

The Lanford–Ruelle theorem for actions of sofic groups

Sebastián Barbieri and Tom Meyerovitch

Abstract

Let Γ be a sofic group, Σ be a sofic approximation sequence of Γ and X be a Γ -subshift with nonnegative sofic topological entropy with respect to Σ . Further assume that X is a shift of finite type, or more generally, that X satisfies the topological Markov property. We show that for any sufficiently regular potential $f: X \rightarrow \mathbb{R}$, any translation-invariant Borel probability measure on X which maximizes the measure-theoretic sofic pressure of f with respect to Σ , is a Gibbs state with respect to f . This extends a classical theorem of Lanford and Ruelle, as well as previous generalizations of Moulin Ollagnier, Pinchon, Tempelman and others, to the case where the group is sofic.

As applications of our main result we present a criterion for uniqueness of an equilibrium measure, as well as sufficient conditions for having that the equilibrium states do not depend upon the chosen sofic approximation sequence. We also prove that for any group-shift over a sofic group, the Haar measure is the unique measure of maximal sofic entropy for every sofic approximation sequence, as long as the homoclinic group is dense.

On the expository side, we present a short proof of Chung’s variational principle for sofic topological pressure.

Keywords: sofic groups, equilibrium states, Gibbs measures, sofic entropy, sofic pressure.

MSC2020: 37B10, 37D35, 82B20.

1 Introduction

The classical Lanford–Ruelle theorem [26] asserts that given a \mathbb{Z}^d -shift of finite type X and an absolutely-summable, translation-invariant interaction Φ on X , any translation-invariant Borel probability measure which maximizes the *measure-theoretic pressure* with respect to Φ (that is, any *equilibrium state* for the interaction Φ) is a *Gibbs state* with respect to Φ .

Various aspects of the Lanford–Ruelle theorem have been generalized by a number of authors. It has been proven that any measure of maximal entropy for a \mathbb{Z}^d -shift of finite type is invariant with respect to the asymptotic relation [11]. Also, the Lanford–Ruelle theorem has been extended to the context where the acting group is any countable amenable group, not only \mathbb{Z}^d [32, 44], and even more generally, to subshifts which satisfy a condition called the *topological Markov property* and with respect to any random *environment* [7]. For an overview of these generalizations, we refer the reader to the introduction of [7].

A fair amount of work related to Gibbs measures over finite graphs and their limits has been carried out by the communities of mathematical physics and probability. In particular, Gibbs measures on tree-like graphs such as d -regular expanders have been considered in [42, 30]. The emergence of sofic entropy [9] and subsequent introduction of sofic pressure [16] provides a natural framework to study equilibrium measures and their relation with Gibbs states in the context of actions of sofic groups. Let us mention that this class of groups includes all countable amenable and residually finite groups, in particular, it includes all free groups.

In this paper we fully generalize the Lanford–Ruelle theorem to the context where the lattice is an arbitrary countable sofic group Γ and the space of configurations X has hard constraints that generalize

the condition of being of finite type. More precisely, our theorem works for any subshift $X \subset \mathcal{A}^\Gamma$ with nonnegative sofic topological entropy which satisfies the topological Markov property.

Theorem A. Let Σ be a sofic approximation sequence for Γ , X be a subshift that satisfies the topological Markov property such that $h_\Sigma(\Gamma \curvearrowright X) \geq 0$, Φ an absolutely-summable interaction on X and μ an equilibrium measure on X for Φ with respect to Σ . Then μ is Gibbs with respect to Φ .

Some authors have already explored results that are related to the Lanford–Ruelle theorem beyond amenable groups. For instance, Grigorchuk and Stepin [21] provide a sketch of proof for the existence of Gibbs measures on the full Γ -shift on residually finite groups for locally-constant potentials. Many years later Alpeev [2, Theorem 1], showed that for any locally-constant potential on the full Γ -shift (that is, no “hard constraints” or “forbidden patterns”) of a sofic group Γ , there exist Gibbs measures and that furthermore, under certain conditions ensuring uniqueness, the value of a variant of sofic topological pressure does not depend upon the sofic approximation sequence. Recent results of Shriver [41] essentially show that the conclusion of the Lanford–Ruelle theorem for the full Γ -shift holds with respect to finite-range interactions. Strictly speaking, Shriver’s paper considers only a certain class of finitely generated sofic groups, although this restriction is presumably just for simplicity of exposition, where the generalization beyond full-shifts and beyond bounded range interactions does require further arguments. We should mention that Shriver’s paper obtains additional related results related to invariance with respect to the Glauber dynamics, which imply Gibbsianness and are outside the scope of this work. Related objects have also been considered in [4], although the relation to sofic entropy in this work is somewhat implicit.

We deduce Theorem A and a close variant of it (Theorem 5.11) as instances of the following slightly more general result:

Theorem B. Let Σ be a sofic approximation sequence for Γ , X be a subshift such that $h_\Sigma(\Gamma \curvearrowright X) \geq 0$, and let $f \in C(X)$ be a $\rho_{\mathcal{T}^0(X)}$ -limit of locally constant functions. Then any equilibrium measure μ on X for f with respect to Σ is étale Gibbs with respect to f .

We remark that Theorem B does not require any structural assumption on the subshift but provides the conclusion of the measure being “étale Gibbs” instead of just Gibbs. Let us explain how this statement is more general than Theorem A and Theorem 5.11:

1. As in [27], the statement of Theorem B applies meaningfully to shifts even without the topological Markov property. The condition of being a Gibbs measure on X with respect to a suitable map $f: X \rightarrow \mathbb{R}$ can be stated in terms of the Radon-Nikodým cocycle of the measure with respect to the relation $\mathcal{T}(X)$ induced by the asymptotic pairs on X . Following ideas from [27], we define a subrelation $\mathcal{T}^0(X)$ which we call the *étale asymptotic relation* (called the topological Gibbs relation in [27]) and define *étale Gibbs measures* as those which satisfy the Gibbs condition with respect to this subrelation. It turns out that both notions of “Gibbs” coincide precisely when the subshift satisfies the topological Markov property (Proposition 2.11). We show that the conclusion of the Lanford–Ruelle theorem holds for any subshift as long as we replace the condition of being Gibbs by being *étale Gibbs*.
2. The statement of Theorem B applies to a larger class of functions. More precisely, a sequence $(f_n)_{n \in \mathbb{N}}$ of locally constant functions converges in the $\rho_{\mathcal{T}^0(X)}$ -metric to $f \in C(X)$ if it converges uniformly, and furthermore, the cocycles induced by the functions f_n on the asymptotic Gibbs relation $\mathcal{T}^0(X)$ converge to the cocycle induced by f on $\mathcal{T}^0(X)$. We show that functions generated by absolutely-summable interactions (Theorem A) and functions which have summable-variation with respect to a filtration of Γ (Theorem 5.11) satisfy this property.

An interesting observation related to our main result is that, while for non-amenable sofic groups the notion of equilibrium measure depends upon a sofic approximation sequence, Gibbs measures do

not. Therefore, if by some means one can show that a subshift $X \subset \mathcal{A}^\Gamma$ with the topological Markov property admits a unique Gibbs measure μ , then necessarily for every sofic approximation sequence Σ such that $h_\Sigma(\Gamma \curvearrowright X) \geq 0$, we have that μ is translation-invariant and is the unique such measure which maximizes the measure-theoretic sofic entropy $h_\Sigma(\Gamma \curvearrowright X) = h_\Sigma(\Gamma \curvearrowright X, \mu)$. We prove this in Theorem 6.1.

In [16, Question 5.4] it was asked whether for every single-site potential there is a unique equilibrium state on the full-shift over a countable sofic group. This question was answered positively in [10, Example 7] and in [40, Corollary 3.6]. Our results also show that the answer is positive but through the fact that there exists a unique translation-invariant Gibbs measure in that case.

Establishing a property of equilibrium measures that is independent of the sofic approximation sequence is particularly relevant given a recent result of Airey, Bowen and Lin [1] which produces a strongly irreducible subshift of finite type (and thus with the topological Markov property) on a sofic group Γ for which two different sofic approximation sequences Σ_1 and Σ_2 of Γ yield two distinct positive values of sofic topological entropy. If one were able to adapt their construction and produce a system with a unique Gibbs measure, then one would actually be able to show that a single measure can produce two distinct positive values of measure-theoretic sofic entropy with respect to distinct sofic approximation sequences, a problem which is still open [1, Remark 2].

Another application of our result concerns the existence of Gibbs measures. For any sofic group, any subshift of finite type which has nonnegative sofic topological entropy with respect to some sofic approximation sequence admits at least one translation-invariant equilibrium state with respect to any translation-invariant, absolutely-summable interaction (Theorem 6.2), in particular, our result shows that these equilibrium states are Gibbsian. The previous statement generalizes the aforementioned results of Grigorchuk and Stepin [21] and Alpeev [2, Theorem 1]. In fact, in our result the existence of translation-invariant Gibbs states holds for any subshift with the topological Markov property, which is more general than the finite type condition. We remark that the assumption of having finite sofic topological entropy holds trivially in the case where the acting group is amenable, but in the non-amenable case, the assumption cannot be completely removed. In particular, for any non-amenable group there exist subshifts of finite type that admit no invariant measures, in which case the sofic topological entropy is equal to $-\infty$ for any sofic approximation sequence.

A further application of our theorem can be given whenever a subshift carries an additional algebraic structure. More precisely, a *group shift* is a subshift whose alphabet is a finite group and which carries a group structure induced by the alphabet. These algebraic subshifts always admit the Haar measure as a translation-invariant Borel probability measure. We are able to show that for any group shift defined on a sofic group for which its homoclinic group is dense (the group of elements in the subshift which are asymptotic to the identity) then the Haar measure is the unique measure of maximal sofic entropy (Theorem 6.3) with respect to every sofic approximation sequence for Γ .

Acknowledgements: The authors wish to thank Raimundo Briceño for drawing our attention to Shriver's work, Hanfeng Li for very helpful comments on a first version of this article and Andrei Alpeev for suggesting reference [21]. We are also grateful to an anonymous referee for several helpful comments. S. Barbieri was supported by the FONDECYT grant 11200037 and T. Meyerovitch was supported by ISF grant 1058/18.

1.1 Organization of the paper

In Section 2 we provide all the necessary definitions and prove a few elementary results which relate the asymptotic relation with the topological Markov property.

In Section 3 we recall the definition of sofic groups and sofic pressure, and provide proofs of a few elementary structural results. We also provide a short self-contained proof of Chung's variational principle (Theorem 3.8).

The proof of our main theorem in the case where the function is locally constant is given in Section 4.

The heart of the argument is an estimate for a “finite approximation of the pressure function on finite models”, conveyed by Lemma 4.3. To handle a certain technical condition regarding periodicity, we use a trick which involves taking a direct product of the original subshift with a full Γ -shift (Lemma 4.8).

In Section 5 we extend the main result of the previous section to functions which are not necessarily locally constant in Theorem 5.3 (= Theorem B). Our argument uses tools from convex analysis and a characterization of étale Gibbs measures in terms of tangent functionals for the sofic pressure function. We also discuss the space of absolutely-summable interactions (Section 5.1) and the space of maps with absolutely-summable variation (Section 5.2) and provide the proofs for Theorem 5.7 (= Theorem A) and Theorem 5.11. In Section 6 we present in more detail the applications of our main theorem discussed above. In section 7 we briefly discuss hypothetical extensions beyond sofic groups.

2 Preliminaries

2.1 Spin systems and shift spaces

Let \mathcal{A} be a finite set, which we think of as “the allowable spins” or “the alphabet”, and let Γ be a countable group, which will play the both the role of “the lattice” and the “spatial symmetries”. The set $\mathcal{A}^\Gamma = \{x: \Gamma \rightarrow \mathcal{A}\}$ is endowed with the product topology, where \mathcal{A} is considered as a discrete topological space. Elements of \mathcal{A}^Γ are called *configurations*. Given a configuration $x \in \mathcal{A}^\Gamma$ we shall denote its value at $g \in \Gamma$ either by $x(g)$ or x_g .

A *pattern* with support $F \Subset \Gamma$ is an element $p \in \mathcal{A}^F$ (we use the notation $A \Subset B$ to denote that A is a finite subset of B). We denote by $[p] = \{x \in \mathcal{A}^\Gamma : x(g) = p(g) \text{ for every } g \in F\}$ the cylinder generated by a pattern p . A *space of configurations* is a compact subset $X \subset \mathcal{A}^\Gamma$. For a space of configurations X and $F \Subset \Gamma$ the set of X -admissible F -patterns is defined as:

$$P_F(X) := \{x|_F : x \in X\}.$$

We denote the set of all X -admissible patterns by $P(X) := \bigcup_{F \Subset \Gamma} P_F(X)$.

The (left) *shift* action $\Gamma \curvearrowright \mathcal{A}^\Gamma$ is given by

$$gx(h) := x(g^{-1}h) \quad \text{for every } g, h \in \Gamma \text{ and } x \in \mathcal{A}^\Gamma.$$

Then $\Gamma \curvearrowright \mathcal{A}^\Gamma$ is a topological dynamical system, which is known as the *full Γ -shift* over \mathcal{A} .

A subset $X \subset \mathcal{A}^\Gamma$ is a *shift space* or *subshift* if and only if it is Γ -invariant and closed in the product topology. Equivalently, X is a subshift if and only if there exists a (possibly infinite) set of patterns \mathcal{F} such that

$$X = \{x \in \mathcal{A}^\Gamma : gx \notin [p], \text{ for every } g \in \Gamma, p \in \mathcal{F}\}.$$

If X satisfies the equation above, we refer to \mathcal{F} as a set of *forbidden patterns* or *hard constraints* of X . Note that the set \mathcal{F} above is not uniquely determined by X . A *shift of finite type (SFT)* is a shift space X that can be defined by a *finite* set of forbidden patterns.

Given two (possibly infinite) subsets $F, F' \subset \Gamma$, $p \in \mathcal{A}^F$ and $q \in \mathcal{A}^{F'}$ such that $p|_{F \cap F'} = q|_{F \cap F'}$, we define their concatenation $p \vee q$ as the map $w: F \cup F' \rightarrow \mathcal{A}$ whose restrictions to F and F' coincide with p and q respectively.

2.2 The (étale) asymptotic relation

Definition 2.1. We say that two patterns $p, q \in \mathcal{A}^F$ are *interchangeable* in a subshift $X \subset \mathcal{A}^\Gamma$ if for every $x \in \mathcal{A}^\Gamma$ we have

$$x|_{\Gamma \setminus F} \vee p \in X \text{ if and only if } x|_{\Gamma \setminus F} \vee q \in X.$$

Where $x|_{\Gamma \setminus F}$ denotes the restriction of x to $\Gamma \setminus F$. In words, p and q are interchangeable if occurrences of p can be interchanged by q on any configuration without introducing any forbidden pattern and vice-versa.

We denote the collection of interchangeable pairs of patterns with support F in $\mathcal{A}^F \times \mathcal{A}^F$ by $\mathcal{I}_F(X)$ and by $\mathcal{I}(X) = \bigcup_{F \in \Gamma} \mathcal{I}_F(X)$ the collection of interchangeable pairs of patterns.

Definition 2.2. For a finite set $F \in \Gamma$ let

$$\mathcal{T}_F(X) := \{(x, x') \in X \times X : x|_{\Gamma \setminus F} = x'|_{\Gamma \setminus F}\}$$

$$\mathcal{T}(X) := \bigcup_{F \in \Gamma} \mathcal{T}_F(X).$$

Then $\mathcal{T}(X)$ is a Borel equivalence relation on X where each equivalence class is countable, which we call *the asymptotic relation on X* . Some authors, for instance in [8, 27, 34], call $\mathcal{T}(X)$ the Gibbs relation on X . We also define:

$$\mathcal{T}_F^0(X) := \{(x, x') \in \mathcal{T}_F(X) : (x|_F, x'|_F) \in \mathcal{I}_F(X)\}.$$

$$\mathcal{T}^0(X) := \bigcup_{F \in \Gamma} \mathcal{T}_F^0(X).$$

The equivalence relation $\mathcal{T}^0(X)$ can be equivalently described as the orbit equivalence relation of the countable group of homeomorphisms of X generated by the maps $I_{p,q}: X \rightarrow X$ which interchange patterns $(p, q) \in \mathcal{I}_F(X)$ for some $F \in \Gamma$, that is

$$I_{p,q}(x) := \begin{cases} q \vee x|_{\Gamma \setminus F} & \text{if } x|_F = p \\ p \vee x|_{\Gamma \setminus F} & \text{if } x|_F = q \\ x & \text{otherwise.} \end{cases}$$

We refer to $\mathcal{T}^0(X)$ as the *étale asymptotic relation*. We remark that in [27] it was called the “topological Gibbs relation”.

We endow both $\mathcal{T}(X)$ and $\mathcal{T}^0(X)$ with the inductive limit topology that comes from viewing $\mathcal{T}(X)$ and $\mathcal{T}^0(X)$ as the direct limit of the compact topological spaces $(\mathcal{T}_F(X))_{F \in \Gamma}$ and $(\mathcal{T}_F^0(X))_{F \in \Gamma}$ respectively. With that topology the spaces $\mathcal{T}(X)$ and $\mathcal{T}^0(X)$ are metrizable and locally compact, but in general not compact. We remark that this topology on $\mathcal{T}^0(X)$ turns it into an *étale equivalence relation* in the sense of [35]. Notice that a function $\Psi: \mathcal{T}^0(X) \rightarrow \mathbb{R}$ is continuous if and only if for every interchangeable pair of patterns (p, q) the map $x \mapsto \Psi(x, I_{p,q}(x))$ is a continuous function with respect to the product topology on X .

Definition 2.3. Given an equivalence relation $\mathcal{R} \subset X \times X$, an \mathcal{R} -cocycle is a function $\Psi: \mathcal{R} \rightarrow \mathbb{R}$ that satisfies

$$\Psi(x, z) = \Psi(x, y) + \Psi(y, z) \text{ for every } (x, y), (y, z) \in \mathcal{R}.$$

From the definition of \mathcal{R} -cocycle it follows immediately that $\Psi(x, x) = 0$ for every $x \in X$ and that $\Psi(y, x) = -\Psi(x, y)$ for every $(x, y) \in \mathcal{R}$.

Definition 2.4. Given an equivalence relation $\mathcal{R} \subset X \times X$ and a function $f: X \rightarrow \mathbb{R}$ one can formally define for $(x, y) \in \mathcal{R}$

$$\Psi_f(x, y) = \sum_{g \in \Gamma} (f(gy) - f(gx)).$$

If the series above is absolutely convergent for every $(x, y) \in \mathcal{R}$ then Ψ_f is an \mathcal{R} -cocycle and is furthermore Γ -invariant.

A function $f: X \rightarrow \mathbb{R}$ is *locally constant* if there exists $F \in \Gamma$ such that $f(x) = f(y)$ whenever $x|_F = y|_F$. We also say that f is F -locally constant if we want to make F explicit. For locally constant functions there are finitely many non-zero terms in the series defining Ψ_f for the relations $\mathcal{T}(X)$ and $\mathcal{T}^0(X)$, so the series is absolutely convergent. Furthermore, whenever f is locally constant, the function $\Psi_f: \mathcal{T}^0(X) \rightarrow \mathbb{R}$ is continuous.

Recall that $\mathcal{T}^0(X)$ is a locally compact topological space. Thus, the space of continuous functions from $\mathcal{T}^0(X)$ to \mathbb{R} can be equipped with the topology of uniform convergence on compact sets, which induces a topology on the space of continuous $\mathcal{T}^0(X)$ -cocycles. There are many metrics that generate this topology. For future reference we write a specific metric $d_{\mathcal{T}^0(X)}$ that generates the topology:

Definition 2.5. The metric $d_{\mathcal{T}^0(X)}$ on the space of $\mathcal{T}^0(X)$ -cocycles is given by:

$$d_{\mathcal{T}^0(X)}(\Phi, \Psi) = \sum_{n=1}^{\infty} \frac{1}{2^n} \min \left\{ 1, \sup_{(x,y) \in \mathcal{T}_{F_n}^0(X)} |\Phi(x, y) - \Psi(x, y)| \right\},$$

where $(F_n)_{n \geq 1}$ is some fixed enumeration of the non-empty finite subsets of Γ .

This metric generates the topology of uniform convergence on compact sets on the space of continuous $\mathcal{T}^0(X)$ -cocycles. Informally speaking, two cocycles Φ and Ψ are close if there exists a large $N \in \mathbb{N}$ such that $|\Psi(x, y) - \Phi(x, y)|$ is uniformly small over $(x, y) \in \mathcal{T}_{F_n}^0(X)$ for every $n \leq N$.

Definition 2.6. Given $f, g \in C(X)$ such that the $\mathcal{T}^0(X)$ -cocycles Ψ_f, Ψ_g are well defined and continuous, let

$$\rho_{\mathcal{T}^0(X)}(f, g) = \|f - g\|_{\infty} + d_{\mathcal{T}^0(X)}(\Psi_f, \Psi_g).$$

The map $\rho_{\mathcal{T}^0(X)}$ is a metric on the space of continuous functions for which the $\mathcal{T}^0(X)$ -cocycles are well defined and continuous. It says that two functions are close if they are uniformly close, and the cocycles they define are close in the topology of uniform convergence on compact sets on $\mathcal{T}^0(X)$.

2.3 Gibbs measures

In this section we formally define Gibbs measures (also called Gibbs states). In the literature it is possible to find several definitions for Gibbs measures, which are equivalent, at least when the setup is sufficiently restricted. Here we essentially follow the ‘‘conformality’’ framework as in [18, 34]. This definition is equivalent to the Dobrušin-Lanford–Ruelle definition of Gibbs measures as shown in Section 3 of [8]. The conformality framework can also be expressed as an invariance property with respect to certain probability kernels as in [2]. In some situations, for instance for the case of finite-range interactions on shift spaces satisfying the pivot property, one can also characterize Gibbs measures by their invariance under the Glauber dynamics as in [41]. In what follows, all measures considered will be Borel probability measures.

We first recall the notion of the Radon-Nikodým cocycle for a measure which is non-singular with respect to a countable Borel equivalence relation in the sense of [20]. A Borel measure μ on a topological space X is *non-singular* with respect to a countable Borel equivalence relation $\mathcal{R} \subset X \times X$ if whenever $\mu(A) = 0$ for some Borel set $A \subset X$ then

$$\mu \left(\bigcup_{x \in A} \{y \in X : (x, y) \in \mathcal{R}\} \right) = 0.$$

If μ is non-singular with respect to a countable Borel equivalence relation \mathcal{R} , then a *Radon-Nikodým \mathcal{R} -cocycle* for μ is a measurable map $\mathcal{D}_{\mu, \mathcal{R}}: \mathcal{R} \rightarrow \mathbb{R}_+$ such that for every Borel bijection $\phi: X \rightarrow X$

satisfying $(x, \phi(x)) \in \mathcal{R}$ for μ -almost every $x \in X$, we have that

$$\frac{d\mu \circ \phi}{d\mu}(x) = \mathcal{D}_{\mu, \mathcal{R}}(x, \phi(x)) \text{ for } \mu\text{-almost every } x \in X.$$

For any Borel Bijection $\phi: X \rightarrow X$ the Radon-Nikodým derivative $\frac{d\mu \circ \phi}{d\mu}$ is uniquely defined up to a μ -null set. Thus the Radon-Nikodým cocycle is “essentially unique” in the following sense: If $\mathcal{D}_1, \mathcal{D}_2: \mathcal{R} \rightarrow \mathbb{R}_+$ are Radon-Nikodým \mathcal{R} -cocycles of μ then there exists a Borel set $X' \subset X$ such that $\mu(X \setminus X') = 0$ and \mathcal{D}_1 and \mathcal{D}_2 coincide on $(X' \times X) \cap \mathcal{R}$.

In what follows, we will say that two \mathcal{R} -cocycles coincide up to a μ -null set if there exists a Borel set $X' \subset X$ with $\mu(X \setminus X') = 0$ such that the cocycles coincide on $(X' \times X) \cap \mathcal{R}$. For $x \in X$ and $F \Subset X$ we denote

$$[x]_F := [x|_F] = \{y \in X : y|_F = x|_F\}.$$

Recall that the countable Borel equivalence relation $\mathcal{T}^0(X)$ is generated by the homeomorphisms $I_{p,q}$ defined in Section 2.2. See also [27, Lemma 2.2]. As explained in [20] the Radon-Nikodým cocycle of a non-singular measure on a countable Borel equivalence relation is “essentially uniquely determined” by the Radon-Nikodým derivatives on a set of Borel bijections that generate the relation. Using this fact it is easy to obtain an explicit formula for a representative of the Radon-Nikodým cocycle of $\mathcal{T}^0(X)$:

Lemma 2.7. *Let μ be a measure on X which is non-singular with respect to $\mathcal{T}^0(X)$. Then there exists $X' \subset X$ such that $\mu(X \setminus X') = 0$ and for every $(x, y) \in (X' \times X) \cap \mathcal{T}^0(X)$ we have*

$$\mathcal{D}_{\mu, \mathcal{T}^0(X)}(x, y) = \lim_{F \nearrow \Gamma} \frac{\mu([y]_F)}{\mu([x]_F)}.$$

Proof. Let $(x, y) \in \mathcal{T}^0(X)$, then for large enough $F \Subset \Gamma$ the involution $I_{x|_F, y|_F}$ satisfies that

$$\mu([y]_F) = (\mu \circ I_{x|_F, y|_F})([x]_F) = \int_{[x]_F} \mathcal{D}_{\mu, \mathcal{T}^0(X)}(z, I_{x|_F, y|_F}(z)) d\mu(z).$$

Consider the filtration $(\sigma(A^F))_{F \Subset \Gamma}$ where the sets are ordered by inclusion and $\sigma(A^F)$ is the σ -algebra generated by A^F . The minimal σ -algebra generated by this filtration is the Borel σ -algebra of A^Γ . The martingale convergence theorem implies that for any $f \in L^1(A^\Gamma)$ the net of conditional expectations $(\mathbb{E}_\mu(f | \sigma(A^F)))_{F \Subset \Gamma}$ converges to f as $F \nearrow \Gamma$ both in $L^1(A^\Gamma)$ and μ -almost surely. By the Radon-Nikodým theorem we have that the map $z \mapsto \mathcal{D}_{\mu, \mathcal{T}^0(X)}(z, I_{x|_F, y|_F}(z))$ is in $L^1(A^\Gamma)$. Notice that for a fixed $F \Subset \Gamma$ and μ -almost every $x \in A^\Gamma$

$$\mathbb{E}_\mu(\mathcal{D}_{\mu, \mathcal{T}^0(X)}(\cdot, I_{x|_F, y|_F}(\cdot)) | \sigma(A^F))(x) = \frac{1}{\mu([x]_F)} \int_{[x]_F} \mathcal{D}_{\mu, \mathcal{T}^0(X)}(z, I_{x|_F, y|_F}(z)) d\mu(z).$$

Letting F be large enough we get that $\mathcal{D}_{\mu, \mathcal{T}^0(X)}(x, y)$ and $\frac{\mu([y]_F)}{\mu([x]_F)}$ are arbitrarily close outside a set of measure 0. \square

Definition 2.8. Let $f: X \rightarrow \mathbb{R}$ be such that the series defining $\Psi_f(x, y)$ is absolutely convergent for every $(x, y) \in \mathcal{T}(X)$. A Borel probability measure μ on a subshift $X \subset A^\Gamma$ is

1. *Gibbs* with respect to f , if μ is non-singular with respect to $\mathcal{T}(X)$ and its Radon-Nikodým $\mathcal{T}(X)$ -cocycle is equal to $\exp(\Psi_f)$ up to a μ -null set.
2. *Étale Gibbs* with respect to f , if μ is non-singular with respect to $\mathcal{T}^0(X)$ and its Radon-Nikodým $\mathcal{T}^0(X)$ -cocycle is equal to $\exp(\Psi_f)$ up to a μ -null set.

As $\mathcal{T}^0(X)$ is a subrelation of $\mathcal{T}(X)$, it follows that for any potential $f: X \rightarrow \mathbb{R}$ any Gibbs measure with respect to f is automatically étale Gibbs with respect to f . The following example shows that the converse does not hold in every subshift.

Example 2.9. Let

$$X_{\leq 1} = \{x \in \{0, 1\}^\Gamma : |\{g \in \Gamma : x(g) = 1\}| \leq 1\}$$

be the *sunny-side up* subshift. It is clear that for an infinite group Γ , the only Γ -invariant measure on $X_{\leq 1}$ is the delta measure supported on the constant configuration 0^Γ . This measure is singular with respect to $\mathcal{T}(X_{\leq 1})$, hence not Gibbs for any interaction. However, this measure is étale Gibbs (with respect to any f), since the étale asymptotic relation is given by $\mathcal{T}^0(X_{\leq 1}) = \{(0^\Gamma, 0^\Gamma)\} \cup \{(x, y) : x, y \in X_{\leq 1} \setminus \{0^\Gamma\}\}$.

2.4 The topological Markov property

In what follows, we shall present the topological property which characterizes the equality of the asymptotic relation $\mathcal{T}(X)$ and the étale asymptotic relation $\mathcal{T}^0(X)$, and therefore ensures that both definitions of Gibbs given above coincide.

Definition 2.10. We say a subshift $X \subset A^\Gamma$ satisfies the *topological Markov property* (TMP) if for every $A \Subset \Gamma$ there exists $B \Subset \Gamma$ with $A \subset B$ such that whenever x, y satisfy that $x|_{B \setminus A} = y|_{B \setminus A}$ we have that the configuration $z \in A^\Gamma$ is in X , where $z(g)$ is given for every $g \in \Gamma$ by

$$z(g) = \begin{cases} x(g) & \text{if } g \in B \\ y(g) & \text{if } g \in \Gamma \setminus B. \end{cases}$$

Any set B satisfying the above is called a *memory set* of A .

The topological Markov property, as we present it here, was defined in [7]. It was proposed as a generalization of the condition of being a subshift of finite type which was sufficient to prove a generalization of the Lanford–Ruelle theorem for actions of amenable groups. The TMP in turn generalizes the more restrictive notion of Topological Markov field, which was explored by several authors [13, 14, 15] as a generalization of subshifts of finite type in the context of Gibbs theory. The related notion of “splicable metric space”, which can be interpreted in the symbolic setting as a bounded variant of the TMP, was used even earlier by Gromov to provide analogues of Ax’s surjunctivity theorem beyond algebraic varieties (see [22, Section 8.C’], and also [12] for an application of the topological Markov property to surjunctivity of general group actions). More generally, an analogous version of the TMP can be defined for arbitrary group actions on compact metrizable spaces, and it naturally generalizes the well known pseudo-orbit tracing property, also called shadowing. See [5] for further background.

The class of subshifts that satisfy the topological Markov property is much larger than the class of shifts of finite type. For instance, it is known that for a fixed group Γ there are countably many subshifts of finite type up to topological conjugacy, whereas for $\Gamma = \mathbb{Z}^2$ there exist uncountably many non-conjugate subshifts with the topological Markov property [14, page 233]. Moreover, every subshift which has a trivial asymptotic relation satisfies the property [5, Proposition 5.3]. Another interesting family of examples is algebraic: every subshift whose alphabet is a finite group and is closed under the pointwise group operation satisfies the property [7, Proposition 5.1], while there are examples with that structure which are not of finite type if the acting group is solvable but not polycyclic-by-finite [38].

It is clear from the definition that if X satisfies the TMP and B is a memory set of A , then all patterns with support B which coincide on $B \setminus A$ and occur in some configuration of X are pairwise interchangeable. This property in fact characterizes the spaces where the étale asymptotic relation coincides with the asymptotic relation as shown in the next proposition.

Proposition 2.11. *Let $X \subset A^\Gamma$ be a subshift. Then X satisfies the TMP if and only if $\mathcal{T}^0(X) = \mathcal{T}(X)$.*

In particular, if X satisfies the TMP and $f: X \rightarrow \mathbb{R}$ is such that the series defining $\Psi_f(x, y)$ is absolutely convergent for every $(x, y) \in \mathcal{T}(X)$, then every Borel probability measure is Gibbs with respect to f if and only if it is étale Gibbs with respect to f .

Proof. Suppose X satisfies the TMP. Obviously, it is always true that $\mathcal{T}^0(X) \subset \mathcal{T}(X)$. Let $(x, y) \in \mathcal{T}(X)$. Then there exists $F \in \Gamma$ such that $(x, y) \in \mathcal{T}_F(X)$. As X satisfies the TMP, there exists a memory set B for F , and therefore the pair of words $x|_B$ and $y|_B$ are interchangeable, hence $(x, y) \in \mathcal{T}_B^0(X)$.

Conversely, suppose X does not satisfy the TMP and let $(F_n)_{n \in \mathbb{N}}$ be an increasing sequence of finite subsets of Γ such that $F_n \nearrow \Gamma$. Then there is $A \in \Gamma$ such that for every F_n which contains A , there are $x^{(n)}, y^{(n)} \in X$ such that $x^{(n)}|_{F_n \setminus A} = y^{(n)}|_{F_n \setminus A}$ but the patterns $p_n = x^{(n)}|_{F_n}$ and $q_n = y^{(n)}|_{F_n}$ are not interchangeable. Let (x', y') be an accumulation point of the sequence $(x^{(n)}, y^{(n)})_{n \in \mathbb{N}}$. By definition, $(x', y') \in \mathcal{T}_A(X)$ and thus $(x', y') \in \mathcal{T}(X)$.

It follows that for each $K \in \Gamma$ which contains A there is $m \in \mathbb{N}$ such that $p_m|_K = x'|_K$ and $q_m|_K = y'|_K$. Thus $(x', y') \notin \mathcal{T}_K^0(X)$ for any large enough K which implies that $(x', y') \notin \mathcal{T}^0(X)$. This shows that $\mathcal{T}^0(X) \neq \mathcal{T}(X)$. \square

3 Sofic groups and sofic pressure

3.1 Sofic groups

For a finite set V we write $\text{Sym}(V)$ for the group of permutations of V . A group Γ is *sofic* if there exist a sequence $(V_i)_{i \in \mathbb{N}}$ of finite sets such that $|V_i| = n_i$ goes to infinity and a sequence $\Sigma = \{\sigma_i: \Gamma \rightarrow \text{Sym}(V_i)\}_{i=1}^\infty$ that is

$$\begin{aligned} \text{asymptotically an action: } & \lim_{i \rightarrow \infty} \frac{1}{n_i} |\{v \in V_i : \sigma_i(st)v = \sigma_i(s)\sigma_i(t)v\}| = 1 & \text{for every } s, t \in \Gamma \\ \text{asymptotically free: } & \lim_{i \rightarrow \infty} \frac{1}{n_i} |\{v \in V_i : \sigma_i(s)v \neq \sigma_i(t)v\}| = 1 & \text{for every } s \neq t \in \Gamma. \end{aligned}$$

In this case we say Σ is a *sofic approximation sequence* of Γ .

Definition 3.1. Given $\sigma: \Gamma \rightarrow \text{Sym}(V)$ and $F \in \Gamma$, we say that $v \in V$ is *F-good* for σ if $\sigma(st)v = \sigma(s)\sigma(t)v$ for all $s, t \in F$ and $\sigma(s)v \neq \sigma(t)v$ for all $s \neq t \in F$.

With this notation, saying that $\Sigma = \{\sigma_i: \Gamma \rightarrow \text{Sym}(V_i)\}_{i=1}^\infty$ is a sofic approximation means that for any $F \in \Gamma$,

$$\lim_{i \rightarrow \infty} \frac{1}{n_i} |\{v \in V_i : v \text{ is } F\text{-good for } \sigma_i\}| = 1.$$

3.2 Pullback names and empirical distributions

For the remainder of this section, we fix a compact metrizable space X , an action $\Gamma \curvearrowright X$. We denote by $\text{Prob}(X)$ the space of Borel probability measures on X , and by $\text{Prob}_\Gamma(X)$ the space of Γ -invariant Borel probability measures on X . We also consider the space X^Γ with the product topology and the natural left action $\Gamma \curvearrowright X^\Gamma$ given by $(gx)(h) = x(g^{-1}h)$ for every $g, h \in \Gamma$ and $x \in X^\Gamma$.

The following notation closely follows Austin's approach to sofic entropy [3]:

Definition 3.2. Let X be a compact metrizable space, V be a finite set, $v \in V$ and $\sigma: \Gamma \rightarrow \text{Sym}(V)$. We define the following maps

1. $\xi_{\sigma,v}: X^V \rightarrow X^\Gamma$ given by

$$\xi_{\sigma,v}(x)(g) = x(\sigma(g^{-1})v) \text{ for every } g \in \Gamma.$$

2. $\xi_\sigma: X^V \rightarrow \text{Prob}(X^\Gamma)$ given by

$$\xi_\sigma(x) = \frac{1}{|V|} \sum_{v \in V} \delta_{\xi_{\sigma,v}(x)}.$$

3. $\zeta_\sigma: \text{Prob}(X^V) \rightarrow \text{Prob}(X^\Gamma)$ given by

$$\zeta_\sigma(\nu)(A) = \int_{X^V} \xi_\sigma(x)(A) d\nu(x) \text{ for every Borel set } A.$$

In [3] $\xi_{\sigma,v}(x)$ was called “the *pullback name* of x via σ at v ”, and $\xi_\sigma(x)$ was called “the *empirical distribution* of x ”.

In the case where X is a discrete finite set (i.e, an alphabet), for $\nu \in \text{Prob}(X^V)$ we can write $\nu = \sum_{x \in X^V} \lambda_x \delta_x$ and thus we may express the map ζ_σ in the following convenient way

$$\zeta_\sigma(\nu) = \sum_{x \in X^V} \lambda_x \xi_\sigma(x) = \sum_{x \in X^V} \frac{\lambda_x}{|V|} \sum_{v \in V} \delta_{\xi_{\sigma,v}(x)}.$$

Recall that a basis for X^Γ is given by the *cylinder sets* i.e. sets $U \subset X^\Gamma$ of the form

$$U = \{x \in X^\Gamma : x(s) \in U_s \text{ for every } s \in S\},$$

where $S \Subset \Gamma$ and $U_s \subset X$ is open for every $s \in S$.

Proposition 3.3. *Let $\Sigma = \{\sigma_i: \Gamma \rightarrow \text{Sym}(V_i)\}_{i=1}^\infty$ be a sofic approximation sequence of Γ . For any $\varepsilon > 0$, $g \in \Gamma$ and cylinder set $U \subset X^\Gamma$, there is $N \in \mathbb{N}$ such that for any $i \geq N$ and $\nu_i \in \text{Prob}(X^{V_i})$ we have $|(\zeta_{\sigma_i}(\nu_i) - g\zeta_{\sigma_i}(\nu_i))(U)| \leq \varepsilon$. In particular, Any weak-* limit of a sequence of measures $(\zeta_{\sigma_i}(\nu_i))_{i \in \mathbb{N}}$ with $\nu_i \in \text{Prob}(X^{V_i})$ is a Γ -invariant probability measure on X^Γ .*

Proof. Let $\varepsilon > 0$ be arbitrary, $S \Subset \Gamma$ be a symmetric set containing the identity, and let $U = \{x \in X^\Gamma : x(s) \in U_s \text{ for every } s \in S\}$, with $U_s \subset X$ open in X for every $s \in S$. Let $g \in \Gamma$. Then for $\nu_i \in \text{Prob}(X^{V_i})$ we may write

$$\begin{aligned} (\zeta_{\sigma_i}(\nu_i) - g\zeta_{\sigma_i}(\nu_i))(U) &= \int_{X^{V_i}} \frac{1}{|V_i|} \sum_{v \in V_i} \left(\delta_{\xi_{\sigma_i,v}(x)}(U) - \delta_{\xi_{\sigma_i,v}(x)}(g^{-1}U) \right) d\nu_i. \\ &= \int_{X^{V_i}} \frac{1}{|V_i|} |\{v \in V_i : \xi_{\sigma_i,v}(x)(s) \in U_s \text{ for every } s \in S\}| d\nu_i \\ &\quad - \int_{X^{V_i}} \frac{1}{|V_i|} |\{v \in V_i : \xi_{\sigma_i,v}(x)(g^{-1}s) \in U_s \text{ for every } s \in S\}| d\nu_i. \end{aligned}$$

Notice that $\xi_{\sigma_i,v}(x)(s) = x(\sigma_i(s^{-1})v)$ and that $\xi_{\sigma_i,v}(x)(g^{-1}s) = x(\sigma_i(s^{-1}g)v)$. As Σ is asymptotically an action, for every large enough $i \in \mathbb{N}$ there exists $V_i'' \subset V_i$ which is $(S \cup gS \cup g^{-1}S)$ -good for σ_i and such that $|V_i''| \geq (1 - \frac{\varepsilon}{4})|V_i|$. Let $V_i' = V_i'' \cap \sigma_i(g^{-1})V_i''$. Then $|V_i'| \geq (1 - \frac{\varepsilon}{2})|V_i|$. We get that that $\xi_{\sigma_i,v}(x) \in U$ if and only if $\xi_{\sigma_i,\sigma_i(g^{-1})v}(x) \in g^{-1}U$ for every $v \in V_i'$. We can thus eliminate vertices in V_i' in the equation above and obtain the following estimate:

$$|(\zeta_{\sigma_i}(\nu_i) - g\zeta_{\sigma_i}(\nu_i))(U)| \leq \int_{X^{V_i}} \frac{2|V_i \setminus V_i'|}{|V_i|} d\nu_i \leq \varepsilon.$$

As the cylinders sets are a basis for the product topology on X^Γ the result follows. \square

When Γ is countable, both the product topology of X^Γ and the space $\text{Prob}(X^\Gamma)$ are metrizable. Let us denote a compatible metric on X^Γ by d_T and a compatible metric on $\text{Prob}(X^\Gamma)$ by d_P (for instance, the Kantorovich metric). The previous lemma can then be restated in the following way: for every $\varepsilon > 0$ and $g \in \Gamma$ there is $N \in \mathbb{N}$ such that for every $i \geq N$ we have $d_P(\zeta_{\sigma_i}(\nu_i), g\zeta_{\sigma_i}(\nu_i)) \leq \varepsilon$ for any measure ν_i on X^{V_i} .

Given a closed Γ -invariant set $Y \subset X^\Gamma$, $\mu \in \text{Prob}(X^\Gamma)$ and $\delta > 0$, we define:

$$N_\delta(Y) := \{\nu \in \text{Prob}(X^\Gamma) : \nu(\{x \in X^\Gamma : \min_{y \in Y} d_T(x, y) \geq \delta\}) \leq \delta\},$$

$$N_\delta(\mu) = \{\nu \in \text{Prob}(X^\Gamma) : d_P(\mu, \nu) \leq \delta\}.$$

Lemma 3.4. *Let Γ be a countable group, $\Sigma = \{\sigma_i : \Gamma \rightarrow \text{Sym}(V_i)\}_{i=1}^\infty$ be a sofic approximation sequence of Γ and $Y \subset X^\Gamma$ a closed Γ -invariant set. For any $\delta > 0$ there exists $\delta' > 0$ and $N \in \mathbb{N}$ such that for any $i \geq N$,*

$$\xi_{\sigma_i}^{-1}(N_{\delta'}(Y)) \subset \bigcup_{\mu \in \text{Prob}_\Gamma(Y)} \xi_{\sigma_i}^{-1}(N_\delta(\mu)).$$

Furthermore, if $\text{Prob}_\Gamma(Y) \neq \emptyset$, then for any $\delta > 0$ there exists $\delta' > 0$, $N \in \mathbb{N}$ and $\mu_1, \dots, \mu_k \in \text{Prob}_\Gamma(Y)$ such that for any $i \geq N$,

$$\xi_{\sigma_i}^{-1}(N_{\delta'}(Y)) \subset \bigcup_{j=1}^k \xi_{\sigma_i}^{-1}(N_\delta(\mu_j)).$$

Proof. By Proposition 3.3, for any $\varepsilon > 0$ and $g \in \Gamma$ there exists $N \in \mathbb{N}$ such that for every $i \geq N$ and $x \in X^{V_i}$, $d_P(\zeta_{\sigma_i}(\delta_x), g\zeta_{\sigma_i}(\delta_x)) = d_P(\xi_{\sigma_i}(x), g\xi_{\sigma_i}(x)) \leq \varepsilon$. For $\delta > 0$ and $g \in \Gamma$ let

$$P(\delta, g) := \{\nu \in \text{Prob}(X^\Gamma) : d_P(\nu, g\nu) \leq \delta\}.$$

It follows that for any $\delta' > 0$ and $F \subseteq \Gamma$ there exists $N \in \mathbb{N}$ such that for any $i \geq N$

$$\xi_{\sigma_i}(A^{V_i}) \subset \bigcap_{g \in F} P(\delta', g).$$

Now, it is clear from the definition that we have

$$\text{Prob}_\Gamma(Y) = \bigcap_{\delta' > 0, F \subseteq \Gamma} \left(N_{\delta'}(Y) \cap \bigcap_{g \in F} P(\delta', g) \right).$$

Since $\text{Prob}_\Gamma(Y) \subset \bigcup_{\mu \in \text{Prob}_\Gamma(Y)} N_\delta(\mu)$ it follows by compactness that for every $\delta > 0$ there exists $\delta' > 0$ and $F \subseteq \Gamma$ such that

$$N_{\delta'}(Y) \cap \bigcap_{g \in F} P(\delta', g) \subset \bigcup_{\mu \in \text{Prob}_\Gamma(Y)} N_\delta(\mu).$$

As $\xi_{\sigma_i}^{-1}(P(\delta', g)) = A^{V_i}$ for every $g \in F$ and $i \geq N$, it follows that

$$\xi_{\sigma_i}^{-1}(N_{\delta'}(Y)) \subset \xi_{\sigma_i}^{-1} \left(\bigcup_{\mu \in \text{Prob}_\Gamma(Y)} N_\delta(\mu) \right) = \bigcup_{\mu \in \text{Prob}_\Gamma(Y)} \xi_{\sigma_i}^{-1}(N_\delta(\mu)).$$

Since $N_{\delta'}(Y) \cap \bigcap_{g \in F} P(\delta', g)$ is compact, and $N_\delta(\mu)$ is open for every $\mu \in \text{Prob}(X)$, it follows that if $\text{Prob}_\Gamma(Y)$ is nonempty, then there exists $\mu_1, \dots, \mu_k \in \text{Prob}_\Gamma(Y)$ such that

$$N_{\delta'}(Y) \cap \bigcap_{g \in F} P(\delta', g) \subset \bigcup_{j=1}^k N_\delta(\mu_j).$$

Therefore we may conclude that

$$\xi_{\sigma_i}^{-1}(N_{\delta'}(Y)) \subset \bigcup_{j=1}^k \xi_{\sigma_i}^{-1}(N_{\delta}(\mu_j)).$$

□

3.3 Topological and measure-theoretic pressure of actions of sofic groups

The notions of topological and measure-theoretic pressure for \mathbb{Z}^d -actions were introduced and studied by Ruelle [36], who also proved a variational principle extending the variational principle for entropy. Both notions and the variational principle were latter extended to actions of countable amenable groups by Ollagnier and Pinchon [33, 31], Stepin and Tagi-Zade [43] and Tempelman [44]. A further generalization of the notions of measure-theoretic and topological pressure for actions of sofic groups on compact metrizable spaces was introduced by Nhan-Phu Chung in [16], who also proved the corresponding variational principle in this setting (building on the landmark papers of Bowen [9] and Kerr and Li [24]). A similar notion is also introduced by Alpeev in [2] for the space A^Γ , albeit with different objectives.

In what follows we will recall Chung's notions of pressure for actions of sofic groups, and then restate them in an equivalent manner using Austin's framework of "empirical measures", which we introduced above. This formalism will turn out to be convenient in our symbolic setting.

Let Γ be a sofic group acting on a compact metrizable space X by homeomorphisms, $\Sigma = \{\sigma_i: \Gamma \rightarrow \text{Sym}(V_i)\}$ a sofic approximation sequence for Γ , ρ a continuous pseudometric on X and $f: X \rightarrow \mathbb{R}$ a continuous map.

Let $i \in \mathbb{N}$, $F \subseteq \Gamma$ and $\delta > 0$. The set $\text{Map}(\rho, F, \delta, \sigma_i)$ consists of all maps $\varphi: V_i \rightarrow X$ such that

$$\max_{s \in F} \left(\frac{1}{|V_i|} \sum_{v \in V_i} (\rho(s\varphi(v), \varphi(\sigma_s(v))))^2 \right)^{1/2} \leq \delta.$$

That is, the maps φ which roughly look like the restriction of an orbit to F for most vertices $v \in V$. For $\varepsilon > 0$, we say $E \subset \text{Map}(\rho, F, \delta, \sigma_i)$ is ε -separated if for every $\varphi_1, \varphi_2 \in E$ we have

$$\max_{v \in V_i} \rho(\varphi_1(v), \varphi_2(v)) \geq \varepsilon.$$

Let $M_{\Sigma, \infty}^\varepsilon(f, X, \Gamma, \rho, F, \delta, \sigma_i)$ denote the supremum over all ε -separated subsets E of $\text{Map}(\rho, F, \delta, \sigma_i)$ of the expression

$$\sum_{\varphi \in E} \exp \left(\sum_{v \in V_i} f(\varphi(v)) \right).$$

Finally, let

$$P_{\Sigma, \infty}^\varepsilon(f, X, \Gamma, \rho, F, \delta, \sigma_i) = \frac{1}{|V_i|} \log (M_{\Sigma, \infty}^\varepsilon(f, X, \Gamma, \rho, F, \delta, \sigma_i))$$

Definition 3.5. The *topological sofic pressure* of $\Gamma \curvearrowright X$ with respect to $f: X \rightarrow \mathbb{R}$ and the sofic approximation sequence Σ is given by

$$P_\Sigma(\Gamma \curvearrowright X, f) = \sup_{\varepsilon > 0} \inf_{F \subseteq \Gamma} \inf_{\delta > 0} \limsup_{i \rightarrow \infty} P_{\Sigma, \infty}^\varepsilon(f, X, \Gamma, \rho, F, \delta, \sigma_i).$$

The value of the sofic pressure function at the zero function $f = 0$ is called the *topological sofic entropy* of $\Gamma \curvearrowright X$ with respect to the sofic approximation sequence Σ . We denote the topological sofic entropy by

$$h_\Sigma(\Gamma \curvearrowright X) = P_\Sigma(\Gamma \curvearrowright X, 0).$$

The topological sofic entropy was defined by Kerr and Li [24] before the the definition of sofic pressure, which can be considered as a generalization of their notion.

If μ is a Γ -equivariant measure on X and L is a finite subset of $C(X)$, we define $\text{Map}_\mu(\rho, F, \delta, \sigma_i, L)$ as the set of $\varphi \in \text{Map}(\rho, F, \delta, \sigma_i)$ which satisfy

$$\left| \frac{1}{|V_i|} \sum_{v \in V_i} h(\varphi(v)) - \int_X h d\mu \right| \leq \delta, \text{ for every } h \in L.$$

Similarly, we let $M_{\Sigma, \infty, \mu}^\varepsilon(f, X, \Gamma, \rho, F, \delta, \sigma_i, L)$ denote the supremum over all ε -separated subsets of $\text{Map}_\mu(\rho, F, \delta, \sigma_i, L)$ of the same expression as above, and let

$$P_{\Sigma, \infty, \mu}^\varepsilon(f, X, \Gamma, \rho, F, \delta, L, \sigma_i) = \frac{1}{|V_i|} \log \left(M_{\Sigma, \infty, \mu}^\varepsilon(f, X, \Gamma, \rho, F, \delta, L, \sigma_i) \right).$$

Definition 3.6. The *measure-theoretic sofic pressure* of $\Gamma \curvearrowright (X, \mu)$ with respect to $f: X \rightarrow \mathbb{R}$ and the sofic approximation sequence Σ is given by

$$P_\Sigma(f, \Gamma \curvearrowright (X, \mu)) = \sup_{\varepsilon > 0} \inf_{F \in \Gamma} \inf_{\delta > 0} \inf_{L \in C(X)} \limsup_{i \rightarrow \infty} P_{\Sigma, \infty, \mu}^\varepsilon(f, X, \Gamma, \rho, F, \delta, L, \sigma_i).$$

The *measure-theoretic sofic entropy* of $\Gamma \curvearrowright (X, \mu)$ with respect to the sofic approximation sequence Σ is the measure-theoretic sofic pressure of $\Gamma \curvearrowright (X, \mu)$ at the zero function $f = 0$. We denote the measure-theoretic sofic entropy by

$$h_\Sigma(\Gamma \curvearrowright X, \mu) = P_\Sigma(\Gamma \curvearrowright X, \mu, 0).$$

A straightforward approximation argument shows that

$$P_\Sigma(\Gamma \curvearrowright X, \mu, f) = h_\Sigma(\Gamma \curvearrowright X, \mu) + \int_X f d\mu.$$

It is consistent with the definition to declare that for $\mu \in \text{Prob}(X)$ which is not Γ -invariant,

$$P_\Sigma(\Gamma \curvearrowright X, \mu, f) = -\infty.$$

It can sometimes be semantically and conceptually convenient to replace the parameters $F \in \Gamma$, $\delta > 0$ and $L \in C(X)$ that appear in the above definition of topological and measure-theoretic sofic pressure by a single parameter. For $\sigma: \Gamma \rightarrow \text{Sym}(V)$, $f: X \rightarrow \mathbb{R}$, $\varepsilon > 0$ and U an open set of probability measures on X^Γ we define

$$P(U, \sigma, \varepsilon, f) = \frac{1}{|V|} \log \left(\sup_{E \in \mathcal{E}(U, \sigma, \varepsilon)} \sum_{\varphi \in E} \exp \left(\sum_{v \in V} f(\varphi(v)) \right) \right),$$

where $\mathcal{E}(U, \sigma, \varepsilon)$ is the set of all collections E of maps $\varphi \in X^{V_i}$ which satisfy that $\xi_\sigma(\varphi) \in U$ and which are ε -separated.

Let $X^* \subset X^\Gamma$ be the space of Γ -orbits, that is

$$X^* = \{y \in X^\Gamma : y_{g^{-1}} = (gy)_{1_\Gamma} \text{ for every } g \in \Gamma\}.$$

Let μ^* be the pushforward measure on X^* induced by a measure μ on X with respect to the map that assigns to each $x \in X$ its orbit under Γ . It is known that the definitions of sofic pressure (both topological and measure-theoretic) do not depend upon the choice of continuous pseudometric, as long

as it is dynamically generating, see [16, Lemma 2.7], from this fact, a computation shows that we may rephrase the definitions above succinctly as

$$P_\Sigma(\Gamma \curvearrowright X, f) = \sup_{\varepsilon > 0} P_\Sigma(\varepsilon, \Gamma \curvearrowright X, f),$$

and

$$P_\Sigma(f, \Gamma \curvearrowright X, \mu, f) = \sup_{\varepsilon > 0} P_\Sigma(\varepsilon, \Gamma \curvearrowright X, \mu, f).$$

Where:

$$P_\Sigma(\varepsilon, \Gamma \curvearrowright X, f) := \inf_{\delta > 0} \limsup_{i \rightarrow \infty} P(\mathcal{N}_\delta(X^*), \sigma_i, \varepsilon, f),$$

and

$$P_\Sigma(\varepsilon, \Gamma \curvearrowright X, \mu, f) := \inf_{\delta > 0} \limsup_{i \rightarrow \infty} P(\mathcal{N}_\delta(\mu^*), \sigma_i, \varepsilon, f).$$

The above equivalent expressions have a number of advantages which will become evident in the symbolic setting. We state following simple observation (essentially equivalent to Proposition 2.2 of [17]):

Proposition 3.7. *For any $\varepsilon > 0$ the function $H_\varepsilon: \text{Prob}_\Gamma(X) \rightarrow \{-\infty\} \cup [0, +\infty)$ given by*

$$H_\varepsilon(\mu) = h_\Sigma(\varepsilon, \Gamma \curvearrowright X, \mu)$$

is upper semi-continuous.

Proof. Fix $\varepsilon > 0$ and $\mu \in \text{Prob}_\Gamma(X)$. Choose any $y > H_\varepsilon(\mu)$, then there exists $\delta > 0$ such that $\limsup_{i \rightarrow \infty} P(\mathcal{N}_\delta(\mu^*), \sigma_i, \varepsilon, 0) < y$. It follows that for any $\nu \in \mathcal{N}_\delta(\mu^*)$, $H_\varepsilon(\nu) < y$. \square

With all the notation above in place, we can provide short proof of the variational principle for sofic pressure. It seems that once we have the definitions in place, the proof below is as short as Misiurewicz's proof of the variational principle for entropy of \mathbb{Z}^d -actions [29], and arguably conceptually simpler.

Theorem 3.8 (N.P. Chung's variational principle for sofic pressure [16]). *Let $\Gamma \curvearrowright X$ be an action of a sofic group Γ on a compact space X and let Σ be a sofic approximation sequence for Γ . For every $\varepsilon > 0$ and $f \in C(X)$ we have:*

$$P_\Sigma(\varepsilon, \Gamma \curvearrowright X, f) = \sup_{\mu \in \text{Prob}_\Gamma(X)} P_\Sigma(\varepsilon, f, \Gamma \curvearrowright X, \mu, f).$$

In particular,

$$P_\Sigma(\Gamma \curvearrowright X, f) = \sup_{\mu \in \text{Prob}_\Gamma(X)} P_\Sigma(\Gamma \curvearrowright X, \mu, f).$$

Proof. Suppose that $\text{Prob}_\Gamma(X) \neq \emptyset$ and fix $\mu \in \text{Prob}_\Gamma(X)$. Let $\delta > 0$ and choose $\delta' > 0$ such that whenever $d_P(\mu^*, \nu) < \delta'$ then for $h \in C(X^\Gamma)$ given by $h(x) = d_T(x, X^*)$ we have

$$\left| \int_{X^\Gamma} h d\mu^* - \int_{X^\Gamma} h d\nu \right| < \delta^2.$$

It follows that

$$\delta > \frac{1}{\delta} \left| \int_{X^\Gamma} h d\mu^* - \int_{X^\Gamma} h d\nu \right| = \frac{1}{\delta} \left| \int_{X^\Gamma} h d\nu \right| \geq \nu(\{x \in X^\Gamma : d_T(x, X^*) \geq \delta\}).$$

Thus we conclude that for each $\delta > 0$, there is $\delta' > 0$ such that $N_{\delta'}(\mu^*) \subset N_\delta(X^*)$. It follows that for any $i \in \mathbb{N}$, $f \in C(X)$ and $\varepsilon > 0$

$$P(N_{\delta'}(\mu^*), \sigma_i, \varepsilon, f) \leq P(N_\delta(X^*), \sigma_i, \varepsilon, f).$$

This immediately yields the ‘‘easy direction’’ of the variational principle:

$$\sup_{\mu \in \text{Prob}_\Gamma(X)} P_\Sigma(\varepsilon, \Gamma \curvearrowright X, \mu, f) \leq P_\Sigma(\varepsilon, \Gamma \curvearrowright X, f).$$

Notice that the above inequality is trivial in the case where $\text{Prob}_\Gamma(X) = \emptyset$.

In order to show the converse we apply Lemma 3.4. The case where $\text{Prob}_\Gamma(X) = \emptyset$ yields that $\xi_{\sigma_i}^{-1}(N_{\delta'}(X^*))$ is empty for all large enough i , and thus $P_\Sigma(f, \Gamma \curvearrowright X) = -\infty$ and the variational principle holds. Suppose that $\text{Prob}_\Gamma(X) \neq \emptyset$, then Lemma 3.4 yields that for any $\delta > 0$ there exists $\delta' > 0$, $N \in \mathbb{N}$ and $\mu_1, \dots, \mu_k \in \text{Prob}_\Gamma(X^*)$ such that for any $i > N$,

$$\xi_{\sigma_i}^{-1}(N_{\delta'}(X^*)) \subset \bigcup_{j=1}^k \xi_{\sigma_i}^{-1}(N_\delta(\mu_j)).$$

It follows that for all $i > N$, $f \in C(X)$ and $\varepsilon > 0$,

$$P(N_{\delta'}(X), \sigma_i, \varepsilon, f) \leq \frac{\log(k)}{|V_i|} + \max_{1 \leq j \leq k} P(N_\delta(\mu_j), \sigma_i, \varepsilon, f).$$

Let $(\delta_n)_{n=1}^\infty$ be a decreasing sequence of positive numbers tending to 0. From the above equation we conclude that there exists another decreasing sequence of positive numbers $(\delta'_n)_{n=1}^\infty$ and a sequence of measures $\mu_n \in \text{Prob}_\Gamma(X^*)$ such that for every $n \in \mathbb{N}$

$$\limsup_{i \rightarrow \infty} P(N_{\delta'_n}(X), \sigma_i, \varepsilon, f) \leq \limsup_{i \rightarrow \infty} P(N_{\delta_n}(\mu_n), \sigma_i, \varepsilon, f).$$

Let $\mu \in \text{Prob}_\Gamma(X^*)$ be a weak-* limit of $(\mu_n)_{n=1}^\infty$. Then for any $\delta > 0$ there exists N such that for all $n > N$ we have $N_{\delta_n}(\mu_n) \subset N_\delta(\mu)$. It follows that for any $\delta > 0$ we have

$$\limsup_{i \rightarrow \infty} P(N_{\delta'_n}(X), \sigma_i, \varepsilon, f) \leq \limsup_{i \rightarrow \infty} P(N_\delta(\mu), \sigma_i, \varepsilon, f),$$

for all sufficiently large n . Taking $n \rightarrow \infty$ and then infimum over $\delta > 0$, we conclude that

$$P_\Sigma(\varepsilon, \Gamma \curvearrowright X, f) \leq P_\Sigma(\varepsilon, \Gamma \curvearrowright X, \mu, f).$$

In particular, we obtain the variational principle for sofic pressure. □

Definition 3.9. We say that $\mu \in \text{Prob}_\Gamma(X)$ is an *equilibrium measure* (also called *equilibrium state*) for $f \in C(X)$ with respect to Σ if

$$P_\Sigma(\Gamma \curvearrowright X, f) = P_\Sigma(\Gamma \curvearrowright X, \mu, f).$$

That is, if it achieves the supremum in the variational principle. In the case $f = 0$ we call μ a *measure of maximal entropy* with respect to Σ .

Let $(W, \|\cdot\|)$ be a Banach space and denote by W^* its continuous dual space. Given a convex function $F: W \rightarrow \mathbb{R}$, we say that a linear functional $\psi \in W^*$ is a *tangent functional* (or *subgradient*) of F at $w \in W$ if for every $u \in W$ we have

$$F(u) - F(w) \geq \psi(u - w).$$

Given a sofic approximation sequence for Γ , and a Γ -action $\Gamma \curvearrowright X$, we can consider the measure-theoretic entropy and the topological pressure as functions on $\text{Prob}(X)$ and $C(X)$:

$$H := h_\Sigma(\Gamma \curvearrowright X, \cdot): \text{Prob}(X) \rightarrow \{-\infty\} \cup [0, +\infty]$$

and

$$\Pi := P_\Sigma(\Gamma \curvearrowright X, \cdot): C(X) \rightarrow [-\infty, +\infty]$$

The variational principle is equivalent to the statement that the topological pressure function Π is precisely equal to the Legendre transform (or Legendre–Fenchel transform) of $-H$.

The following properties of the topological pressure are either classical (at least in the amenable case) or trivial, some of them can be found in [16, Proposition 6.1]. The straightforward verification of these properties extends verbatim to the case of actions of sofic groups. We provide short self-contained proofs for completeness.

Proposition 3.10. *Let Σ be a sofic approximation sequence for Γ , and let $\Gamma \curvearrowright X$ be a Γ -action such that $h_\Sigma(\Gamma \curvearrowright X)$ is finite. Then the sofic topological pressure of $\Gamma \curvearrowright X$ with respect to Σ is finite for every $f \in C(X)$, and the function $\Pi: C(X) \rightarrow \mathbb{R}$ given by $\Pi(f) = P_\Sigma(f, \Gamma \curvearrowright X)$ satisfies the following properties:*

1. Π is monotonically non-decreasing: $\Pi(f) \geq \Pi(g)$ whenever $f - g \geq 0$.
2. For every $c \in \mathbb{R}$ and $f \in C(X)$, $\Pi(f + c) = \Pi(f) + c$.
3. Π is 1-Lipschitz with respect to the $\|\cdot\|_\infty$ -norm on $C(X)$. In particular, it is continuous.
4. Π is well defined on Γ -cohomology classes. Namely, for every $f, f_0 \in C(X)$ and $g \in \Gamma$,

$$\Pi(f_0 + f - f \circ g) = \Pi(f_0).$$

In particular, Π is Γ -invariant in the sense that $\Pi(f) = \Pi(f \circ g)$ for every $f \in C(X)$ and $g \in \Gamma$.

5. Π is a convex function.
6. If $\mu \in \text{Prob}_\Gamma(X)$ is an equilibrium state for $f_0 \in C(X)$ with respect to Σ , then μ is a tangent functional of Π at f_0 .
7. Conversely, a tangent functional of Π at any $f \in C(X)$, is unital, positive and Γ -invariant. If the sofic measure-theoretic entropy is an upper semi-continuous and concave function on $\text{Prob}(X)$, any tangent functional of Π at any $f \in C(X)$ is furthermore an equilibrium state for $f_0 \in C(X)$ with respect to Σ .

Proof. If $h_\Sigma(\Gamma \curvearrowright X)$ is finite, from the variational principle we have

$$|P_\Sigma(f, \Gamma \curvearrowright X)| \leq h_\Sigma(\Gamma \curvearrowright X) + \sup_{\mu \in \text{Prob}_\Gamma(X)} \left| \int f d\mu \right| \leq h_\Sigma(\Gamma \curvearrowright X) + \|f\|_\infty.$$

From this inequality it follows that $P_\Sigma(f, \Gamma \curvearrowright X)$ is finite for every $f \in C(X)$. Let us remark that in the proof of properties 1 – 3 below we do not need to assume that $P_\Sigma(f, \Gamma \curvearrowright X) \in \mathbb{R}$. In fact, the finiteness of $P_\Sigma(f, \Gamma \curvearrowright X)$ also follows from the assumption that $h_\Sigma(\Gamma \curvearrowright X) \in \mathbb{R}$ together with property 3.

1. Monotonicity of Π follows directly from monotonicity of $f \mapsto P(U, \sigma, \varepsilon, f)$.

2. This follows directly from the fact that for any $E \subset X^{V_i}$, $f \in C(X)$ and $c \in \mathbb{R}$ we have

$$\sum_{\varphi \in E} \exp \left(\sum_{v \in V_i} f(\varphi(v)) + c \right) = e^{|V_i|c} \sum_{\varphi \in E} \exp \left(\sum_{v \in V_i} f(\varphi(v)) \right),$$

thus for any open set U of probability measures on X^Γ ,

$$P(U, \sigma_i, \varepsilon, f + c) = P(U, \sigma_i, \varepsilon, f) + c.$$

The result follows by taking $i \rightarrow \infty$, $\delta \rightarrow 0$ and $\varepsilon \rightarrow 0$.

3. By the two previous properties, for $f_1, f_2 \in C(X)$

$$\Pi(f_1) - \Pi(f_2) \leq \Pi(f_2 + \|f_1 - f_2\|_\infty) - \Pi(f_2) = \|f_1 - f_2\|_\infty.$$

Interchanging the roles of f_1 and f_2 we conclude that $|\Pi(f_1) - \Pi(f_2)| \leq \|f_1 - f_2\|_\infty$.

4. Suppose $f, f_0, \in C(X)$ and $g \in \Gamma$. Since $\int f_0 d\mu = \int (f_0 + f - f \circ g) d\mu$ for any $\mu \in \text{Prob}_\Gamma(X)$, it follows that

$$P_\Sigma(\Gamma \curvearrowright X, \mu, f_0) = P_\Sigma(\Gamma \curvearrowright X, \mu, f_0 + f - f \circ g)$$

So by the variational principle

$$\Pi(f_0 + f - f \circ g) = \Pi(f_0).$$

5. Convexity of Π follows easily from Hölder's inequality (see [16, Proposition 6.1]). Alternatively it follows directly from the variational principle, which shows the value of Π is a pointwise supremum of affine functions.

6. Suppose that $\mu \in \text{Prob}_\Gamma(X)$ is an equilibrium state for f with respect to Σ . This means that

$$\Pi(f_0) = h_\Sigma(\Gamma \curvearrowright X, \mu) + \int_X f_0 d\mu.$$

On the other hand, by the variational principle we have that for any $f \in C(X)$

$$\Pi(f) \geq h_\Sigma(\Gamma \curvearrowright X, \mu) + \int_X f d\mu.$$

It follows that for any $f \in C(X)$

$$\Pi(f) - \Pi(f_0) \geq \int_X f - f_0 d\mu.$$

This shows that μ is a tangent functional of Π at f_0 .

7. Suppose that $\eta \in C(X)^*$ is a tangent functional of Π at f_0 . This means that for every $f \in C(X)$

$$\Pi(f) - \Pi(f_0) \geq \eta(f - f_0).$$

In particular, set $f = f_0 \pm 1$ and apply the property 2 to conclude that $\eta(1) = 1$. Because Π is monotone it follows that for any nonnegative $g \in C(X)$

$$\Pi(f_0 - g) - \Pi(f_0) \leq 0,$$

So $\eta(-g) \leq 0$ so $\eta(g) \geq 0$. This shows η corresponds to a positive measure. To see that η is Γ -invariant, use the fact that $\Pi(f_0 + f - f \circ g) = \Pi(f_0)$.

Now further assume that $H = h_\Sigma(\Gamma \curvearrowright X, \cdot): \text{Prob}(X) \rightarrow \mathbb{R}$ is upper semi-continuous and concave. By hypothesis we have that $h_\Sigma(\Gamma \curvearrowright X)$ is finite and thus $-H$ is proper. It follows by the Fenchel–Moreau theorem that $-H$ is equal on $\text{Prob}(X)$ to the Legendre transform of its Legendre transform, which is the sofic topological pressure function Π . This means that

$$h_\Sigma(\Gamma \curvearrowright X, \eta) = \inf_{f \in C(X)} P_\Sigma(\Gamma \curvearrowright X, f) - \eta(f).$$

In particular,

$$h_\Sigma(\Gamma \curvearrowright X, \eta) \geq P_\Sigma(\Gamma \curvearrowright X, f_0) - \eta(f_0).$$

So

$$P_\Sigma(\Gamma \curvearrowright X, \eta, f_0) \geq P_\Sigma(\Gamma \curvearrowright X, f_0).$$

By the variational principle the above inequality is in fact an equality and thus η is an equilibrium measure for f_0 with respect to Σ .

□

By the previous proposition, whenever the measure-theoretic sofic entropy is an upper semi-continuous and concave function of the measure, equilibrium measures are precisely the tangent functionals of the convex continuous function Π . Let us discuss some sufficient conditions for these hypotheses to hold.

With regards to concavity of measure-theoretic sofic entropy: It is well known that when Γ is amenable, the measure-theoretic entropy H , when restricted to the invariant measures $\text{Prob}_\Gamma(X)$, is an affine function. Recall that $H(\mu) = -\infty$ for $\mu \notin \text{Prob}_\Gamma(X)$, so for any action of an amenable group Γ , we have that $-H = -h_\Sigma(\Gamma \curvearrowright X, \cdot)$ is convex. When Γ is non-amenable, it can happen that $\mu_1, \mu_2 \in \text{Prob}_\Gamma(X)$ satisfy $h_\Sigma(\Gamma \curvearrowright X, \mu_i) \geq 0$ for $i = 1, 2$ but $h_\Sigma(\Gamma \curvearrowright X, \frac{1}{2}\mu_1 + \frac{1}{2}\mu_2) = -\infty$, see for instance Proposition 3.1 of [10].

With regards to upper semi-continuity of the measure-theoretic sofic entropy, we now recall that *expansiveness* is a sufficient condition. Recall that an action $\Gamma \curvearrowright X$ is *expansive* if there exists $\varepsilon > 0$ such that for any $x_1 \neq x_2$ in X there exists $g \in \Gamma$ such that the distance between gx_1 and gx_2 is at least ε . A positive constant $\varepsilon > 0$ as above is called an *expansive constant* for $\Gamma \curvearrowright X$. Whether a particular $\varepsilon > 0$ is an expansive constant depends on the choice of metric, but the existence of an expansive constant does not. If $\Gamma \curvearrowright X$ is expansive then with respect to any sofic approximation the measure-theoretic sofic entropy is an upper semi-continuous function of the measure, see Theorem 2.1 of [17]. This result can be obtained as a consequence of the fact that when $\Gamma \curvearrowright X$ is expansive then there exists $\varepsilon > 0$ such that for all $f \in C(X)$ and $\mu \in \text{Prob}_\Gamma(X)$

$$P_\Sigma(\Gamma \curvearrowright X, \mu, f) = P_\Sigma(\varepsilon, \Gamma \curvearrowright X, \mu, f).$$

Upper semi-continuity of the measure-theoretic entropy for expansive actions now follows using proposition 3.7.

3.4 Sofic pressure for shift spaces

As discussed above, the sofic topological pressure with respect to Σ does not depend upon the choice of a continuous dynamically generating pseudometric. In particular, for a subshift $X \subset A^\Gamma$ we may choose $\rho: X \times X \rightarrow \mathbb{R}$ given by

$$\rho(x, y) = \begin{cases} 0 & \text{if } x(1_\Gamma) = y(1_\Gamma), \\ 1 & \text{if } x(1_\Gamma) \neq y(1_\Gamma). \end{cases}$$

It is clear that with this pseudometric, the notion of being ε -separated is exactly the same for every value $\varepsilon < 1$. Namely, it is equivalent to the statement that for every $\varphi_1, \varphi_2 \in X^{V_i}$ there is $v \in V_i$ such that $\varphi_1(v)_{1_\Gamma} \neq \varphi_2(v)_{1_\Gamma}$. Thus for subshifts both definitions simplify to

$$P_\Sigma(f, \Gamma \curvearrowright X) = \inf_{\delta > 0} \limsup_{i \rightarrow \infty} \frac{1}{|V_i|} \log \left(P(f, N_\delta(X^*), \sigma_i, \frac{1}{2}) \right).$$

$$P_\Sigma(f, \Gamma \curvearrowright (X, \mu)) = \inf_{\delta > 0} \limsup_{i \rightarrow \infty} \frac{1}{|V_i|} \log \left(P(f, N_\delta(\mu^*), \sigma_i, \frac{1}{2}) \right).$$

We can further simplify the formulae for pressure by replacing X^{V_i} by the finite set A^{V_i} and writing for U an open set of probability measures on X

$$P(f, U, \sigma) = \frac{1}{|V|} \log \left(\sum_{w \in A^{V_i}: \xi_\sigma(w) \in U} \exp \left(\sum_{v \in V} f(\xi_{\sigma, v}(w)) \right) \right).$$

The advantage of the previous definition is that we no longer need to consider a supremum over ε -separated sets nor make computations on the space of orbits (X^*, μ^*) . It is clear that any $f \in C(X)$ and $\mu \in \text{Prob}_\Gamma(X)$ we have

$$P_\Sigma(f, \Gamma \curvearrowright X) = \inf_{\delta > 0} \limsup_{i \rightarrow \infty} P(f, N_\delta(X), \sigma_i).$$

$$P_\Sigma(f, \Gamma \curvearrowright (X, \mu)) = \inf_{\delta > 0} \limsup_{i \rightarrow \infty} P(f, N_\delta(\mu), \sigma_i).$$

For the second equality, we use the fact that for every $w \in A^V$ we have,

$$\sum_{v \in V} f(\xi_{\sigma, v}(w)) = |V| \int_X f \, d(\xi_\sigma(w)).$$

For convenience in what follows, if W is a set of probability measures on A^Γ define:

$$P_\Sigma(f, W, \Gamma \curvearrowright X) = \inf_{U \supset W} \limsup_{i \rightarrow \infty} P(f, U, \sigma_i),$$

where the infimum is over all open sets of probability measures U that contain W .

4 Proof of the sofic Lanford–Ruelle theorem for locally constant functions

For the remainder of this section, we fix a subshift $X \subset A^\Gamma$. Recall that we denote by $\mathcal{T}(X)$ the asymptotic relation of X and by $\mathcal{T}^0(X)$ the étale asymptotic relation of X .

For $\delta > 0$, a Borel subset $A \subset X$ and $t \in [0, 1]$ let

$$N_\delta(A, t) := \{ \mu \in N_\delta(X) : |\mu(A) - t| \leq \delta \}.$$

and for $x, y \in X$, $F \in \Gamma$ and $r \geq 0$,

$$N_\delta[(x, y), t, r, F] := \left\{ \mu \in N_\delta([x]_F \cup [y]_F, t) : \left| \frac{\mu([y]_F)}{\mu([x]_F)} - r \right| \leq \delta \right\}.$$

For $r \geq 0$, $C \in \mathbb{R}$ let $r^* = \frac{r}{1+r} \in [0, 1)$ and

$$\hat{p}(r, C) := H(r^*) + r^* C$$

where $H: [0, 1] \rightarrow \mathbb{R}$ is defined by $H(r^*) = -r^* \log(r^*) - (1-r^*) \log(1-r^*)$ with the usual convention that $0 \log(0) = 0$. Note that for fixed $C \in \mathbb{R}$, $\hat{p}(r, C)$ admits a unique global (and local) maximum which is attained at $r = \exp(C)$.

Definition 4.1. Let $F_1, F_2 \in \Gamma$. We say that a subset $X_0 \subset A^\Gamma$ has *trivial F_1 -overlaps within F_2* if

$$(gy)|_{F_2 \cap gF_2} \neq x|_{F_2 \cap gF_2} \text{ whenever } g \in F_2 F_1^{-1} \setminus \{1_\Gamma\} \text{ and } x, y \in X_0.$$

Suppose that $F_1 \subset F_2$. In this case the condition above is just a local way to state that for any $z \in A^\Gamma$ and $g \in \Gamma \setminus \{1_\Gamma\}$ such that $z|_{F_2} = x|_{F_2}$ and $(gz)|_{F_2} = y|_{F_2}$, we must necessarily have that $F_2 \cap gF_1 = \emptyset$. In other words, the patterns $x|_{F_2}$ and $y|_{F_2}$ can only occur in such a way that the “ F_1 -center” of y does not intersect the “ F_2 -center” of x . Next we will show that this condition provides a way to define an endomorphism of X which replaces all occurrences of a pattern $y|_{F_2}$ by another pattern $x|_{F_2}$ whenever they are exchangeable and $x|_{F_2 \setminus F_1} = y|_{F_2 \setminus F_1}$.

Lemma 4.2. Let $F_1 \subset F_2 \in \Gamma$, $(x, y) \in \mathcal{T}_{F_1}^0(X)$ and suppose that $\{x, y\}$ has trivial F_1 -overlaps within F_2 . There exists a continuous Γ -equivariant map $\pi: A^\Gamma \rightarrow A^\Gamma$ such that $\pi(X) \subset X$ and for every $z \in A^\Gamma$,

1. If $z|_{F_2} \in \{x|_{F_2}, y|_{F_2}\}$, then $\pi(z)|_{F_2} = x|_{F_2}$.
2. $\pi(z)|_{F_2} \neq y|_{F_2}$.

Proof. Since $(x, y) \in \mathcal{T}_{F_1}^0(X)$, it follows that $x|_{F_2 \setminus F_1} = y|_{F_2 \setminus F_1}$. Let us define $\pi: A^\Gamma \rightarrow A^\Gamma$ as follows. For $z \in A^\Gamma$,

$$\pi(z)(g) = \begin{cases} x(h) & \text{if there is } h \in F_1 \text{ such that } (hg^{-1}z)|_{F_2} = y|_{F_2}, \\ z(g) & \text{otherwise.} \end{cases}$$

Let us first show that π is well defined. Let $z \in A^\Gamma$ and suppose there is $g \in \Gamma$ and $h, h' \in F_1$ such that $(hg^{-1}z)|_{F_2} = y|_{F_2} = (h'g^{-1}z)|_{F_2}$. Let $z' = h'g^{-1}z$ and $\gamma = h'h^{-1}$, then we have $(\gamma^{-1}z')|_{F_2} = z'|_{F_2} = y|_{F_2}$. From $F_1 \subset F_2$ we get that $\gamma \in F_1 F_1^{-1} \subset F_2 F_2^{-2}$, while from the previous relation we get that $y|_{F_2 \cap \gamma F_2} = (\gamma y)|_{F_2 \cap \gamma F_2}$. As $\{x, y\}$ (and in particular $\{y\}$) has trivial F_1 -overlaps within F_2 we conclude that $\gamma = 1_\Gamma$ and thus $h = h'$.

Now let $z \in A^\Gamma$. If $z|_{F_2} \in \{x|_{F_2}, y|_{F_2}\}$ it follows by definition that $\pi(z)|_{F_1} = x|_{F_1}$. Now let $g \in F_2 \setminus F_1$, if there were $h \in F_1$ such that $(hg^{-1}z)|_{F_2} = y|_{F_2}$, we would have that $(gh^{-1}y)|_{F_2 \cap gh^{-1}F_2} \in \{x|_{F_2 \cap gh^{-1}F_2}, y|_{F_2 \cap gh^{-1}F_2}\}$ which cannot hold as $gh^{-1} \in F_2 F_1^{-1} \setminus \{1_\Gamma\}$ and $\{x, y\}$ has trivial F_1 -overlaps within F_2 . From the argument above we deduce that we must have $\pi(z)(g) = z(g)$ for $g \in F_2 \setminus F_1$ and thus $\pi(z)|_{F_2} = x|_{F_2}$.

Now suppose $\pi(z)|_{F_2} = y|_{F_2}$. By the previous argument, we couldn't have had $z|_{F_2} = y|_{F_2}$ and thus there must be $g \in F_2$ such that $\pi(z)(g) \neq z(g)$. This implies that there is $h \in F_1$ such that $(hg^{-1}z)|_{F_2} = y|_{F_2}$ and thus $(hg^{-1}\pi(z))|_{F_2} = x|_{F_2}$. This again cannot occur as $gh^{-1} \in F_2 F_1^{-1}$ and $\{x, y\}$ has trivial F_1 -overlaps within F_2 .

It is clear from the definition that π is continuous and Γ -equivariant. Now let $z \in X$. Since $(x, y) \in \mathcal{T}_{F_1}^0(X)$ and $F_1 \subset F_2$ it follows that $\xi: A^\Gamma \rightarrow A^\Gamma$ given by

$$\xi(z) = \begin{cases} z|_{\Gamma \setminus F_2} \vee x|_{F_2} & \text{if } z|_{F_2} = y|_{F_2} \\ z & \text{otherwise,} \end{cases}$$

is a continuous self-map on X . It is clear that the fact that $\{x, y\}$ has trivial F_1 -overlaps within F_2 implies that the collection of maps $\{g^{-1}\xi g\}_{g \in \Gamma}$ pairwise commute. Thus for any finite $F \subset \Gamma$ the composition $\xi_F := \prod_{g \in F} g^{-1}\xi g$ is a well defined self-map of X (independent of the order of composition). For every $x \in X$ and any finite $F \subset \Gamma$, $\pi(x)|_F$ coincides with $\xi_{F'}(x)|_F$ for sufficiently big F' . From here it follows that $\pi(X) \subset X$. \square

Let us recall that a function $f: X \rightarrow \mathbb{R}$ is said to be F -locally constant if $f(x) = f(y)$ for every $x, y \in X$ such that $x|_F = y|_F$. In the following lemma we shall use the usual little o notation $o(g(n))$ to denote a function which goes to zero as n goes to infinity when divided by $g(n)$.

Lemma 4.3. *Let $F_1 \subset F_2 \in \Gamma$ such that $F_1 F_1^{-1} \subset F_2$, $(x, y) \in \mathcal{T}_{F_1}^0(X)$, $t \in [0, 1]$ and $f: X \rightarrow \mathbb{R}$ an F_1 -locally constant function. Suppose that $\{x, y\}$ has trivial F_1 -overlaps within F_2 and that for all $\delta > 0$ there are infinitely many $i \in \mathbb{N}$ such that*

$$\xi_{\sigma_i}^{-1}(N_\delta([x]_{F_2} \cup [y]_{F_2}, t)) \neq \emptyset.$$

Then for all $r_1, r_2 \geq 0$ we have

$$\begin{aligned} \inf_{\delta > 0} \limsup_{i \rightarrow \infty} (P(f, N_\delta[(x, y), t, r_1, F_2], \sigma_i) - P(f, N_\delta[(x, y), t, r_2, F_2], \sigma_i)) \\ = t(\hat{p}(r_1, \Psi_f(x, y)) - \hat{p}(r_2, \Psi_f(x, y))). \end{aligned}$$

Proof. Let $F_1, F_2 \in \Gamma$, $(x, y) \in \mathcal{T}_{F_1}^0(X)$, $t \in [0, 1]$ and $f: X \rightarrow \mathbb{R}$ as in the statement. Lemma 4.2 provides the existence of a continuous Γ -equivariant map $\pi: A^\Gamma \rightarrow A^\Gamma$ such that $\pi(X) \subset X$ and which replaces any occurrence of the pattern $y|_{F_2}$ by the pattern $x|_{F_2}$ and which contains no occurrences of the pattern $y|_{F_2}$.

For every $i \in \mathbb{N}$ let $\pi_i: A^{V_i} \rightarrow A^{V_i}$ be the map given by

$$\pi_i(w)(v) = (\pi(\xi_{\sigma_i, v}(w)))(1_\Gamma).$$

For $i \in \mathbb{N}$, $\delta > 0$ and $r \geq 0$ let

$$\begin{aligned} \Omega_i(\delta) &= \xi_{\sigma_i}^{-1}(N_\delta([x]_{F_2} \cup [y]_{F_2}, t)), \\ \Omega_{\pi_i}(\delta) &= \pi_i(\xi_{\sigma_i}^{-1}(N_\delta([x]_{F_2} \cup [y]_{F_2}, t))), \\ \Omega_i(\delta, r) &= \xi_{\sigma_i}^{-1}(N_\delta[(x, y), t, r, F_2]). \end{aligned}$$

For $w \in A^{V_i}$, let us denote by $S_i(f, w) = \sum_{v \in V_i} f(\xi_{\sigma_i, v}(w))$. Our main goal is to estimate the expression

$$P(f, N_\delta[(x, y), t, r, F_2], \sigma_i) = \frac{1}{|V_i|} \log \left(\sum_{w \in \Omega_i(\delta, r)} \exp(S_i(f, w)) \right).$$

In order to do that, let us first fix $w \in \Omega_i(\delta, r)$. We shall first show that the difference $S_i(f, w) - S_i(f, \pi_i(w))$ is close to $r^*t|V_i|\Psi_f(x, y)$. Intuitively, this happens because if $v \in V_i$ is a sufficiently good approximation (in the sense that the configurations $\xi_{\sigma_i, v}(w)$ and $\xi_{\sigma_i, v}(\pi_i(w))$ “look like elements of X ” restricted to some large set), then the term $f(\xi_{\sigma_i, v}(w)) - f(\xi_{\sigma_i, v}(\pi_i(w)))$ is non-zero only if $\xi_{\sigma_i, v}(w)|_{F_2} = y|_{F_2}$, which occurs with probability close to r^*t due to w being in $\Omega_i(\delta, r)$.

Let us now proceed formally. As Σ is a sofic approximation sequence for Γ , the proportion of vertices $v \in V_i$ which are not $F_1 F_2$ -good is $o(|V_i|)$ and thus

$$\sum_{v \in V_i \text{ is not } F_1 F_2 \text{ good}} (f(\xi_{\sigma_i, v}(w)) - f(\xi_{\sigma_i, v}(\pi_i(w)))) \leq 2o(|V_i|)\|f\|_\infty = o(|V_i|).$$

Suppose that $v \in V_i$ is $F_1 F_2$ -good for σ_i . Using the fact that f is F_1 -local and the definition of π_i , we obtain that, as long as there is no $g \in F_1$ such that $\xi_{\sigma_i, \sigma_i(g)v}(w)|_{F_2} = y|_{F_2}$, then

$$f(\xi_{\sigma_i, v}(w)) - f(\xi_{\sigma_i, v}(\pi_i(w))) = 0.$$

Therefore we only need to consider the v for which there is some $g \in F_1$ such that $\xi_{\sigma_i, \sigma_i(g)v}(w)|_{F_2} = y|_{F_2}$. Note that as v is $F_1 F_2$ -good and $\{x, y\}$ have trivial F_1 -overlaps within F_2 , it follows that such a g is unique and we have

$$\sum_{h \in F_1 F_1^{-1}} (f(\xi_{\sigma_i, \sigma_i(gh)v}(w)) - f(\xi_{\sigma_i, \sigma_i(gh)v}(\pi(w)))) = \Psi_f(x, y).$$

Given $\delta > 0$ as above, let $r^* = \frac{r}{1+r}$ and define δ^- and δ^+ by

$$\delta^- = r^* - \frac{r - \delta}{1 + r - \delta}, \quad \delta^+ = \frac{r + \delta}{1 + r + \delta} - r^*.$$

As $w \in \Omega_i(\delta, r)$, it follows that the number N_i of $v \in V_i$ such that $\xi_{\sigma_i, \sigma_i(gh)v}(w)|_{F_2} = y|_{F_2}$ is bounded above and below as follows

$$(r^* - \delta^-)(t - \delta)(|V_i| - o(V_i)) \leq N_i \leq (r^* + \delta^+)(t + \delta)|V_i|.$$

Putting all of the previous estimates together, we obtain that

$$o(|V_i|) + (r^* - \delta^-)(t - \delta)|V_i|\Psi_f(x, y) \leq S_i(f, w) - S_i(f, \pi_i(w)) \leq o(|V_i|) + (r^* + \delta^+)(t + \delta)|V_i|\Psi_f(x, y).$$

For $w' \in \Omega_{\pi_i, i}(\delta)$, let us define $K_{w'}(\delta, r) = |\pi_i^{-1}(\{w'\}) \cap \Omega_i(\delta, r)|$ to be the number of $w \in \Omega_i(\delta, r)$ such that $\pi_i(w) = w'$. Note that we may always write

$$\sum_{w \in \Omega_i(\delta, r)} \exp(S_i(f, w)) = \sum_{w' \in \Omega_{\pi_i, i}(\delta)} \left(\sum_{w \in \Omega_i(\delta, r) : \pi_i(w) = w'} \exp(S_i(f, w') + S_i(f, w) - S_i(f, \pi_i(w))) \right).$$

Using the bounds we obtained for $S_i(f, w) - S_i(f, \pi_i(w))$, we get that

$$\begin{aligned} \sum_{w \in \Omega_i(\delta, r)} \exp(S_i(f, w)) &\leq \sum_{w' \in \Omega_{\pi_i, i}(\delta)} K_{w'}(\delta, r) \exp(S_i(f, w') + o(|V_i|) + (r^* + \delta^+)(t + \delta)|V_i|\Psi_f(x, y)) \\ &\leq \sum_{w' \in \Omega_{\pi_i, i}(\delta)} \exp(S_i(f, w') + \log(K_{w'}(\delta, r)) + o(|V_i|) + (r^* + \delta^+)(t + \delta)|V_i|\Psi_f(x, y)), \end{aligned}$$

and

$$\begin{aligned} \sum_{w \in \Omega_i(\delta, r)} \exp(S_i(f, w)) &\geq \sum_{w' \in \Omega_{\pi_i, i}(\delta)} K_{w'}(\delta, r) \exp(S_i(f, w') + o(|V_i|) + (r^* - \delta^-)(t - \delta)|V_i|\Psi_f(x, y)) \\ &\geq \sum_{w' \in \Omega_{\pi_i, i}(\delta)} \exp(S_i(f, w') + \log(K_{w'}(\delta, r)) + o(|V_i|) + (r^* - \delta^-)(t - \delta)|V_i|\Psi_f(x, y)). \end{aligned}$$

Our next goal is to provide an estimate for $\log(K_{w'}(\delta, r))$. Let $G(w', x) := \{v \in V_i : \xi_{\sigma_i, v}(w')|_{F_2} = x|_{F_2}\}$. By our assumptions we have that for infinitely many values of i ,

$$(t - \delta)|V_i| + o(|V_i|) \leq |G(w', x)| \leq (t + \delta)|V_i| + o(|V_i|).$$

Notice that any $w \in A^{V_i}$ such that $\pi_i(w) = w'$ is uniquely determined by the subset of vertices in $G(w', x)$, which consists of the positions that initially held the pattern $y|_{F_2}$ and were erased by the map π_i . Using again that $w \in \Omega_i(\delta, r)$ we obtain that

$$\left(\begin{array}{c} G(w', x) \\ \lfloor r^* G(w', x) \rfloor \end{array} \right) \leq K_{w'}(\delta, r) \leq \sum_{k = \lceil (r^* - \delta^-)G(w', x) \rceil}^{\lfloor (r^* + \delta^+)G(w', x) \rfloor} \binom{G(w', x)}{k}.$$

It is easy to show using Stirling's approximation that whenever $n \in \mathbb{N}$ and $\alpha \in (0, \frac{1}{2})$, then

$$\log \binom{n}{\lfloor \alpha n \rfloor} = (1 + o(1))H(\alpha)n.$$

This immediately yields the lower bound

$$K_{w'}(\delta, r) \geq \binom{G(w', x)}{\lfloor r^* G(w', x) \rfloor} = \exp((1 + o(1))H(r^*)(t - \delta)|V_i| + o(|V_i|)).$$

For the upper bound, let $s \in (-\delta^-, \delta^+)$ such that the binomial coefficient

$$\binom{G(w', x)}{\lfloor (r^* + s)G(w', x) \rfloor}$$

is largest. We obtain the following upper bound for small enough δ ,

$$\begin{aligned} K_{w'}(\delta, r) &\leq \lceil (\delta^+ - \delta^-) |G(w', x)| \rceil \exp((1 + o(1))H(r^* + s)(t + \delta)|V_i| + o(|V_i|)) \\ &\leq |V_i| \exp((1 + o(1))H(r^* + s)(t + \delta)|V_i| + o(|V_i|)). \end{aligned}$$

Putting these two bounds together and taking logarithms, we obtain

$$o(|V_i|) + (t - \delta)H(r^*)|V_i| \leq \log(K_{w'}(\delta, r)) \leq o(|V_i|) + (t + \delta)H(r^* + s)|V_i|.$$

Now we can refine our previous bounds. Putting this last computation back in the previous formulas we get

$$\begin{aligned} \frac{1}{|V_i|} \log \left(\sum_{w \in \Omega_i(\delta, r)} \exp(S_i(f, w)) \right) &\leq \frac{1}{|V_i|} \log \left(\sum_{w' \in \Omega_{\pi_i}(\delta)} \exp(S_i(f, w')) \right) \\ &\quad + \frac{o(|V_i|)}{|V_i|} + (t + \delta) (H(r^* + s) + (r^* + \delta^+) \Psi_f(x, y)). \end{aligned}$$

and

$$\begin{aligned} \frac{1}{|V_i|} \log \left(\sum_{w \in \Omega_i(\delta, r)} \exp(S_i(f, w)) \right) &\geq \frac{1}{|V_i|} \log \left(\sum_{w' \in \Omega_{\pi_i}(\delta)} \exp(S_i(f, w')) \right) \\ &\quad + \frac{o(|V_i|)}{|V_i|} + (t - \delta) (H(r^*) + (r^* - \delta^-) \Psi_f(x, y)). \end{aligned}$$

Let now $r_1, r_2 \geq 0$ and notice that the first term in both bounds does not depend on r and thus disappears when plugging the above bounds into the difference

$$P(f, N_\delta[(x, y), t, r_1, F_2], \sigma_i) - P(f, N_\delta[(x, y), t, r_2, F_2], \sigma_i).$$

Furthermore, the term $o(|V_i|)/|V_i|$ vanishes when taking the limsup as $i \rightarrow \infty$, also, the terms δ^-, δ^+ and s all go to 0 uniformly as δ goes to 0. We obtain thus that for all $r_1, r_2 \geq 0$ we have

$$\begin{aligned} \inf_{\delta > 0} \limsup_{i \rightarrow \infty} (P(f, N_\delta[(x, y), t, r_1, F_2], \sigma_i) - P(f, N_\delta[(x, y), t, r_2, F_2], \sigma_i)) \\ = t (H(r_1^*) + r_1^* \Psi_f(x, y) - (H(r_2^*) + r_2^* \Psi_f(x, y))) \\ = t (\hat{p}(r_1, \Psi_f(x, y)) - \hat{p}(r_2, \Psi_f(x, y))). \end{aligned}$$

Which is what we wanted to show. \square

Definition 4.4. Given $x \in A^\Gamma$ and $F_1 \subseteq F_2 \in \Gamma$, we say that x is (F_1, F_2) -self-overlapping if

$$(gx)|_{(F_2 \setminus F_1) \cap g(F_2 \setminus F_1)} = x|_{(F_2 \setminus F_1) \cap g(F_2 \setminus F_1)} \text{ for some } g \in F_2 F_1^{-1} \setminus \{1_\Gamma\}.$$

We say that x is *non self-overlapping* if for every $F_1 \in \Gamma$ there exists $F_2 \in \Gamma$ such that $F_1 \subseteq F_2$ and x is not (F_1, F_2) -self-overlapping.

Notice that the condition of being (F_1, F_2) -self-overlapping is satisfied automatically if there is $g \in F_2 F_1^{-1} \setminus \{1_\Gamma\}$ for which $(F_2 \setminus F_1) \cap g(F_2 \setminus F_1) = \emptyset$. Thus in the definition of non self-overlapping it is understood that the sets F_2 must satisfy that $(F_2 \setminus F_1) \cap g(F_2 \setminus F_1) \neq \emptyset$ for every $g \in F_2 F_1^{-1} \setminus \{1_\Gamma\}$. The usefulness of this condition is justified in the following lemma.

Lemma 4.5. *Let $x \in A^\Gamma$ be non self-overlapping. For every $F_1 \in \Gamma$ and $y \in A^\Gamma$ such that $(x, y) \in \mathcal{T}_{F_1}(X)$, there exists $F_2 \in \Gamma$ with $F_1 \subset F_2$ such that $\{x, y\}$ has trivial F_1 -overlaps within F_2 .*

Proof. As x is non self-overlapping, there is $F_2 \in \Gamma$ such that $F_1 \subseteq F_2$ and x is not (F_1, F_2) -self-overlapping. Suppose that $\{x, y\}$ does not have trivial F_1 -overlaps within F_2 , then there is $g \in F_2 F_1^{-1} \setminus \{1_\Gamma\}$ and $x', y' \in \{x, y\}$ such that

$$(gy')|_{F_2 \cap gF_2} = x'|_{F_2 \cap gF_2}.$$

In particular, as $(F_2 \setminus F_1) \cap g(F_2 \setminus F_1) \neq \emptyset$, we have

$$(gy')|_{(F_2 \setminus F_1) \cap g(F_2 \setminus F_1)} = x'|_{(F_2 \setminus F_1) \cap g(F_2 \setminus F_1)}.$$

Finally, as $(x, y) \in \mathcal{T}_{F_1}(X)$, it follows that $x|_{F_2 \setminus F_1} = y|_{F_2 \setminus F_1}$ and $(gx)|_{g(F_2 \setminus F_1)} = (gy)|_{g(F_2 \setminus F_1)}$. We obtain that

$$(gx)|_{(F_2 \setminus F_1) \cap g(F_2 \setminus F_1)} = x|_{(F_2 \setminus F_1) \cap g(F_2 \setminus F_1)}.$$

Which contradicts the fact that x is not (F_1, F_2) -self-overlapping. \square

Lemma 4.6. *Let Σ be a sofic approximation for Γ such that $h_\Sigma(\Gamma \curvearrowright X) \geq 0$ and let μ be an equilibrium measure for a subshift $X \subset A^\Gamma$ with respect to an F -locally constant function $f: X \rightarrow \mathbb{R}$ and Σ . Suppose also that $(x, y) \in \mathcal{T}^0(X)$, x is in the support of μ and x is non self-overlapping.*

Then y is also in the support of μ and there exists a sequence of finite subsets $F_n \in \Gamma$ increasing to Γ so that

$$\lim_{n \rightarrow \infty} \frac{\mu([y]_{F_n})}{\mu([x]_{F_n})} = \exp(\Psi_f(x, y)).$$

Proof. Choose a sequence of finite subsets $F_1^{(n)} \in \Gamma$ which is increasing to Γ and such that $F \subset F_1^{(n)}$ and $(x, y) \in \mathcal{T}_{F_1^{(n)}}^0(X)$ for all $n \in \mathbb{N}$. By Lemma 4.5, for each $n \in \mathbb{N}$ there exists $F_2^{(n)} \in \Gamma$ such that $F_1^{(n)} \subset F_2^{(n)}$ and $\{x, y\}$ has trivial $F_1^{(n)}$ -overlaps within $F_2^{(n)}$. Also, because x is in the support of μ , we can define

$$t = \mu([x|_{F_2^{(n)}}] \cup [y|_{F_2^{(n)}}]) > 0 \text{ and } r = \frac{\mu([y|_{F_2^{(n)}}])}{\mu([x|_{F_2^{(n)}}])} \geq 0.$$

For now fix $n \in \mathbb{N}$ and denote $F_2 := F_2^{(n)}$.

Note that $\mu \in N_\delta[(x, y), t, r, F_2] \subset N_\delta([x|_{F_2}] \cup [y|_{F_2}], t)$ for every $\delta > 0$. From the assumption that $h_\Sigma(\Gamma \curvearrowright X) \geq 0$ and the variational principle for sofic pressure (Theorem 3.8) we obtain that $h_\Sigma(\Gamma \curvearrowright (X, \mu)) \geq 0$. As $\mu \in N_\delta[(x, y), t, r, F_2]$, it follows that there is an infinite sequence of $i \in \mathbb{N}$ such that $\xi_{\sigma_i}^{-1}(N_\delta[(x, y), t, r, F_2]) \neq \emptyset$.

From Lemma 4.3 it follows that for every $r_1, r_2 \geq 0$ we have

$$\begin{aligned} & \lim_{\delta \rightarrow 0} \limsup_{i \rightarrow \infty} (P(f, N_\delta[(x, y), t, r_1, F_2], \sigma_i) - P(f, N_\delta[(x, y), t, r_2, F_2], \sigma_i)) \\ & = t(\hat{p}(r_1, \Psi_f(x, y)) - \hat{p}(r_2, \Psi_f(x, y))). \end{aligned}$$

As μ is an equilibrium measure for f , for every $\varepsilon > 0$ we can find $\delta_2 > 0$ such that for all sufficiently big $i \in \mathbb{N}$ we have

$$|P_\Sigma(f, \Gamma \curvearrowright (X, \mu)) - P(f, N_{\delta_2}[(x, y), t, r, F_2], \sigma_i)| < \varepsilon.$$

Fix any $\varepsilon > 0$. It follows that for any $r' \geq 0$ and sufficiently big $i \in \mathbb{N}$ we have

$$P(f, N_{\delta_2}[(x, y), t, r', F_2], \sigma_i) - P(f, N_{\delta_2}[(x, y), t, r, F_2], \sigma_i) \leq \varepsilon.$$

From the above relation we obtain that

$$\hat{p}(r', \Psi_f(x, y)) - \hat{p}(r, \Psi_f(x, y)) \leq \frac{\varepsilon}{t}.$$

Recall that for any $C \in \mathbb{R}$, the expression $\hat{p}(r', C)$ as a function of $r' \geq 0$ has a unique local maximum at $r' = \exp(C)$. We obtain that

$$\hat{p}(\exp(\Psi_f(x, y)), \Psi_f(x, y)) - \hat{p}(r, \Psi_f(x, y)) \leq \frac{\varepsilon}{t}.$$

The restriction to the first coordinate of the function \hat{p} which maps $r' \mapsto \hat{p}(r', \Psi_f(x, y))$ attains a unique maximum at $r' = \exp(\Psi_f(x, y))$. Moreover, its inverse has two continuous branches which meet on the maximum. It follows that for every $\eta > 0$ we can choose $\varepsilon > 0$ such that whenever the above equation holds, then $|r - \exp(\Psi_{f_0}(x, y))| \leq \eta$. As ε was arbitrary, we obtain that

$$\frac{\mu([y]_{F_2^{(n)}})}{\mu([x]_{F_2^{(n)}})} = r = \exp(\Psi_f(x, y)).$$

As $n \in \mathbb{N}$ was arbitrary, we obtain that

$$\lim_{n \rightarrow \infty} \frac{\mu([y]_{F_2^{(n)}})}{\mu([x]_{F_2^{(n)}})} = \exp(\Psi_f(x, y)),$$

as we required. This equality automatically implies that both x and y are in the support of μ . \square

The remainder of the proof consists essentially on getting rid of the non self-overlapping hypothesis. In order to do this, we will first show that the condition occurs almost surely with respect to the uniform Bernoulli measure on the full Γ -shift, and then we use a trick involving a direct product to take advantage of the above property on an arbitrary subshift.

Lemma 4.7. *Let μ be the uniform Bernoulli measure on $\{0, 1\}^\Gamma$. There exists a Borel subset $X_0 \subset \{0, 1\}^\Gamma$ with $\mu(X_0) = 1$ such that every $x \in X_0$ is non self-overlapping.*

Proof. For any $F_1 \subseteq F_2 \in \Gamma$ and $g \in F_2 F_1^{-1} \setminus \{1_\Gamma\}$

$$\mu(\{x \in \{0, 1\}^\Gamma : (gx)|_{(F_2 \setminus F_1) \cap g(F_2 \setminus F_1)} = x|_{(F_2 \setminus F_1) \cap g(F_2 \setminus F_1)}\}) = \left(\frac{1}{2}\right)^{|(F_2 \setminus F_1) \cap g(F_2 \setminus F_1)|}.$$

By a simple union bound, it follows that

$$\begin{aligned} \mu(\{x \in \{0,1\}^\Gamma : x \text{ is } (F_1, F_2)\text{-self-overlapping}\}) &\leq \sum_{g \in F_2 F_1^{-1} \setminus \{1_\Gamma\}} \left(\frac{1}{2}\right)^{|(F_2 \setminus F_1) \cap g(F_2 \setminus F_1)|} \\ &\leq |F_1| |F_2| \max_{g \in F_2 F_1^{-1}} 2^{-|(F_2 \setminus F_1) \cap g(F_2 \setminus F_1)|} \\ &\leq |F_1| 4^{|F_1|} |F_2| \max_{g \in F_2 F_1^{-1}} 2^{-|F_2 \cap g F_2|}. \end{aligned}$$

As Γ is countably infinite, in order to prove the lemma it suffices to show that for any fixed $F_1 \in \Gamma$ there exists a sequence $(F_2^{(n)})_{n \geq 1}$, with $F_1 \subseteq F_2^{(n)} \in \Gamma$ for all $n \geq 1$, and with the property that μ -almost surely the number of n 's for which x is $(F_1, F_2^{(n)})$ -self-overlapping is at most finite. By the Borel–Cantelli lemma, it suffices in turn to find for each $F_1 \in \Gamma$ a sequence $(F_2^{(n)})_{n \geq 1}$ with $F_1 \subseteq F_2^{(n)} \in \Gamma$ for all $n \geq 1$ such that

$$\sum_{n=1}^{\infty} \mu(\{x \in \{0,1\}^\Gamma : x \text{ is } (F_1, F_2^{(n)})\text{-self-overlapping}\}) < \infty.$$

By the previous calculation, it suffices to have that

$$\sum_{n=1}^{\infty} |F_2^{(n)}| \max_{g \in F_2^{(n)} F_1^{-1}} 2^{-|F_2^{(n)} \cap g F_2^{(n)}|} < \infty.$$

Let $F_1 \in \Gamma$ be an arbitrary non-empty finite set. If Γ is locally finite, then we can choose a sequence $(F_2^{(n)})_{n \geq 1}$ of finite subgroups of Γ , each containing F_1 with $|F_n| > n$. In that case $F_2^{(n)} F_1^{-1} = F_2^{(n)}$ and for any $g \in F_2^{(n)}$ we have $F_2^{(n)} \cap g F_2^{(n)} = F_2^{(n)}$. We deduce that in this case

$$\sum_{n=1}^{\infty} |F_2^{(n)}| \max_{g \in F_2^{(n)} F_1^{-1}} 2^{-|F_2^{(n)} \cap g F_2^{(n)}|} \leq \sum_{n=1}^{\infty} |F_2^{(n)}| 2^{-|F_2^{(n)}|} \leq \sum_{n=1}^{\infty} n 2^{-n} < \infty.$$

If Γ is not locally finite, there exists $S \in \Gamma$ which is a finite symmetric set with $\{1_\Gamma\} \cup F_1 \subset S$ and so that the group generated by S is infinite. For any $m \geq 1$ and $g \in S^{3m} F_1^{-1}$ we can choose $h \in \Gamma$ such that $h S^m \subset S^{3m} \cap g S^{3m}$ and thus we have $|S^{3m} \cap g S^{3m}| \geq |S^m|$. As $|S^{3n}| \leq |S|^{3n}$ for all n it follows that there exists a strictly increasing sequence of integers $(k_n)_{n=1}^{\infty}$ such that $|S^{3^{k_n-1}}| \geq |S^{3^{k_n}}|^{\frac{1}{4}}$ for all $n \in \mathbb{N}$. For a sequence as above, let $F_2^{(n)} = S^{3^{k_n}}$. It follows that for each $n \in \mathbb{N}$ and every $g \in F_2^{(n)} F_1^{-1}$,

$$|F_2^{(n)} \cap g F_2^{(n)}| \geq |S^{3^{k_n-1}}| \geq |S^{3^{k_n}}|^{\frac{1}{4}} = |F_2^{(n)}|^{\frac{1}{4}},$$

and in particular

$$\sum_{n=1}^{\infty} |F_2^{(n)}| \max_{g \in F_2^{(n)} F_1^{-1}} 2^{-|F_2^{(n)} \cap g F_2^{(n)}|} \leq \sum_{n=1}^{\infty} |F_2^{(n)}| 2^{-|F_2^{(n)}|^{\frac{1}{4}}} \leq \sum_{n=1}^{\infty} n 2^{\frac{1}{4} \sqrt{n}} < \infty.$$

□

Lemma 4.8. *Let Σ be a sofic approximation sequence for Γ . Let $f \in C(X)$ and μ an equilibrium measure on X for f . Consider $\tilde{X} = X \times \{0,1\}^\Gamma$, $\tilde{f}: \tilde{X} \rightarrow \mathbb{R}$ given by $\tilde{f}(x,y) = f(x)$ for every $x \in X$ and $y \in \{0,1\}^\Gamma$ and let $\tilde{\mu}$ be the product of μ and the uniform Bernoulli measure on $\{0,1\}^\Gamma$. We have that $\tilde{\mu}$ is an equilibrium measure for \tilde{f} on \tilde{X} .*

Proof. Let ν denote the the uniform Bernoulli measure on $\{0, 1\}^\Gamma$. As $(\{0, 1\}^\Gamma, \nu)$ is a Bernoulli shift, it follows by Theorem 6.3 of [10] that the measure-theoretic sofic entropy of the product system can be expressed as follows,

$$h_\Sigma(\Gamma \curvearrowright (\tilde{X}, \tilde{\mu})) = h_\Sigma(\Gamma \curvearrowright (X, \mu)) + h_\Sigma(\Gamma \curvearrowright (\{0, 1\}^\Gamma, \nu)) = h_\Sigma(\Gamma \curvearrowright (X, \mu)) + \log(2).$$

Let μ' be an arbitrary Γ -invariant measure on X . We argue that for every Γ -invariant measure $\tilde{\mu}'$ on \tilde{X} with marginal μ' on X we have

$$h_\Sigma(\Gamma \curvearrowright (\tilde{X}, \tilde{\mu}')) \leq h_\Sigma(\Gamma \curvearrowright (X, \mu')) + \log(2).$$

This follows from the fact that the number of $w \in (A \times \{0, 1\})^{V_i}$ for which $\xi_{\sigma_i}(w) \in N_\delta(\tilde{\mu}')$ is bounded by $2^{|V_i|}$ times the number of microstates $u \in A^{V_i}$ such that $\xi_{\sigma_i}(u) \in N_\delta(\mu')$.

Using the definition of \tilde{f} , we get that for every Γ -invariant measure $\tilde{\mu}'$ with marginal μ' on X , we have

$$\int \tilde{f} d\tilde{\mu}' = \int f d\mu'.$$

Putting together all the previous computations, we get

$$\begin{aligned} P_\Sigma(\tilde{f}, \Gamma \curvearrowright (\tilde{X}, \tilde{\mu}')) &= h_\Sigma(\Gamma \curvearrowright (\tilde{X}, \tilde{\mu}')) + \int \tilde{f} d\tilde{\mu}' \\ &= h_\Sigma(\Gamma \curvearrowright (\tilde{X}, \tilde{\mu}')) + \int f d\mu' \\ &\leq \log(2) + h_\Sigma(\Gamma \curvearrowright (X, \mu')) + \int f d\mu' \\ &\leq \log(2) + h_\Sigma(\Gamma \curvearrowright (X, \mu)) + \int f d\mu \\ &= h_\Sigma(\Gamma \curvearrowright (\tilde{X}, \tilde{\mu})) + \int \tilde{f} d\tilde{\mu} \\ &= P_\Sigma(\tilde{f}, \Gamma \curvearrowright (\tilde{X}, \tilde{\mu})) \end{aligned}$$

And therefore $\tilde{\mu}$ is an equilibrium measure for \tilde{f} on \tilde{X} . \square

The following result is the main theorem in the case where the function f is locally constant. We shall first show it in this case and then generalize it to larger classes of functions in Section 5.

Theorem 4.9. *Let Σ be a sofic approximation sequence for Γ . Consider a subshift $X \subset A^\Gamma$ with $h_\Sigma(\Gamma \curvearrowright X) \geq 0$. Let $F \Subset \Gamma$ and suppose $f: X \rightarrow \mathbb{R}$ is an F -locally constant function and μ an equilibrium measure on X for f . Then μ is étale Gibbs with respect to f .*

Proof. Let $(x, y) \in \mathcal{T}^0(X)$ such that x is in the support of μ . Let \tilde{X}, \tilde{f} and $\tilde{\mu}$ as in Lemma 4.8. Then $\tilde{\mu}$ is an equilibrium measure for \tilde{f} on \tilde{X} . Now let $(x', y') \in \mathcal{T}(\{0, 1\}^\Gamma) = \mathcal{T}^0(\{0, 1\}^\Gamma)$ such that x' is non self-overlapping (its existence assured by Lemma 4.7).

Then it follows that $((x, x'), (y, y')) \in \mathcal{T}^0(\tilde{X})$, (x, x') is in the support of $\tilde{\mu}$, and (x, x') is non self-overlapping. By Lemma 4.6 there exists a sequence of finite sets $F_n \Subset \Gamma$ increasing to Γ such that

$$\lim_{n \rightarrow \infty} \frac{\tilde{\mu}([(y, y')]_{F_n})}{\tilde{\mu}([(x, x')]_{F_n})} = \exp(\Psi_{\tilde{f}}((x, x'), (y, y'))).$$

Let ν be the uniform Bernoulli measure on $\{0, 1\}^\Gamma$. By definition we have $\tilde{\mu} = \mu \times \nu$, therefore on the one hand we get for any $K \Subset \Gamma$

$$\frac{\tilde{\mu}([(y, y')]_K)}{\tilde{\mu}([(x, x')]_K)} = \frac{\mu([y]_K)\nu([y']_K)}{\mu([x]_K)\nu([x']_K)} = \frac{\mu([y]_K)}{\mu([x]_K)}.$$

On the other hand, using that $\tilde{f}(x, x') = f(x)$ for every $(x, x') \in \tilde{X}$, we get

$$\exp(\Psi_{\tilde{f}}((x, x'), (y, y'))) = \exp(\Psi_f(x, y)).$$

Hence we conclude that

$$\lim_{n \rightarrow \infty} \frac{\mu([y]_{F_n})}{\mu([x]_{F_n})} = \exp(\Psi_f(x, y)).$$

For every x in the support of μ . By Lemma 2.7 it follows that there is a Borel set $X' \subset X$ with $\mu(X \setminus X') = 0$ such that for every $(x, y) \in X \times X' \cap \mathcal{T}^0(X)$ we have

$$D_{\mu, \mathcal{T}^0(X)}(x, y) = \lim_{F \nearrow \Gamma} \frac{\mu([y]_F)}{\mu([x]_F)}.$$

From the above we obtain that up to a μ -null set

$$D_{\mu, \mathcal{T}^0(X)}(x, y) = \exp(\Psi_f(x, y)).$$

And thus μ is étale Gibbs for f . □

Before closing up this section, we show that the condition that $h_\Sigma(\Gamma \curvearrowright X) \geq 0$ cannot be removed.

Example 4.10. Let F_2 be the free group on two generators a, b and consider the subshift of finite type $X \subset (\mathbb{Z}/4\mathbb{Z})^{F_2}$ given by the algebraic condition

$$X = \{x \in (\mathbb{Z}/4\mathbb{Z})^{F_2} : x(gs) = x(g) + 1 \pmod{2}, \text{ for } s \in \{a, b\}\}.$$

This action is topologically conjugate to direct product of the full F_2 -shift on two symbols times the action on $\mathbb{Z}/2\mathbb{Z}$ on which both generators act non-trivially. It is known [10, Theorem 4.1] that if Σ is a sofic approximation sequence which is “far from bipartite” (for instance one obtained by choosing the homomorphisms $\sigma_i: F_2 \rightarrow \text{Sym}(V_i)$ to be generated by $\sigma_i(a)$ and $\sigma_i(b)$ and choosing those uniformly at random) then $F_2 \curvearrowright \mathbb{Z}/2\mathbb{Z}$ admits no models for large enough i and thus $h_\Sigma(\Gamma \curvearrowright X) = -\infty$. In this setting, every invariant measure is an equilibrium measure for Σ (with respect to $f = 0$) but the unique Gibbs measure of the system is the product of the uniform Bernoulli measure on the full F_2 -shift and the uniform measure on $\mathbb{Z}/2\mathbb{Z}$.

5 Lanford–Ruelle theorem beyond locally constant functions

In this section we extend the conclusion of Theorem 4.9 beyond locally constant functions, and in particular provide a proof of Theorem A, and more generally of Theorem B. The argument is essentially an application of abstract functional analysis of convex functions on topological vector spaces. Our basic strategy here closely follows the original approach of Lanford and Ruelle (see Appendix A of [26]), with a slight deviation which we discuss below.

By a continuous norm on a topological vector space V we mean a function $\|\cdot\|: V \rightarrow [0, \infty)$ satisfying the axioms of a norm which is continuous with respect to the topology that makes V a topological vector space. Note that the topology induced on V by the norm $\|\cdot\|$ need not coincide with the given topology of V . A typical example of a topological vector space that admits a continuous norm is $C^\infty([0, 1])$, where the uniform norm is one example for a continuous norm. The following functional-analytic result is the essence of our argument:

Proposition 5.1. *Let V be a topological vector space which admits a continuous norm. Let \bar{V} denote the completion of V with respect to that norm. Let $f: \bar{V} \rightarrow \mathbb{R}$ be a continuous convex function, and $\psi: V \rightarrow \bar{V}$ be a continuous function. Suppose that there exists a dense subset $A \subset V$ such that if $\bar{v} \in A$, then there exists a tangent functional $\bar{w} \in \bar{V}^*$ of f at \bar{v} with norm at most 1, such that $\bar{w}(\psi(\bar{v})) \geq 0$. Then for every $v \in V$ and every tangent functional $w \in \bar{V}^*$ of f at v we have $w(\psi(v)) \geq 0$.*

Proof. The first step in the proof is to show that the set $V_0 \subset V$ consisting of those $v \in V$ for which there exists a tangent functional $w \in \overline{V}^*$ of f at v with norm at most 1 such that $w(\psi(v)) \geq 0$, is closed. Since $A \subset V_0$ and A is dense in V , it would follow that $V_0 = V$.

Fix $v \in V$. Let $(v_\alpha)_{\alpha \in \mathcal{D}}$ be a net which takes values on A and converges to v . For each $\alpha \in \mathcal{D}$, choose a tangent functional $w_\alpha \in \overline{V}^*$ of f at v_α with norm at most 1, such that $w_\alpha(\psi(v_\alpha)) \geq 0$. By weak-* compactness of the unit ball in \overline{V}^* we can pass to a subnet and assume that $(w_\alpha)_{\alpha \in \mathcal{D}}$ converges in the weak-* topology to some $w \in \overline{V}^*$. Then w is a tangent functional of f at v whose norm is bounded by 1.

Fix $\varepsilon > 0$. By continuity of ψ , it follows that $(\psi(v_\alpha))_{\alpha \in \mathcal{D}}$ converges to $\psi(v)$. Since all w_α 's have norm at most 1, $(w_\alpha(\psi(v_\alpha)) - w_\alpha(\psi(v)))_{\alpha \in \mathcal{D}}$ converges to 0. Furthermore, as $w_\alpha(\psi(v_\alpha)) \geq 0$, it follows that $\liminf_\alpha w_\alpha(\psi(v)) \geq 0$. Finally, as $(w_\alpha)_{\alpha \in \mathcal{D}}$ converges to w in the weak-* topology, it follows that $(w_\alpha(\psi(v)))_{\alpha \in \mathcal{D}}$ converges to $w(\psi(v))$, so $w(\psi(v)) \geq 0$ and thus $v \in V_0$. This shows that V_0 is closed.

Now fix some arbitrary $v \in V$ and suppose $\omega \in \overline{V}^*$ is a tangent functional for f at v . Our goal is to prove that $\omega(\psi(v)) \geq 0$. Let $\varepsilon > 0$ be arbitrary. It suffices to prove that $\omega(\psi(v)) + \varepsilon > 0$. Choose $v' \in V$ such that $\|\psi(v) - v'\| < \frac{\varepsilon}{2\|\omega\|^*}$. Since $|\omega(\psi(v)) - \omega(v')| \leq \|\omega\|^* \|\psi(v) - v'\| < \frac{\varepsilon}{2}$, it suffices to show that $\omega(v') + \frac{\varepsilon}{2} \geq 0$.

The function $F: \mathbb{R} \rightarrow \mathbb{R}$ given by $F(t) = f(v + tv')$ is convex. Now let $w_t \in \overline{V}^*$ be a tangent functional at $v + tv'$, it follows that for any $t' \in \mathbb{R}$ we have $F(t+t') - F(t) \geq w_t(t'v')$. In particular, if F is differentiable at t then for every tangent w_t of f at $v + tv'$ we have $w_t(v') = F'(t)$. By convexity of F it follows that the left derivative of F at 0, denoted by $F'(0^-)$, satisfies $F'(t) \leq F'(0^-)$ for every $t < 0$ where F is differentiable, which occurs for all but a countable set of t 's. Since $w(v') \geq F'(0^-)$ we have in particular:

$$w(v') \geq \liminf_{t \rightarrow 0^-} w_t(v')$$

Using the result from the first part of the proof, for any $t \in \mathbb{R}$ we may choose a tangent w_t of f at $v + tv'$ with norm at most 1, such that $w_t(\psi(v + tv')) \geq 0$. We have the following estimates:

$$|w_t(\psi(v + tv')) - w_t(v')| \leq \|\psi(v + tv') - v'\| \leq \|\psi(v + tv') - \psi(v)\| + \frac{\varepsilon}{2}.$$

The inequality on the left follows because $\|w_t\|^* \leq 1$ for every $t \in \mathbb{R}$. As both $\psi: V \rightarrow \overline{V}$ and the norm $\|\cdot\|$ on V are continuous, it follows that $\|\psi(v + tv') - \psi(v)\|$ converges to 0 as t converges to 0. From here we obtain that $\liminf_{t \rightarrow 0^-} w_t(v') + \frac{\varepsilon}{2} \geq 0$. Thus, $w(v') + \frac{\varepsilon}{2} \geq 0$. \square

The original argument of Lanford and Ruelle from [26] implicitly proves a rather similar statement. A fundamental difference is that in the argument involved in their article, the topological vector space V has additional properties that guarantee that the convex function f is differentiable at a residual set of points. Next a variant of Mazur's lemma proven in [25] is invoked to show that every tangent functional of f at any $v \in V$ is in the closed convex-hull of tangent functionals at differentiable points, from which the conclusion follows. In contrast, in Proposition 5.1 above we do not assume much knowledge about the topological and functional-analytic properties of the topological vector space V itself.

In order to apply Proposition 5.1 to our setting, we need the following characterization of measures which are étale Gibbs:

Lemma 5.2. *Let $f: X \rightarrow \mathbb{R}$ be such that the $\mathcal{T}(X)$ -cocycle Ψ_f is well defined and continuous. A probability measure μ is étale Gibbs for f if and only if for any homeomorphism $\gamma: X \rightarrow X$ such that $(x, \gamma(x)) \in \mathcal{T}^0(X)$ for μ -almost every $x \in X$, and any $h \in C(X)$ we have*

$$\int_X h(\gamma(x)) - \exp(\Psi_f(x, \gamma(x)))h(x)d\mu \geq 0.$$

Proof. On the one hand, if μ is étale Gibbs for f , it follows that

$$D_{\mu, \mathcal{T}^0(X)}(x, y) = \exp(\Psi_f(x, y)) \text{ up to a } \mu\text{-null set.}$$

Therefore for any homeomorphism γ of X which preserves $\mathcal{T}^0(X)$ and $h \in C(X)$ we have

$$\int_X h(\gamma(x))d\mu = \int_X h(x)d(\mu \circ \gamma) = \int_X \exp(\Psi_f(x, \gamma(x)))h(x)d\mu.$$

From here the inequality follows (as an equality). Conversely, replacing h by $-h$, for every $h \in C(X)$ and homeomorphism γ of X which preserves $\mathcal{T}^0(X)$, we obtain the equality

$$\int_X h(\gamma(x))d\mu = \int_X \exp(\Psi_f(x, \gamma(x)))h(x)d\mu.$$

Because μ is a Borel measure on X , the Radon-Nikodým derivative of a Borel measure $\nu \ll \mu$ is characterized (up to a μ -null set) by the property that $\int h d\nu = \int h \frac{d\nu}{d\mu} d\mu$ for any $h \in C(X)$. It follows that for any homeomorphism γ of X which preserves $\mathcal{T}^0(X)$ we have $\frac{d\mu \circ \gamma}{d\mu}(x) = \exp(\Psi_f(x, \gamma(x)))$ for μ -almost every $x \in X$. Because the equivalence relation $\mathcal{T}^0(X)$ is the orbit equivalence of the homeomorphisms $I_{p,q}$, it follows that $\exp(\Psi_f(x, y))$ is the Radon-Nikodým cocycle of μ with respect to the equivalence relation $\mathcal{T}^0(X)$ and thus that μ is étale Gibbs for f . \square

Let us recall that the metric $\rho_{\mathcal{T}^0(X)}$ is defined on all $f, g \in C(X)$ such that the $\mathcal{T}^0(X)$ -cocycles Ψ_f, Ψ_g are well defined and continuous. It is given by

$$\rho_{\mathcal{T}^0(X)}(f, g) = \|f - g\|_\infty + d_{\mathcal{T}^0(X)}(\Psi_f, \Psi_g).$$

Where $d_{\mathcal{T}^0(X)}$ is the metric from Definition 2.5.

The definition of $\rho_{\mathcal{T}^0(X)}$ implicitly sets the stage for a space of potentials where our extension scheme applies.

Theorem 5.3. (Theorem B) *Let Σ be a sofic approximation sequence for Γ , X be a subshift such that $h_\Sigma(\Gamma \curvearrowright X) \geq 0$, and let $f \in C(X)$ be a $\rho_{\mathcal{T}^0(X)}$ -limit of locally constant functions. Then any equilibrium measure μ on X for f with respect to Σ is étale Gibbs with respect to f .*

Proof. Let V denote the space of functions $f \in C(X)$ that are $\rho_{\mathcal{T}^0(X)}$ -limit of locally constant functions. Then $\rho_{\mathcal{T}^0(X)}$ induces a metric on V which makes it a topological vector space. It is clear that the locally constant functions are dense in V and that the uniform norm is continuous on V . Since the locally constant functions are also dense in $C(X)$, it follows that the metric completion \bar{V} of V with respect to the uniform norm is (up to a natural isomorphism) $C(X)$. Also the map $f \mapsto \Psi_f$ is well defined on V and defines a continuous linear function from V to the space of continuous $\mathcal{T}^0(X)$ -cocycles.

By Proposition 3.10, the map $\Pi: C(X) \rightarrow \mathbb{R}$ given by $\Pi(f) = P_\Sigma(\Gamma \curvearrowright X, f)$ is a convex continuous map and the equilibrium measures for $f \in C(X)$ are all tangent functionals of Π at f . Choose $f_0 \in V$ and let γ be a homeomorphism of X such that $(x, \gamma(x)) \in \mathcal{T}^0(X)$ for every $x \in X$, and let $h \in C(X)$.

Let A be the subspace of V given by the locally constant functions and $\psi: V \rightarrow C(X)$ be given by $\psi(f)(x) = h(\gamma(x)) - \Psi_f(x, \gamma(x))h(x)$. Continuity of ψ follows from continuity of the map $f \mapsto \Psi_f$ on V . Since any subshift is expansive the measure-theoretic sofic entropy is an upper semi-continuous function of the measure (as in [17, Theorem 2.1]). In particular for any $f \in C(X)$ there exists an equilibrium measure on X (with respect to Σ). Fix $f_1 \in A$, by Theorem 4.9, it follows that any equilibrium measure μ_1 at f_1 is étale Gibbs with respect to Σ . By Proposition 3.10 any equilibrium measure at f_1 is a tangent functional μ_1 for the topological sofic pressure function $\Pi: C(X) \rightarrow [-\infty, +\infty]$. By Lemma 5.2 this implies that an equilibrium measure μ_1 at f_1 satisfies

$$\mu_1(\psi(f_1)) = \int_X h(\gamma(x)) - \Psi_{f_1}(x, \gamma(x))h(x)d\mu_1 \geq 0.$$

Applying Proposition 5.1, it follows that for every $f_0 \in V$ and every tangent functional μ for Π at f_0 we have

$$\mu(\psi(f_0)) = \int_X h(\gamma(x)) - \Psi_{f_0}(x, \gamma(x))h(x)d\mu \geq 0.$$

It follows, again by Lemma 5.2, that every equilibrium measure μ at $f_0 \in V$ is étale Gibbs for f_0 with respect to Σ . \square

5.1 Absolutely summable interactions

Definition 5.4. An *interaction* on a space of configurations X is a real valued function on the X -admissible patterns $P(X)$. Given an interaction $\Phi: P(X) \rightarrow \mathbb{R}$ and $F \in \Gamma$ we denote by Φ_F the restriction of Φ to $P_F(X)$. An interaction Φ on X a shift space X is *translation-invariant* if $\Phi(p) = \Phi(g(p))$ for every $g \in \Gamma$ and $p \in P(X)$. An interaction Φ is called *absolutely-summable* (also called *norm-summable*) if

$$\|\Phi\| := \sum_{1_\Gamma \in F \in \Gamma} \|\Phi_F\|_\infty < +\infty,$$

where $\|\Phi_F\|_\infty = \max_{p \in P_F(X)} |\Phi(p)|$ is the uniform norm on the (finite-dimensional) space $\mathbb{R}^{P_F(X)}$.

A translation-invariant interaction Φ is of *finite-range* if there exists $F_0 \in \Gamma$ such that Φ_F is identically zero whenever F is not contained in gF_0 for some $g \in \Gamma$. When Γ is a finitely generated group, this is equivalent to the statement that Φ is supported on patterns whose supports have bounded diameter.

The space of absolutely-summable interactions is a Banach space with respect to the above norm [6, Proposition 2.19]. The subspace of finite-range interactions is dense.

We now discuss a procedure to obtain a continuous map $f_\Phi: X \rightarrow \mathbb{R}$ from a absolutely-summable interaction Φ , following Ruelle [37] (in the context of \mathbb{Z}^d -subshifts). Let $\mathcal{P}_c(\Gamma)$ denote the space of finite non-empty subsets of Γ . Let $\widetilde{\mathcal{P}}_c(\Gamma) \subset \mathcal{P}_c(\Gamma)$ be a subset containing exactly one representative in each Γ -orbit. Namely, for every $F \in \mathcal{P}_c(\Gamma)$ the set $\{gF : g \in \Gamma\}$ contains exactly one element of $\widetilde{\mathcal{P}}_c(\Gamma)$. Let $\text{Fix}_\Gamma(F)$ be the set of $g \in \Gamma$ such that $gF = F$. Define $f_\Phi: X \rightarrow \mathbb{R}$ as

$$f_\Phi(x) := \sum_{F \in \widetilde{\mathcal{P}}_c(\Gamma)} \frac{1}{|\text{Fix}_\Gamma(F)|} \Phi(x|_F).$$

Let us remark that in the case where Γ has no torsion (as in the case of \mathbb{Z}^d in the work of Ruelle), then $\text{Fix}_\Gamma(F)$ is trivial and can be removed from the equation above.

An alternative, and perhaps more natural, way of defining the potential is to just average over all translates of a finite set which contain the identity, namely, we can define $h_\Phi: X \rightarrow \mathbb{R}$ by

$$h_\Phi(x) := \sum_{1_\Gamma \in F \in \Gamma} \frac{1}{|F|} \Phi(x|_F).$$

Let us notice that these two definitions are in general different, and it is not even clear that they define the same class of functions. However, they do define the same $\mathcal{T}(X)$ -cocycle whenever it is well-defined. Indeed,

$$\begin{aligned}
\Psi_{f_\Phi}(x, y) &= \sum_{g \in \Gamma} (f_\Phi(gy) - f_\Phi(gx)) \\
&= \sum_{g \in \Gamma} \sum_{F \in \widetilde{\mathcal{P}}_c(\Gamma)} \frac{1}{|\text{Fix}_\Gamma(F)|} (\Phi(y|_{g^{-1}F}) - \Phi(x|_{g^{-1}F})) \\
&= \sum_{F \in \Gamma} (\Phi(y|_F) - \Phi(x|_F)) \\
&= \sum_{g \in \Gamma} \sum_{g \in F \in \Gamma} \frac{1}{|F|} (\Phi(y|_F) - \Phi(x|_F)) \\
&= \sum_{g \in \Gamma} \sum_{1_\Gamma \in F \in \Gamma} \frac{1}{|F|} (\Phi(y|_{g^{-1}F}) - \Phi(x|_{g^{-1}F})) \\
&= \sum_{g \in \Gamma} (h_\Phi(gy) - h_\Phi(gx)) = \Psi_{h_\Phi}(x, y)
\end{aligned}$$

Consequently, any Gibbs measure for f_Φ is a Gibbs measure for h_Φ and vice versa. In what follows we will use Ruelle's definition, f_Φ , because it has the advantage that all locally constant maps can be obtained from finite-range interactions.

Given any interaction Φ for which f_Φ can be defined, we will say that a measure on X is (étale) Gibbs with respect to Φ if it is (étale) Gibbs with respect to f_Φ . This definition is consistent with the original definition of Lanford and Ruelle [26].

Remark 5.5. Any finite-range interaction Φ induces a locally constant function $f_\Phi : X \rightarrow \mathbb{R}$. Therefore Theorem 4.9 holds for finite-range interactions.

The potential f_Φ is continuous. Furthermore, the map $\Phi \mapsto f_\Phi$ is a bounded linear map from the Banach space of absolutely-summable interactions to the space of continuous functions with the uniform norm.

To simplify the notation, let us denote the $\mathcal{T}(X)$ -cocycle by Ψ_Φ instead of Ψ_{f_Φ} and notice that for every $(x, y) \in \mathcal{T}(X)$ we have

$$\Psi_\Phi(x, y) = \sum_{F \in \Gamma} (\Phi(y|_F) - \Phi(x|_F)).$$

Whenever Φ is absolutely-summable and translation-invariant, one can show that for any $(x, y) \in \mathcal{T}_F(X)$ we have that $|\Psi_\Phi(x, y)| \leq 2|F|\|\Phi\|$ (see Proposition 3.1 of [6]) and therefore Ψ_Φ is a well-defined Γ -invariant $\mathcal{T}(X)$ -cocycle. A direct computation shows that Ψ_Φ is continuous.

Let

$$\text{NS}(X) := \{f_\Phi : \Phi \text{ is an absolutely-summable, translation-invariant interaction on } X\}.$$

Let us endow $\text{NS}(X)$ with the norm given by

$$\|f\|_{\text{NS}} := \inf\{\|\Phi\| : f = f_\Phi\}.$$

It is straightforward to verify that $\|\cdot\|_{\text{NS}}$ is a semi-norm. The fact that it is a norm follows from the fact that $\|f_\Phi\|_\infty \leq \|f_\Phi\|_{\text{NS}}$, which is direct from the definition of f_Φ .

First, let us show that $f_\Phi \mapsto \Psi_\Phi$ is a continuous linear map from $\text{NS}(X)$ to the topological vector space of continuous, translation-invariant $\mathcal{T}^0(X)$ -cocycles. Let $f \in \text{NS}(X)$ and choose Φ such that $f = f_\Phi$. Let also $F \in \Gamma$. By Proposition 3.1 of [6] we have that for any $(x, y) \in \mathcal{T}_F(X)$ we have

$$|\Psi_f(x, y)| = |\Psi_\Phi(x, y)| \leq 2|F|\|\Phi\|.$$

In particular, the bound is still valid for $(x, y) \in \mathcal{T}_F^0(X)$. Taking infimum over all Φ such that $f = f_\Phi$ we get that

$$|\Psi_f(x, y)| \leq 2|F|\|f\|_{\text{NS}} \text{ for every } (x, y) \in \mathcal{T}_F^0(X).$$

Therefore the map $f_\Phi \mapsto \Psi_\Phi$ is continuous.

Proposition 5.6. *Any $f \in \text{NS}(X)$ is a $\rho_{\mathcal{T}^0(X)}$ -limit of locally constant functions.*

Proof. Let Φ be a translation-invariant, absolutely summable interaction with $f = f_\Phi$. Let $(F_n)_{n \in \mathbb{N}}$ be an enumeration of all finite subsets of Γ modulo translation. For each $n \geq 1$ consider the interaction $\Phi^{(n)}$ which coincides with Φ on all patterns whose supports is a translation of F_k for $k \leq n$, and is 0 everywhere else. It follows that $\Phi^{(n)}$ is a finite-range, translation-invariant interaction and that $f_n := f_{\Phi^{(n)}}$ is a locally constant function. To complete the proof, it suffices to show that $\lim_{n \rightarrow \infty} \rho_{\mathcal{T}^0(X)}(f_n, f) = 0$. Indeed, as Φ is absolutely-summable, we have that

$$\lim_{n \rightarrow \infty} \|\Phi - \Phi^{(n)}\| = \sum_{k=n+1}^{\infty} \|\Phi_{F_k}\|_{\infty} = 0.$$

From this it follows that

$$d_{\mathcal{T}^0(X)}(\Psi_{f_n}, \Psi_f) \leq \sum_{j=1}^{\infty} \frac{1}{2^j} \min \left\{ 1, 2|F_j| \|\Phi - \Phi^{(n)}\| \right\} \rightarrow 0 \text{ as } n \rightarrow \infty.$$

Also, $\|f - f_n\|_{\infty} \leq \|f - f_n\|_{\text{NS}} \rightarrow 0$ as $n \rightarrow \infty$. This shows that $\lim_{n \rightarrow \infty} \rho_{\mathcal{T}^0(X)}(f_n, f) = 0$. \square

Theorem 5.7. (Theorem A) *Let Σ be a sofic approximation sequence for Γ , X be a subshift that satisfies the topological Markov property such that $h_{\Sigma}(\Gamma \curvearrowright X) \geq 0$, Φ an absolutely-summable interaction on X and μ an equilibrium measure on X for Φ with respect to Σ . Then μ is Gibbs with respect to Φ .*

Proof. The proof follows from putting together Theorem B, Proposition 5.6 and Proposition 2.11. Indeed, by Proposition 5.6 we have that any map induced by a translation-invariant, absolutely-summable interaction is a $\rho_{\mathcal{T}^0(X)}$ -limit of locally constant functions, and thus from Theorem B it follows that every equilibrium measure is étale Gibbs. Finally, as the space satisfies the topological Markov property, Proposition 2.11 tells us that $\mathcal{T}(X) = \mathcal{T}^0(X)$ and thus we may replace “étale Gibbs” by “Gibbs”. \square

5.2 Functions with \mathbb{F} -summable variation

Some authors rather than using a space of interactions, directly use a space of continuous functions $f: X \rightarrow \mathbb{R}$ to model “the potential energy”, for instance [27, 23]. In this setting, we introduce the notion of functions with summable variation with respect to a *filtration* of Γ , that is, an increasing sequence of finite subsets of Γ which cover it. This notion of convergence generalizes the concept of d -summable variation of [28], the notion of “regular local energy functions” of [23] and the notion of “shell-regular potentials” of [8]. We show that every function with this property satisfies the equivalent statement of the Lanford–Ruelle theorem.

We now define a class of potentials which generalizes the space of functions having “ d -summable variation” in the case $\Gamma = \mathbb{Z}^d$ as in [27] and show that they are $\rho_{\mathcal{T}^0(X)}$ -limits of locally constant functions, thus recovering [27, Theorem 3.1] and generalizing it the case where the acting group is an arbitrary sofic group.

Definition 5.8. We call a sequence $\mathbb{F} = (F_n)_{n=1}^\infty$ of finite subsets of Γ a *filtration* of Γ if $F_n \subset F_{n+1}$ for all n and $\Gamma = \bigcup_{n=1}^\infty F_n$. Given $f: X \rightarrow \mathbb{R}$ and $S \in \Gamma$ denote

$$\text{Var}_S(f) := \sup \{|f(x) - f(y)| : x, y \in X \text{ and } x|_S = y|_S\},$$

Given a filtration $\mathbb{F} = (F_n)_{n=1}^\infty$ of Γ and $S \in \Gamma$, let

$$\|f\|_{\text{sv}(\mathbb{F}), S} := \sum_{n=1}^\infty |F_{n+1}S \setminus F_nS| \text{Var}_{F_n}(f).$$

A function $f: X \rightarrow \mathbb{R}$ has \mathbb{F} -*summable variation* if $\|f\|_{\text{sv}(\mathbb{F}), S} < \infty$ for all $S \in \Gamma$. We denote the space of functions on X with \mathbb{F} -summable variation by $\text{SV}_{\mathbb{F}}(X)$.

One may verify that any function with \mathbb{F} -summable variation is continuous, and that any locally constant function has \mathbb{F} -summable variation with respect to any filtration of Γ . We endow the vector space $\text{SV}_{\mathbb{F}}(X)$ with the topology induced by the countable collection of semi-norms $(\|\cdot\|_{\text{sv}(\mathbb{F}), S})_{S \in \Gamma}$ together with the uniform norm $\|\cdot\|_\infty$. In general, $\text{SV}_{\mathbb{F}}(X)$ is not a Banach space, but it may be verified that it is always a Fréchet space (we shall not make use of this property).

Particularly, when $\Gamma = \mathbb{Z}^d$ and $\mathbb{F} = (F_n)_{n=1}^\infty$ is the filtration of \mathbb{Z}^d given by $F_n = \{-n, \dots, n\}^d$, then $\|f\|_{\text{sv}(\mathbb{F}), S} < \infty$ for all $S \in \mathbb{Z}^d$ if and only if $\sum_{n=1}^\infty n^{d-1} \text{Var}_{F_n}(f) < +\infty$, so in that case $\text{SV}_{\mathbb{F}}(X)$ is exactly the space of function with d -summable variation as defined in [27] or the space of “regular local energy functions” as defined in [23, Section 5]. More generally, if Γ is a finitely generated group with “bounded sphere ratios” and \mathbb{F} is the filtration corresponding to balls with respect to some finite symmetric generating set, then $\text{SV}_{\mathbb{F}}(X)$ is exactly the space of “shell-regular potentials” as defined in [8, Section 5].

For the rest of this section let $\mathbb{F} = (F_n)_{n=1}^\infty$ be an arbitrary filtration on the countable group Γ . We shall prove that any $f \in \text{SV}_{\mathbb{F}}(X)$ is a $\rho_{\mathcal{T}^0(X)}$ -limit of locally constant functions. It is clear that convergence in $\text{SV}_{\mathbb{F}}(X)$ implies $\|\cdot\|_\infty$ -convergence, so we only need to verify that locally constant functions are dense, and that the map $f \mapsto \Psi_f$ from $\text{SV}_{\mathbb{F}}(X)$ to the space of continuous $\mathcal{T}^0(X)$ -cocycles is continuous. Let us mention that these two proofs extend almost verbatim from the well known “classical case” $\Gamma = \mathbb{Z}^d$, and $\mathbb{F} = (\{-n, \dots, n\}^d)_{n=1}^\infty$.

Lemma 5.9. *The set of locally constant functions is dense in $\text{SV}_{\mathbb{F}}(X)$.*

Proof. Let $f \in \text{SV}_{\mathbb{F}}(X)$. For every $n \geq 1$ consider the locally constant function $f_n: X \rightarrow \mathbb{R}$ given by

$$f_n(x) := \sup\{f(y) : y|_{F_n} = x|_{F_n}\}.$$

We claim that $(f_n)_{n=1}^\infty$ converges to f in $\text{SV}_{\mathbb{F}}(X)$. Indeed, as f is continuous it is clear that $\|f - f_n\|_\infty$ converges to 0. Therefore it suffices to check that $\|f - f_n\|_{\text{sv}(\mathbb{F}), S}$ converges to 0 for every $S \in \Gamma$. This follows directly from the fact that $\text{Var}_{F_m}(f - f_n) \leq 2 \text{Var}_{F_m}(f)$ for every $m \leq n$. \square

Lemma 5.10. *For every function $f: X \rightarrow \mathbb{R}$ with \mathbb{F} -summable variation, $S \in \Gamma$ and $(x, y) \in \mathcal{T}_S(X)$, the series defining $\Psi(x, y)$ in Definition 2.4 is absolutely convergent. Hence, Ψ_f is a well-defined Γ -invariant $\mathcal{T}(X)$ -cocycle and is furthermore continuous. Moreover, the map $f \mapsto \Psi_f$ from $\text{SV}_{\mathbb{F}}(X)$ to the space of continuous $\mathcal{T}^0(X)$ -cocycles is continuous. Consequently, convergence in $\text{SV}_{\mathbb{F}}(X)$ implies $\rho_{\mathcal{T}^0(X)}$ -convergence.*

Proof. Let $f \in \text{SV}_{\mathbb{F}}(X)$, $S \in \Gamma$ and $(x, y) \in \mathcal{T}_S(X)$. Let us write $F_0 = \emptyset$. As \mathbb{F} is a filtration we may write Γ as the disjoint union

$$\Gamma = \bigcup_{n=0}^\infty (F_{n+1}S^{-1} \setminus F_nS^{-1}).$$

Therefore we may write

$$\sum_{g \in \Gamma} |f(gy) - f(gx)| = \sum_{n=0}^{\infty} \sum_{g \in F_{n+1}S^{-1} \setminus F_nS^{-1}} |f(gy) - f(gx)|.$$

Now for every $g \in \Gamma$, we have the trivial bound $|f(gy) - f(gx)| \leq 2\|f\|_{\infty}$. Moreover, as $(x, y) \in \mathcal{T}_S(X)$ it follows that if $g \in \Gamma \setminus F_nS^{-1}$ then $(gx)|_{F_n} = (gy)|_{F_n}$ and thus $|f(gy) - f(gx)| \leq \text{Var}_{F_n}(f)$, so

$$\begin{aligned} \sum_{g \in \Gamma} |f(gy) - f(gx)| &\leq 2|F_1S^{-1}|\|f\|_{\infty} + \sum_{n=1}^{\infty} |F_{n+1}S^{-1} \setminus F_nS^{-1}| \text{Var}_{F_n}(f) \\ &= 2|F_1S^{-1}|\|f\|_{\infty} + \|f\|_{\text{sv}(\mathbb{F}), S^{-1}}. \end{aligned}$$

We conclude that the series defining $\Psi_f(x, y)$ is indeed absolutely convergent for every $S \in \Gamma$ and $(x, y) \in \mathcal{T}_S(X)$.

Let us now show that $\Psi_f: \mathcal{T}(X) \rightarrow \mathbb{R}$ is continuous. As explained in Section 2.2, it suffices to show that $\Psi_f|_{\mathcal{T}_S(X)}$ is continuous for every $S \in \Gamma$. More explicitly, we need to show that for every $S \in \Gamma$ and $\varepsilon > 0$ there exists $K \in \Gamma$ such that if $(x, y), (x', y') \in \mathcal{T}_S(X)$ are such that $x|_K = x'|_K$ and $y|_K = y'|_K$ then $|\Psi_f(x, y) - \Psi_f(x', y')| \leq \varepsilon$.

Fix $S \in \Gamma$ and $\varepsilon > 0$. Let $N, M \in \mathbb{N}$ and choose $K \in \Gamma$ such that $SF_N^{-1}F_M \in K$. Let $(x, y), (x', y') \in \mathcal{T}_S(X)$ such that $x|_K = x'|_K$ and $y|_K = y'|_K$. Then it follows that for every $g \in F_NS^{-1}$ we have $g(x)|_{F_M} = g(x')|_{F_M}$ and $g(y)|_{F_M} = g(y')|_{F_M}$ and thus $|f(gy) - f(gy')| \leq \text{Var}_{F_M}(f)$ and $|f(x) - f(x')| \leq \text{Var}_{F_M}(f)$. Since $(x, x') \in \mathcal{T}_S(X)$ and $(y, y') \in \mathcal{T}_S(X)$, for every $n \in \mathbb{N}$ and $g \in \Gamma \setminus F_nS^{-1}$ we have $|f(gy) - f(gx)| \leq \text{Var}_{F_n}(f)$ and $|f(gy') - f(gx')| \leq \text{Var}_{F_n}(f)$.

Putting all the above bounds together yields:

$$\begin{aligned} |\Psi_f(x, y) - \Psi_f(x', y')| &\leq \sum_{g \in F_NS^{-1}} (|f(gy) - f(gy')| + |f(gx) - f(gx')|) \\ &\quad + \sum_{n=N}^{\infty} \sum_{g \in F_{n+1}S^{-1} \setminus F_nS^{-1}} (|f(gy) - f(gx)| + |f(gy') - f(gx')|) \\ &\leq 2|F_NS^{-1}| \text{Var}_{F_M}(f) + 2 \sum_{n=N}^{\infty} |F_{n+1}S^{-1} \setminus F_nS^{-1}| \text{Var}_{F_n}(f). \end{aligned}$$

Since $f \in \text{SV}_{\mathbb{F}}(X)$, there exists $N \in \mathbb{N}$ such that $2 \sum_{n=N}^{\infty} |F_{n+1}S^{-1} \setminus F_nS^{-1}| \text{Var}_{F_n}(f) \leq \frac{\varepsilon}{2}$, and there exists $M \in \mathbb{N}$ such that $2|F_NS^{-1}| \text{Var}_{F_M}(f) \leq \frac{\varepsilon}{2}$. Now for any $K \in \Gamma$ satisfying $SF_N^{-1}F_M \subset K$, it follows that if $(x, y), (x', y') \in \mathcal{T}_S(X)$ are such that $x|_K = x'|_K$ and $y|_K = y'|_K$ then $|\Psi_f(x, y) - \Psi_f(x', y')| \leq \varepsilon$. This proves that Ψ_f is continuous.

Finally, we show that the map $f \mapsto \Psi_f$ from $\text{SV}_{\mathbb{F}}(X)$ to the space of continuous $\mathcal{T}^0(X)$ -cocycles is continuous. Since $f \mapsto \Psi_f$ is linear, it suffices to prove continuity at 0 for every restriction of the domain of the image to $\mathcal{T}_S^0(X)$. This follows directly from the estimate

$$|\Psi_f(x, y)| \leq 2|F_1S^{-1}|\|f\|_{\infty} + \|f\|_{\text{sv}(\mathbb{F}), S^{-1}} \text{ for every } (x, y) \in \mathcal{T}_S^0(X).$$

□

Theorem 5.11. *Let Σ be a sofic approximation sequence for Γ , X be a subshift with the topological Markov property such that $h_{\Sigma}(\Gamma \curvearrowright X) \geq 0$, $f: X \rightarrow \mathbb{R}$ a function with \mathbb{F} -summable variation with respect to a filtration \mathbb{F} of Γ and μ an equilibrium measure on X for f with respect to Σ . Then μ is Gibbs with respect to f .*

Proof. The combination of Lemmas 5.9 and 5.10 along with continuity of the uniform norm shows any function with \mathbb{F} -summable variation is a $\rho_{\mathcal{T}^0(X)}$ -limit of locally constant functions, so the result follows from Theorem B and Proposition 2.11. □

6 Applications

In this brief section we present a few applications of our main theorem. Our first result concerns dynamical systems which admit a unique Gibbs measure with respect to some $f \in C(X)$. There are several results in the literature ensuring uniqueness of Gibbs measures, see for instance [19, 45].

Theorem 6.1. *Let Γ be a sofic group and $X \subset \mathcal{A}^\Gamma$ be a subshift with the topological Markov property which admits a unique Gibbs measure μ with respect to $f \in C(X)$ which is a $\rho_{\mathcal{T}^0(X)}$ -limit of locally constant functions. For any sofic approximation sequence Σ of Γ such that $h_\Sigma(\Gamma \curvearrowright X) \geq 0$, we have that μ is translation-invariant and furthermore, it is the unique equilibrium measure on X for f with respect to Σ .*

Proof. Let us fix Σ such that $h_\Sigma(\Gamma \curvearrowright X) \geq 0$. By Proposition 3.7 it follows that the entropy function is upper semi-continuous and thus there exists a (translation-invariant) equilibrium measure on X for f . Theorem B implies every equilibrium measure on X for f is Gibbs with respect to f and thus we conclude that any such measure must coincide with the unique Gibbs measure μ with respect to f . \square

A consequence of Theorem 6.1, which has already been mentioned in the introduction, is uniqueness of the equilibrium measure, and thus independence of the equilibrium measure on the sofic approximation sequence, for single-site potentials over a full shift: suppose $f: \mathcal{A}^\Gamma \rightarrow \mathbb{R}$ is a single-site potential on \mathcal{A}^Γ , that is, $f(x) = f(y)$ if $x(1_\Gamma) = y(1_\Gamma)$ (and thus we may identify f with a function $f: \mathcal{A} \rightarrow \mathbb{R}$). Then there is a unique Gibbs measure μ_f with respect to f which is Bernoulli and given by

$$\mu_f(\{x \in \mathcal{A}^\Gamma : x(g) = a\}) = \frac{\exp(f(a))}{\sum_{b \in \mathcal{A}} \exp(f(b))} \text{ for every } g \in \Gamma.$$

Theorem 6.1 yields that μ_f is indeed the unique equilibrium measure on \mathcal{A}^Γ for the single-site potential f . As mentioned in the introduction, the question of uniqueness in this setting was asked in [16, Question 5.4] and answered via a direct argument in [10, Example 7]. It also received a positive answer via the theory of Rokhlin entropy [40, Corollary 3.6].

Theorem 6.1 can be combined with more sophisticated criteria for uniqueness of Gibbs measures to obtain uniqueness of equilibrium measures on more complicated systems. For instance, van der Berg's percolation criterion in [45, Corollary 1] implies that for any finitely generated sofic group, the Ising model has a unique Gibbs measure at sufficiently high temperature. Theorem 6.1 then implies uniqueness of the equilibrium measure.

A second application is the existence of Gibbs measures for any subshift $X \subset \mathcal{A}^\Gamma$ with the topological Markov property for which there is some sofic approximation sequence with nonnegative sofic topological entropy. This generalizes a result of Alpeev [2].

Theorem 6.2. *Let Γ be a sofic group and $X \subset \mathcal{A}^\Gamma$ be a subshift with the topological Markov property for which there exists a sofic approximation sequence Σ of Γ with $h_\Sigma(\Gamma \curvearrowright X) \geq 0$. For any $f \in C(X)$ which is a $\rho_{\mathcal{T}^0(X)}$ -limit of locally constant functions there exists a Gibbs measure on X with respect to f .*

Proof. Let us fix Σ such that $h_\Sigma(\Gamma \curvearrowright X) \geq 0$. By Proposition 3.7 the entropy function is upper semi-continuous and thus there exists an equilibrium measure μ on X for f with respect to Σ . By Theorem B it follows that μ must be Gibbs with respect to f . \square

In what follows we consider an algebraic variant of shift spaces. Let H be a finite group and consider a subshift $X \subset H^\Gamma$. We say that X is a *group shift* if X forms a group under the operation induced by H operating pointwise on every $g \in \Gamma$. In [7, Proposition 5.1] it was shown that every group shift satisfies the topological Markov property (in fact, every algebraic action satisfies the non-symbolic variant of the topological Markov property, see [5, Proposition 4.1]). Also, it is clear that for any sofic

approximation sequence Σ we have $h_\Sigma(\Gamma \curvearrowright X) \geq 0$ as X contains a constant configuration, namely its identity configuration e which satisfies $e(g) = 1_H$ for every $g \in \Gamma$.

Given a group shift, the *homoclinic group* $\Delta(X)$ of X is the set of all $y \in X$ such that $(y, e) \in \mathcal{T}(X)$. Is is not hard to check (see [7, Proposition 5.2]) that translation-invariant Gibbs measures for $f \equiv 0$ in X are precisely the translation-invariant Borel probability measures which are invariant under multiplication by any element of $\Delta(X)$. From the arguments above and Theorem 4.9 we obtain that equilibrium measures on X are invariant under the action of $\Delta(X)$. Using Theorem 6.1, we have the following extension of [7, Corollary 5.4] to groups shifts over sofic groups:

Theorem 6.3. *Let Γ be a sofic group and let $X \subset H^\Gamma$ be a group shift. Suppose the homoclinic group $\Delta(X)$ is dense in X , then the Haar measure on X is the unique measure of maximal entropy with respect to every sofic approximation sequence Σ .*

Proof. Let ν be a Gibbs measure on X . By [7, Proposition 5.2], ν is invariant under multiplication by any element of $\Delta(X)$ which is dense, and thus is invariant under multiplication by any element of X . It follows that the unique Gibbs measure is the Haar measure. By Theorem 6.1 it follows that there is a unique measure of maximal entropy with respect to every Σ and it coincides with the Haar measure. \square

7 Beyond sofic groups

To the knowledge of the authors, the existence of a non-sofic group is still open. Despite this fact, our main result suggests that it is reasonable to ask whether there is a meaningful version of the Lanford–Ruelle theorem that holds beyond sofic groups, perhaps for actions of all countable groups. In this short final section we point out that a hypothetical consequence of such a generalization would be a positive answer to Gottschalk’s conjecture stating that every injective, continuous and Γ -equivariant map on A^Γ is surjective. The core of the argument is rather old, and comes from the notion of *intrinsic ergodicity*.

To formulate a Lanford–Ruelle theorem for group actions, one needs a reasonable notion of equilibrium measure, which implicitly depends on a notion of topological pressure which satisfies a variational principle. For this theoretical application we need only use the potential 0, and thus we shall rather speak of topological and measure-theoretic entropy.

Suppose Γ is a topological group and $\Gamma \curvearrowright (X, \mu)$ is a probability measure preserving action. A *measure-theoretic entropy theory* for Γ is an assignment

$$\Gamma \curvearrowright (X, \mu) \mapsto h(\Gamma \curvearrowright X, \mu) \in [-\infty, +\infty].$$

which is invariant under measure-theoretic isomorphism. A *topological entropy theory* for Γ is an assignment

$$\Gamma \curvearrowright X \mapsto h(\Gamma \curvearrowright X) \in [-\infty, +\infty].$$

which is invariant under topological isomorphism. Given a measure-theoretic entropy theory, one can obtain a topological entropy theory by imposing a variational principle as follows:

$$h(\Gamma \curvearrowright X) = \sup_{\mu \in \text{Prob}_\Gamma(X)} h(\Gamma \curvearrowright X, \mu).$$

A topological entropy theory which satisfies a variational principle with respect to a measure-theoretic entropy theory is always monotone with respect to subsystems in the sense that for any closed (in fact, measurable) Γ -invariant set $X_0 \subset X$ we have $h(\Gamma \curvearrowright X_0) \leq h(\Gamma \curvearrowright X)$.

A Lanford–Ruelle theorem for an entropy theory would state that for every Γ -SFT $X \subset A^\Gamma$ (hopefully, any subshift with the topological Markov property), any $\mu \in \text{Prob}_\Gamma(X)$ such that $h(\Gamma \curvearrowright X) =$

$h(\Gamma \curvearrowright X, \mu)$ would necessarily be a Gibbs state for $f = 0$. Then uniqueness of the Gibbs measure on A^Γ for $f = 0$ would imply that the uniform Bernoulli measure is the unique measure of maximal entropy for A^Γ . As this measure is not supported in any strict subset of A^Γ , it follows that any strict subsystem X_0 of X has strictly lower topological entropy. This precludes the possibility of equivariantly and continuously embedding A^Γ into a proper subsystem, thus settling Gottschalk's conjecture for the group Γ .

One available measure-theoretic entropy which works in any group is *Rokhlin entropy* [39]. The Rokhlin entropy of a free, ergodic, probability measure preserving action $\Gamma \curvearrowright (X, \mu)$ is given by

$$h_R(\Gamma \curvearrowright X, \mu) := \inf_{\mathcal{P}} H_\mu(\mathcal{P}),$$

where \mathcal{P} runs over all countable generating partitions of $\Gamma \curvearrowright X$, and $H_\mu(\mathcal{P})$ is the Shannon entropy of the partition.

It is obvious from the definitions that $h_R(\Gamma \curvearrowright X, \mu)$ is invariant under measure-theoretic isomorphism. As a consequence, for any closed (in fact, measurable) Γ -invariant set $X_0 \subset X$ and any $\mu \in \text{Prob}_\Gamma(X)$ satisfying $\mu(X_0) = 1$ we have $h_R(\Gamma \curvearrowright X_0, \mu|_{X_0}) = h_R(\Gamma \curvearrowright X, \mu)$, since $(\Gamma \curvearrowright X_0, \mu|_{X_0})$ is measure-theoretically isomorphic to $(\Gamma \curvearrowright X, \mu)$.

In the case Γ is a countable amenable group, Rokhlin entropy coincides with the usual notion for free ergodic actions. For free actions of sofic groups, Rokhlin entropy dominates the sofic entropy with respect to any sofic approximation sequence of Γ .

Question 7.1. *Does Rokhlin entropy satisfy a Lanford–Ruelle theorem for every group Γ ?*

It seems the answer to the question above might not be easy because any group which satisfies a Lanford–Ruelle theorem for Rokhlin entropy must also admit an essentially free ergodic action with finite nonzero Rokhlin entropy, which is an open question. For a group Γ , one can define $h_R^{sup}(\Gamma)$ as the supremum over all values of h_R on essentially free ergodic actions with finite Rokhlin entropy. It can be proven (see [40]) that if there is a group Γ for which $h_R^{sup}(\Gamma) < +\infty$ then for every group Γ' which is the direct product of Γ with an infinite and locally finite group we have $h_R^{sup}(\Gamma') = 0$. This means that for Γ' every essentially free ergodic action has Rokhlin entropy zero, thus no Lanford–Ruelle theorem for Rokhlin entropy can hold on Γ' . On the other hand, having $h_R^{sup}(\Gamma) = +\infty$ for every group Γ already is enough to settle Gottschalk's conjecture [40] (see also [10]).

S. Barbieri, DEPARTAMENTO DE MATEMÁTICA Y CIENCIA DE LA COMPUTACIÓN, UNIVERSIDAD DE SANTIAGO DE CHILE, SANTIAGO. CHILE.

E-mail address: `sebastian.barbieri@usach.cl`

T. Meyerovitch, DEPARTMENT OF MATHEMATICS, BEN-GURION UNIVERSITY OF THE NEGEV, BE'ER SHEVA, ISRAEL

E-mail address: `mtom@bgu.ac.il`

References

- [1] D. Airey, L. Bowen, and F. Lin. A topological dynamical system with two different positive sofic entropies. *arXiv:1911.08272*, 2019.
- [2] A. Alpeev. The entropy of Gibbs measures on sofic groups. *Journal of Mathematical Sciences*, 215(6):649–658, 2016.
- [3] T. Austin. Additivity properties of sofic entropy and measures on model spaces. *Forum of Mathematics, Sigma*, 4:e25, 2016.

- [4] T. Austin and M. Podder. Gibbs measures over locally tree-like graphs and percolative entropy over infinite regular trees. *Journal of Statistical Physics*, 170(5):932–951, 2018.
- [5] S. Barbieri, F. García-Ramos, and H. Li. Markovian properties of continuous group actions: Algebraic actions, entropy and the homoclinic group. *Advances in Mathematics*, 397:108196, 2022.
- [6] S. Barbieri, R. Gómez, B. Marcus, T. Meyerovitch, and S. Taati. Gibbsian representations of continuous specifications: the theorems of Kozlov and Sullivan revisited. *Communications in Mathematical Physics*, 382(2):1111–1164, 2021.
- [7] S. Barbieri, R. Gómez, B. Marcus, and S. Taati. Equivalence of relative Gibbs and relative equilibrium measures for actions of countable amenable groups. *Nonlinearity*, 33(5):2409–2454, 2020.
- [8] L. Borsato and S. MacDonald. Conformal measures and the Dobrušin-Lanford-Ruelle equations. *Proceedings of the American Mathematical Society*, 149(10):4355–4369, 2021.
- [9] L. Bowen. Measure conjugacy invariants for actions of countable sofic groups. *Journal of the American Mathematical Society*, 23(1):217–245, 2010.
- [10] L. Bowen. Examples in the entropy theory of countable group actions. *Ergodic Theory and Dynamical Systems*, 40(10):2593–2680, 2020.
- [11] R. Burton and J. E. Steif. Non-uniqueness of measures of maximal entropy for subshifts of finite type. *Ergodic Theory and Dynamical Systems*, 14(2):213–235, 1994.
- [12] T. Ceccherini-Silberstein, M. Coornaert, and H. Li. Expansive actions with specification of sofic groups, strong topological markov property, and surjunctivity. *arXiv:2107.12047*, 2021.
- [13] N. Chandgotia. Markov random fields and measures with nearest neighbour Gibbs potential. Master’s thesis, University of British Columbia, 2011.
- [14] N. Chandgotia, G. Han, B. Marcus, T. Meyerovitch, and R. Pavlov. One-dimensional Markov random fields, Markov chains and topological Markov fields. *Proceedings of the American Mathematical Society*, 142(1):227–242, 2014.
- [15] N. Chandgotia and T. Meyerovitch. Markov random fields, Markov cocycles and the 3-colored chessboard. *Israel Journal of Mathematics*, 215(2):909–964, 2016.
- [16] N.-P. Chung. Topological pressure and the variational principle for actions of sofic groups. *Ergodic Theory and Dynamical Systems*, 33(5):1363–1390, 2013.
- [17] N.-P. Chung and G. Zhang. Weak expansiveness for actions of sofic groups. *Journal of Functional Analysis*, 268(11):3534–3565, 2015.
- [18] M. Denker and M. Urbański. On the existence of conformal measures. *Transactions of the American Mathematical Society*, 328(2):563–587, 1991.
- [19] R. L. Dobrušin. Description of a random field by means of conditional probabilities and conditions for its regularity. *Akademiya Nauk SSSR. Teoriya Veroyatnostei i ee Primeneniya*, 13:201–229, 1968.
- [20] J. Feldman and C. C. Moore. Ergodic equivalence relations, cohomology, and von Neumann algebras. I. *Trans. Amer. Math. Soc.*, 234(2):289–324, 1977.
- [21] R. I. Grigorchuk and A. M. Stëpin. Gibbs states on countable groups. *Akademiya Nauk SSSR. Teoriya Veroyatnostei i ee Primeneniya*, 29(2):351–354, 1984.

- [22] M. Gromov. Endomorphisms of symbolic algebraic varieties. *Journal of the European Mathematical Society*, 1(2):109–197, 1999.
- [23] G. Keller. *Equilibrium States in Ergodic Theory*. Cambridge University Press, 1998.
- [24] D. Kerr and H. Li. Entropy and the variational principle for actions of sofic groups. *Inventiones mathematicae*, 186(3):501–558, 2011.
- [25] O. E. Lanford, III and D. W. Robinson. Statistical mechanics of quantum spin systems. III. *Communications in Mathematical Physics*, 9:327–338, 1968.
- [26] O. E. Lanford, III and D. Ruelle. Observables at infinity and states with short range correlations in statistical mechanics. *Communications in Mathematical Physics*, 13:194–215, 1969.
- [27] T. Meyerovitch. Gibbs and equilibrium measures for some families of subshifts. *Ergodic Theory and Dynamical Systems*, 33(3):934–953, 2013.
- [28] T. Meyerovitch. Pseudo-orbit tracing and algebraic actions of countable amenable groups. *Ergodic Theory and Dynamical Systems*, pages 1–22, 2018.
- [29] M. Misiurewicz. A short proof of the variational principle for a \mathbb{Z}_+^n action on a compact space. In *International conference on dynamical systems in mathematical physics*, number 40 in Astérisque. Société mathématique de France, 1976.
- [30] A. Montanari, E. Mossel, and A. Sly. The weak limit of Ising models on locally tree-like graphs. *Probability Theory and Related Fields*, 152(1-2):31–51, 2012.
- [31] J. Moulin Ollagnier. *Ergodic Theory and Statistical Mechanics*. Lecture Notes in Mathematics, 1115. Springer-Verlag, Berlin, 1985.
- [32] J. Moulin Ollagnier and D. Pinchon. Mesures de Gibbs invariantes et mesures d’équilibre. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 55(1):11–23, 1981.
- [33] J. Moulin Ollagnier and D. Pinchon. The variational principle. *Studia Mathematica*, 72:151–159, 1982.
- [34] K. Petersen and K. Schmidt. Symmetric Gibbs measures. *Transactions of the American Mathematical Society*, 349(7):2775–2811, 1997.
- [35] I. F. Putnam. *Cantor minimal systems*, volume 70 of *University Lecture Series*. American Mathematical Society, 2018.
- [36] D. Ruelle. Statistical mechanics on a compact set with Z^v action satisfying expansiveness and specification. *Transactions of the American Mathematical Society*, 187:237–251, 1973.
- [37] D. Ruelle. *Thermodynamic Formalism*. Cambridge University Press, 2nd edition, 2004.
- [38] V. Salo. When are group shifts of finite type? *arXiv:1807.01951*, 2018.
- [39] B. Seward. Krieger’s finite generator theorem for actions of countable groups I. *Inventiones Mathematicae*, 215(1):265–310, 2019.
- [40] B. Seward. Krieger’s finite generator theorem for actions of countable groups II. *Journal of Modern Dynamics*, 15:1–39, 2019.
- [41] C. Shriver. Free energy, Gibbs measures, and Glauber dynamics for nearest-neighbor interactions on trees. *arXiv:2011.00653*, 2020.

- [42] A. Sly. Computational transition at the uniqueness threshold. In *2010 IEEE 51st Annual Symposium on Foundations of Computer Science*. IEEE, 2010.
- [43] A. M. Stepin and A. T. Tagi-Zade. Variational characterization of topological pressure of the amenable groups of transformations. *Doklady Akademii Nauk*, 254(3):545–549, 1980.
- [44] A. Tempelman. Specific characteristics and variational principle for homogeneous random fields. *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte Gebiete*, 65(3):341–365, 1984.
- [45] J. van den Berg. A uniqueness condition for Gibbs measures, with application to the 2-dimensional Ising antiferromagnet. *Communications in Mathematical Physics*, 152(1):161–166, 1993.