argument of Heurteaux [21].

Theorem 4.5. *Let ν be a σ-invariant ergodic quasi-Bernoulli measure on Ω and let $f, g\colon \Gamma_{\mathbf{A}} \mapsto \mathbb{R}$ be a continuous potentials with summable variation. Then the function $p\colon t \mapsto P((1-t)g + tf)$ is differentiable at $t = 0$. In particular,*

$$p'(0) = \int (f - g)d\mu_g.$$

Proof. It is clear by the bounded distortion (4.6) that there exists a constant $C > 0$ such that for every $t \in \mathbb{R}$ and every $(\mathbf{w},\mathbf{i}) \in \Gamma_{\mathbf{A},*}$

$$C^{-1}Y(tf + (1-t)g,\mathbf{w},\mathbf{i}) \le Y(f,\mathbf{w},\mathbf{i})^t Y(g,\mathbf{w},\mathbf{i})^{1-t} \le CY(tf + (1-t)g,\mathbf{w},\mathbf{i}).$$

Let μ_f and μ_g be the unique ergodic measures defined in Theorem 4.4. Then for every $t \in \mathbb{R}$ and every $(\mathbf{w},\mathbf{i}) \in \Gamma_{\mathbf{A},*}$

$$C^{-2}\frac{Y(tf + (1-t)g,\mathbf{w},\mathbf{i})}{W(f,\mathbf{w})^t W(g,\mathbf{w})^{1-t}}\nu(\mathbf{w}) \le \mu_f([\mathbf{w},\mathbf{i}])^t \mu_g([\mathbf{w},\mathbf{i}])^{1-t}$$

$$\le C^2\frac{Y(tf + (1-t)g,\mathbf{w},\mathbf{i})}{W(f,\mathbf{w})^t W(g,\mathbf{w})^{1-t}}\nu(\mathbf{w}).$$

Hence,

$$P_\nu((1-t)g + ft)$$

$$= (1-t)P_\nu(g) + tP_\nu(f) + \lim_{n\to\infty}\frac{1}{n}\sum_{\mathbf{w}\in\Omega_n}\nu([\mathbf{w}])\log\sum_{\mathbf{i}\in\Sigma_{\mathbf{A},n}}\frac{\mu_f([\mathbf{w},\mathbf{i}])^t\mu_g([\mathbf{w},\mathbf{i}])^{1-t}}{\nu([\mathbf{w}])}.$$

Thus, it is enough to show that

$$H(t) = \lim_{n\to\infty}\frac{1}{n}\sum_{\mathbf{w}\in\Omega_n}\nu([\mathbf{w}])\log\sum_{\mathbf{i}\in\Sigma_{\mathbf{A},n}}\frac{\mu_f([\mathbf{w},\mathbf{i}])^t\mu_g([\mathbf{w},\mathbf{i}])^{1-t}}{\nu([\mathbf{w}])}$$

is differentiable.

Claim: There exists a constant $C > 0$ such that the sequence

$$\overline{H}_n(t) = \sum_{\mathbf{w}\in\Omega_n}\nu([\mathbf{w}])\log\sum_{\mathbf{i}\in\Sigma_{\mathbf{A},n}}\frac{C\mu_f([\mathbf{w},\mathbf{i}])^t\mu_g([\mathbf{w},\mathbf{i}])^{1-t}}{\nu([\mathbf{w}])}$$

is submultiplicative $\overline{H}_{n+m}(t) \le \overline{H}_n(t) + \overline{H}_m(t)$ and

$$\underline{H}_n(t) = \sum_{\mathbf{w}\in\Omega_n}\nu([\mathbf{w}])\log\sum_{\mathbf{i}\in\Sigma_{\mathbf{A},n}}\frac{C^{-1}\mu_f([\mathbf{w},\mathbf{i}])^t\mu_g([\mathbf{w},\mathbf{i}])^{1-t}}{\nu([\mathbf{w}])}$$

is supermultiplicative $\underline{H}_{n+m}(t) \ge \underline{H}_n(t) + \underline{H}_m(t)$.

Proof of the Claim. By Theorem 4.4 and the equations (4.7)-(4.11), we have that the measures μ_f and μ_g are quasi-Bernoulli, and hence, there exists a constant $C > 0$ such that

$$\sum_{\mathbf{ij}\in\Sigma_{\mathbf{A},n+m}} \mu_f([\mathbf{w},\mathbf{ij}])^t \mu_g([\mathbf{w},\mathbf{ij}])^{1-t}$$

$$\leq C \sum_{\mathbf{ij}\in\Sigma_{\mathbf{A},n+m}} \mu_f([\mathbf{w}|_n,\mathbf{i}])^t \mu_f([\sigma^n\mathbf{w},\mathbf{j}])^t \mu_g([\mathbf{w}|_n,\mathbf{i}])^t \mu_g([\sigma^n\mathbf{w},\mathbf{j}])^t$$

$$\leq C \sum_{\substack{\mathbf{i}\in\Sigma_{\mathbf{A},n} \\ \mathbf{j}\in\Sigma_{\mathbf{A},m}}} \mu_f([\mathbf{w}|_n,\mathbf{i}])^t \mu_f([\sigma^n\mathbf{w},\mathbf{j}])^t \mu_g([\mathbf{w}|_n,\mathbf{i}])^t \mu_g([\sigma^n\mathbf{w},\mathbf{j}])^t.$$

On the other hand,

$$\sum_{\mathbf{ij}\in\Sigma_{\mathbf{A},n+m}} \mu_f([\mathbf{w},\mathbf{ij}])^t \mu_g([\mathbf{w},\mathbf{ij}])^{1-t}$$

$$\geq C^{-1} \sum_{\mathbf{ij}\in\Sigma_{\mathbf{A},n+m}} \mu_f([\mathbf{w}|_n,\mathbf{i}])^t \mu_f([\sigma^n\mathbf{w},\mathbf{j}])^t \mu_g([\mathbf{w}|_n,\mathbf{i}])^t \mu_g([\sigma^n\mathbf{w},\mathbf{j}])^t$$

$$C^{-1}C' \sum_{\substack{\mathbf{i}\in\Sigma_{\mathbf{A},n-r} \\ \mathbf{j}\in\Sigma_{\mathbf{A},m-r}}} \mu_f([\mathbf{w}|_{n-r},\mathbf{i}])^t \mu_f([\sigma^{n+2r}\mathbf{w},\mathbf{j}])^t \mu_g([\mathbf{w}|_{n-r},\mathbf{i}])^{1-t} \mu_g([\sigma^{n+2r}\mathbf{w},\mathbf{j}])^{1-t}$$

$$\geq C^{-1}C'K^{-2r} \sum_{\substack{\mathbf{i}\in\Sigma_{\mathbf{A},n} \\ \mathbf{j}\in\Sigma_{\mathbf{A},m}}} \mu_f([\mathbf{w}|_n,\mathbf{i}])^t \mu_f([\sigma^n\mathbf{w},\mathbf{j}])^t \mu_g([\mathbf{w}|_n,\mathbf{i}])^{1-t} \mu_g([\sigma^n\mathbf{w},\mathbf{j}])^{1-t}.$$

$\square$

Since $H(0) = 0$ and $\overline{H}_n(t)$ is differentiable for every n, we get for every $n \geq 1$

$$\limsup_{t\to 0} \frac{H(t)}{t} \leq \limsup_{t\to 0} \frac{\overline{H}_n(t)}{nt}$$

$$= \frac{1}{n} \sum_{\mathbf{w}\in\Omega_n} \nu([\mathbf{w}]) \frac{\sum_{\mathbf{i}\in\Sigma_{\mathbf{A},n}} \frac{\mu_f([\mathbf{w},\mathbf{i}])^t \mu_g([\mathbf{w},\mathbf{i}])^{1-t}(\log\mu_f([\mathbf{w},\mathbf{i}])-\log\mu_g([\mathbf{w},\mathbf{i}]))}{\nu([\mathbf{w}])}}{\sum_{\mathbf{i}\in\mathcal{A}^n} \frac{\mu_f([\mathbf{w},\mathbf{i}])^t \mu_g([\mathbf{w},\mathbf{i}])^{1-t}}{\nu([\mathbf{w}])}} \Bigg|_{t=0}$$

$$= \frac{1}{n} \sum_{\mathbf{w}\in\Omega_n} \nu([\mathbf{w}]) \sum_{\mathbf{i}\in\Sigma_{\mathbf{A},n}} \frac{\mu_g([\mathbf{w},\mathbf{i}])(\log\mu_f([\mathbf{w},\mathbf{i}]) - \log\mu_g([\mathbf{w},\mathbf{i}]))}{\nu([\mathbf{w}])}$$

$$= \frac{1}{n} \sum_{\substack{\mathbf{w}\in\Omega_n \\ \mathbf{i}\in\Sigma_{\mathbf{A},n}}} \mu_g([\mathbf{w},\mathbf{i}])(\log\mu_f([\mathbf{w},\mathbf{i}]) - \log\mu_g([\mathbf{w},\mathbf{i}]))$$

$$\leq \frac{C}{n} + \frac{1}{n} \sum_{\substack{\mathbf{w}\in\Omega_n \\ \mathbf{i}\in\Sigma_{\mathbf{A},n}}} \mu_g([\mathbf{w},\mathbf{i}])(\log\frac{Y(f,\mathbf{w},\mathbf{i})\nu([\mathbf{w}])}{W(f,\mathbf{w})} - \log\mu_g([\mathbf{w},\mathbf{i}]))$$

$$\to \int f d\mu_g - h_\nu - P_\nu(f) + h_{\mu_g} \text{ as } n \to \infty,$$

where we applied again Theorem 4.4. The other inequality,

$$\liminf_{t\to 0} \frac{H(t)}{t} \geq \int f d\mu_g - h_\nu - P_\nu(f) + h_{\mu_g} \text{ as } n \to \infty$$

is similar. Hence,

$$p'(0) = -P_\nu(g) + P_\nu(f) + \int f d\mu_g - h_\nu - P_\nu(f) + h_{\mu_g} = \int f d\mu_g - \int g d\mu_g.$$

$\square$

4.3. **Weighted Birkhoff average.** For $\alpha, \underline{p} \in \mathbb{R}^d$, let us consider the potential $f_{\underline{p}} \colon \Gamma_{\mathbf{A}} \mapsto \mathbb{R}$ defined as

$$f_{\underline{p}} := \langle \underline{p}, f - \alpha \rangle.$$

First, we show the upper bound in Theorem 2.4.

Lemma 4.6. *For every* $\mathbf{w} \in \Omega$ *and* $\alpha \in \mathbb{R}^d$

$$h_{\text{top}}(E_{\mathbf{w}}(\alpha)) \le \inf_{\underline{p} \in \mathbb{R}^d} P(f_{\underline{p}}, \mathbf{w}).$$

Proof. The proof is standard, but for completeness, we give it here.

Let $s > s_0 > \inf_{\underline{p} \in \mathbb{R}^d} P(f_{\underline{p}}, \mathbf{w})$. Hence, there exists $\underline{p} \in \mathbb{R}^d$ such that $s_0 > P(f_{\underline{p}}, \mathbf{w})$. Thus there exists $N' \ge 1$ such that for every $n \ge N'$

$$\sum_{\mathbf{i} \in \Sigma_{\mathbf{A},n}} e^{\langle \underline{p}, S_n f - n\alpha \rangle} < e^{s_0 n}.$$

By definition,

$$E_{\mathbf{w}}(\alpha) = \bigcap_{M=1}^{\infty} \bigcup_{N=1}^{\infty} \bigcap_{n \ge N} \left\{ \mathbf{i} \in X : \left| \frac{1}{n} S_n f(\mathbf{w}, \mathbf{i}) - \alpha \right| < \frac{1}{M} \right\}. \qquad (4.14)$$

Since $f(\mathbf{w}, \cdot) \colon \Sigma_{\mathbf{A}} \mapsto \mathbb{R}^d$ is continuous over a compact set, we get that it is uniformly continuous. Thus, for every $M \ge 1$ there exists $C > 0$ such that for every $n \ge 1$, $\mathbf{i} \in \Sigma_{\mathbf{A},n}$ and every $\mathbf{j} \in [\mathbf{i}]$

$$\left| S_n f(\mathbf{w}, \mathbf{j}) - \sup_{\mathbf{j} \in [\mathbf{i}]} S_n f(\mathbf{w}, \mathbf{j}) \right| \le \frac{Cn}{M}.$$

Choose $M \ge 1$ such that $|\underline{p}| \frac{1+C}{M} < (s - s_0)/2$. By (4.14), we get that for every N sufficiently large

$$\mathcal{H}_N^s(E_{\mathbf{w}}(\alpha)) \le \sum_{n=N}^{\infty} \sum_{\substack{\mathbf{i} \in \Sigma_{\mathbf{A},n} \\ \left| \sup_{\mathbf{j} \in [\mathbf{i}]} S_n f(\mathbf{w}, \mathbf{j}) - n\alpha \right| < (1+C)n/M}} e^{-ns}$$

$$\le \sum_{n=N}^{\infty} e^{-n(s-s_0)/2} \sum_{\substack{\mathbf{i} \in \Sigma_{\mathbf{A},n} \\ \left| \sup_{\mathbf{j} \in [\mathbf{i}]} S_n f(\mathbf{w}, \mathbf{j}) - n\alpha \right| < (1+C)n/M}} e^{-ns_0 + \langle \underline{p}, S_n f - n\alpha \rangle}$$

$$\le \sum_{n=N}^{\infty} e^{-n(s-s_0)/2} \to 0 \text{ as } N \to \infty.$$

$\square$

Recall that

$$\mathcal{P}_{\mathbf{A}} = \{ \alpha \in \mathbb{R}^d : \text{ there exists } \mu \in \mathcal{M}_\nu(\Gamma_{\mathbf{A}}) \text{ such that } \int f d\mu = \alpha \}. \qquad (4.15)$$

It is easy to see that $\mathcal{P}_{\mathbf{A}}$ is a closed and convex set. Moreover, without loss of generality, we may assume that $\mathcal{P}_{\mathbf{A}}$ has an interior point. Indeed, if $\mathcal{P}_{\mathbf{A}}$ does not

contain interior point then there exists a d'-dimensional hyperplane V such that $\mathcal{P}_{\mathbf{A}} \subset V$. By changing coordinates, we may assume that $f \colon \Gamma_{\mathbf{A}} \mapsto \mathbb{R}^{d'}$. Also, for ν-almost every $\mathbf{w}$,

$$\mathcal{P}_{\mathbf{A}} = \{\alpha \in \mathbb{R}^d : \text{ there exists } \mathbf{i} \in \Sigma_{\mathbf{A}} \text{ such that } \lim_{n \to \infty} \frac{1}{n} S_n f(\mathbf{w}, \mathbf{i}) = \alpha\}.$$

Indeed, take the sequence $\mu_n = \frac{1}{n} \sum_{k=0}^{n} \delta_{\sigma^k \mathbf{w}, \sigma^k \mathbf{i}}$ and let μ be an accumulation point of the sequence μ_n in the weak*-topology, we get $\int f d\mu = \lim_{k \to \infty} \int f d\mu_{n_k} = \alpha$ and for every $g \in L^1(\Omega)$, $\int g d\Pi_* \mu = \lim_{k \to \infty} \int g \circ \Pi d\mu_{n_k} = \lim_{k \to \infty} \frac{1}{n_k} \sum_{\ell=0}^{n_k} g(\sigma^\ell \mathbf{w}) = \int g d\nu$. Moreover, since $\sigma_* \mu_n = \mu_n - \frac{1}{n} \delta_{\mathbf{w}, \mathbf{i}} + \frac{1}{n} \delta_{\sigma^{n+1}\mathbf{w}, \sigma^{n+1}\mathbf{i}}$, we get that μ is σ-invariant.

Theorem 4.4 implies that for every $\underline{p} \in \mathbb{R}^d$ there exists a σ-invariant ergodic measure $\mu_{\underline{p}}$ such that

$$P_\nu(f_{\underline{p}}) = h_{\mu_{\underline{p}}}^\xi + \int f_{\underline{p}} d\mu_{\underline{p}}. \tag{4.16}$$

Lemma 4.7. *The conditional pressure $\underline{p} \mapsto P_\nu(f_{\underline{p}})$ is convex.*

Proof. Let $\beta_1, \beta_2 > 0$ be with $\beta_1 + \beta_2 = 1$ and $\underline{p}_1, \underline{p}_2 \in \mathbb{R}^d$. Then there exist a measure $\mu = \mu_{\beta_1 \underline{p}_1 + \beta_2 \underline{p}_2} \in \mathcal{E}_\nu(\Gamma_{\mathbf{A}})$ such that

$$\begin{aligned}
P_\nu(f_{\beta_1 \underline{p}_1 + \beta_2 \underline{p}_2}) &= h_\mu^\xi + \int f_{\beta_1 \underline{p}_1 + \beta_2 \underline{p}_2} d\mu \\
&= \beta_1 h_\mu^\xi + \beta_2 h_\mu^\xi + \beta_1 \int f_{\underline{p}_1} d\mu + \beta_2 \int f_{\underline{p}_2} d\mu \\
&\leq \beta_1 P_\nu(f_{\underline{p}_1}) + \beta_2 P_\nu(f_{\underline{p}_2}).
\end{aligned}$$

$\square$

Lemma 4.8. *For every $\alpha \in \mathcal{P}_{\mathbf{A}}^o$, there there exists $\underline{p}^* \in \mathbb{R}^d$ such that $\inf_{\underline{p}} P_\nu(f_{\underline{p}}) = P_\nu(f_{\underline{p}^*})$, where $\mathcal{P}_{\mathbf{A}}^o$ denotes the interior of $\mathcal{P}_{\mathbf{A}}$.*

Proof. Suppose that $\alpha \in \mathcal{P}_{\mathbf{A}}^o$. Then there exists an $\eta > 0$ such that for every $\underline{p} \in \mathbb{R}^d$ with $|\underline{p}| = 1$ there exists $\mu \in \mathcal{M}_\nu(\Gamma_{\mathbf{A}})$ such that $\int f d\mu - \alpha = \eta \underline{p}$. Thus, for every $c > 0$

$$P_\nu(f_{c\underline{p}}) \geq h_\mu^\xi + \int \langle c\underline{p}, f - \alpha \rangle d\mu \geq c\eta |\underline{p}|^2 = c\eta.$$

Thus, $\lim_{|\underline{p}| \to \infty} P_\nu(f_{\underline{p}}) = \infty$ and by the convexity of the conditional pressure Lemma 4.7, we get the statement. $\square$

Lemma 4.9. *Let $\underline{p}^* \in \mathbb{R}^d$ be such that $\inf_{\underline{p}} P_\nu(f_{\underline{p}}) = P_\nu(f_{\underline{p}^*})$ and let $\mu_{\underline{p}^*}$ be the conditional equilibrium defined in Theorem 4.4. Then*

$$\int \phi d\mu_{\underline{p}^*} = \alpha.$$

Proof. Let us argue by contradiction. Suppose that $\int \phi d\mu_{\underline{p}^*} \neq \alpha$. Let $\underline{q} = \frac{\int \phi d\mu_{\underline{p}^*} - \alpha}{|\int \phi d\mu_{\underline{p}^*} - \alpha|}$.

Observe that for any $\underline{p}_1, \underline{p}_2 \in \mathbb{R}^d$ and $t \in \mathbb{R}$, $t f_{\underline{p}_1} + (1-t) f_{\underline{p}_2} = f_{t\underline{p}_1 + (1-t)\underline{p}_2}$. Hence, by Theorem 4.5, the function $p \colon t \mapsto P_\nu(f_{(1-t)\underline{p}^* + (\underline{p}^* + \underline{q})t})$ is differentiable at $t = 0$, moreover,

$$p'(0) = \int f_{\underline{p}^* + \underline{q}} - f_{\underline{p}^*} d\mu_{\underline{p}^*}.$$

But p has a minimum at $t = 0$ so

$$0 = p'(0) = \int f_{\underline{p}^* + \underline{q}} - f_{\underline{p}^*} d\mu_{\underline{p}^*} = \langle \underline{q}, \int \phi d\mu_{\underline{p}^*} - \alpha \rangle = \left| \int \phi d\mu_{\underline{p}^*} - \alpha \right|,$$

which is a contradiction. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ □

Proof of Theorem 2.4. It is enough to show that for every $\alpha \in \mathcal{P}_{\mathbf{A}}^o$ and ν-almost every $\mathbf{w}$

$$h_{\text{top}}(E_{\mathbf{w}}(\alpha)) \geq h_{\mu_{\underline{p}^*}} - h_{\nu},$$

where $\mu_{\underline{p}^*}$ is the conditional equilibrium of $P_{\nu}(f_{\underline{p}^*}) = \inf_{\underline{p} \in \mathbb{R}^d} P_{\nu}(f_{\underline{p}})$ defined in Theorem 4.4. Indeed, Theorem 4.2, Lemma 4.9 and Theorem 4.4 imply that

$$h_{\mu_{\underline{p}^*}} - h_{\nu} = h_{\mu_{\underline{p}^*}}^{\xi} = h_{\mu_{\underline{p}^*}}^{\xi} + \int f_{\underline{p}^*} d\mu_{\underline{p}^*} = P_{\nu}(f_{\underline{p}^*}) = \inf_{\underline{p} \in \mathbb{R}^d} P_{\nu}(f_{\underline{p}}).$$

The upper bound follows by equation (4.5) and Lemma 4.6.

Let $\mu_{\mathbf{w}}^{\xi}$ be the family of conditional measures with respect to the partition ξ and $\mu_{\underline{p}^*}$ defined by Rohlin's Disintegration Theorem. By Theorem 4.2,

$$\lim_{n \to \infty} \frac{-1}{n} \log \mu_{\mathbf{w}}^{\xi}([\mathbf{i}|_n]) = h_{\mu_{\underline{p}^*}} - h_{\nu} \text{ for } \mu_{\underline{p}^*}\text{-almost every } (\mathbf{w}, \mathbf{i}) \in \Gamma_{\mathbf{A}}.$$

By Egorov's theorem, for every $\varepsilon > 0$ there exists a set $J_1 \subset \Gamma_{\mathbf{A}}$ and a constant $C > 0$ such that $\mu_{\underline{p}^*}(J_1) > 1 - \varepsilon$ and for every $(\mathbf{w}, \mathbf{i}) \in J_1$ and $n \geq 1$

$$\mu_{\mathbf{w}}^{\xi}([\mathbf{i}|_n]) \leq C e^{-n(h_{\mu_{\underline{p}^*}} - h_{\nu} - \varepsilon)}.$$

Since $1 - \varepsilon < \mu_{\underline{p}^*}(J_1) = \int \mu_{\mathbf{w}}^{\xi}(J_1) d\nu(\mathbf{w})$, by Markov's inequality, we get that

$$\nu(\{\mathbf{w} \in \Omega : \mu_{\mathbf{w}}^{\xi}(J_1 \cap \xi(\mathbf{w})) > 1 - \sqrt{\varepsilon}) > 1 - \sqrt{\varepsilon}.$$

By Birkhoff's Ergodic Theorem and Lemma 4.9,

$$\lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} f(\sigma^k \mathbf{w}, \sigma^k \mathbf{i}) = \alpha.$$

Hence, there exists $J \subset J_1$ such that $\nu(J_1 \setminus J) = 0$ and for every $\mathbf{w} \in J$, $\mu_{\mathbf{w}}^{\xi}(E_{\mathbf{w}}(\alpha) \cap J_1) > 1 - \sqrt{\varepsilon}$. Thus, by Lemma 2.1 for every $\mathbf{w} \in J$

$$h_{\text{top}}(E_{\mathbf{w}}(\alpha)) \geq h_{\text{top}}(E_{\mathbf{w}}(\alpha) \cap J_1) \geq h_{\mu_{\underline{p}^*}} - h_{\nu} - \varepsilon.$$

Since $\varepsilon > 0$ was arbitrary, the statement follows.

Finally, let $\widetilde{\mu}$ be the ergodic σ-invariant measure on $\Sigma_{\mathbf{A}}$ such that $h_{\widetilde{\mu}} = h_{\text{top}}(\Sigma_{\mathbf{A}})$. Then for $\alpha_0 = \iint f(\mathbf{w}, \mathbf{i}) d\widetilde{\mu}(\mathbf{i}) d\nu(\mathbf{w})$ we get $h_{\text{top}}(E_{\mathbf{w}}(\alpha_0)) \geq h_{\text{top}}(\Sigma_{\mathbf{A}})$ for ν-almost every $\mathbf{w}$. $\qquad\qquad\qquad\qquad$ □

Proof of Theorem 2.5. Let I be the domain of the map

$$p \colon \alpha \mapsto \inf_{p \in \mathbb{R}} P_{\nu}(p \cdot (f - \alpha)) = \inf_{p \in \mathbb{R}} (P_{\nu}(p \cdot f) - p\alpha).$$

If I is empty or a single point then there is nothing to prove, so we might assume that I has non-empty interior. By Theorem 4.5, the map $p \mapsto P_{\nu}(p \cdot f)$ is differentiable and by Lemma 4.7, the derivative $p \mapsto P_{\nu}'(p \cdot f) = \int f d\mu_p$ is increasing. Hence, $I = [\lim_{p \to -\infty} P_{\nu}'(pf), \lim_{p \to \infty} P_{\nu}'(pf)]$. Moreover, the map $\alpha \mapsto p(\alpha)$ is concave and continuous over I.

By Theorem 2.4, for every $\alpha \in I^o$ and ν-almost every $\mathbf{w}$, $h_{\text{top}}(E_{\mathbf{w}}(\alpha)) = p(\alpha)$. Then by Fubini's Theorem, for ν-almost every $\mathbf{w}$ and Lebesgue almost every $\alpha \in I^o$, $h_{\text{top}}(E_{\mathbf{w}}(\alpha)) = p(\alpha)$.

Using Theorem 2.2 with the choice $\phi_i(\mathbf{i}) := f(\sigma^i \mathbf{w}, \mathbf{i})$, the map $\alpha \mapsto h_{\text{top}}(E_{\mathbf{w}}(\alpha))$ is continuous for every $\mathbf{w} \in \Omega$. This together with the continuity of the map $\alpha \mapsto p(\alpha)$ implies that $p(\alpha) \equiv h_{\text{top}}(E_{\mathbf{w}}(\alpha))$ over I for ν-almost every $\mathbf{w}$. □

5. **Frequency regular sequences.** In the rest of the paper, we assume that $\Sigma_{\mathbf{A}} = \Sigma$, that is, we need to work on the full shift. In this section, we establish the connection between ν-typical and frequency regular sequences and prove Theorem 2.6. The proof of our main theorem relies on the following construction, which first appeared in Rams [25].

Let $\mathbf{w}, \mathbf{w}' \in \Omega$ be two q-frequency regular sequences with the same frequency. We define a permutation γ on $\mathbb{N}$ such that

$$\gamma(k) = \ell \text{ if } \omega_k \text{ is the } n\text{th appearance of the symbol of } \omega_k \text{ in } \mathbf{w}$$
$$\text{then } \ell \text{ is the position of the } n\text{th appearance of } \omega_k \text{ in } \mathbf{w}'. \tag{5.1}$$

More precisely, let

$$M_{n,\lambda_i}(\mathbf{w}) = \min\{k \geq 1 : \#\{1 \leq j \leq k : w_j = \lambda_i\} = n\}$$

and

$$P_k(\mathbf{w}) = \#\{1 \leq i \leq k : w_i = w_k\}.$$

Then

$$\gamma(k) = M_{P_k(\mathbf{w}), w_k}(\mathbf{w}').$$

By the definition of $\gamma(k)$, we have $w_k = w'_{\gamma(k)}$. Finally, we set the map

$$G_{\mathbf{w}, \mathbf{w}'}(\mathbf{i}) := (i_{\gamma(1)}, i_{\gamma(2)}, \ldots). \tag{5.2}$$

Lemma 5.1. *For $\mathbf{w}, \mathbf{w}' \in \Omega$ as above, for every $\alpha < 1$ there exists $C > 0$ such that for every $\mathbf{i}, \mathbf{j} \in \Sigma$*

$$d(G_{\mathbf{w}, \mathbf{w}'}(\mathbf{i}), G_{\mathbf{w}, \mathbf{w}'}(\mathbf{j})) \leq Cd(\mathbf{i}, \mathbf{j})^{\alpha}. \tag{5.3}$$

Moreover, $G_{\mathbf{w}, \mathbf{w}'} \circ G_{\mathbf{w}', \mathbf{w}}(\mathbf{i}) = \mathbf{i}$.

Proof. The construction clearly implies that $G_{\mathbf{w}, \mathbf{w}'} \circ G_{\mathbf{w}', \mathbf{w}}$ is the identity map on Σ.

Since $\mathbf{w}, \mathbf{w}' \in \Omega$ are frequency regular sequences, we have that λ_i appears infinitely often in $\mathbf{w}, \mathbf{w}'$. Thus, for every $n \geq 1$ we can define m_n such that m_n is the smallest positive integer such that $\{1, \ldots, n\} \subseteq \{\gamma(1), \ldots, \gamma(m_n)\}$. Hence, for every $n \geq 1$

$$\text{if } d(\mathbf{i}, \mathbf{j}) = e^{-m_n - 1} \text{ then } d(G_{\mathbf{w}, \mathbf{w}'}(\mathbf{i}), G_{\mathbf{w}, \mathbf{w}'}(\mathbf{j})) = e^{-n - 1}.$$

Thus, to prove (5.3), it is enough to show that

$$\lim_{n \to \infty} \frac{m_n}{n} = 1. \tag{5.4}$$

Clearly $m_n \geq n$, so $\liminf_{n \to \infty} \frac{m_n}{n} \geq 1$. By the definition of m_n, for every $i = 1, \ldots, N$,

$$\#\{1 \leq k \leq m_n : w_k = \lambda_i\} \geq \#\{1 \leq k \leq n : w'_k = \lambda_i\}, \tag{5.5}$$

and there exists (at least one) $j = j(n)$ such that

$$\#\{1 \leq k \leq m_n : w_k = \lambda_j\} = \#\{1 \leq k \leq n : w'_k = \lambda_j\}. \tag{5.6}$$

By frequency regularity, for every $0 < \varepsilon < \min_i q_i/2$ there exists $N \geq 1$ such that for every $n \geq N$

$$\left|\frac{\#\{1 \le k \le n : w_k = \lambda_i\}}{n} - q_i\right|, \left|\frac{\#\{1 \le k \le n : w_k' = \lambda_i\}}{n} - q_i\right| < \varepsilon.$$

Hence, by (5.6) for every $n \ge N$

$$\frac{m_n}{n}(q_{j(n)} - \varepsilon) \le \frac{m_n}{n} \frac{\#\{1 \le k \le m_n : w_k = \lambda_{j(n)}\}}{m_n}$$

$$= \frac{\#\{1 \le k \le n : w_k' = \lambda_{j(n)}\}}{n} \le q_{j(n)} + \varepsilon.$$

Thus, for every $n \ge N$, $\frac{m_n}{n} \le 1 + 4\varepsilon$. $\hfill\square$

Proposition 5.2. *For every two $\underline{q}$-frequency regular sequences $\mathbf{w}, \mathbf{w}'$ with the same frequency*

$$h_{\mathrm{top}}(E_{\mathbf{w}}(\alpha)) = h_{\mathrm{top}}(E_{\mathbf{w}'}(\alpha)). \tag{5.7}$$

Proof. Let $\mathbf{w}, \mathbf{w}'$ be $\underline{q}$-frequency regular sequences. Let $G_{\mathbf{w}, \mathbf{w}'}$ be the map defined in (5.2). It is enough to show that

$$G_{\mathbf{w}, \mathbf{w}'}(E_{\mathbf{w}'}(\alpha)) \subseteq E_{\mathbf{w}}(\alpha). \tag{5.8}$$

Indeed, by (5.3),

$$h_{\mathrm{top}}(E_{\mathbf{w}'}(\alpha)) = h_{\mathrm{top}}(G_{\mathbf{w}', \mathbf{w}} \circ G_{\mathbf{w}, \mathbf{w}'}(E_{\mathbf{w}'}(\alpha)) \le h_{\mathrm{top}}(G_{\mathbf{w}, \mathbf{w}'}(E_{\mathbf{w}'}(\alpha)) \le h_{\mathrm{top}}(E_{\mathbf{w}}(\alpha)).$$

The other inequality follows by symmetry.

Let $\gamma \colon \mathbb{N} \mapsto \mathbb{N}$ be the map defined in (5.1). Let us define p_n as the largest non-negative integer such that $\{1, \ldots, p_n\} \subseteq \{\gamma(1), \ldots, \gamma(n)\}$. In other words, $p_n = \min\{k \ge 1 : k \notin \{\gamma(1), \ldots, \gamma(n)\}\} - 1$. Similarly to (5.4) one can show that

$$\lim_{n \to \infty} \frac{p_n}{n} = 1. \tag{5.9}$$

Let $\mathbf{i} \in E_{\mathbf{w}'}(\alpha)$. Then by (5.9)

$$\frac{1}{n}\sum_{k=0}^{n-1} f(\sigma^k \mathbf{w}, \sigma^k G_{\mathbf{w}, \mathbf{w}'}(\mathbf{i})) = \frac{1}{n}\sum_{k=0}^{n-1} f_{w_k, i_{\gamma(k)}}$$

$$= \frac{1}{n}\sum_{k=0}^{n-1} f_{w'_{\gamma(k)}, i_{\gamma(k)}}$$

$$= \frac{1}{n}\sum_{k=0}^{p_n-1} f_{w'_k, i_k} + \frac{1}{n}\sum_{\substack{k=0 \\ \gamma(k) > p_n}}^{n-1} f_{w'_{\gamma(k)}, i_{\gamma(k)}}$$

$$\le \frac{p_n}{n}\frac{1}{p_n}\sum_{k=0}^{p_n-1} f_{w'_k, i_k} + \frac{n - p_n}{n}\max_{i,j} f_{i,j} \to \alpha,$$

as $n \to \infty$. Similarly,

$$\frac{1}{n}\sum_{k=0}^{n-1} w_k \phi(\sigma^k G_{\mathbf{w}, \mathbf{w}'}(\mathbf{i})) \ge \frac{p_n}{n}\frac{1}{p_n}\sum_{k=0}^{p_n-1} f_{w'_k, i_k} + \frac{n - p_n}{n}\min_{i,j} f_{i,j} \to \alpha$$

as $n \to \infty$. Hence, $G_{\mathbf{w}, \mathbf{w}'}(\mathbf{i}) \in E_{\mathbf{w}}(\alpha)$ which verifies (5.8). $\hfill\square$

Proof of Theorem 2.6. Let ν be the Bernoulli measure associated to the weights $\underline{q} = (q_1, \ldots, q_N)$. Simple calculations show that the conditional pressure $P_\nu(\langle f - \alpha, \underline{p} \rangle)$ defined in (2.5) equals to $P_{\underline{q}}(\langle f - \alpha, \underline{p} \rangle)$ in (2.8).

Hence, by applying Theorem 2.4 we get that for every α and ν-almost every $\mathbf{w}$

$$h_{\text{top}}(E_\mathbf{w}(\alpha)) = \sup\{h_\mu : \mu \in \mathcal{E}_\nu(\Gamma) \text{ and } \sum_{i,j}^{K,N} f_{j,i}\mu([j,i]) = \alpha\} - h_\nu$$

$$= \inf_{\underline{p} \in \mathbb{R}^d} P_{\underline{q}}(\langle \underline{p}, f - \alpha \rangle).$$

By convexity, $\inf_{\underline{p} \in \mathbb{R}^d} P_{\underline{q}}(\langle \underline{p}, f - \alpha \rangle)$ is attained at $\underline{p}^*$. By Theorem 4.4, we know that the measure $\mu_{\underline{p}^*}$ where the supremum is attained can be chosen such that

$$C^{-1} \frac{Y(f_{\underline{p}^*}, \mathbf{w}, \mathbf{i})}{W_{|\mathbf{w}|}(f_{\underline{p}^*}, \mathbf{w})} \nu([\mathbf{w}]) \leq \mu_{\underline{p}^*}([\mathbf{w}, \mathbf{i}]) \leq C \frac{Y(f_{\underline{p}^*}, \mathbf{w}, \mathbf{i})}{W_{|\mathbf{w}|}(f_{\underline{p}^*}, \mathbf{w})} \nu([\mathbf{w}]),$$

hold for some uniform constant $C > 0$, where $f_{\underline{p}^*}(\mathbf{w}, \mathbf{i}) = \langle \underline{p}, \lambda_{w_0}\phi_{i_0} - \alpha \rangle$. However, in this case,

$$\eta([\mathbf{w}, \mathbf{i}]) = \frac{Y(f_{\underline{p}^*}, \mathbf{w}, \mathbf{i})}{W_{|\mathbf{w}|}(f_{\underline{p}^*}, \mathbf{w})} \nu([\mathbf{w}]) = \prod_{k=0}^{|\mathbf{w}|-1} \frac{q_{w_k} e^{\langle \underline{p}^*, \lambda_{w_k}\phi_{i_k} - \alpha \rangle}}{\sum_{i=1}^{K} e^{\langle \underline{p}^*, \lambda_{w_k}\phi_i - \alpha \rangle}}$$

is clearly an ergodic Bernoulli measure on Γ, since $\mu_{\underline{p}^*}$ is equivalent to η, we have $\eta = \mu_{\underline{p}^*}$. This shows that the supreme is attained at Bernoulli measures.

Finally, since ν-almost every sequence $\mathbf{w}$ is $\underline{q}$-frequency regular, the statement follows by Proposition 5.2. $\square$

Proof of Theorem 2.7. Since the function $g(i) = \sum_{j=1}^{N} q_j f_{j,i}$ is not constant by assumption, the possible values of α, for which $\sum_{i,j} p_{j,i} f_{j,i} = \alpha$ and $\sum_i p_{j,i} = q_j$ form a non-trivial closed interval. Hence, the statement follows by Theorem 2.3. $\square$

Now we finish the paper by showing the necessity of the frequency regular condition to have non-degenerate spectrum. Example 2.11 follows by the next example.

Example 5.3. *There exists a sequence $\mathbf{w} \in \{0,1\}^\mathbb{N}$, which is not frequency regular, such that the following holds: For every continuous potential $\varphi \colon \{0,1\}^\mathbb{N} \mapsto \mathbb{R}$, $E_\mathbf{w}(\alpha) = \emptyset$ for every $\alpha \in \mathbb{R} \setminus \{0\}$.*

Moreover, if φ depends only the first symbol then $E_\mathbf{w}(0) \neq \emptyset$ if and only if $\varphi_0 \varphi_1 \leq 0$, moreover if additionally $\varphi_0 \neq -\varphi_1$ then $h_{\text{top}}(E_\mathbf{w}(0)) < \log 2$.

Proof of Example 5.3. First, let us define the sequence $\mathbf{w} \in \{0,1\}^\mathbb{N}$. Let $\{M_n\}_{n=0}^\infty$ be a fast increasing sequence, that is, suppose that $2M_n < M_{n+1}$ for every $n \geq 0$ and $\lim_{n \to \infty} \frac{\sum_{j=1}^n M_j}{M_{n+1}} = 0$. Let $\mathbf{w} := (w_0, w_1, \ldots)$, where

$$w_k = \begin{cases} 0 & \text{if } 2M_{n-1} < k \leq M_n, \\ 1 & \text{if } M_n < k \leq 2M_n. \end{cases}$$

Clearly, $\mathbf{w}$ is not frequency regular. Moreover, since for every $\mathbf{i} \in \Sigma$

$$\left| \frac{1}{M_n} \sum_{k=0}^{M_n} w_k \varphi(\sigma^k \mathbf{i}) \right| \leq \frac{\max_{\mathbf{i} \in \Sigma} |\varphi(\mathbf{i})| \sum_{\ell=0}^{n-1} M_\ell}{M_n} \to 0 \text{ as } n \to \infty,$$

we get that $E_{\mathbf{w}}(\alpha) = \emptyset$ for every $\alpha \neq 0$. On the other hand, if $\min_{\mathbf{i} \in \Sigma} \varphi(\mathbf{i}) > 0$ then

$$\frac{1}{2M_n} \sum_{k=0}^{2M_n} w_k \varphi(\sigma^k \mathbf{i}) \geq \frac{\min_{\mathbf{i} \in \Sigma} \varphi(\mathbf{i}) \sum_{\ell=0}^{n} M_\ell}{2M_n} \to \frac{\min_{\mathbf{i} \in \Sigma} \varphi(\mathbf{i})}{2} \text{ as } n \to \infty,$$

so $E_{\mathbf{w}}(0) = \emptyset$ as well. Similarly, $E_{\mathbf{w}}(0) = \emptyset$ also in the case if $\max_{\mathbf{i} \in \Sigma} \varphi(\mathbf{i}) < 0$.

Now, suppose that $\varphi(\mathbf{i}) = \varphi_{i_0}$. Using the previous calculations if $\varphi_0 \varphi_1 > 0$ then $E_{\mathbf{w}}(\alpha) = \emptyset$ for every $\alpha \in \mathbb{R}$. So we may assume that $\varphi_0 \varphi_1 \leq 0$. If $\varphi_0 = 0$ then the sequence $(0, 0, \ldots)$ belongs to $E_{\mathbf{w}}(0)$, so let us assume $\varphi_0 < 0 < \varphi_1$. Then let us define the sequence $\mathbf{i}$ inductively by the rule $i_{m+1} = 0$ if and only if $A_m(\mathbf{i}) > 0$. Thus, $\mathbf{i}$ belongs to $E_{\mathbf{w}}(0)$.

Additionally, suppose that $\varphi_0 \neq -\varphi_1$. For every m, let n_m be such that $M_{n_m} < m \leq M_{n_m+1}$. By the definition of $\mathbf{w}$ we get that

$$A_m(\mathbf{i}) = \frac{1}{m} \sum_{k=0}^{m} w_k \varphi_{i_k} = \frac{M_{n_m} \sum_{k=0}^{M_{n_m}} w_k \varphi_{i_k}}{m M_{n_m}} + \frac{\sum_{k=M_{n_m}+1}^{\min\{m, 2M_{n_m}\}} \varphi_{i_k}}{m}.$$

Since $\dfrac{\sum_{k=0}^{M_{n_m}} w_k \varphi_{i_k}}{M_{n_m}} \to 0$ as $m \to \infty$ and $\frac{M_{n_m}}{m}$ is bounded, we get $A_m(\mathbf{i}) \to 0$ if and only if

$$\frac{\sum_{\ell=0,1} \#\{M_{n_m} < k \leq \min\{m, 2M_{n_m}\} : i_k = \ell\} \varphi_\ell}{m} \to 0.$$

In particular, $A_m(\mathbf{i}) \to 0$ implies that

$$\frac{\#\{M_n < k \leq 2M_n : i_k = 0\}}{M_n} \to \frac{\varphi_1}{\varphi_1 - \varphi_0} \text{ as } n \to \infty. \tag{5.10}$$

Denote F the set of all $\mathbf{i} \in \Sigma$, which satisfy (5.10). Then $h_{\text{top}}(E_{\mathbf{w}}(0)) \leq h_{\text{top}}(F)$.

For short, let $p = \frac{\varphi_1}{\varphi_1 - \varphi_0}$. Well known (for example, it is an application of Stirling's formula) that there exists $K(p) > |H'(p)|$, where $H(p) = -p \log p - (1 - p) \log(1 - p)$ such that for every $\varepsilon > 0$ there exists $L \geq 1$ such that for every $n \geq L$

$$\#\left\{\mathbf{i} \in \{0,1\}^n : \left|\frac{\#\{0 < k \leq n : i_k = 0\}}{n} - p\right| < \varepsilon\right\} \leq e^{(-p \log p - (1-p) \log(1-p) + K(p)\varepsilon)n}$$

and by (2.2),

$$h_{\text{top}}(F) \leq \liminf_{n \to \infty} \frac{1}{2M_n} \log \# \{\mathbf{i} \in \Sigma_{2M_n} : F \cap [\mathbf{i}] \neq \emptyset\}$$

$$\leq \lim_{n \to \infty} \frac{1}{2M_n} \log \prod_{k=1}^{n} 2^{M_k - 2M_{k-1}} e^{(-p \log p - (1-p) \log(1-p) + K(p)\varepsilon)M_k}$$

$$= \frac{\log 2 - p \log p - (1 - p) \log(1 - p) + K(p)\varepsilon}{2}.$$

Since $\varepsilon > 0$ was arbitrary and by assumption $p \neq 1/2$, we get $h_{\text{top}}(F) < \log 2$, which completes the proof. $\qquad\square$

REFERENCES

[1] E. H. El Abdalaoui, J. Kułaga-Przymus, M. Lemańczyk and T. de la Rue, The Chowla and the Sarnak conjectures from ergodic theory point of view, *Discrete Contin. Dyn. Syst.*, **37** (2017), 2899–2944.

[2] L. Barreira, B. Saussol and J. Schmeling, Distribution of frequencies of digits via multifractal analysis, *J. Number Theory*, **97** (2002), 410–438.

[3] L. Barreira, B. Saussol and J. Schmeling, Higher-dimensional multifractal analysis, *J. Math. Pures Appl.*, **81** (2002), 67–91.

[4] A. S. Besicovitch, On the sum of digits of real numbers represented in the dyadic system, *Math. Ann.*, **110** (1935), 321–330.

[5] C. J. Bishop and Y. Peres, *Fractals in Probability and Analysis,* Cambridge Studies in Advanced Mathematics, vol. 162, Cambridge University Press, Cambridge, 2017.

[6] F. Cellarosi and Y. Sinai, The Möbius function and statistical mechanics, *Bull. Math. Sci.*, **1** (2011), 245–275.

[7] H. G. Eggleston, The fractional dimension of a set defined by decimal properties, *Quart. J. Math.*, Oxford Ser, **20** (1949), 31–36.

[8] T. Eisner, B. Farkas, M. Haase and R. Nagel, *Operator Theoretic Aspects of Ergodic Theory*, Graduate Texts in Mathematics, 272. Springer-Verlag, New York, 2015.

[9] A. Fan, Multifractal analysis of infinite products, *J. Statist. Phys.*, **86** (1997), 1313–1336.

[10] A. Fan, Weighted Birkhoff ergodic theorem with oscillating weights, *Ergodic Theory Dynam. Systems*, **39** (2019), 1275–1289.

[11] A. Fan, Multifractal analysis of weighted ergodic averages, *Adv. Math.*, **377** (2021), 107488.

[12] A. Fan and D. Feng, Analyse multifractale de la recurrence sur l'espace symbolique, *C. R. Acad. Sci. Paris Sér. I Math.*, **327** (1998), 629–632.

[13] A. Fan and D. Feng, On the distribution of long-term time averages on symbolic space, *J. Statist. Phys.*, **99** (2000), 813–856.

[14] A. Fan, D. Feng and W. Jun, Recurrence, dimension and entropy, *J. London Math. Soc.*, **64** (2001), 229–244.

[15] A. Fan and Y. Jiang, Oscillating sequences, MMA and MMLS flows and Sarnak's conjecture, *Ergodic Theory Dynam. Systems*, **38** (2018), 1709–1744.

[16] A. Fan, L. Liao and J. Peyriere, Generic points in systems of specification and Banach valued Birkhoff ergodic average, *Discrete Contin. Dyn. Syst.*, **21** (2008), 1103–1128.

[17] D. J. Feng, Equilibrium states for factor maps between subshifts, *Adv. Math.*, **226** (2011), 2470–2502.

[18] D. Feng, K. Lau and J. Wu, Ergodic limits on the conformal repellers, *Adv. Math.*, **169** (2002), 58–91.

[19] S. Ferenczi, J. Kułaga-Przymus and M. Lemańczyk, Sarnak's conjecture: What's new, *Ergodic Theory and Dynamical Systems in Their Interactions with Arithmetics and Combinatorics*, **2213** (2018), 163–235.

[20] K. Gelfert and M. Rams, The Lyapunov spectrum of some parabolic systems, *Ergodic Theory Dynam. Systems*, **29** (2009), 919–940.

[21] Y. Heurteaux, Estimations de la dimension inférieure et de la dimension supérieure des mesures, *Ann. Inst. H. Poincaré Probab. Statist.*, **34** (1998), 309–338.

[22] H. B. Keynes and J. B. Robertson, Generators for topological entropy and expansiveness, *Math. Systems Theory*, **3** (1969), 51–59.

[23] F. Ledrappier and P. Walters, A relativised variational principle for continuous transformations, *J. London Math. Soc.*, **16** (1977), 568–576.

[24] E. Olivier, Analyse multifractale de fonctions continues, *C. R. Acad. Sci. Paris Sér. I Math.*, **326** (1998), 1171–1174.

[25] M. Rams, On some non-conformal fractals, *Nonlinearity*, **23** (2010), 2423–2428.

[26] V. A. Rohlin, Lectures on the entropy theory of measure-preserving transformations, *Russian Math. Surveys*, **22** (1967), 1–52.

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

[28] J. Schmeling, On the completeness of multifractal spectra, *Ergodic Theory Dynam. Systems*, **19** (1999), 1595–1616.

[29] F. Takens and E. Verbitskiy, On the variational principle for the topological entropy of certain non-compact sets, *Ergodic Theory Dynam. Systems*, **23** (2003), 317–348.

[30] P. Walters, *An Introduction to Ergodic Theory*, Graduate Texts in Mathematics, 79. Springer-Verlag, New York-Berlin, 1982.

E-mail address: balubsheep@gmail.com
E-mail address: rams@impan.pl
E-mail address: ruxi.shi@upmc.fr