*Ergod. Th. & Dynam. Sys.* (2020), **40**, 2593–2680 © *Cambridge University Press, 2019*
doi:10.1017/etds.2019.18

# SURVEY

# Examples in the entropy theory of countable group actions

LEWIS BOWEN

*University of Texas at Austin, Mathematics, 1 University Station C1200, Austin, Texas, 78712, USA*
*(e-mail: lpbowen@math.utexas.edu)*

(*Received* 14 *January* 2019 *and accepted in revised form* 14 *January* 2019)

*Abstract*. Kolmogorov–Sinai entropy is an invariant of measure-preserving actions of the group of integers that is central to classification theory. There are two recently developed invariants, sofic entropy and Rokhlin entropy, that generalize classical entropy to actions of countable groups. These new theories have counterintuitive properties such as factor maps that increase entropy. This survey article focusses on examples, many of which have not appeared before, that highlight the differences and similarities with classical theory.

Key words: sofic group, entropy, Ornstein theory, Benjamini–Schramm convergence
2010 Mathematics Subject Classification: 37A35 (Primary)

*Contents*

1 Introduction 2596
1.1 Classical entropy theory 2596
1.2 General groups and naive entropy 2597
1.3 The Ornstein–Weiss example 2598
1.4 The $f$-invariant 2599
1.5 Sofic groups via Benjamini–Schramm convergence 2601
1.6 Sofic entropy: a probabilistic approach 2602
1.7 Rokhlin entropy 2603
1.8 What's in this article? 2603
2 Preliminaries 2604
2.1 Notation and conventions 2604
2.2 Sofic groups 2605
2.2.1 Soficity via maps into symmetric groups . . . . . 2605

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

2.2.2 Soficity via ultraproducts . . . 2606
2.2.3 Which groups are sofic? . . . 2606
2.2.4 The space of sofic approximations . . . 2607
2.3 Topological sofic entropy 2608
2.3.1 An application: Gottschalk's conjecture and Kaplansky's conjecture 2610
2.4 Measure sofic entropy 2611
2.4.1 The pseudometric definition . . . 2611
2.4.2 The partition definition . . . 2613
2.5 The $f$-invariant 2613
2.5.1 Other formulations and other groups . . . 2615
2.5.2 The sofic interpretation of the $f$-invariant . . . 2616
2.6 Rokhlin entropy 2617
2.6.1 Applications to the classification of Bernoulli shifts . . . 2618
2.7 Naive entropy 2619
2.7.1 Applications to Gottschalk's conjecture . . . 2620
2.7.2 Topological naive entropy . . . 2620
3 Special classes of actions 2621
3.1 Trivial actions 2621
3.1.1 Expanders . . . 2621
3.1.2 Diffuse sofic approximations . . . 2622
3.1.3 The $f$-invariant . . . 2623
3.2 Bernoulli shifts 2623
3.3 Markov chains 2623
3.4 Algebraic dynamics 2625
3.4.1 The case $\Gamma = \mathbb{Z}$ . . . 2625
3.4.2 Group rings . . . 2628
3.4.3 Principal algebraic actions . . . 2628
3.4.4 Yuzvinskii's addition formula . . . 2629
3.4.5 Further results . . . 2630
3.5 Gaussian actions 2631
3.6 Distal actions 2631
3.7 Smooth actions 2632
3.8 Non-free actions 2633
4 Perturbations 2634
4.1 Perturbing the sofic approximation 2634
4.2 Subgroups 2635
4.3 Co-induction 2636
4.4 Perturbing the partition 2637
4.5 Perturbing the measure 2638
4.6 Orbit-equivalence 2638
4.6.1 Weakly compact actions . . . 2638
4.6.2 Orbit equivalence and relative entropy . . . 2640
4.6.3 Integrable orbit-equivalence . . . 2640

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

5 Factors and extensions 2641
5.1 A variant of the Ornstein–Weiss example 2641
5.2 Bernoulli factors of Bernoulli shifts 2642
5.3 Zero entropy extensions 2643
5.4 Finite-to-one factors 2644
6 Combinations 2648
6.1 Ergodic decomposition 2648
6.2 Direct products 2649
6.2.1 Direct products and topological entropy . . . . . 2650
6.2.2 Subadditivity . . . . . 2650
6.2.3 Direct products and the $f$-invariant . . . . . 2650
6.2.4 A counterexample to additivity . . . . . 2651
6.2.5 Variants of sofic entropy and direct products . . . . . 2651
7 Ornstein theory 2654
7.1 The isomorphism theorem 2654
7.1.1 Non-Bernoulli factors of Bernoulli shifts . . . . . 2655
7.2 Krieger's generator theorem 2657
7.3 Sinai's theorem 2657
7.3.1 Spectral theory implications . . . . . 2658
7.3.2 Markov chains factor onto Bernoulli shifts . . . . . 2659
7.4 The $\bar{d}$ metric 2660
8 The variational principle 2662
8.1 Measures of maximal entropy: existence 2663
8.2 Measures of maximal entropy: uniqueness 2663
9 Gibbs measures, pressure and equilibrium states 2665
9.1 Motivation: finite graphs 2665
9.2 Pressure 2665
9.3 Pressure in symbolic systems 2666
9.4 Equilibrium states 2666
9.5 The Ising model 2667
9.6 Gibbs measures 2668
10 Relative entropy 2668
10.1 The Abramov–Rokhlin formula 2669
11 Outer/extension entropy 2670
11.1 Outer sofic entropy 2670
11.1.1 Outer sofic Pinsker algebras . . . . . 2671
11.2 Outer Rokhlin entropy 2671
11.2.1 Outer Rokhlin Pinsker algebras . . . . . 2672
11.3 Completely positive outer entropy 2672
11.4 Uniformly mixing 2672
11.4.1 Markov chains . . . . . 2673
11.4.2 A spectral criterion for CPE . . . . . 2674
Acknowledgements 2675
References 2675

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

## 1. *Introduction*

Sections 1.1–1.3 of this introduction set notation, give a brief review of classical entropy theory and motivate entropy theory for actions of general countable groups. Sections 1.4–1.7 provide an intuitive approach to the $f$-invariant, sofic groups, sofic and Rokhlin entropy. The last §1.8 summarizes the contents of this article.

1.1. *Classical entropy theory.* To set notation, let $(X, \mu)$, $(Y, \nu)$ denote standard probability spaces. An *automorphism* of $(X, \mu)$ is a measurable map $T : X \to X$ with measurable inverse that preserves the measure $\mu$. Two such maps $T : X \to X$, $S : Y \to Y$ are *measurably conjugate* or *isomorphic* if there exists a measure-space isomorphism $\Phi : X \to Y$ such that $\Phi \circ T = S \circ \Phi$ almost everywhere. The main motivating problem of this article is to classify automorphisms (and more generally, group actions) up to measure-conjugacy.

A special type of automorphism, called a Bernoulli shift, plays a central role. To define it, let $K$ denote a standard Borel space and $K^{\mathbb{Z}}$ the infinite direct power. An element $x \in K^{\mathbb{Z}}$ is a sequence $x = \{x_n\}_{n\in\mathbb{Z}}$ with values $x_n \in K$. Let $\sigma : K^{\mathbb{Z}} \to K^{\mathbb{Z}}$ denote the *shift map* defined by $\sigma(x)_n = x_{n-1}$. If $\kappa$ is a probability measure on $K$ then the shift map preserves the product measure $\kappa^{\mathbb{Z}}$. The triple $(\sigma, K^{\mathbb{Z}}, \kappa^{\mathbb{Z}})$ is called the *Bernoulli shift over the integers with base space* $(K, \kappa)$. In the early days of ergodic theory, von Neumann asked for a classification of these Bernoulli shifts. At the time, it was known that all Bernoulli shifts are spectrally isomorphic (that is, the induced operators on $L^2(K^{\mathbb{Z}}, \kappa^{\mathbb{Z}})$ are unitarily isomorphic). However not a single pair of Bernoulli shifts was known to be non-isomorphic and there were no non-trivial tools for proving isomorphism.

Motivated by this problem, Kolmogorov introduced dynamical entropy theory [**Kol58**, **Kol59**]. Here is an intuitive explanation: suppose that a system is under observation. At each unit of time, a measurement is made and recorded. The measuring device can only take on a finite number of distinct values. The entropy of the system is the amount of new information gained per unit time, on average and in the long run.

In this interpretation, $(X, \mu)$ represents all possible states of the system and $T : X \to X$ represents time evolution. The measuring device is represented by a measurable partition $\mathcal{P}$ of $X$. The *Shannon entropy* of $\mathcal{P}$ is defined by

$$H_\mu(\mathcal{P}) := \sum_{P\in\mathcal{P}} \mu(P) I_\mu(P) = -\sum_{P\in\mathcal{P}} \mu(P) \log(\mu(P)).$$

To motivate the above, suppose $x \in X$ is random. The amount of information gained by learning which part $P$ of $\mathcal{P}$ contains $x$ is defined by $I_\mu(P) := -\log \mu(P)$. So $H_\mu(\mathcal{P})$ is the average amount of information gained from learning which part of $\mathcal{P}$ contains $x$. (The definition of $I_\mu(P)$ is chosen so that if $P, Q \subset X$ are independent events then $I_\mu(P \cap Q) = I_\mu(P) + I_\mu(Q)$.)

The coarsest common refinement of two partitions $\mathcal{P}$, $\mathcal{Q}$ is denoted $\mathcal{P} \vee \mathcal{Q}$. The *entropy rate of $T$ with respect to* $\mathcal{P}$ is

$$h_\mu(T, \mathcal{P}) := \lim_{n\to\infty} \frac{1}{n} H_\mu\left(\bigvee_{i=0}^{n} T^{-i}\mathcal{P}\right).$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

This is the average quantity of information gained per unit time (represented by $T$) when observing the itinerary of a $\mu$-random point $x$ through the partition $\mathcal{P}$.

The *entropy rate of* $T$ is defined by

$$h_\mu(T) := \sup_{\mathcal{P}} h_\mu(T, \mathcal{P}).$$

A partition $\mathcal{P}$ is *generating* if the smallest sigma-algebra containing $T^{-n}\mathcal{P}$ for all $n \in \mathbb{Z}$ is the sigma-algebra of all measurable sets, up to sets of measure zero. Kolmogorov proved the crucially important result that if $\mathcal{P}$ is any generating partition with finite Shannon entropy then the entropy rate of $T$ is $h_\mu(T, \mathcal{P})$. Therefore, to compute entropy of $T$ one can choose any convenient generating partition. In the special case of the Bernoulli shift with base $(K, \kappa)$, if $K$ is countable then the *time* 0 *partition* $\mathcal{P} = \{P_k : k \in K\}$ defined by $P_k = \{x \in K^{\mathbb{Z}} : x_0 = k\}$ is generating. The entropy rate of the Bernoulli shift coincides with the Shannon entropy of $\mathcal{P}$. The latter is also called the *Shannon entropy of* $(K, \kappa)$:

$$H(K, \kappa) := -\sum_{k \in K} \kappa(\{k\}) \log \kappa(\{k\})$$

if $\kappa$ is purely atomic and $H(K, \kappa) := +\infty$ otherwise. This shows that Bernoulli shifts with different base space entropies are not measurably conjugate.

In 1970, Ornstein proved the converse: two Bernoulli shifts with the same entropy are isomorphic. Moreover, he developed a deep set of tools for determining whether a given automorphism is Bernoulli. With these tools, he and co-authors proved that many automorphisms are isomorphic to Bernoulli shifts including mixing Markov chains, hyperbolic toral automorphisms, the time-one map of geodesic flow on a hyperbolic surface and more.

In 1964, Sinai proved that every ergodic automorphism is a zero-entropy extension of a Bernoulli shift (which may be trivial). This explains why Bernoulli shifts are so important to the classification of automorphisms in general.

1.2. *General groups and naive entropy.* Now let $\Gamma$ be a countable group. An *action* of $\Gamma$ on a set $X$ is a collection $T = (T^g)_{g \in \Gamma}$ of transformations $T^g : X \to X$ satisfying $T^{gh} = T^g T^h$, $T^{g^{-1}} = (T^g)^{-1}$. For convenience we may write $gx$ for $T^g x$ when there is only one action of $\Gamma$ on $X$ under consideration. An action $T$ on a probability space $(X, \mu)$ is *probability-measure-preserving* (*pmp*) if each $T^g$ preserves $\mu$. We also denote an action by $\Gamma \curvearrowright^T (X, \mu)$ or simply $\Gamma \curvearrowright (X, \mu)$ when $T$ is implicit. A pmp action $T$ of $\Gamma$ on $(X, \mu)$ *factors onto* an action $S$ of $\Gamma$ on $(Y, \nu)$ if there is a measurable map $\Phi : X \to Y$ that pushes $\mu$ forward to $\nu$ and intertwines the action (so $\Phi T^g = S^g \Phi$ up to measure zero). Such a map is called a *factor map*. If, in addition, $\Phi$ is invertible with measurable inverse then the two actions are said to be *measurably conjugate* or *isomorphic*. The main motivating problem is to classify actions up to measure-conjugacy and determine which actions factor onto which.

The Bernoulli shifts described above generalize to this context. To be precise, let $K$ be a standard Borel space and $K^\Gamma$ be the set of all functions $x : \Gamma \to K$. When convenient we may represent $x \in K^\Gamma$ as a collection $x = (x_g)_{g \in \Gamma}$ of elements $x_g \in K$. The *shift action* of $\Gamma$ is denoted $S = (S^g)_{g \in \Gamma}$ where $S^g : K^\Gamma \to K^\Gamma$ is the transformation

$$(S^g x)_h = x_{g^{-1}h}.$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

If $\kappa$ is a probability measure on $K$ then $S$ preserves the product measure $\kappa^\Gamma$. The action $\Gamma \curvearrowright (K, \kappa)^\Gamma$ is called the *Bernoulli shift over* $\Gamma$ *with base space* $(K, \kappa)$.

There is a large difference between the entropy theory of amenable group actions and that of non-amenable groups. So it is worthwhile to review the definitions. A countable group $\Gamma$ is *amenable* if there exists a sequence $\{F_n\}$ of non-empty finite subsets of $\Gamma$ satisfying

$$\lim_{n\to\infty} \frac{|K F_n \cap F_n|}{|F_n|} = 1$$

for every non-empty finite $K \subset \Gamma$. Such a sequence is called a *Følner sequence*. There are many other equivalent definitions of amenability [**KL16**, **BdlHV08**]. For example, abelian, nilpotent and solvable groups are amenable while non-abelian free groups, $\mathrm{SL}(n, \mathbb{Z})$ $(n \geq 2)$, mapping class groups (with a few exceptions) and fundamental groups of closed hyperbolic $n$-manifolds $(n \geq 2)$ are not.

Now suppose $\Gamma$ is amenable with Følner sequence $\{F_n\}$ and $T$ is a pmp action of $\Gamma$ on $(X, \mu)$. The standard definition of entropy is

$$h_\mu(T, \mathcal{P}) = \lim_{n\to\infty} |F_n|^{-1} H_\mu\Big(\bigvee_{f\in F_n} T^{f^{-1}}\mathcal{P}\Big),$$
$$h_\mu(T) = \sup_{\mathcal{P}} h_\mu(T, \mathcal{P}),$$

where $\mathcal{P}$ is an arbitrary countable measurable partition of $X$ with finite Shannon entropy.

By a subadditivity argument it can be shown that the limit defining $h_\mu(T, \mathcal{P})$ exists, does not depend on the choice of Følner sequence and moreover, if $\mathcal{P}$ is a *generating* partition (this means that the smallest $\Gamma$-invariant sigma-algebra containing $\mathcal{P}$ is the sigma-algebra of all measurable sets up to measure zero) with finite Shannon entropy then $h_\mu(T, \mathcal{P}) = h_\mu(T)$. In particular, the entropy of the Bernoulli shift action $\Gamma \curvearrowright (K, \kappa)^\Gamma$ equals the Shannon entropy of the base space $H(K, \kappa)$. Entropy for amenable groups was first considered in [**Kie75**]. See also [**MO85**, **OW80**].

Moreover, the above entropy coincides with the so-called *naive entropy* defined by

$$h_\mu^{\mathrm{naive}}(T, \mathcal{P}) := \inf_{F\Subset\Gamma} |F|^{-1} H_\mu\Big(\bigvee_{f\in F} T^{f^{-1}}\mathcal{P}\Big),$$
$$h_\mu^{\mathrm{naive}}(T) := \sup_{\mathcal{P}} h_\mu^{\mathrm{naive}}(T, \mathcal{P}),$$

where $F \Subset \Gamma$ means $F$ is a non-empty *finite* subset of $\Gamma$. This definition makes sense for arbitrary countable groups $\Gamma$. However, if $\Gamma$ is non-amenable then Theorem 2.13 below shows that $h_\mu^{\mathrm{naive}}(T) \in \{0, \infty\}$ so naive entropy cannot distinguish Bernoulli shifts in this case.

1.3. *The Ornstein–Weiss example.* The next example convinced many researchers that entropy theory could not be extended to non-amenable groups. To explain, first suppose $\Gamma$ is amenable. Consider the *full* $n$*-shift* over $\Gamma$; this is the Bernoulli shift with base space $(\mathbb{Z}/n, u_n)$ where $\mathbb{Z}/n$ is the cyclic group of order $n$ and $u_n$ is the uniform probability measure. The entropy of the full $n$-shift is $\log(n)$. Because entropy is non-increasing under

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

factor maps, the full two-shift cannot factor onto the full four-shift. However, Ornstein and Weiss showed in [**OW87**] that when $\mathbb{F}_2 = \langle a, b\rangle$ is the rank-two free group, the full two-shift does indeed factor onto the full four-shift. Their example is as follows: define $\phi : (\mathbb{Z}/2)^{\mathbb{F}_2} \to (\mathbb{Z}/2 \times \mathbb{Z}/2)^{\mathbb{F}_2}$ by

$$\phi(x)_g = (x_g + x_{ga}, x_g + x_{gb}).$$

The spaces $(\mathbb{Z}/2)^{\mathbb{F}_2}$ and $(\mathbb{Z}/2 \times \mathbb{Z}/2)^{\mathbb{F}_2}$ are compact abelian groups under pointwise addition. With this interpretation, $\phi$ is a group homomorphism. It is a good exercise to show that it is surjective. Surjectivity implies that $\phi$ takes Haar measure to Haar measure. Therefore, it is indeed a factor map. In fact the kernel consists of the constants, so it is a two-one factor map. In the setting of $\mathbb{Z}$-actions, entropy is preserved under finite-to-one maps. So this example led some researchers to speculate that the two-shift and the four-shift over $\mathbb{F}_2$ could be isomorphic. The $f$-invariant (and later, sofic entropy) was developed to prove that they are not.

In work in progress [**Bow17**], the author has shown that if $\Gamma$ is any non-amenable group then all Bernoulli shifts over $\Gamma$ factor onto each other. This is based on a generalization of the Gaboriau–Lyons theorem [**GL09**] and is reviewed in §5.2. So there does not exist a monotone invariant that distinguishes Bernoulli shifts over a non-amenable group.

1.4. *The $f$-invariant.* The $f$-invariant is a measure-conjugacy invariant for actions of free groups that distinguishes Bernoulli shifts. To explain, it will be convenient to use probabilistic notation as follows. Suppose $X$ is a random variable that takes values in a Borel space $K$. Let $\mathrm{Prob}(K)$ denote the set of all Borel probability measures on $K$. The *law* of $X$ is the probability measure $\mathrm{Law}(X) \in \mathrm{Prob}(K)$ satisfying

$$\mathrm{Law}(X)(E) = P(X \in E)$$

for $E \subset K$ where $P(\cdot)$ denotes probability. If $Y$ is also a random variable taking values in a Borel space $L$, then $\mathrm{Law}(X, Y) \in \mathrm{Prob}(K \times L)$ is the probability measure satisfying

$$\mathrm{Law}(X, Y)(E \times F) = P(X \in E, Y \in F).$$

Also $\mathrm{Law}(X|Y)$ is the random variable taking values in $\mathrm{Prob}(K)$ defined by

$$\mathrm{Law}(X|Y)(E) = P(X \in E|Y).$$

In general, a *stationary $\Gamma$-process* is a $\Gamma$-indexed family of random variables $\mathbf{X} = (X_g)_{g\in\Gamma}$ such that each $X_g$ takes values in some Borel space $K$ and the law of $\mathbf{X}$ is invariant under left-multiplication of the indices. So the law of $\mathbf{X}$ is a $\Gamma$-invariant Borel probability measure on $K^\Gamma$ and stationarity means that the law of $(X_g)_{g\in\Gamma}$ is the same as the law of $(X_{hg})_{g\in\Gamma}$ for any $h \in \Gamma$.

The $f$-invariant is motivated by way of Markov chains over free groups. Before getting into that, it makes sense to recall Markov chains over the integers. A stationary $\mathbb{Z}$-process $\mathbf{X} = (X_i)_{i\in\mathbb{Z}}$ is *Markov* if

$$\mathrm{Law}(X_0|X_{-1}, X_{-2}, \ldots) = \mathrm{Law}(X_0|X_{-1}).$$

The law of a stationary Markov process is uniquely determined by the *time-0 distribution* $\mathrm{Law}(X_0)$ and the transition probabilities $P(X_0 = k_0|X_{-1} = k_{-1})$ (for $k_{-1}, k_0 \in K$).

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

Moreover, the Markov process uniquely maximizes entropy over all stationary processes that have the same time-zero distribution and transition probabilities.

If $\mathbf{X} = (X_i)_{i\in\mathbb{Z}}$ is an arbitrary stationary process with values in a finite or countable set $K$ then its *nth step Markov approximation* is the Markov process $\mathbf{Y}^{(n)} = (Y_i^{(n)})_{i\in\mathbb{Z}}$ taking values in the Cartesian power $K^n$ satisfying

$$\mathrm{Law}(Y_0^{(n)}) = \mathrm{Law}(X_0, \ldots, X_{n-1})$$

$$\begin{aligned}\mathrm{Law}(Y_0^{(n)}|Y_{-1}^{(n)} &= (x_{-1}, x_0, \ldots, x_{n-2}))\\ &= \mathrm{Law}(X_0, \ldots, X_{n-1}|(X_{-1}, \ldots, X_{n-2}) = (x_{-1}, x_0, \ldots, x_{n-2}))\end{aligned}$$

for any $x_{-1}, \ldots, x_{n-2} \in K$. The entropy rate of $\mathbf{X}$ satisfies

$$h(\mathbf{X}) = \lim_{n\to\infty} h(\mathbf{Y}^{(n)}) = \lim_{n\to\infty} H(Y_0^{(n)}|Y_{-1}^{(n)}).$$

Recall that the Shannon entropy of a random variable $X$ conditioned on another random variable $Y$ that takes on only countably many values is

$$H(X|Y) := -\sum_{x,y} P(X = x, Y = y)\log(P(X = x|Y = y)).$$

Since

$$\begin{aligned}H(Y_0^{(n)}|Y_{-1}^{(n)}) &= H(X_0, \ldots, X_{n-1}|X_{-1}, \ldots, X_{n-2})\\ &= H(X_{n-1}|X_{n-2}, \ldots, X_{-1}) = H(X_0|X_{-1}, \ldots, X_{-n})\end{aligned}$$

we arrive at the familiar formula

$$h(\mathbf{X}) = H(X_0|X_{-1}, X_{-2}, \ldots).$$

This generalizes to free groups as follows: let $\mathbb{F}_r = \langle s_1, \ldots, s_r\rangle$ be a free group of rank $r$ and $\mathbf{X} = (X_g)_{g\in\mathbb{F}_r}$ a stationary process over $\mathbb{F}_r$. The process $\mathbf{X}$ is *Markov* if for every $1 \le i \le r$

$$\mathrm{Law}(X_e|(X_g)_{g\in\mathrm{Pre}(s_i)}) = \mathrm{Law}(X_e|X_{s_i}),$$

where $\mathrm{Pre}(s_i) \subset \mathbb{F}_r$ is the set of all elements with prefix $s_i$. In other words, $g \in \mathrm{Pre}(s_i)$ if and only if $|s_i^{-1}g| < |g|$ where $|\cdot| : \mathbb{F}_r \to \mathbb{N}$ denote the word length (so $|\gamma|$ is the smallest natural number $n$ such that $\gamma$ can be written as a product of $n$ elements of $\{s_1, \ldots, s_r, s_1^{-1}, \ldots, s_r^{-1}\}$).

Define

$$F(\mathbf{X}) := -(r-1)H(X_e) + \sum_{i=1}^{r} H(X_e|X_{s_i}).$$

The intuition for this formula is as follows: the term $H(X_e|X_{s_i})$ measures the entropy in the $s_i$-direction. The sum $\sum_{i=1}^r H(X_e|X_{s_i})$ 'counts' the entropy at the identity $r$ times. To compensate for this, substract $(r-1)H(X_e)$ to obtain the formula above.

Now suppose $\mathbf{X} = (X_g)_{g\in\mathbb{F}_r}$ is an arbitrary stationary $\mathbb{F}_r$-process taking values in a finite or countable set $K$. The *nth Markov approximation* to $\mathbf{X}$ is the Markov process $\mathbf{Y}^{(n)} = (Y_g^{(n)})_{g\in\mathbb{F}_r}$ taking values in $K^{B(n)}$ (where $B(n) \subset \mathbb{F}_r$ is the ball of radius $n$) satisfying

$$\mathrm{Law}(Y_e^{(n)}) = \mathrm{Law}((X_g)_{g\in B(n)}),$$

$$\mathrm{Law}(Y_e^{(n)}|Y_{s_i}^{(n)} = x) = \mathrm{Law}((X_g)_{g\in B(n)}|(X_{s_i g})_{g\in B(n)} = x)$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

for any function $x : B(n) \to K$. Define

$$f(\mathbf{X}) := \lim_{n\to\infty} F(\mathbf{Y}^{(n)}).$$

In §2.5 a proof is sketched that this does indeed define a measure-conjugacy invariant, called the $f$-invariant. Moreover, the $f$-invariant of the Bernoulli shift $\mathbb{F}_r \curvearrowright (K^\Gamma, \kappa^\Gamma)$ is $H(K, \kappa)$. This proves the two-shift over $\mathbb{F}_r$ is not isomorphic to the four-shift.

The $f$-invariant is a particularly nice invariant: it can be computed exactly for Markov chains (§3.3), it satisfies an ergodic decomposition formula (§6.1), a subgroup formula (§4.2), an Abramov–Rokhlin-type formula (§10.1), a Yuzvinskii-type addition formula (§3.4.4) and is additive under direct products. However, it can increase under factor maps and it can take on negative values (§3.1.3). It can be generalized to some other groups (§2.5.1) and is related to sofic entropy (§2.5.2).

1.5. *Sofic groups via Benjamini–Schramm convergence.* The sofic concept provides a new perspective on the $f$-invariant and extends entropy theory beyond amenable groups. The most intuitive definition of sofic groups is based on Benjamini–Schramm convergence [**BS01**]. Only what is needed for sofic group theory will be explained here; for the more general theory see [**HLS14**].

Suppose $\Gamma$ has a finite symmetric generating set $S \subset \Gamma$. The *Cayley graph* of $(\Gamma, S)$ is a directed graph with edge labels in $S$, denoted by $\mathrm{Cay}(\Gamma, S)$. Its vertex set is $\Gamma$ and for every $s \in S$ and $g \in \Gamma$ there is an $s$-labeled directed edge from $g$ to $gs$. These are all of the edges. Because the edges of the Cayley graph are directed and labeled, the group can be recovered from the Cayley graph. This would not be true otherwise since there are groups with Cayley graphs that are isomorphic as unlabeled graphs.

Now consider a *finite* $S$-edge-labeled directed graph $G = (V, E)$. For $r > 0$, let $\mathscr{V}_r(G)$ be the set of all vertices $v \in V$ such that there exists a graph isomorphism from the ball of radius $r$ centered at $v$ to the ball of radius $r$ centered at the identity element in the Cayley graph $\mathrm{Cay}(\Gamma, S)$. This isomorphism is required to map $v$ to the identity, preserves edge directions and preserves labels. The graph $G$ is called an $(r, \epsilon)$-*sofic approximation* to $(\Gamma, S)$ if $|\mathscr{V}_r(G)| \geq (1 - \epsilon)|V|$. So with probability $\geq 1 - \epsilon$, a uniformly random vertex's radius $r$-neighborhood looks the same as the radius $r$-neighborhood of the identity in the Cayley graph $\mathrm{Cay}(\Gamma, S)$.

By definition, a sequence $\{G_i\}$ of finite $S$-edge-labeled directed graphs *Benjamini–Schramm converges* to the Cayley graph $\mathrm{Cay}(\Gamma, S)$ if

$$\lim_{i\to\infty} \frac{|\mathscr{V}_r(G_i)|}{|V_i|} = 1$$

for every $r > 0$. Such a sequence is called a *sofic approximation* to $\Gamma$. The group $\Gamma$ is *sofic* if there exists a sofic approximation to $\Gamma$.

*Exercise 1.* Show that soficity does not depend on the choice of generating set $S$.

*Exercise 2.* Show that $\mathbb{Z}^d$ and finitely generated free groups are sofic.

For example, if $C_n$ is the directed $n$-cycle then $C_n$ Benjamini–Schramm converges to the standard Cayley graph of $\mathbb{Z}$ as $n \to \infty$. For another example, suppose $\Gamma$ is a finitely

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

generated residually finite group. Residual finiteness means there exists a decreasing sequence $N_i \lhd \Gamma$ of finite-index normal subgroups with $\cap_i N_i = \{e\}$. Let $S$ be a finite generating set for $\Gamma$ and consider the associated Cayley graphs $G_n$ with vertex set $\Gamma/N_n$ and edges $\{(gN_n, gsN_n) : s \in S\}$. This sequence Benjamini–Schramm converges to the Cayley graph of $\Gamma$ with respect to $S$. Therefore all residually finite groups are sofic.

The above definition is the most intuitive. However, it has the unfortunate drawback that it applies only to finitely generated groups. Section 2.2.1 presents a more general definition based on maps from $\Gamma$ into the symmetric group.

It is a major open problem whether all countable groups are sofic. For further reading, there are several surveys on sofic groups [**Pes08**, **PK09**, **CL15a**].

1.6. *Sofic entropy: a probabilistic approach.* A more thorough account of sofic entropy is presented in §§2.3–2.4. Here is an intuitive approach under simplifying conditions. Let $\mathbf{X} = (X_g)_{g\in\Gamma}$ be a $\Gamma$-stationary process taking values in a finite set $K$. Assume $\Gamma$ has a finite generating set $S$. Fix a sequence $\Sigma = \{G_i\}_{i=1}^{\infty}$ of finite $S$-labeled directed graphs $G_i = (V_i, E_i)$ that Benjamini–Schramm converge to the Cayley graph $\mathrm{Cay}(\Gamma, S)$. The sofic entropy of $\mathbf{X}$ with respect to $\Sigma$, denoted $h_\Sigma(\mathbf{X})$, is the exponential rate of growth of the number of microstates for $\mathbf{X}$ on $G_i$. Intuitively, a microstate is a function $\phi : V_i \to K$ that approximates $\mathbf{X}$ in a local statistical sense. To be precise, fix a radius $r > 0$ and let vertex $v \in \mathcal{V}_r(G_i)$ be a uniformly random vertex. Consider the restriction of $\phi$ to the ball of radius $r$ centered at $v$. Because $\Sigma$ is a sofic approximation, this ball is isomorphic to the ball $B_r(\Gamma, S)$ of radius $r$ centered at the identity in $\mathrm{Cay}(\Gamma, S)$ with high probability. So the law of the restriction $\phi \restriction B_r(v)$ determines a (sub-)probability measure on the set $K^{B_r(\Gamma,S)}$ of all functions from $B_r(\Gamma, S)$ to $K$. If this law is $\delta$-close in total variation distance to the law of $(X_g)_{g\in B_r(\Gamma,S)}$ then $\phi$ is said to be an $(r, \delta)$-*microstate* of $\mathbf{X}$.

The *sofic entropy* of $\mathbf{X}$ is

$$h_\Sigma(\mathbf{X}) = \inf_{\delta>0} \inf_{r>0} \limsup_{i\to\infty} |V_i|^{-1} \log(\#(r, \delta)\text{-microstates on } G_i).$$

This is also called the $\Sigma$-*entropy*. In [**Bow10b**] it was shown that this entropy is invariant under measure-conjugacy and the entropy of an independent and identically distributed (i.i.d) process is the Shannon entropy $H(X_{1_\Gamma})$.

A few words about the definition of $h_\Sigma(\mathbf{X})$: if $\Gamma$ is amenable then the sofic entropy agrees with classical entropy. However, if $\Gamma$ is non-amenable then it is possible that there are no $(r, \delta)$-microstates for any graph $G_i$ in the sofic approximation. In this case, $h_\Sigma(\mathbf{X}) = -\infty$. Examples of such behavior are presented in §3.1. Also the lim sup can be replaced with lim inf or by an ultralimit. These changes give *a priori* different invariants (when $\Gamma$ is non-amenable). In §§3.1 and 4.1 examples are presented of sofic approximations $\Sigma$, $\Sigma'$ to a group $\Gamma$ and an explicit action such that the $\Sigma$-entropy is $-\infty$ but the $\Sigma'$ entropy is non-negative. The following is a major open problem.

*Problem 1.* Suppose $\Sigma$, $\Sigma'$ are two sofic approximations to $\Gamma$ and the $\Sigma$-entropy and $\Sigma'$-entropy of $\mathbf{X}$ are both positive. Are they necessarily equal?

Other expositions of sofic entropy include [**Wei15**, **Gab17**, **KL16**].

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

The concept of sofic approximation can be generalized by replacing the finite graphs with random finite graphs. That is, the $i$th approximating graph $G_i$ is allowed to be random; but the number of vertices $|V_i|$ is required to be determined. Now define

$$h_\Sigma(\mathbf{X}) = \inf_{\delta>0} \inf_{r>0} \limsup_{i\to\infty} |V_i|^{-1} \log \mathbb{E}[(\#(r,\delta)\text{-microstates on } G_i)].$$

In the special case in which $\Gamma = \langle s_1, \ldots, s_r\rangle$ is a free group, this leads to a new interpretation of the $f$-invariant:

$$f(\mathbf{X}) = h_\Sigma(\mathbf{X}),$$

where $\Sigma = \{G_i\}_{i=1}^\infty$ is the 'permutation model' of the random $2r$-regular graphs. To be precise, let $\sigma : \Gamma \to \mathrm{Sym}(n)$ be a homomorphism chosen at random uniformly amongst all $(n!)^r$ homomorphisms where $\mathrm{Sym}(n)$ is the symmetric group on $[n] := \{1, \ldots, n\}$. Then $G_n$ is the graph with vertex $V_n = [n]$ and $s_i$-labeled edges $(p, \sigma(s_i)p)$ for $1 \le i \le r$ and $p \in [n]$. This is Theorem 2.10 below.

1.7. *Rokhlin entropy.* Suppose $T = (T^g)_{g\in\Gamma}$ is a pmp action of $\Gamma$ on $(X, \mu)$ and $\mathcal{P}$ is a generating partition for the action. It follows immediately that the sofic entropy of $T$ is bounded by the Shannon entropy $H_\mu(\mathcal{P})$. This leads to the following idea: let $h^{\mathrm{Rok}}(T)$ denote the infimum of $H_\mu(\mathcal{P})$ over all generating partitions $\mathcal{P}$. If $T$ is ergodic then this is called the *Rokhlin entropy* of the action (the non-ergodic case is slightly different; see §2.6 for details). Some basic facts are as follows.

- Rokhlin entropy is a measure-conjugacy invariant. Moreover, it is well defined for every action of every countable group (even non-sofic ones, if they exist!).
- Rokhlin entropy is an upper bound for sofic entropy.
- Rokhlin entropy agrees with classical entropy for amenable groups [**STD16**].

Moreover, in recent groundbreaking work, the following results have been obtained.

- [**STD16**] For every $\epsilon > 0$, every essentially free ergodic action with positive Rokhlin entropy admits a factor that is essentially free and has Rokhlin entropy $< \epsilon$.
- [**Sew19**] Every ergodic action with Rokhlin entropy $< \log(n)$ admits a generating partition with $n$ parts.
- [**Sew18**] Every essentially free ergodic action with positive Rokhlin entropy admits a Bernoulli factor.

See §§2.6 and 7 for more details.

*Problem 2.* Suppose $T$ is ergodic and essentially free. If the $\Sigma$-entropy of $T$ is not $-\infty$ then is it equal to the Rokhlin entropy?

1.8. *What's in this article?* Section 2 reviews the fundamental aspects of sofic groups, sofic entropy, the $f$-invariant, Rokhlin entropy and naive entropy.

Section 3 covers a list of examples in which entropy has been computed. Perhaps the most interesting cases are the Bernoulli shifts, Markov chains over free groups and principal algebraic actions (in which the entropy is related to the Fuglede–Kadison determinant). There are also degenerate cases in which the entropy is non-positive. This includes (under mild conditions) trivial actions, distal, smooth and non-free actions.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

One surprising case is that of lattices $\Gamma$, $\Lambda$ in a totally disconnected locally compact group $G$ such that $\Gamma \curvearrowright G/\Lambda$ has positive entropy (Example 4 in §3.3).

About §§4 and 6: much of the usefulness of classical entropy theory derives from a list of formulas and inequalities expressing how entropy changes as the system is perturbed or combined with other systems. This includes: inducing to a subgroup, co-inducing a subgroup action, continuity or semi-continuity in the measure, the partition or the action or passing to an orbit-equivalent action. It also includes direct products, ergodic decomposition, relative entropy and inverse limits. In the sofic case, we usually have an inequality where in the classical case an equality holds. Moreover, there are counterexamples. For example, in §6.2 we present an example showing that sofic entropy need not be additive under direct products.

Section 5.1 presents a finite-to-one factor map from a zero entropy action to a Bernoulli shift, §5.2 sketches a proof that if $\Gamma$ is non-amenable then all Bernoulli shifts factor onto each other, §5.3 sketches a proof that if $\Gamma$ is non-amenable then every free ergodic action has a zero-entropy extension, §5.4 explores how entropy varies under finite-to-one factor maps.

Section 7 covers generalizations of Ornstein theory for non-amenable groups including the isomorphism theorem, Krieger's generator theorem and Sinai's factor theorem. It also contains counterexamples such as Popa's example of a non-Bernoulli factor of a Bernoulli shift, and a non-Bernoulli $\bar{d}$-limit of Bernoulli shifts.

Section 8 sketches a proof of the variational principle for sofic entropy. This naturally leads to the question of whether measures of maximal entropy exist and whether or not they are unique. The existence problem is similar to that of the classical case: namely, existence occurs under weak forms of expansitivity that imply upper semi-continuity of entropy with respect to weak* topology on the space of measures. Regarding uniqueness: an example is presented in §8.2 of a mixing Markov chain over the free group with multiple measures of maximum $f$-invariant.

Section 9 defines sofic pressure and equilibrium states (measures) for actions of sofic groups and relates them to Gibbs measures on random regular graphs.

Section 10 is a short section on relative entropy. This includes an Abramov–Rokhlin formula for actions of free groups.

Section 11 defines and explores outer sofic and Rokhlin entropy. For example, outer sofic entropy of a factor map is the exponential rate of growth of the number of microstates for the target action that lift to microstates for the source action. When $\Gamma$ is amenable, this is just the entropy of the target. However, when $\Gamma$ is non-amenable it can be different; for example the outer sofic entropy of the Ornstein–Weiss map is $\log(2)$, not $\log(4)$. Using outer entropy, we define outer Pinsker algebra and completely positive outer entropy. For example, Bernoulli shifts and a large class of algebraic actions are known to have completely positive outer entropy. This notion is also related to uniform model mixing, which is a generalization of uniform mixing to the sofic context.

## 2. *Preliminaries*

2.1. *Notation and conventions.* Throughout this article, all measure spaces are standard and all maps measurable unless otherwise specified. We often ignore measure zero

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

phenomena without explicit mention. Also, $\Gamma$ denotes a countable group, $(X, \mu)$, $(Y, \nu)$ probability spaces and $\Gamma \curvearrowright^T (X, \mu)$, $\Gamma \curvearrowright^S (Y, \nu)$ are probability-measure-preserving (pmp) actions of $\Gamma$ (unless otherwise specified). This means $T = (T^g)_{g\in\Gamma}$ is a collection of measure-space automorphisms of $(X, \mu)$ such that $T^{gh} = T^g T^h$ and $T^{g^{-1}} = (T^g)^{-1}$ almost everywhere.

A *factor map* between these actions is a measurable map $\Phi : X \to Y$ such that $\Phi_*\mu = \nu$ and $\Phi(gx) = g\Phi(x)$ for almost every (a.e.) $x$ and every $g \in \Gamma$. There is a natural correspondence between factors of the action $\Gamma \curvearrowright (X, \mu)$ and $\Gamma$-invariant sigma-algebras of $X$ (up to measure zero sets). Namely, if $\Phi : X \to Y$ is a factor map and $\mathcal{B}_Y$ is the Borel sigma-algebra on $Y$ then $\Phi^{-1}(\mathcal{B}_Y)$ is a $\Gamma$-invariant sigma-algebra of $X$. We will call this the *sigma-algebra associated with* $\Phi$. Conversely, if we are given a $\Gamma$-invariant sigma-algebra $\mathcal{F} \subset \mathcal{B}_X$ (where $\mathcal{B}_X$ is the Borel sigma-algebra on $X$) then by the Mackey realization theorem there is a Borel space $Y$ and a factor map $\Phi : X \to Y$ such that $\mathcal{F} = \Phi^{-1}(\mathcal{B}_Y)$.

If $\mathcal{P}$ is a measurable partition of $X$ then the *factor associated to* $\mathcal{P}$ is the factor associated to the smallest $\Gamma$-invariant sigma-algebra containing $\mathcal{P}$.

The notation $X \Subset Y$ means that $X$ is a finite subset of $Y$.

2.2. *Sofic groups.* The next definition might be less intuitive than the definition of soficity in the introduction; however it is the most useful.

2.2.1. *Soficity via maps into symmetric groups.* Let $\Gamma$ be a countable group. Throughout, $V$ denotes a finite set and $\mathrm{Sym}(V)$ the group of bijections from $V$ to itself.

*Definition 1.* Given $F \subset G$ and $\delta > 0$ we say that a map $\sigma : \Gamma \to \mathrm{Sym}(V)$ is:

- $(F, \delta)$-*multiplicative* if
$$1 - \delta < |V|^{-1}\#\{p \in V : \sigma(g)\sigma(h)p = \sigma(gh)p\} \quad \text{for all } g, h \in F;$$
- $(F, \delta)$-*trace-preserving* if
$$\delta > |V|^{-1}\#\{p \in V : \sigma(g)p = p\} \quad \text{for all } g \in F\backslash\{e\}.$$

A *sofic approximation* of a group $\Gamma$ is a sequence $\Sigma = \{\sigma_i\}_{i\in\mathbb{N}}$ of set maps $\sigma_i : \Gamma \to \mathrm{Sym}(V_i)$ such that for every finite $F \subset \Gamma$ and $\delta > 0$ there exists $I$ such that $i > I$ implies $\sigma_i$ is $(F, \delta)$-multiplicative, $(F, \delta)$-trace-preserving and $\lim_{i\to\infty} |V_i| = +\infty$. A group $\Gamma$ is *sofic* if it admits a sofic approximation.

*Exercise 3.* Show that the definition above is equivalent to the definition in §1.5 in case $\Gamma$ is finitely generated.

*Exercise 4.* Suppose that $\Gamma$ is residually finite. So there exist finite-index normal subgroups $\Gamma \geq N_1 \geq N_2 \geq \cdots$ such that $\cap_i N_i$ is trivial. Show that the canonical homomorphisms $\Gamma \to \mathrm{Sym}(G/N_i)$ form a sofic approximation.

*Exercise 5.* Suppose that $\Gamma$ is amenable and $\{F_i\}$ is a Følner sequence. Show that for every $i$ there is a map $\sigma_i : \Gamma \to \mathrm{Sym}(F_i)$ such that $\sigma_i(g)f = gf$ whenever $gf \in F_i$. Show that these maps form a sofic approximation.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

2.2.2. *Soficity via ultraproducts.* Suppose $\Sigma$ is a sofic approximation to $\Gamma$ as above. Let $\mathcal{U}$ be a non-principal ultrafilter on $\mathbb{N}$. Let $\prod_{i\to\mathcal{U}} \mathrm{Sym}(V_i)$ denote the ultraproduct of symmetric groups. To be precise, $\prod_{i\to\mathcal{U}} \mathrm{Sym}(V_i)$ is the direct product $\prod_i \mathrm{Sym}(V_i)$ modulo the equivalence relation $(x_i) \sim (y_i)$ if and only if $\{i \in \mathbb{N} : x_i = y_i\} \in \mathcal{U}$. This is a group. Let $N \leq \prod_{i\to\mathcal{U}} \mathrm{Sym}(V_i)$ be the set of all $\mathcal{U}$-equivalence classes of sequences $(\pi_1, \pi_2, \ldots)$ (with $\pi_i \in \mathrm{Sym}(V_i)$) such that

$$\lim_{i\to\mathcal{U}} |V_i|^{-1}|\mathrm{Fix}(\pi_i)| = 1$$

where $\mathrm{Fix}(\pi_i)$ is the set of fixed points of $\pi_i$ in $V_i$. Then $N$ is a normal subgroup of $\prod_{i\to\mathcal{U}} \mathrm{Sym}(V_i)$ and the map

$$g \in \Gamma \mapsto (\sigma_1(g), \sigma_2(g), \ldots)$$

determines an injective homomorphism from $\Gamma$ into the quotient group $\prod_{i\to\mathcal{U}} \mathrm{Sym}(V_i)/N$.

*Exercise 6.* Prove that $\Gamma$ is sofic if and only if for some (any) increasing sequence $\{V_i\}_{i=1}^{\infty}$ of finite sets, $\Gamma$ admits an injective homomorphism into $\prod_{i\to\mathcal{U}} \mathrm{Sym}(V_i)/N$.

From this description it can be shown that the group von Neumann algebra of $\Gamma$ satisfies Connes' embedding conjecture (see [**ES05**]). This point of view is elaborated on in [**Pes08**, **PK09**, **CL15a**].

2.2.3. *Which groups are sofic?*

THEOREM 2.1. *The class of sofic groups is closed under:*

(1) *subgroups, direct products, direct limits, inverse limits (so a residually sofic group is sofic), free products;*
(2) *extensions by amenable groups;*
(3) *free products with amalgamation over an amenable subgroup;*
(4) *certain graph products and wreath products.*

For detailed proofs of (1–2) see [**ES06**]. There are three different proofs of (3) in [**DKP14**, **Pău11**, **ES11**]. Soficity of graph products is studied in [**CHR14**] and wreath products in [**HS16**, **HS18**].

*Proof sketch of (1).* Suppose $\Sigma = \{\sigma_i\}$ is a sofic approximation to a group $\Gamma$ as in Definition 2.2.1. Restricting to a subgroup $\Lambda \leq \Gamma$ yields a sofic approximation to $\Lambda$. This shows soficity is closed under subgroups. If $\Sigma' = \{\sigma_i'\}$ is a sofic approximation to a group $\Gamma'$ then the direct product $\sigma_i \times \sigma_i' : \Gamma \times \Gamma' \to \mathrm{Sym}(V_i) \times \mathrm{Sym}(V_i') \leq \mathrm{Sym}(V_i \times V_i')$ gives a sofic approximation to $\Gamma \times \Gamma'$. So soficity is closed under direct products. Diagonalization arguments show that soficity is preserved under direct limits and inverse limits.

To see that soficity is preserved under free products, consider $\Gamma$, $\Gamma'$ as above. Suppose $V_i = V_i'$ and let $\pi_i \in \mathrm{Sym}(V_i)$ be a uniformly random permutation. Define $\sigma_i^{\pi_i} : \Gamma \to \mathrm{Sym}(V_i)$ by conjugation: $\sigma_i^{\pi_i}(g) = \pi_i \sigma_i(g) \pi_i^{-1}$. Now we define $\sigma_i^{\pi_i} * \sigma_i' : \Gamma * \Gamma' \to \mathrm{Sym}(V_i)$ by

$$\sigma_i^{\pi_i} * \sigma_i'(g_1 g_1' \cdots g_n g_n') = \sigma_i^{\pi_i}(g_1)\sigma_i'(g_1') \cdots \sigma_n^{\pi_i}(g_n)\sigma_n'(g_n')$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

if $g_1, \ldots, g_n \in \Gamma\backslash\{e\}$ and $g_1', \ldots, g_n' \in \Gamma'\backslash\{e\}$ for example. It can be shown that, with probability 1, $\{\sigma_i^{\pi_i} * \sigma_i'\}$ is a sofic approximation to $\Gamma * \Gamma'$. □

*Remark 1.* Mal'cev proved that all finitely generated linear groups are residually finite [**Mal40**] and therefore they are sofic. Because soficity is preserved under direct limits it follows that all countable linear groups are sofic.

It is open whether all countable groups are sofic. However the soficity of the following groups is unknown: free Burnside groups (this was pointed out by Weiss [**Wei00**]), Tarski monsters, $\mathrm{SL}(3, \mathbb{Z}) *_{F_1=F_2} \mathrm{SL}(3, \mathbb{Z})$ where $F_1, F_2 \leq \mathrm{SL}(3, \mathbb{Z})$ are isomorphic non-abelian free groups, and the Burger–Mozes groups from [**BM97**, **BM00**]. On the other hand, Thom constructed a non-residually finite property (T) sofic group [**Tho10**] and de Cornulier constructed a sofic group that is not a limit of amenable groups in the space of marked groups [**Cor11**]. Elek and Szabo show that there exists a non-amenable simple sofic group [**ES05**]. There are several recent surveys on sofic groups [**Pes08**, **PK09**, **CL15a**].

2.2.4. *The space of sofic approximations.*

*Problem 3.* For a given interesting group $\Gamma$, describe the set of all sofic approximations to $\Gamma$.

Here we will make the above problem more precise and explain some partial results and specific questions.

To begin we observe that it is possible to perturb a sofic approximation in an inessential way. To be precise, let $\Sigma = \{\sigma_i\}$, $\Sigma' = \{\sigma_i'\}$ be two sofic approximations to $\Gamma$ and suppose that

$$\sigma_i : \Gamma \to \mathrm{Sym}(V_i), \quad \sigma_i' : \Gamma \to \mathrm{Sym}(V_i').$$

In the special case that $V_i = V_i'$ for all $i$ we can define the *edit-distance* between $\Sigma$ and $\Sigma'$ with respect to a finite set $F \subset \Gamma$ by

$$d^F(\Sigma, \Sigma') = \limsup_{i\to\infty} |V_i|^{-1} \#\{v \in V_i : \exists f \in F, \sigma_i(f)v \neq \sigma_i'(f)v\}.$$

Strictly speaking this is a pseudodistance since it is entirely possible that two different sofic approximations satisfy $d^F(\Sigma, \Sigma') = 0$ for all $F \subset \Gamma$. If $d^F(\Sigma, \Sigma') = 0$ for every finite $F \subset \Gamma$ then an exercise shows that the sofic entropy with respect to $\Sigma$ equals the sofic entropy with respect to $\Sigma'$. So we call two sofic approximations that have this property *equivalent*.

If $\Gamma$ is amenable then in [**ES11**] it is shown that every sofic approximation to $\Gamma$ is equivalent to one obtained from finite unions of Følner sets in a natural way. This completely describes all sofic approximations to $\Gamma$.

We will say that $\Sigma$ is *by homomorphisms* if each $\sigma_i : \Gamma \to \mathrm{Sym}(V_i)$ is a homomorphism. For example, if $\Sigma = \{\sigma_i\}_{i=1}^\infty$ is any sofic approximation to a free group $\Gamma = \langle S \rangle$ and $\Sigma' = \{\sigma_i'\}_{i=1}^\infty$ is the sofic approximation defined by $\sigma_i' : \Gamma \to \mathrm{Sym}(V_i)$ is the unique homomorphism satisfying

$$\sigma_i'(s) = \sigma_i(s) \quad \text{for all } s \in S$$

then $\Sigma'$ is by homomorphisms and it is equivalent to $\Sigma$.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*Problem 4.* If $\Gamma$ is an interesting group, such as the fundamental group of a surface, $\mathbb{F}_2 \times \mathbb{Z}$, $\mathbb{F}_2 \times \mathbb{F}_2$, $\mathrm{SL}(2, \mathbb{Z}) \rtimes \mathbb{Z}^2$ or $\mathrm{SL}(3, \mathbb{Z})$, is every sofic approximation to $\Gamma$ equivalent to one by homomorphisms?

2.3. *Topological sofic entropy.* Given a countable group $\Gamma$, a sofic approximation $\Sigma$ to $\Gamma$, a compact metrizable space $X$, and an action $T = (T^g)_{g\in\Gamma}$ on $X$ by homeomorphisms, we will define the topological sofic entropy $h_\Sigma(T)$. In a nutshell, the entropy is the exponential rate of growth of the number of approximate partial orbits that can be distinguished up to some small scale.

First, we recall some basic concepts. A *pseudometric* on a space $X$ is a function $\rho : X \times X \to [0, \infty)$ satisfying all of the properties of a metric with one exception: it is possible that $\rho(x, y) = 0$ even if $x \neq y$. If $\rho$ is a pseudometric then a subset $S \subset X$ is $(\rho, \epsilon)$-*separated* if $\rho(s_1, s_2) \geq \epsilon$ for all $s_1, s_2 \in S$ with $s_1 \neq s_2$. Let $N_\epsilon(S, \rho)$ denote the maximum cardinality of a $(\rho, \epsilon)$-separated subset of $S$. We also let $\rho_2$ and $\rho_\infty$ denote the pseudometrics on $X^d$ (for any integer $d \geq 1$) defined by

$$\rho_\infty(x, y) = \max_i \rho(x_i, y_i), \quad \rho_2(x, y) = \left(\frac{1}{d}\sum_i \rho(x_i, y_i)^2\right)^{1/2},$$

where $x = (x_1, \ldots, x_d)$, $y = (y_1, \ldots, y_d) \in X^d$.

Given an action $T = (T^g)_{g\in\Gamma}$ on $X$, a map $\sigma : \Gamma \to \mathrm{Sym}(d)$, a finite subset $F \subset \Gamma$ and $\delta > 0$, let $\mathrm{Map}(T, \rho, F, \delta, \sigma)$ denote the set of all $x \in X^d$ such that

$$\rho_2(T^f x, x \circ \sigma(f)) < \delta \quad \text{for all } f \in F,$$

where $(T^f x)_i = T^f x_i$ and $(x \circ \sigma(f))_i = x_{\sigma(f)i}$ for all $i$.

In the literature, an element $x \in \mathrm{Map}(T, \rho, F, \delta, \sigma)$ has been referred to as a *microstate*, a *good model* or a *good map*. These terms will be used informally and will not be defined rigorously. The *entropy of $T$ with respect to $\rho$* is

$$h_\Sigma(T, \rho) = \sup_{\epsilon>0} \inf_{F\Subset\Gamma} \inf_{\delta>0} \limsup_{i\to\infty} |V_i|^{-1} \log(N_\epsilon(\mathrm{Map}(T, \rho, F, \delta, \sigma_i), \rho_\infty)),$$

where $F \Subset \Gamma$ means that $F$ is a finite subset of $\Gamma$. We will also write

$$h_\Sigma(\Gamma \curvearrowright X, \rho) = h_\Sigma(T, \rho)$$

if $T$ is implicit.

*Exercise 7.* Suppose $\Gamma$ is residually finite and has finite-index normal subgroups $\Gamma \geq N_1 \geq N_2 \geq \cdots$ with $\cap_i N_i = \{e\}$. Let $\sigma_i : \Gamma \to \mathrm{Sym}(\Gamma/N_i)$ be the canonical homomorphisms. By Exercise 4, $\Sigma = \{\sigma_i\}$ is a sofic approximation. Now suppose that $z \in X$ is stabilized by $(T^g)_{g\in N_i}$ (so $z$ is $(T, N_i)$-*periodic*). Show that if $x \in X^{\Gamma/N_i}$ is defined by $x_{gN_i} = T^g z$ then $x \in \mathrm{Map}(T, \rho, F, \delta, \sigma_i)$ for every $F, \delta$. It follows that the $\Sigma$-entropy of $T$ is at least the exponential rate of growth of the $(T, N_i)$-periodic points.

A pseudometric $\rho$ on $X$ is *generating* for the action if for every $x, y \in X$ with $x \neq y$ there exists $g \in \Gamma$ with $\rho(gx, gy) > 0$.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

THEOREM 2.2. *Let $\Gamma \curvearrowright X$ be an action by homeomorphisms on a compact metrizable space. If $\rho_1$, $\rho_2$ are continuous generating pseudometrics on $X$ then*

$$h_\Sigma(T, \rho_1) = h_\Sigma(T, \rho_2).$$

This theorem (and the definition of topological sofic entropy) is due to Kerr–Li [**KL11b**]. See also [**KL13b**, Proposition 2.4] and [**KL16**] for a simplified exposition.

*Proof sketch.* The first step is showing that we can replace a generating pseudometric $\rho$ with a metric $\rho'$. To be precise, let $\phi \in \ell^1(\Gamma)$ be a strictly positive function. Define

$$\rho'(x, y) := \sum_{g \in \Gamma} \rho(T^g x, T^g y)\phi(g).$$

Then $\rho'$ is a continuous metric on $X$ and

$$h_\Sigma(T, \rho) = h_\Sigma(T, \rho').$$

Informally, this is because any microstate for $\Gamma \curvearrowright X$ with respect to $\rho$ is a microstate with respect to $\rho'$ and *vice versa*, although the parameters $F$ and $\delta$ may change.

We can now assume that $\rho_1$ and $\rho_2$ are metrics. The statement can now be derived from the observation that for any $\epsilon > 0$ there is a $\delta > 0$ such that $\rho_1(x, y) < \delta \Rightarrow \rho_2(x, y) < \epsilon$ and *vice versa*. □

*Definition 2.* The *$\Sigma$-entropy of $T$* is $h_\Sigma(T) := h_\Sigma(T, \rho)$ where $\rho$ is any continuous generating pseudometric.

*Remark 2.* The $\rho_\infty$ appearing in the formula for $h_\Sigma(T, \rho)$ can be replaced with $\rho_2$ without affecting the definition of $h_\Sigma(T, \rho)$. Also the lim sup can be replaced by a lim inf or an ultralimit; however these replacements can lead to different invariants because sofic entropy depends on the choice of sofic approximation in general (see §4.1).

*Exercise 8.* (Symbolic dynamics) Suppose $A$ is a finite set. An element $\mathbf{x} \in A^\Gamma$ is written as either a collection $\mathbf{x} = (x_g)_{g \in \Gamma}$ or a function $\mathbf{x} : \Gamma \to A$. Let $T = (T^g)_{g \in \Gamma}$ be the *shift action* on $A^\Gamma$ defined by $T^g \mathbf{x}(f) = \mathbf{x}(g^{-1} f)$.

(1) Let $\rho$ be the pseudometric on $A^\Gamma$ given by $\rho(\mathbf{x}, \mathbf{y}) = 1$ if $\mathbf{x}_e \neq \mathbf{y}_e$ and $\rho(\mathbf{x}, \mathbf{y}) = 0$ otherwise. Show that $\rho$ is generating for the shift-action.

(2) Suppose $X \subset A^\Gamma$ is closed and shift-invariant. Given $\mathbf{x} \in A^{V_i}$, its *pullback name* is

$$\Pi_v^{\sigma_i}(\mathbf{x}) \in A^\Gamma, \quad \Pi_v^{\sigma_i}(\mathbf{x})(g) = \mathbf{x}(\sigma_i(g)^{-1} v).$$

Also let

$$P_{\mathbf{x}}^{\sigma_i} = |V_i|^{-1} \sum_{v \in V_i} \delta_{\Pi_v^{\sigma_i}(\mathbf{x})} \in \mathrm{Prob}(A^\Gamma)$$

be its *empirical distribution*. Show that the entropy of the restriction of $T$ to $X$ simplifies to

$$h_\Sigma(T \restriction X) = \inf_{\mathcal{O}} \inf_{\delta > 0} \limsup_{i \to \infty} |V_i|^{-1} \log \#\{\mathbf{x} \in A^{V_i} : P_{\mathbf{x}}^{\sigma_i}(\mathcal{O}) > 1 - \delta\},$$

where the first infimum is over all open neighborhoods $\mathcal{O}$ of $X$ in $A^\Gamma$.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

(3) Show that $h_\Sigma(T) = \log |A|$.
(4) Show that if $X \subset A^\Gamma$ is not $A^\Gamma$ then $h_\Sigma(T \restriction X) < \log |A|$.
(5) Suppose $\Gamma$ is residually finite and has finite-index normal subgroups $\Gamma \geq N_1 \geq N_2 \geq \cdots$ with $\cap_i N_i = \{e\}$. Let $\sigma_i : \Gamma \to \mathrm{Sym}(\Gamma/N_i)$ be the canonical homomorphisms. Let $\mathcal{O}$ be an open neighborhood of $X$ in $A^\Gamma$ and $\delta > 0$. Show that $h_\Sigma(T \restriction X)$ is at most the exponential rate of growth of the number of $(T, N_i)$-periodic points $z \in A^\Gamma$ such that

$$\#\{gN_i \in \Gamma/N_i : gz \in \mathcal{O}\} \geq (1 - \delta)|\Gamma/N_i|.$$

Compare with the lower bound in Exercise 7.

THEOREM 2.3. [**KL13b**] *If $\Gamma$ is amenable then topological sofic entropy agrees with classical topological entropy.*

*Remark 3.* The main tool involved in the proof of this theorem is a Rokhlin lemma for sofic approximations of countable amenable groups. This lemma allows us to express any sofic approximation to an amenable group $\Gamma$ as essentially equivalent to a Følner sequence.

Topological sofic entropy can also be defined in terms of open covers [**Zha12**] (in a manner similar to the original definition of topological entropy [**AKM65**]) or in terms of sequences of continuous functions [**KL11b**].

2.3.1. *An application: Gottschalk's conjecture and Kaplansky's conjecture.* Any self-map of a finite set satisfies the following property: if it is injective then it must be surjective. This fundamental property is called *surjunctivity*. It has been generalized to algebraic varieties and regular maps [**Ax68**] and proalgebraic varieties satisfying a soficity condition [**Gro99**].

*Conjecture 1.* (Gottschalk's surjunctivity conjecture) Suppose $A$ is a finite set (called an alphabet), $\Gamma$ a countable group and $\phi : A^\Gamma \to A^\Gamma$ a continuous $\Gamma$-equivariant map. If $\phi$ is injective then it must be surjective.

It was this conjecture that inspired Gromov to invent sofic groups (although the name 'sofic', derived from the Hebrew word for finite, was coined by Weiss [**Wei00**]). Gromov proved the conjecture holds for all sofic groups. A new proof, obtained by Kerr and Li [**KL11b**] goes as follows: assuming $\phi$ is injective,

$$h_\Sigma(\Gamma \curvearrowright \phi(A^\Gamma)) = h_\Sigma(\Gamma \curvearrowright A^\Gamma) = \log |A|.$$

However, if $\phi$ is not surjective then $h_\Sigma(\Gamma \curvearrowright \phi(A^\Gamma)) < \log |A|$ (because the image $\phi(A^\Gamma)$ has trivial intersection with some finite cylinder set). This implies the conjecture.

Now suppose that $A$ is a finite field and $\phi$ is $A$-linear. In this case, we can think of $\phi$ as an element of the group ring $A\Gamma$. The theorem implies that the group $A\Gamma$ is *directly finite*: that is $xy = 1$ implies $yx = 1$ for all $x, y \in A\Gamma$. More generally, because all fields can be embedded into an ultraproduct of finite fields, the same result holds when $A$ is an arbitrary field. This proves Kaplansky's direct finiteness conjecture for sofic groups. Actually, more is true: $A\Gamma$ is directly finite whenever $A$ is a matrix algebra over a division ring [**ES04**].

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

2.4. *Measure sofic entropy.* There are two equivalent definitions of measure sofic entropy: one via pseudometrics (similar to topological entropy) and one via partitions.

2.4.1. *The pseudometric definition.* Suppose $X$ is a compact metrizable space, $T = (T^g)_{g\in\Gamma}$ is an action on $X$ by homeomorphisms and $\mu$ is an invariant Borel probability measure on $X$. Let $\mathrm{Prob}(X)$ denote the space of Borel probability measures on $X$. Recall that the weak* topology on $\mathrm{Prob}(X)$ is defined as follows: a sequence $\{\mu_n\}_{n\in\mathbb{N}}$ converges to a measure $\mu_\infty$ if and only if for every continuous function $f : X \to \mathbb{C}$,

$$\int f\, d\mu_n \to \int f\, d\mu_\infty$$

as $n \to \infty$. By the Banach–Alaoglu theorem, $\mathrm{Prob}(X)$ is compact in the weak* topology.

Given a pseudometric $\rho$ on $X$, a finite subset $F \subset \Gamma$, $\delta > 0$ and $\sigma : \Gamma \to \mathrm{Sym}(V)$, define $\mathrm{Map}(T, \rho, F, \delta, \sigma)$ as in §2.3. In addition, if $\mathcal{O} \subset \mathrm{Prob}(X)$ is an open neighborhood of $\mu$ then let $\mathrm{Map}(T, \rho, \mathcal{O}, F, \delta, \sigma)$ denote the set of all $x \in \mathrm{Map}(T, \rho, F, \delta, \sigma)$ such that $x_* u_V \in \mathcal{O}$ where $u_V$ denotes the uniform probability measure on $V$. These are the microstates that are approximately equidistributed.

The *sofic entropy of $T$ with respect to $\rho$ and $\Sigma$* is

$$h_{\Sigma,\mu}(T, \rho) = \sup_{\epsilon>0} \inf_{\mathcal{O}} \inf_{F\in\Gamma} \inf_{\delta>0} \limsup_{i\to\infty} |V_i|^{-1} \log(N_\epsilon(\mathrm{Map}(T, \rho, \mathcal{O}, F, \delta, \sigma_i), \rho_\infty)).$$

Intutively, this measures the exponential rate of growth of the number of microstates for the action that are approximately equidistributed with respect to $\mu$.

As in the topological case, a pseudometric $\rho$ on $X$ is *generating* with respect to $T$ if for every $x, y \in X$ with $x \neq y$ there exists $g \in \Gamma$ with $\rho(T^g x, T^g y) > 0$.

THEOREM 2.4. [**KL11b**] *For $i = 1, 2$ let $T_i$ be pmp actions of $\Gamma$ by homeomorphisms on compact metrizable spaces $X_i$ and $\rho_i$ be continuous generating pseudometrics on $X_i$. If these actions are measurably conjugate then*

$$h_{\Sigma,\mu_1}(T_1, \rho_1) = h_{\Sigma,\mu_2}(T_2, \rho_2).$$

*Proof sketch.* As in the proof of Theorem 2.2, we can assume, without loss of generality that $\rho_1$ and $\rho_2$ are metrics, not just pseudometrics. Let $\Phi : X_1 \to X_2$ be a measure-conjugacy. By Lusin's theorem, for every $\eta > 0$ there exists a compact set $Y_1 \subset X_1$ such that $\Phi$ restricted to $Y_1$ is uniformly continuous and $\mu_1(Y_1) > 1 - \eta$.

Recall that a subset $Z \subset X_1$ is a *continuity set* if $\mu_1(\partial Z) = 0$ where $\partial Z = \bar{Z} \cap \overline{X_1 - Z}$. For simplicity suppose that Lusin's set $Y_1$ defined above is a continuity set and that its image $\Phi(Y_1) =: Y_2$ is also a continuity set. This does not have to be true but since the continuity sets form an algebra that is dense in the measure algebra it is approximately true.

The Portmanteau theorem states that a sequence $\{\nu_n\}$ of Borel probability measures in $\mathrm{Prob}(X_1)$ (say) converges to $\nu_\infty$ in the weak* topology if and only if $\lim_n \nu_n(Z) = \nu_\infty(Z)$ for every continuity set $Z \subset X_1$. It follows that any microstate for $\Gamma \curvearrowright (X_1, \mu_1)$ pushes forward under $\Phi$ to a microstate for $\Gamma \curvearrowright (X_2, \mu_2)$ although the parameters qualifying how good (or bad) the microstate is may change. The theorem follows from this. □

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*Remark 4.* The proof sketch above is very different from the proofs in [**KL11b**], which are operator-algebraic. See also [**KL13b**, Proposition 3.4].

*Remark 5.* In [**Hay18**] Hayes relaxes the condition that $X$ is compact to being merely completely metrizable and separable assuming the measure satisfies a 'tightness' condition.

*Definition 3.* (Measure sofic entropy) The *measure sofic entropy* of the action $T$ with respect to $\Sigma$ is

$$h_{\Sigma,\mu}(T) = h_{\Sigma,\mu_1}(T_1, \rho_1),$$

where $\Gamma \curvearrowright^{T_1} (X_1, \mu_1)$ is any compact topological model for $\Gamma \curvearrowright^{T} (X, \mu)$ and $\rho_1$ is any generating pseudometric on $X_1$.

THEOREM 2.5. *If $\Gamma$ is amenable then measure sofic entropy agrees with classical Kolmogorov–Sinai entropy.*

*Remark 6.* There are two very different proofs of this result. The one in [**KL13b**] is based on the sofic Rokhlin's lemma for amenable groups. The other in [**Bow12b**] is based on a sofic-version of the Rudolph–Weiss theorem that relative entropy is preserved under orbit-equivalence with respect to the orbit-change sigma-algebra.

*Remark 7.* As in the topological case, the $\rho_\infty$ appearing in the formula for $h_{\Sigma,\mu}(T, \rho)$ can be replaced with $\rho_2$ without affecting its value. Also the lim sup can be replaced by a lim inf or an ultralimit; however these replacements can lead to different invariants because sofic entropy depends on the choice of sofic approximation in general (see §4.1).

THEOREM 2.6. *If $\Gamma \curvearrowright (K, \kappa)^\Gamma$ is a Bernoulli shift and $\Sigma$ is an arbitrary sofic approximation to $\Gamma$ then*

$$h_{\Sigma,\kappa^\Gamma}(\Gamma \curvearrowright K^\Gamma) = H(K, \kappa).$$

The case in which $H(K, \kappa) < \infty$ is proven in [**Bow10b**]. The infinite entropy case is handled in [**KL11a**].

*Proof sketch.* The lower bound is obtained as follows: fix an open neighborhood $\mathcal{O}$, finite $F \subset \Gamma$ and $\delta > 0$. Let $\phi : V_i \to K$ be a random map with law equal to the product measure $\kappa^{V_i}$. Let $\widetilde{\phi} : V_i \to K^\Gamma$ be the 'pullback' defined by

$$\widetilde{\phi}(v)(g) = \phi(\sigma_i(g)^{-1} v).$$

Then using Chebyshev's inequality it is shown that with high probability, when $i$ is large, $\widetilde{\phi} \in \mathrm{Map}(T, \rho, \mathcal{O}, F, \delta, \sigma_i)$. Here $\rho$ is the pseudometric given by $\rho(x, y) = \rho_K(x(e), y(e))$ where $\rho_K$ is an arbitrary metric on $K$ (which may be assumed to be a compact metrizable space). The law of large numbers now gives the lower bound.

In the case $H(K, \kappa) < \infty$, the upper bound is shown as follows. Let $\phi : V_i \to K^\Gamma$ be any microstate. Let $\pi : K^\Gamma \to K$ be projection to the identity coordinate. Then if $\phi$ is a good enough microstate the composition $\pi \circ \phi$ pushes the uniform measure $u_{V_i}$ forward to a measure on $K$ that is close to $\kappa$ in total variation distance. On the other hand, observe that $\phi$ is essentially determined by $\pi \circ \phi$. So it suffices to observe that the number of maps $\phi' : V_i \to K$ such that $\phi'_* u_{V_i}$ is close to $\kappa$ is approximately $\exp(H(K, \kappa)|V_i|)$. This is an application of elementary combinatorics and Stirling's formula. □

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

2.4.2. *The partition definition.*

*Definition 4.* If $\Sigma_1, \Sigma_2$ are sigma-algebras on sets $X_1, X_2$ respectively then a *homomorphism* between them is a map $\phi : \Sigma_1 \to \Sigma_2$ such that for all $A, B \in \Sigma_1$,

$$\phi(A \cup B) = \phi(A) \cup \phi(B), \quad \phi(A \cap B) = \phi(A) \cap \phi(B), \quad \phi(\emptyset) = \emptyset, \quad \phi(X_1) = X_2.$$

Suppose $\Gamma \curvearrowright^T (X, \mu)$ is a pmp action. For simplicity, $T^g x$ is denoted by $gx$ for $g \in \Gamma$, $x \in X$.

*Definition 5.* Let $\mathcal{P}$ be a finite measurable partition of $X$, $F \subset \Gamma$ a finite set with $1_\Gamma \in F$, and $\delta > 0$. Let $\mathcal{P}^F = \bigvee_{f \in F} f^{-1}\mathcal{P}$ be the coarsest partition containing $f^{-1}\mathcal{P}$ for $f \in F$. If $\mathcal{Q}$ is any partition, let $\sigma$-alg$(\mathcal{Q})$ be the smallest sigma-algebra containing $\mathcal{Q}$. Also let $2^V$ denote the sigma-algebra of all subsets of $V$ and $u_V$ be the uniform probability measure on $V$.

Given $\sigma : \Gamma \to \mathrm{Sym}(V)$, let $\mathrm{Hom}_\mu(\mathcal{P}, F, \delta, \sigma)$ be the set of all homomorphisms $\phi : \sigma\text{-alg}(\mathcal{P}^F) \to 2^V$ such that:

(1) $\sum_{P \in \mathcal{P}} u_V(\sigma_s \phi(P) \triangle \phi(sP)) < \delta$ for all $s \in F^{-1}$ and

(2) $\sum_{A \in \mathcal{P}^F} |u_V(\phi(A)) - \mu(A)| < \delta$.

The sofic entropy is defined as the exponential rate of growth of the number of such homomorphisms that can be extended to a more refined partition. To be precise, if $\mathcal{Q} \leq \mathcal{P}$ is a partition coarser than $\mathcal{P}$ and $\phi : \sigma\text{-alg}(\mathcal{P}^F) \to 2^V$ is a homomorphism then let $\phi \restriction \mathcal{Q}$ be the restriction of $\phi$ to $\mathcal{Q}$. Let $|\mathrm{Hom}_\mu(\mathcal{P}, F, \delta, \sigma)|_\mathcal{Q}$ be the cardinality of the set of restrictions $\{\phi \restriction \mathcal{Q} : \phi \in \mathrm{Hom}_\mu(\mathcal{P}, F, \delta, \sigma)\}$. Finally, for a sigma-sub-algebra $\mathcal{S} \subset \mathcal{B}$ define

$$h_{\Sigma,\mu}(T, \mathcal{S}) = h_{\Sigma,\mu}(\Gamma \curvearrowright X, \mathcal{S}) = \sup_\mathcal{Q} \inf_\mathcal{P} \inf_{F \subset \Gamma} \inf_{\delta > 0} \limsup_{i \to \infty} \frac{1}{|V_i|} \log |\mathrm{Hom}_\mu(\mathcal{P}, F, \delta, \sigma_i)|_\mathcal{Q},$$

where the sup is over all finite partitions $\mathcal{Q} \subset \mathcal{S}$, the first inf is over finite partitions $\mathcal{P}$ with $\mathcal{Q} \leq \mathcal{P} \subset \mathcal{S}$ and the second inf is over all finite subsets of $\Gamma$.

Recall that a sigma-sub-algebra $\mathcal{S} \subset \mathcal{B}_X$ is *generating* for $T$ if $\mathcal{B}_X$ is the smallest $T(\Gamma)$-invariant sigma-algebra containing $\mathcal{S}$ up to sets of measure zero.

THEOREM 2.7. **[Ker13]** *If $\mathcal{S} \subset \mathcal{B}_X$ is a generating sigma-sub-algebra then $h_{\Sigma,\mu}(T, \mathcal{S}) = h_{\Sigma,\mu}(T)$.*

*Remark 8.* In the special case in which $\mathcal{S}$ is the sigma-algebra generated by a finite partition, the definition above is easily seen to be equivalent to the one given in the introduction.

2.5. *The $f$-invariant.* Let $S$ be a finite or countable set and $\Gamma = \langle S \rangle$ the free group generated by $S$. Let $\Gamma \curvearrowright^T (X, \mu)$ be a pmp action and $\mathcal{P}$ a measurable partition of $X$ with finite Shannon entropy. Define

$$F_\mu(T, \mathcal{P}) := H_\mu(\mathcal{P}) + \sum_{s \in S} (H_\mu(\mathcal{P} \vee s\mathcal{P}) - 2H_\mu(\mathcal{P})),$$

$$f_\mu(T, \mathcal{P}) := \inf_{W \Subset \Gamma} F_\mu(T, \mathcal{P}^W),$$

where $\mathcal{P}^W = \bigvee_{w \in W} w^{-1}\mathcal{P}$. For simplicity, we have written $s\mathcal{P}$ instead of $T^s\mathcal{P}$.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

THEOREM 2.8. [**Bow10d**] *If $\mathcal{P}, \mathcal{Q}$ are generating partitions each with finite Shannon entropy then $f_\mu(T, \mathcal{P}) = f_\mu(T, \mathcal{Q})$.*

The $f$-invariant of the action is defined by $f_\mu(T) = f_\mu(T, \mathcal{P})$ where $\mathcal{P}$ is any generating partition with finite Shannon entropy. The $f$-invariant is undefined if no such partition exists.

*Remark 9.* This was proven in [**Bow10d**] under the assumption that $S$ is finite. The proof when $S$ is countable is essentially the same.

*Proof sketch.* Given two partitions $\mathcal{P}, \mathcal{Q}$, their *Rokhlin distance* is defined by

$$d(\mathcal{P}, \mathcal{Q}) := H_\mu(\mathcal{P}|\mathcal{Q}) + H_\mu(\mathcal{Q}|\mathcal{P}).$$

Partitions that agree up to measure zero sets are identified. With this convention, the Rokhlin distance really is a distance function on the space of all partitions with finite Shannon entropy, which is denoted by $\mathrm{Part}(X, \mu)$.

Two partitions $\mathcal{P}, \mathcal{Q}$ are *combinatorially equivalent* if there exist finite subsets $F, K \subset \Gamma$ such that $\mathcal{Q} \leq \mathcal{P}^F$ and $\mathcal{P} \leq \mathcal{Q}^K$. The first step is showing that if $\mathcal{P} \in \mathrm{Part}(X, \mu)$ is a generating partition then its combinatorial equivalence class is dense in the subspace of all generating partitions with finite Shannon entropy. Since $F$ is continuous on $\mathrm{Part}(X, \mu)$, $f$ is upper semi-continuous. It now suffices to show that if $\mathcal{P}, \mathcal{Q}$ are combinatorially equivalent then $f_\mu(T, \mathcal{P}) = f_\mu(T, \mathcal{Q})$.

The partition $\mathcal{Q}$ is a *simple splitting* of $\mathcal{P}$ if there is an element $s \in S \cup S^{-1}$ and a partition $\mathcal{R} \leq \mathcal{P}$ such that $\mathcal{Q} = \mathcal{P} \vee s\mathcal{R}$. We say $\mathcal{Q}$ is a *splitting* of $\mathcal{P}$ if there is a sequence $\mathcal{P} = \mathcal{Q}_0, \mathcal{Q}_1, \ldots, \mathcal{Q}_n = \mathcal{Q}$ such that $\mathcal{Q}_{i+1}$ is a simple splitting of $\mathcal{Q}_i$ for $0 \leq i < n$. The second step is showing that if $\mathcal{P}_1, \mathcal{P}_2$ are combinatorially equivalent then there exists a common splitting $\mathcal{Q}$ of both of them. Moreover, splittings preserve the combinatorial equivalence class. Therefore, it suffices to show the following: if $\mathcal{Q}$ is a simple splitting of $\mathcal{P}$ then $F_\mu(T, \mathcal{Q}) \leq F_\mu(T, \mathcal{P})$. This fact follows from a short calculation. For simplicity, assume $\mathcal{R} \leq \mathcal{P}$, $t \in S$ and $\mathcal{Q} = \mathcal{P} \vee t\mathcal{R}$. Then

$$\begin{aligned}
F_\mu(T, \mathcal{Q}) &= H_\mu(\mathcal{Q}) + \sum_{s \in S} H_\mu(\mathcal{Q} \vee s\mathcal{Q}) - 2H_\mu(\mathcal{Q}) \\
&= H_\mu(\mathcal{P}) + H_\mu(\mathcal{Q}|\mathcal{P}) + \sum_{s \in S} H_\mu(\mathcal{P} \vee s\mathcal{P}) - 2H_\mu(\mathcal{P}) \\
&\quad + H_\mu(\mathcal{Q} \vee s\mathcal{Q}|\mathcal{P} \vee s\mathcal{P}) - 2H_\mu(\mathcal{Q}|\mathcal{P}) \\
&= F_\mu(T, \mathcal{P}) + H_\mu(\mathcal{Q}|\mathcal{P}) + \sum_{s \in S} H_\mu(\mathcal{Q} \vee s\mathcal{Q}|\mathcal{P} \vee s\mathcal{P}) - 2H_\mu(\mathcal{Q}|\mathcal{P}) \\
&= F_\mu(T, \mathcal{P}) + H_\mu(\mathcal{Q}|\mathcal{P}) + \sum_{s \in S} H_\mu(\mathcal{Q}|\mathcal{P} \vee s\mathcal{P}) \\
&\quad + H_\mu(s\mathcal{Q}|\mathcal{P} \vee s\mathcal{P} \vee \mathcal{Q}) - 2H_\mu(\mathcal{Q}|\mathcal{P}) \\
&= F_\mu(T, \mathcal{P}) + (H_\mu(\mathcal{Q}|\mathcal{P} \vee t\mathcal{P}) + H_\mu(t\mathcal{Q}|\mathcal{P} \vee t\mathcal{P} \vee \mathcal{Q}) - H_\mu(\mathcal{Q}|\mathcal{P})) \\
&\quad + \sum_{s \in S - \{t\}} H_\mu(\mathcal{Q}|\mathcal{P} \vee s\mathcal{P}) + H_\mu(s\mathcal{Q}|\mathcal{P} \vee s\mathcal{P} \vee \mathcal{Q}) - 2H_\mu(\mathcal{Q}|\mathcal{P}).
\end{aligned}$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

Observe that $H_\mu(\mathcal{Q}|\mathcal{P} \vee s\mathcal{P}) + H_\mu(s\mathcal{Q}|\mathcal{P} \vee s\mathcal{P} \vee \mathcal{Q}) - 2H_\mu(\mathcal{Q}|\mathcal{P}) \leq 0$ for every $s \in S$ and moreover if $s = t$ then $H_\mu(\mathcal{Q}|\mathcal{P} \vee s\mathcal{P}) = 0$. So

$$F_\mu(T, \mathcal{Q}) \leq F_\mu(T, \mathcal{P}) + H_\mu(t\mathcal{Q}|\mathcal{P} \vee t\mathcal{P} \vee \mathcal{Q}) - H_\mu(\mathcal{Q}|\mathcal{P}) \leq F_\mu(T, \mathcal{P}). \qquad \square$$

*Remark 10.* The proof shows a little more: partitions can be partially ordered by $\mathcal{P} \prec \mathcal{Q}$ if $\mathcal{Q}$ is a splitting of $\mathcal{P}$. Then $f_\mu(T, \mathcal{P})$ is the limit of $F_\mu(T, \mathcal{Q})$ as $\mathcal{Q}$ tends to infinity in this partial order. Moreover,

$$f_\mu(T, \mathcal{P}) = \lim_{n\to\infty} F_\mu(T, \mathcal{P}^{W_n}),$$

where $W_n$ is any increasing sequence of finite subsets of $\Gamma$ such that (1) the induced subgraph of $W_n$ is connected in the Cayley graph of $(\Gamma, S)$ and (2) $\cup_n W_n = \Gamma$. Using this last fact, a direct computation shows that the $f$-invariant of the Bernoulli shift $\Gamma \curvearrowright (K, \kappa)^\Gamma$ is the Shannon entropy $H(K, \kappa)$.

Why do we care about the $f$-invariant? In contrast to sofic entropy, the $f$-invariant tends to be easy to compute. For example, for Markov processes $f = F$ (see §3.3). Morever, it is additive under direct products, satisfies an ergodic decomposition formula, a subgroup formula, has a relative entropy theory and satisfies some cases of Yuzvinskii's formula. These results do not hold for sofic entropy in general.

2.5.1. *Other formulations and other groups.* If $\mathcal{P}$ is a Markov partition then $F_\mu(T, \mathcal{P}) = f_\mu(T, \mathcal{P})$. Using this fact, the following alternative formula for the $f$-invariant was found in [**Bow10c**]. Define

$$F^*_\mu(T, \mathcal{P}) := H_\mu(\mathcal{P}) + \sum_{s\in S}(h_\mu(T^s, \mathcal{P}) - H_\mu(\mathcal{P})).$$

THEOREM 2.9. [**Bow10c**]

$$f_\mu(T, \mathcal{P}) = \inf_{W\Subset\Gamma} F^*_\mu(T, \mathcal{P}^W) = \lim_{n\to\infty} F^*_\mu(T, \mathcal{P}^{W_n}),$$

*where the limit is with respect to any increasing sequence of finite subsets of* $\Gamma$ *such that (1) the induced subgraph of* $W_n$ *is connected in the Cayley graph of* $(\Gamma, S)$ *and (2)* $\cup_n W_n = \Gamma$.

The theorem above leads to the following idea: suppose for $i = 1, 2$, $\Gamma_i$ are amenable groups, $A_i \leq \Gamma_i$ are subgroups and $\phi : A_1 \to A_2$ is an isomorphism. Let $\Gamma = \Gamma_1 *_\phi \Gamma_2$ be the amalgamated free product. Let $\Gamma \curvearrowright^T (X, \mu)$ be a pmp action and $\mathcal{P}$ a partition of $X$ with finite Shannon entropy. Define

$$F_\mu(T, \mathcal{P}) = h_\mu(\Gamma_1 \curvearrowright X, \mathcal{P}) + h_\mu(\Gamma_2 \curvearrowright X, \mathcal{P}) - h_\mu(A \curvearrowright X, \mathcal{P}),$$

where $A \leq \Gamma$ is the subgroup corresponding to $A_1$, $A_2$ and, for example, $h_\mu(\Gamma_i \curvearrowright X, \mathcal{P})$ is the classical entropy rate of $\mathcal{P}$ with respect to $\Gamma_i$-action. Also let

$$f_\mu(T, \mathcal{P}) = \inf_{W\Subset\Gamma} F_\mu(T, \mathcal{P}^W).$$

As above, it can be shown that if $\mathcal{P}$, $\mathcal{Q}$ are generating partitions with finite Shannon entropy then $f_\mu(T, \mathcal{P}) = f_\mu(T, \mathcal{Q})$ and so this determines a measure-conjugacy invariant for $\Gamma$-actions.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*Problem 5.* This idea has not appeared yet in the literature and is worthy of further exploration: can one extend it to other graphs of groups? Does it depend on how the group is represented as a graph of groups? Can one obtain results for such invariants similar to the results for the $f$-invariant of free group actions (for example, the sofic interpretation, the ergodic decomposition formula, the subgroup formula, and so on)? For example, the fundamental group of a closed surface of genus $g \geq 2$ can be written as a free product of free groups amalgamated over an infinite cyclic subgroup.

2.5.2. *The sofic interpretation of the $f$-invariant.* In order to interpret the $f$-invariant as a kind of sofic entropy, we introduce random sofic approximations and their sofic entropies. Let $\{V_i\}_i$ be a sequence of finite sets and for each $i$, let $\mathbb{P}_i$ be a probability measure on the space of maps $\Gamma \to \mathrm{Sym}(V_i)$. The sequence $\mathbb{P} := \{\mathbb{P}_i\}_{i=1}^{\infty}$ is a *random sofic approximation* to $\Gamma$ if for every finite set $F \subset \Gamma$ and $\delta > 0$,

- $$1 = \lim_{i\to\infty} \mathbb{P}_i\{\sigma : \Gamma \to \mathrm{Sym}(V_i) : \sigma \text{ is } (F, \delta)\text{-trace preserving}\};$$
- there exists $I$ such that $i > I$ implies $\mathbb{P}_i$-a.e. $\sigma$ is $(F, \delta)$-multiplicative;
- $\lim_i |V_i| = +\infty$.

Each definition of $\Sigma$-entropy given above can be generalized to $\mathbb{P}$-entropy by replacing $N_\epsilon(\cdot)$ or $|\mathrm{Hom}_\mu(\cdot)|$ with its expectation with respect to $\mathbb{P}_i$. For example, suppose $\Gamma \curvearrowright^T X$ is a continuous action on a compact space, $\mu$ an invariant probability measure on $X$ and $\rho$ a continuous generating pseudometric. Then the topological $\mathbb{P}$-entropy and measure $\mathbb{P}$-entropy are

$$h_{\mathbb{P}}(T, \rho) := \sup_{\epsilon>0} \inf_{F\Subset\Gamma} \inf_{\delta>0} \limsup_{i\to\infty} |V_i|^{-1} \log \mathbb{E}_i(N_\epsilon(\mathrm{Map}(T, \rho, F, \delta, \sigma_i), \rho_\infty)),$$

$$h_{\mathbb{P},\mu}(T, \rho) := \sup_{\epsilon>0} \inf_{\mathcal{O}} \inf_{F\Subset\Gamma} \inf_{\delta>0} \limsup_{i\to\infty} |V_i|^{-1} \log \mathbb{E}_i(N_\epsilon(\mathrm{Map}(T, \rho, \mathcal{O}, F, \delta, \sigma_i), \rho_\infty)),$$

where $\mathbb{E}_i$ denotes expectation with respect to $\mathbb{P}_i$. The obvious analogs of Theorems 2.2, 2.4 and 2.7 remain true and the proofs are essentially the same.

Now let $\Gamma = \langle S \rangle$ be a free group where $S$ is finite or countable. The set of homomorphisms from $\Gamma$ to $\mathrm{Sym}(n)$ is naturally identified with the set $\mathrm{Sym}(n)^S$ of all maps from $S$ to $\mathrm{Sym}(n)$. Let $\pi_n$ be the uniform probability measure on $\mathrm{Sym}(n)$ and $\mathbb{P}_n = \pi_n^S$ be the product measure on $\mathrm{Sym}(n)^S$.

*Exercise 9.* $\mathbb{P} = \{\mathbb{P}_n\}_{n=1}^{\infty}$ is a random sofic approximation to $\Gamma$.

THEOREM 2.10. *If $\mathcal{P}$ is a generating partition with finite Shannon entropy then*

$$f_\mu(T) = h_{\mathbb{P},\mu}(T).$$

*That is, the $f$-invariant of the action is the same as the sofic entropy with respect to the random sofic approximation $\mathcal{P}$.*

*Remark 11.* This is proven in [**Bow10a**] when $S$ is finite. The case of countably infinite $S$ is similar.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*Proof sketch.* For simplicity, assume $S$ is finite, $\mathcal{P}$ is finite and $\mu(P \cap T^s Q)$ is rational for every $P, Q \in \mathcal{P}$ and $s \in S$. Given $\sigma : \Gamma \to \mathrm{Sym}(n)$, let $\mathrm{Part}(\sigma, \mathcal{P})$ be the set of all maps $\phi : [n] \to \mathcal{P}$ such that

$$n\mu(P \cap s^{-1}Q) = \#\{v \in [n] : \phi(v) = P, \phi(\sigma(s)v) = Q\}.$$

By direct combinatorial arguments, one can obtain an exact formula for $\mathbb{E}[|\mathrm{Part}(\sigma, \mathcal{P})|]$ where $\sigma : \Gamma \to \mathrm{Sym}(n)$ is uniformly random. An application of Stirling's formula shows

$$F_\mu(T, \mathcal{P}) = \limsup_{n\to\infty} n^{-1} \log \mathbb{E}[|\mathrm{Part}(\sigma, \mathcal{P})|].$$

To handle the case in which $\mu(P \cap T^s Q)$ is irrational, let $\mathrm{Part}(\sigma, \mathcal{P}, \epsilon)$ be the set of all maps $\phi : [n] \to \mathcal{P}$ such that

$$|\mu(P \cap s^{-1}Q) - n^{-1}\#\{v \in [n] : \phi(v) = P, \phi(\sigma(s)v) = Q\}| < \epsilon.$$

A perturbation argument and Stirling's formula implies

$$\begin{aligned} F_\mu(T, \mathcal{P}) &= \lim_{\epsilon \searrow 0} \limsup_{n\to\infty} n^{-1} \log \mathbb{E}[|\mathrm{Part}(\sigma, \mathcal{P}, \epsilon)|] \\ &= \lim_{\epsilon \searrow 0} \liminf_{n\to\infty} n^{-1} \log \mathbb{E}[|\mathrm{Part}(\sigma, \mathcal{P}, \epsilon)|]. \end{aligned}$$

The theorem follows by replacing $\mathcal{P}$ with $\mathcal{P}^W$ and taking the infimum over finite $W \subset \Gamma$. □

*Problem 6.* Unlike the $f$-invariant, $h_{\mathbb{P},\mu}(T)$ is well defined even if $\Gamma \curvearrowright^T (X, \mu)$ does not have a generating partition with finite Shannon entropy. The $f$-invariant satisfies many useful identities: it is additive under direct products, satisfies an ergodic decomposition formula, a subgroup formula and possesses a relative version. Can any of these results be extended to $h_{\mathbb{P},\mu}(T)$?

2.6. *Rokhlin entropy.* Let $\mathcal{B}_X$ denote the sigma-algebra of measurable subsets of $X$. For any subcollection $\mathcal{F} \subset \mathcal{B}_X$, let $\sigma\text{-alg}(\mathcal{F}) \subset \mathcal{B}_X$ denote the sub-sigma-algebra generated by $\mathcal{F}$ and, if $\Gamma \curvearrowright^T X$ is a measurable action then let $\sigma\text{-alg}(T, \mathcal{F})$ denote the smallest sub-sigma-algebra containing $T^g F$ for every $g \in \Gamma$ and $F \in \mathcal{F}$. We do not distinguish between sigma-algebras that agree up to null sets. Thus we write $\mathcal{F}_1 = \mathcal{F}_2$ if $\mathcal{F}_1$ and $\mathcal{F}_2$ agree up to null sets.

*Definition 6.* (Rokhlin entropy) The *Rokhlin entropy* of an ergodic pmp action $\Gamma \curvearrowright^T (X, \mu)$ is defined by

$$h^{\mathrm{Rok}}(T) = \inf_{\mathcal{P}} H_\mu(\mathcal{P})$$

where the infimum is over all partitions $\mathcal{P}$ with $\sigma\text{-alg}(T, \mathcal{P}) = \mathcal{B}_X$. For any $\mathcal{F} \subset \mathcal{B}_X$ the *relative Rokhlin entropy* is defined by

$$h^{\mathrm{Rok}}(T|\mathcal{F}) = \inf_{\mathcal{P}} H(\mathcal{P}|\mathcal{F})$$

where the infimum is over all partitions $\mathcal{P}$ such that $\sigma\text{-alg}(T, \mathcal{P} \cup \mathcal{F}) = \mathcal{B}_X$. If $T$ is non-ergodic then the Rokhlin entropy is defined by

$$h^{\mathrm{Rok}}(T) = h^{\mathrm{Rok}}(T|\mathcal{I}_T),$$

where $\mathcal{I}_T$ denotes the sigma-algebra of $T(\Gamma)$-invariant Borel sets.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

Rokhlin entropy is clearly a measure-conjugacy invariant. Moreover, in case $\Gamma$ is amenable, it agrees with Kolmogorov–Sinai entropy [**STD16**] (the special case in which $A = \mathbb{Z}$ was handled earlier by Rokhlin [**Roh67**]). However, it can be difficult to compute. For example, it is not known whether every countable group has an ergodic essentially free action with positive Rokhlin entropy. The only known lower bound is the following proposition.

PROPOSITION 2.11. *For any pmp action $\Gamma \curvearrowright^T (X, \mu)$ and any sofic approximation $\Sigma$, $h^{\mathrm{Rok}}(T) \geq h_{\Sigma,\mu}(T)$.*

*Exercise 10.* Use the partition definition of $\Sigma$-entropy to prove Proposition 2.11.

*Question 1.* Suppose $T$ is an essentially free ergodic pmp action. Does $h_{\Sigma,\mu}(T) \neq -\infty$ necessarily imply $h_{\Sigma,\mu}(T) = h^{\mathrm{Rok}}(T)$?

*Question 2.* Suppose $\Gamma$ is a finitely generated free group, $T$ is an essentially free ergodic pmp action with a finite generating partition $\mathcal{P}$. Is the difference

$$f_\mu(T) - h^{\mathrm{Rok}}(T)$$

an invariant of the weak equivalence class of the action? The notions of weak containment and equivalence for group actions were introduced by Kechris as analogs of weak containment and equivalence for unitary representations [**Kec10**, II.10 (C)].

2.6.1. *Applications to the classification of Bernoulli shifts.* By Theorem 2.6, if $\Gamma$ is sofic then the sofic entropy of the Bernoulli shift $\Gamma \curvearrowright (K, \kappa)^\Gamma$ is the Shannon entropy of the base $H(K, \kappa)$. It is clear that the Rokhlin entropy of this shift is $\leq H(K, \kappa)$ since the partition $\mathcal{P} = \{P_k : k \in K\}$ defined by $P_k = \{x \in K^\Gamma : x(e) = k\}$ is generating and $H_\mu(\mathcal{P}) = H(K, \kappa)$. So in this case at least, the Rokhlin entropy agrees with the sofic entropy.

What if $\Gamma$ is non-sofic? Of course, we do not know whether non-sofic groups exist but even in this case there are some very interesting results. To describe them, let $h^{\mathrm{Rok}}_{\mathrm{sup}}(\Gamma)$ be the supremum of $h^{\mathrm{Rok}}(T)$ over all essentially free, ergodic actions $T$ of $\Gamma$ with finite Rokhlin entropy. Of course, when $\Gamma$ is sofic then $h^{\mathrm{Rok}}_{\mathrm{sup}}(\Gamma) = +\infty$. We do not know whether or not $h^{\mathrm{Rok}}_{\mathrm{sup}}(\Gamma) = +\infty$ for all countable groups. However, the following is shown.

THEOREM 2.12. [**Sew15b**] $h^{\mathrm{Rok}}(\Gamma \curvearrowright (K, \kappa)^\Gamma) = \min(H(K, \kappa), h^{\mathrm{Rok}}_{\mathrm{sup}}(\Gamma))$.

*Remarks on the proof.* The full proof is quite intricate and the reader is encouraged to see [**Sew15b**] for details. It is a proof by contradiction. Assuming the result is false, there exists an essentially free ergodic action $T$ with

$$h^{\mathrm{Rok}}(\Gamma \curvearrowright (K, \kappa)^\Gamma) < h^{\mathrm{Rok}}(T) < H(K, \kappa).$$

From this, one constructs a $\Gamma$-equivariant Borel map $\Phi : X \to K^\Gamma$ such that (1) $\Phi$ is an isomorphism onto its image, (2) $\Phi_* \mu$ is close to $\kappa^\Gamma$ in the weak* topology. Using upper semi-continuity of Rokhlin entropy this implies a contradiction. The construction of $\Phi$ is highly non-trivial. It combines tools from Seward's generalization of Krieger's generator theorem [**Sew19**] with an Abert–Weiss factor map [**AW13**]. □

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

It follows that if $h_{\sup}^{\mathrm{Rok}}(\Gamma) = +\infty$ then Rokhlin entropy distinguishes Bernoulli shifts. Moreover, Seward proves in [**Sew15b**] that if $h_{\sup}^{\mathrm{Rok}}(\Gamma) < \infty$ and $H$ is any infinite locally finite group then $h_{\sup}^{\mathrm{Rok}}(\Gamma \times H) = 0$. This is a most interesting condition! It implies that all ergodic actions of $\Gamma \times H$ have Rokhlin entropy zero, even the Bernoulli shift with base space ([0, 1], Leb)! Thus if it is true that for every countable group $\Gamma$ there exists some ergodic essentially free action with positive Rokhlin entropy, then Bernoulli shifts are distinguished by Rokhlin entropy.

2.7. *Naive entropy.*

*Definition 7.* Let $\Gamma \curvearrowright^T (X, \mu)$ be a pmp action and $\mathcal{P}$ a partition of $X$. The *naive entropy* of $\mathcal{P}$ is

$$h_\mu^{\mathrm{naive}}(T, \mathcal{P}) = \inf_{W \Subset \Gamma} |W|^{-1} H(\mathcal{P}^W),$$

where $\mathcal{P}^W = \bigvee_{w \in W} (T^w)^{-1}\mathcal{P}$. The *naive entropy* of $T$ is

$$h_\mu^{\mathrm{naive}}(T) = \sup_{\mathcal{P}} h_\mu^{\mathrm{naive}}(T, \mathcal{P}),$$

where the supremum is over all finite-entropy partitions $\mathcal{P}$.

It is an exercise to show that if $\Gamma$ is amenable then naive entropy coincides with Kolmogorov–Sinai entropy. However if $\Gamma$ is non-amenable the situation is very different.

THEOREM 2.13. *If $\Gamma$ is non-amenable and $\Gamma \curvearrowright^T (X, \mu)$ is a pmp action then $h_\mu^{\mathrm{naive}}(T) \in \{0, +\infty\}$.*

*Proof.* Suppose there is a finite-entropy partition $\mathcal{P}$ of $X$ with $h_\mu^{\mathrm{naive}}(T, \mathcal{P}) > 0$. Let $W \subset \Gamma$ be finite. Then

$$h_\mu^{\mathrm{naive}}(T, \mathcal{P}^W) = \inf_{F \Subset \Gamma} |F|^{-1} H_\mu(\mathcal{P}^{WF}) = \inf_{F \Subset \Gamma} \frac{H_\mu(\mathcal{P}^{WF})}{|WF|} \frac{|WF|}{|F|}.$$

Since $\Gamma$ is non-amenable, for every real number $r > 0$ there is a finite $W \subset \Gamma$ such that

$$\inf_{F \Subset \Gamma} \frac{|WF|}{|F|} > r.$$

Hence $\sup_{W \Subset \Gamma} h_\mu^{\mathrm{naive}}(T, \mathcal{P}^W) = +\infty$ proving the theorem. □

In many respects naive entropy behaves better than sofic entropy. For example, it is monotone under factor maps and additive under direct products. Moreover, the naive Pinsker algebra is naturally defined as the collection of all measurable subsets $A \subset X$ such that the partition $\mathcal{P} = \{A, X - A\}$ has zero naive entropy. An exercise shows that this really is a sigma-sub-algebra that contains all zero-naive-entropy factors. Moreover, the following proposition holds.

PROPOSITION 2.14. *For ergodic actions, naive entropy is an upper bound for Rokhlin entropy.*

The proof of this requires the following surprising result due to Seward.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

THEOREM 2.15. [**Sew15b**] *Let $T$ be an ergodic pmp action with a generating partition $\mathcal{P}$ with finite Shannon entropy. Then*

$$h^{\mathrm{Rok}}(T) \leq h_{\mu}^{\mathrm{naive}}(T, \mathcal{P}).$$

*Proof of Proposition 2.14.* Let $T$ be an ergodic pmp action with naive entropy zero. It suffices to show that, for every $\epsilon > 0$, there exists a generating partition with Shannon entropy $< \epsilon$.

Let $\mathcal{P}$ be a countable generating partition for $T$. Such a partition always exists by a general result due to Rokhlin [**Roh67**]. It might, however, be the case that $\mathcal{P}$ has infinite Shannon entropy. In any case, there exist finite partitions $\mathcal{Q}_1 \leq \mathcal{Q}_2 \leq \cdots \leq \mathcal{P}$ such that $\mathcal{P} = \bigvee_i \mathcal{Q}_i$. We apply Theorem 2.15 to the factor generated by $\mathcal{Q}_i$. So there exists a partition $\mathcal{R}_i$ such that $H_{\mu}(\mathcal{R}_i) < \epsilon/2^i$ and $\mathcal{R}_i$ generates the same factor as $\mathcal{Q}_i$. It follows that if $\mathcal{R}_{\infty} := \bigvee_i \mathcal{R}_i$, then $H_{\mu}(\mathcal{R}_{\infty}) < \epsilon$ and $\mathcal{R}_{\infty}$ is generating. □

2.7.1. *Applications to Gottschalk's conjecture.*

THEOREM 2.16. [**Sew15b**] *If $h_{\mathrm{sup}}^{\mathrm{Rok}}(\Gamma) = +\infty$ then Gottshalk's conjecture (from §2.3.1) is true for $\Gamma$.*

*Proof sketch.* Let $A$ be a finite alphabet and let $u_A$ denote the uniform probability measure on $A$. Assuming $\Phi : A^{\Gamma} \to A^{\Gamma}$ is continuous, $\Gamma$ equivariant and injective, Theorem 2.12 implies

$$h^{\mathrm{Rok}}(\Gamma \curvearrowright (A, u_A)^{\Gamma}) = \log |A| = h^{\mathrm{Rok}}(\Gamma \curvearrowright (A^{\Gamma}, \Phi_* u_A^{\Gamma})).$$

Using Theorem 2.15 with the canonical partition, it can be shown that if $\Phi$ is not surjective then

$$h^{\mathrm{Rok}}(\Gamma \curvearrowright (A^{\Gamma}, \Phi_* u_A^{\Gamma})) < \log |A|.$$

This contradiction proves the theorem. □

2.7.2. *Topological naive entropy.* Recent work of Burton [**Bur17**] introduced the following topological counterpart.

*Definition 8.* (Topological naive entropy) Let $\Gamma \curvearrowright^T X$ be a continuous action on a compact metrizable space. Given an open cover $\mathcal{U}$ of $X$, let $N(\mathcal{U})$ be the smallest cardinality of a subcover. The *naive entropy* of $\mathcal{U}$ is

$$h^{\mathrm{naive}}(T, \mathcal{U}) = \inf_{F \Subset \Gamma} |F|^{-1} \log(N(\mathcal{U}^F)),$$

where $\mathcal{U}^F$ is the open cover $\vee_{f \in F} (T^f)^{-1} \mathcal{U}$. The *naive entropy* of $T$ is

$$h^{\mathrm{naive}}(T) = \sup_{\mathcal{U}} h^{\mathrm{naive}}(T, \mathcal{U}),$$

where the supremum is over all finite open covers.

Burton shows that topological naive entropy provides an upper bound for measure naive entropy and that distal systems have zero naive entropy in both measure and topological senses when $\Gamma$ has an element of infinite order. He also shows that the generic action of the free group by homeomorphisms on the Cantor set has zero topological naive entropy. However the following question appears to be open.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*Question 3.* Does every countable group admit an essentially free pmp action with zero naive entropy?

## 3. *Special classes of actions*

3.1. *Trivial actions.* It might come as a surprise that the trivial action of $\Gamma$ on $(X, \mu)$ is interesting, from the point of view of sofic entropy theory. To explain, fix a sofic approximation $\Sigma = \{\sigma_n\}_{n\in\mathbb{N}}$ to $\Gamma$ and a standard probability space $(X, \mu)$ (which may be atomic). The trivial action $\tau_X = (\tau_X^g)_{g\in\Gamma}$ on $X$ is defined by $\tau_X^g y = y$ for all $y \in X$ and $g \in \Gamma$. We will show that $h_{\Sigma,\mu}(\tau_X) \in \{-\infty, 0\}$ and that both cases occur. The upper bound $h_{\Sigma,\mu}(\tau_X) \leq 0$ can be derived directly.

3.1.1. *Expanders.* For simplicity, suppose $\Gamma$ is finitely generated and let $S \subset \Gamma$ be a finite generating set. Let $\Sigma = \{\sigma_n : \Gamma \to \mathrm{Sym}(V_n)\}$ be a sofic approximation. Let $G_n = (V_n, E_n)$ be the graph with edges $\{v, \sigma_n(s)v\}$ for $v \in V_n$, $s \in S$. The sequence of graphs $\{G_n\}_n$ is an *expander sequence* if there is an $\epsilon > 0$ such that for every subset $A \subset V_n$ with $|A_n| \leq |V_n|/2$,

$$|\partial A_n| \geq \epsilon |A_n|,$$

where $\partial A_n$ is the set of edges $e \in E_n$ with one endpoint in $A_n$ and one endpoint not in $A_n$. In this case, we say $\Sigma$ is *a sofic approximation by expanders*.

*Remark 12.* (Actions on ultraproduct spaces) The set $V_n$ can be thought of as a probability space endowed with the uniform probability measure. The ultraproduct of the $V_n$'s forms a non-standard probability measure space on which $\Gamma$ acts by measure-preserving transformations [**ES12**]. If $\Sigma$ is by expanders then the action $\Gamma$ on this ultraproduct is ergodic. On the other hand, if the action is ergodic then there is an equivalent sofic approximation $\Sigma'$ such that $\Sigma'$ is by expanders. In this case, $\Sigma$ is said to be an *ergodic sofic approximation* [**Hay17b**].

*Example 1.* Margulis showed that if $\Gamma$ has property (T) and $N_i \lhd \Gamma$ is a decreasing sequence of finite-index normal subgroups of $\Gamma$ then the Schreier coset graphs of $\Gamma/N_i$ form an expander sequence. Therefore, the sofic approximation $\Sigma = \{\sigma_i : \Gamma \to \mathrm{Sym}(\Gamma/N_i)\}_i$ given by the canonical actions of $\Gamma$ on $\Gamma/N_i$ by left translation is by expanders. Kun recently proved every sofic approximation of a property (T) group is equivalent (in the sense of §2.2.4) to one that is a disjoint union of expanders [**Kun16**].

PROPOSITION 3.1. *If $\Sigma$ is a sofic approximation by expanders and $(X, \mu)$ is non-trivial (this means, for every $x \in X$, $\mu(\{x\}) < 1$) then the trivial action of $\Gamma$ on $(X, \mu)$ has $h_{\Sigma,\mu}(\tau_X) = -\infty$.*

*Proof sketch.* Let $A \subset X$ have $0 < \mu(A) \leq 1/2$. If $\phi : V_n \to X$ is a microstate then $\phi^{-1}(A)$ should have cardinality approximately $\mu(A)|V_n|$. Moreover, the boundary of $\phi^{-1}(A)$ should have small cardinality relative to $|V_n|$ because $A$ is an invariant set. However this contradicts the expander property. So no such microstates exist. More precisely, $\mathrm{Map}(\tau_X, \rho, \mathcal{O}, F, \delta, \sigma_n)$ is empty if $F \subset \Gamma$ contains a generating set, $\delta > 0$ is sufficiently small, $\mathcal{O}$ is sufficiently small and $n$ is sufficiently large. Here $\rho$ is any metric on $X$. □

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

3.1.2. *Diffuse sofic approximations.* Next we define a sufficient condition on a sofic approximation implying the trivial action has sofic entropy zero.

*Definition 9.* Let $\Sigma = \{\sigma_n\}$ be a sofic approximation to $\Gamma$. The *density* of a sequence $\{A_n\}$ of subsets $A_n \subset V_n$ is

$$\text{density}(\{A_n\}) := \lim_{n\to\infty} \frac{|A_n|}{|V_n|}$$

provided the limit exists.

A sequence $A_n \subset V_n$ is *asymptotically invariant* if for every $g \in \Gamma$ the sequence of symmetric differences $\{A_n \triangle \sigma_n(g)A_n\}$ has density zero. If $\{A_n\}$ is asymptotically invariant and has positive density then there exist maps $\sigma'_n : \Gamma \to \text{Sym}(A_n)$ satisfying

$$\lim_{n\to\infty} \frac{|\{v \in A_n : \sigma'_n(g)v = \sigma_n(g)v\}|}{|A_n|} = 1$$

for every $g \in \Gamma$. The sequence $\Sigma' = \{\sigma'_n\}$ is a *sub-sofic approximation* of $\Sigma$. It is well defined only up to edit distance zero (see §2.2.4). It is called *proper* if the density of $\{A_n\}$ is strictly less than 1. The sequence $\Sigma$ is *diffuse* if every sub-sofic approximation $\Sigma'$ admits a proper sub-sofic approximation.

*Example 2.* Sofic approximations can be amplified as follows. Let $\{W_n\}$ be a sequence of finite sets. Given $\Sigma = \{\sigma_n\}$ as above, define

$$\sigma'_n : \Gamma \to \text{Sym}(V_n \times W_n)$$

by $\sigma'_n(g)(v, w) = (\sigma_n(g)v, w)$. If $|W_n| \to \infty$ as $n \to \infty$ then $\Sigma' = \{\sigma'_n\}$ is a diffuse sofic approximation.

PROPOSITION 3.2. *If $\Sigma$ is diffuse then the trivial action of $\Gamma$ on $(X, \mu)$ has zero $\Sigma$-entropy.*

*Proof sketch.* The upper bound can be obtained directly. To prove the lower bound, let $G$ be the set of all numbers $p \in [0, 1]$ such that $p$ is the density of $\{A_n\}$ for some asymptotically invariant $\{A_n\}$. We claim that $G = [0, 1]$. To see this, let $p \in [0, 1]$ be arbitrary and let $a = \sup\{x \in G : x \leq p\}$. A diagonalization argument shows $G$ is closed. Therefore, $a \in G$ and there exists asymptotically invariant $\{A_n\}$ with density $a$.

Another diagonalization argument shows there exists an asymptotically invariant sequence $\{B_n\}$ with $A_n \subset B_n$ and $\text{density}(\{B_n\}) = b$ where $b \geq p$ is minimal subject to these conditions. Note that $B_n \setminus A_n$ is asymptotically invariant. Because $\Sigma$ is diffuse, either $a = b$ (in which case $a = p \in G$) or there exists an asymptotically invariant sequence $\{C_n\}$ with $A_n \subset C_n \subset B_n$ and

$$a < \text{density}(\{C_n\}) < b.$$

This contradicts the choice of $a, b$. So $G = [0, 1]$ as claimed.

Now let $\mathcal{P} = \{P_1, \ldots, P_r\}$ be a finite partition of $X$ with $0 < \mu(P_i) < 1$ for all $i$. By the claim above there exists an asymptotically invariant sequence $\{Q_{1,n}\}_n$ with density $\mu(P_1)$. Apply the claim again to the complement of $Q_{1,n}$ in $V_n$ to obtain an asymptotically invariant sequence $\{Q_{2,n}\}_n$ with $Q_{1,n} \cap Q_{2,n} = \emptyset$ and density $\mu(P_2)$. Continue in this fashion to obtain a sequence $\{\mathcal{Q}_n\}$ of partitions $\mathcal{Q}_n = \{Q_{1,n}, \ldots, Q_{r,n}\}$ such that each sequence $\{Q_{i,n}\}_n$ is asymptotically invariant and has density $\mu(P_i)$. This shows there exist

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

microstates (with respect to the partition definition) for the trivial action of $\Gamma$ on $X$ and therefore $h_{\Sigma,\mu}(\tau_X) \geq 0$. □

3.1.3. *The $f$-invariant.*

THEOREM 3.3. *Let $(X, \mu)$ be a probability space with finite Shannon entropy $H(X, \mu)$. Then $f_\mu(\tau_X) = -(r-1)H(X, \mu)$.*

*Proof.* Because $H(X, \mu) < \infty$ we may assume $X$ is countable. Let $\mathcal{P}$ be the partition into points. Then $f_\mu(\tau_X) = F_\mu(\tau_X, \mathcal{P}) = -(r-1)H_\mu(\mathcal{P})$. □

3.2. *Bernoulli shifts.*

THEOREM 3.4. *For any countably infinite group $\Gamma$, if $(K, \kappa)$, $(L, \lambda)$ are probability space with the same Shannon entropy then the corresponding Bernoulli shifts $\Gamma \curvearrowright (K, \kappa)^\Gamma$ and $\Gamma \curvearrowright (L, \lambda)^\Gamma$ are measurably conjugate. In particular, if $\Gamma$ is sofic or $h_{\sup}^{\mathrm{Rok}}(\Gamma) = +\infty$ then Bernoulli shifts over $\Gamma$ are completely classified up to measure-conjugacy by Shannon entropy of the base.*

The first statement is outlined in §7.1. The second statement follows from Theorems 2.6 and 2.12. It is unknown whether the second statement holds for all countably infinite groups.

3.3. *Markov chains.* Recall the definition of a Markov chain over a free group from §1.4.

THEOREM 3.5. *Let $\mathbf{X} = (X_g)_{g\in\Gamma}$ be a stationary process over the free group $\Gamma = \langle s_1, \ldots, s_r\rangle$. Suppose the state space $K$ is countable and $H(X_e) < \infty$. Let $\mu = \mathrm{Law}(\mathbf{X})$ and $\Gamma \curvearrowright K^\Gamma$ the shift action. Then*

$$f_\mu(\Gamma \curvearrowright K^\Gamma) \leq F_\mu(\mathcal{P}) = H(X_e) + \sum_{s\in S} H(X_e|X_s) - H(X_e),$$

*where $\mathcal{P}$ is the canonical 'time zero' partition of $K^\Gamma$. Moreover, equality holds if and only if $\mathbf{X}$ is Markov.*

*Remark 13.* In the case rank $r = 1$, this formula reduces to the well-known formula for the entropy of a Markov chain as the entropy of the present conditioned on the immediate past: $h(\mathbf{X}) = H(X_0|X_{-1})$.

*Proof sketch of Theorem 3.5 (details in* [**Bow10c**]*).* Since $f_\mu(\Gamma \curvearrowright K^\Gamma)$ is the infimum of $F_\mu(\mathcal{Q})$ over all splittings $\mathcal{Q}$ of $\mathcal{P}$, the inequality $\leq$ is immediate. So it suffices to show that if $\mathcal{Q}$ is a simple splitting of $\mathcal{P}$ then $F_\mu(\mathcal{Q}) = F_\mu(\mathcal{P})$ if and only if $\mathbf{X}$ is Markov in the direction of the simple splitting. This can be achieved by direct computation following the steps in the proof sketch of Theorem 2.8. □

*Example 3.* (The Ising model) Let $K = \{-1, 1\}$. Given $\epsilon > 0$ we consider the Markov chain $\mathbf{X} = (X_g)_{g\in\Gamma}$ over $\Gamma = \langle s_1, \ldots, s_r\rangle$ with state space $K$ satisfying

$$P(X_e = -1) = P(X_e = 1) = 1/2$$
$$P(X_e = k|X_s = k) = 1 - \epsilon, \quad P(X_e \neq k|X_s = k) = \epsilon$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

for every $k \in K$ and $s \in \{s_1, \ldots, s_r, s_1^{-1}, \ldots, s_r^{-1}\}$. The $f$-invariant of this process, denoted $f(\mathbf{X})$, is

$$f(\mathbf{X}) = -(r-1)\log(2) - r\epsilon\log(\epsilon) - r(1-\epsilon)\log(1-\epsilon).$$

In the limiting case $\epsilon = 0$, the law of $\mathbf{X}$ is equally distributed on only two atoms: all 0's and all 1's. In this case, $f(\mathbf{X}) = -(r-1)\log(2) < 0$. From the sofic interpretation of the $f$-invariant §2.5.2, this means that for most homomorphisms $\sigma : \Gamma \to \mathrm{Sym}(n)$, there are no good models/microstates for this process. This is because the graph $(V, E)$ defined by $V = [n]$ and $E = \{(v, \sigma(s_i)v) : v \in [n], 1 \le i \le r\}$ is typically an expander. As in §3.1 expansitivity implies there are no good models for a non-ergodic action.

If $\epsilon$ is positive but small and $r > 1$ then the $f$-invariant is still negative and so it cannot be measurably conjugate to a Bernoulli shift. However the Markov chain is mixing. By contrast, every mixing Markov chain over the integers is isomorphic to a Bernoulli shifts [**FO70**].

*Problem 7.* Classify mixing Markov chains up to measure-conjugacy. If two mixing Markov chains are spectrally isomorphic and have the same $f$-invariant are they measurably conjugate?

While it is straightforward to compute the $f$-invariant for Markov chains, there are no known methods for computing the sofic entropy with respect to a given sofic approximation. In particular the following is open.

*Question 4.* Does the sofic entropy of the Ising model depend on the choice of sofic approximation? To avoid trivialities, we assume the approximation is such that the sofic entropy is non-negative.

*Remark 14.* Markov chains are used for counterexamples. In §11.4.2 it is shown that the Ising model with small transitive probabilities is uniformly mixing but does not have completely positive entropy (CPE). By contrast it was shown in [**RW00**] that for $\mathbb{Z}$-actions, CPE and uniformly mixing are equivalent properties. In §8.2 a topological Markov chain is constructed that has multiple measures of maximal $f$-invariant. This is impossible if $\Gamma = \mathbb{Z}$.

*Remark 15.* Markov chains have also been used to obtain positive general results by using that an arbitrary invariant measure can be approximated (in the weak* sense) by measures that are multi-step Markov chains. These results include the following: a relative entropy theory for the $f$-invariant §10, a partial Yuzvinskii's formula §3.4, a formula for the restriction to a subgroup §4.2 and an ergodic decomposition formula §6.1.

*Example 4.* (Tree lattices) Let $T_{2d}$ be the $2d$-regular tree and $\Lambda < \mathrm{Aut}(T_{2d})$ a lattice. Because the free group $\mathbb{F}_d$ of rank $d$ is also a lattice in $\mathrm{Aut}(T_{2d})$, it acts by measure-preserving transformations on the quotient $\mathrm{Aut}(T_{2d})/\Lambda$. In general, it is an open problem to compute either the sofic entropy or the $f$-invariant for these homogeneous actions. However there is at least one special case in which the action is isomorphic to a Markov chain and therefore its $f$-invariant can be computed explicitly. For simplicity, let us assume $d = 2$. We consider the case $\Lambda = \mathbb{F}_2 = \langle a, b \rangle$.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

We assume $\mathbb{F}_2$ acts simply transitively on the vertices of $T_4$. By fixing a vertex $v_0$, we identify $\mathbb{F}_2$ with the set of vertices via the map $g \mapsto gv_0$. Let $\vec{E}(T_4)$ be the set of directed edges of $T_4$ and $L_0 : \vec{E}(T_4) \to S$ the standard labeling: $L_0((gv_0, gsv_0)) = s$ for $s \in S$. Note that if $e \in \vec{E}(T_4)$ and $\check{e}$ denotes the same edge with the opposite orientation then $L_0(\check{e}) = L_0(e)^{-1}$.

More generally, a *legal labeling* of $T_4$ is any map $L : \vec{E}(T_4) \to S$ satisfying the following.

- For $v \in V(T_4)$, let $N^+(v)$ and $N^-(v)$ denote the set of edges directed out and directed into $v$ (respectively). Then $L$ is one-one on $N^+(v)$ and on $N^-(v)$.
- For every $e \in \vec{E}(T_4)$, $L(\check{e}) = L(e)^{-1}$.

$\mathrm{Aut}(T_4)$ acts on the set of legal labelings, denoted $\mathcal{L}$, by $gL = L \circ g^{-1}$. This action is transitive and the stabilizer of $L_0$ is $\mathbb{F}_2$. Therefore we can identify $\mathcal{L}$ with $\mathrm{Aut}(T_4)/\mathbb{F}_2$.

Let $K = \mathrm{Sym}(S)$. For $L \in \mathcal{L}$, let $x_L \in K^\Gamma$ be the map

$$x_L(g) = L_0 \circ (L \restriction N^+(gv_0))^{-1}.$$

It is straightforward to verify that $x_{gL} = gx_L$. So the map $L \mapsto x_L$ gives an embedding of $\mathcal{L}$ into $K^\Gamma$. Moreover, it pushes forward the Haar measure on $\mathrm{Aut}(T_4)/\mathbb{F}_2 = \mathcal{L}$ onto a Markov measure on $K^\Gamma$ denoted by $\mu$. If $\mathbf{X} = (X_g)_{g\in\Gamma}$ is a stationary process with law $\mu$ then for any $s \in S$, the pair $X_e$ and $X_s$ is uniformly distributed over the set of pairs of permutations $(\pi_1, \pi_2) \in K \times K$ with the property that if $t \in S$ is such that $\pi_1(t) = s$ then $\pi_2(t^{-1}) = s^{-1}$. The number of such pairs is 4!3!. So $H(X_e, X_s) = \log(4!3!)$. Moreover, $X_e$ is uniformly distributed over $K$. So $H(X_e) = \log(4!)$. So

$$f(\mathbb{F}_2 \curvearrowright \mathrm{Aut}(T_4)/\mathbb{F}_2) = f(\mathbf{X}) = -3\log(4!) + 2\log(4!3!) = \log(3/2) > 0.$$

It is an open problem whether $\mathbb{F}_2 \curvearrowright \mathrm{Aut}(T_4)/\mathbb{F}_2$ is isomorphic to a Bernoulli shift or to a factor of Bernoulli shift.

3.4. *Algebraic dynamics.* Let $X$ denote a compact group and $\mathrm{Aut}(X)$ the group of all automorphisms of $X$. Naturally, any automorphism preserves the Haar measure $\mu$. Therefore any homomorphism $\Gamma \to \mathrm{Aut}(X)$ induces a measure-preserving action $\Gamma \curvearrowright (X, \mu)$. The goal of algebraic dynamics is to relate dynamical properties of the action $\Gamma \curvearrowright X$ to algebraic and analytic properties of the homomorphism $\Gamma \to \mathrm{Aut}(X)$. Such actions have been explored in great detail when $\Gamma = \mathbb{Z}^d$ (see the book [**Sch95**]).

A general procedure for computing the entropy of algebraic $\mathbb{Z}$-actions was obtained by Yuzvinskii in the 1960s [**Juz65a**, **Juz65b**] and extended to $\mathbb{Z}^d$ by Lind–Schmidt–Ward [**LSW90**]. This procedure rests on Yuzvinskii's addition formula and the special case of principal algebraic actions. These are explained next along with (partial) results in the non-amenable case. We also present recent results on Pinsker algebras of algebraic actions, CPE, mixing properties and the coincidence of measure entropy and topological entropy. This area is rapidly developing!

3.4.1. *The case* $\Gamma = \mathbb{Z}$. This section is meant to motivate the results for more general groups by explaining the case $\Gamma = \mathbb{Z}$ in detail.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

THEOREM 3.6. *Suppose*

$$1 \to Y \to X \to X/Y \to 1$$

*is an exact sequence of compact metrizable groups and* $T_X \in \mathrm{Aut}(X)$ *is an automorphism that leaves* $Y$ *invariant. Then*

$$h(T_X) = h(T_Y) + h(T_{X/Y}),$$

*where* $T_Y \in \mathrm{Aut}(Y)$, $T_{X/Y} \in \mathrm{Aut}(X/Y)$ *are the induced automorphisms and the entropy is either topological or with respect to Haar measure (these two cases coincide).*

This theorem was first proven by Yuzvinskii [**Juz65b**].

We will apply the above theorem to principal algebraic actions. If $f = \sum_{i=0}^{s} c_i x^i \in \mathbb{Z}[x]$ is a polynomial and $z \in \mathbb{T}^{\mathbb{Z}}$ (where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is the 1-torus as an additive group) then the convolution $f * z \in \mathbb{T}^{\mathbb{Z}}$ is defined by

$$(f * z)_n = \sum_{m \in \mathbb{Z}} c_m z_{n-m},$$

where we set $c_m = 0$ if $m < 0$ or $m > s$. Let

$$X_f = \{(z_i) \in \mathbb{T}^{\mathbb{Z}} : f^* * z = 0\}.$$

Note that $X_f$ is a compact abelian group and the shift action $T_f : X_f \to X_f$, $T_f(x)_i = x_{i+1}$ is an automorphism.

If $f, g \in \mathbb{Z}[x]$ are non-zero polynomials then there is an exact sequence

$$0 \to X_f \to X_{gf} \to X_g \to 0,$$

where the map $X_{gf} \to X_g$ is convolution with $f$. So Yuzvinskii's addition formula implies $h(T_{fg}) = h(T_f) + h(T_g)$. In other words, the map $f \mapsto \exp(h(T_f))$ is multiplicative. There are only a few multiplicative functions on polynomials. For example, the leading coefficient is one. The product of all roots contained in some fixed subset of $\mathbb{C}$ is another. So perhaps it is not too suprising that

$$h(T_f) = \log |c_s| + \sum_{i=1}^{s} \max(0, \log |r_i|),$$

where $c_s$ is the leading coefficient of $f$ and $r_1, \ldots, r_s$ are its roots.

This formula is most easily confirmed in the special case that $c_s = 1$ and $c_0 = \pm 1$ (where $c_0$ is the constant term of $f$). In this case the map

$$(x_i) \in X_f \mapsto (x_0, \ldots, x_{s-1}) \in \mathbb{T}^s$$

induces a measure-conjugacy between the shift $T_f$ and the linear map

$$(x_0, \ldots, x_{s-1}) \mapsto \left(x_1, \ldots, x_{s-1}, -\sum_{i=0}^{s-1} c_i x_i\right).$$

The latter map has entropy equal to $\sum_{i=1}^{s} \max(0, \log |\lambda_i|)$ where $\lambda_1, \ldots, \lambda_s$ are its eigenvalues (by Pesin's entropy formula). These eigenvalues are exactly the roots of $f$ because $f$ is its characteristic polynomial.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

Using Jensen's formula, we can rewrite the above formula as

$$h(T_f) = \int_0^1 \log |f(\exp(2\pi i x))|\, dx.$$

We will now show how to generalize the formula on the right to arbitrary countable groups.

We view $f$ as an element of the group ring $\mathbb{C}\mathbb{Z}$ (by identifying $f$ with $\sum_n c_n n$). We can view $\mathbb{C}\mathbb{Z}$ as a subring of $B(\ell^2(\mathbb{Z}))$, the algebra of bounded operators on $\ell^2(\mathbb{Z})$, via

$$\sum_n c_n n \mapsto c_n \sigma^n,$$

where $\sum_n c_n n \in \mathbb{C}\mathbb{Z}$ is a formal sum and $\sigma : \ell^2(\mathbb{Z}) \to \ell^2(\mathbb{Z})$ is the shift-operator $\sigma(x)_i = x_{i-1}$. The von Neumann algebra of $\mathbb{Z}$, denoted $L\mathbb{Z}$, is the weak operator closure of $\mathbb{C}\mathbb{Z}$ in $B(\ell^2(\mathbb{Z}))$. The trace on $L\mathbb{Z}$ is defined by

$$\mathrm{tr}(x) = \langle x\delta_0, \delta_0 \rangle,$$

where $\delta_0 \in \ell^2(\mathbb{Z})$ is the Dirac function supported on $0 \in \mathbb{Z}$.

The Fourier transform $\mathcal{F} : \ell^2(\mathbb{Z}) \to L^2(\mathbb{T})$ is the continuous linear map such that $\mathcal{F}(\delta_n)$ is the function $z \mapsto z^n \in \mathbb{T}$ where $\mathbb{T} = \{z \in \mathbb{C} : |z| = 1\}$ is the unit circle. This is an isomorphism. Next we show that $\mathcal{F}$ conjugates $L\mathbb{Z}$ to $L^\infty(\mathbb{T})$.

For $\phi \in L^\infty(\mathbb{T})$, define the multiplication operator $M_\phi : L^2(\mathbb{T}) \to L^2(\mathbb{T})$ by

$$M_\phi(f)(z) = \phi(z) f(z).$$

The map $\phi \mapsto M_\phi$ embeds $L^\infty(\mathbb{T})$ into algebra of bounded operators $B(L^2(\mathbb{T}))$. Moreover, $\mathcal{F}\sigma\mathcal{F}^{-1} = M_z$. So $\mathcal{F}(\mathbb{C}\mathbb{Z})\mathcal{F}^{-1}$ is the algebra of Laurent polynomials. Its weak operator closure is $L^\infty(\mathbb{T})$. Thus $\mathcal{F}L\mathbb{Z}\mathcal{F}^{-1}$ is naturally identified with $L^\infty(\mathbb{T})$. Moreover, for any $x \in L\mathbb{Z}$, the trace of $x$, $\mathrm{tr}(x)$ equals the integral of $\mathcal{F}x\mathcal{F}^{-1}$ over $\mathbb{T}$. In particular, if $f \in \mathbb{Z}[x]$ is a polynomial, then

$$\int_{\mathbb{T}} \log |f(z)|\, dz = \mathrm{tr}(\log |f|),$$

where $\log |f| \in L\mathbb{Z}$ is defined via spectral calculus. The *Fuglede–Kadison determinant* of $f$ is defined by

$$\det(f) := \exp(\mathrm{tr}(\log |f|)) = \exp\left(\int_{[0,\infty)} \log(t)\, d\zeta_{|f|}(t)\right),$$

where $\zeta_{|f|}$ is the spectral measure of $|f|$. If $|f|$ has a non-trivial kernel then the determinant is 0. So we modify the definition slightly: the *positive Fuglede–Kadison determinant* of $f$ is

$$\det^+(f) := \exp\left(\int_{(0,\infty)} \log(t)\, d\zeta_{|f|}(t)\right).$$

In this way we are naturally led to the conjecture that, for general groups $\Gamma$ and $f \in \mathbb{Z}\Gamma$, the entropy of $\Gamma \curvearrowright X_f$ should be the logarithmic positive Fuglede–Kadison determinant of $f$.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

3.4.2. *Group rings.* To discuss algebraic actions, we will need some background on certain group rings. Let $\mathbb{C}\Gamma$ denote the complex group ring of $\Gamma$. Formally, $\mathbb{C}\Gamma$ is the set of all sums $\sum_{g\in\Gamma} c_g g$ with $c_g \in \mathbb{C}$ such that all but finitely many of the $c_g$'s are zero. Addition, multiplication and the adjoint operator are given by

$$\sum_{g\in\Gamma} c_g g + \sum_{g\in\Gamma} c'_g g = \sum_{g\in\Gamma} (c_g + c'_g) g$$
$$\left(\sum_{g\in\Gamma} c_g g\right)\left(\sum_{g\in\Gamma} c'_g g\right) = \sum_{g\in\Gamma}\sum_{h\in\Gamma} c_g c'_h gh.$$
$$\left(\sum_{g\in\Gamma} c_g g\right)^* = \sum_{g\in\Gamma} \overline{c_g} g^{-1},$$

where $\overline{c_g}$ is the complex-conjugate of $c_g$.

We view $\mathbb{Z}\Gamma$ as a subring of $\mathbb{C}\Gamma$. Also consider $\mathbb{T}^\Gamma$ where $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ is the additive group of the circle. If $x \in \mathbb{T}^\Gamma$ and $f = \sum_g f_g g \in \mathbb{Z}\Gamma$ then $fx, xf \in \mathbb{T}^\Gamma$ are well defined via

$$(fx)_g = \sum_{h\in\Gamma} f_{gh} x_{h^{-1}}, \quad (xf)_g = \sum_{h\in\Gamma} x_{gh} f_{h^{-1}}.$$

These are finite sums. The inner product of $f$ with $x$ is defined by

$$\langle f, x\rangle = \sum_g f_g x_g \in \mathbb{T}.$$

Note

$$\langle af, x\rangle = \langle a, xf^*\rangle = \langle f, a^* x\rangle.$$

(By linearity, it suffices to check this formula when each of $a, x, f$ is supported on a single element.) This inner product allows us to identify $\mathrm{Hom}(\mathbb{Z}\Gamma, \mathbb{T})$ with $\mathbb{T}^\Gamma$.

3.4.3. *Principal algebraic actions.* There is a simple procedure for associating to any $f \in \mathbb{Z}\Gamma$ a dynamical system. First consider the left ideal $\mathbb{Z}\Gamma f$ and the associated quotient $\mathbb{Z}\Gamma/\mathbb{Z}\Gamma f$. We consider the latter as a countable abelian group on which $\Gamma$ acts by automorphisms (namely, left multiplication). Let

$$X_f = \widehat{\mathbb{Z}\Gamma/\mathbb{Z}\Gamma f} = \mathrm{Hom}(\mathbb{Z}\Gamma/\mathbb{Z}\Gamma f, \mathbb{T})$$

be the Pontryagin dual. By identifying $\mathrm{Hom}(\mathbb{Z}\Gamma, \mathbb{T})$ with $\mathbb{T}^\Gamma$ as above, we can identify $X_f$ with the subgroup

$$X_f = \{x \in \mathbb{T}^\Gamma : xf^* = 0\}.$$

For $g \in \Gamma$, let $T_f^g : X_f \to X_f$ be the automorphism $(T_f^g x)(p) = x(g^{-1}p)$. Then $T_f = (T_f^g)_{g\in\Gamma}$ is an action on $X_f$ by automorphisms. It preserves the Haar measure, which we denote by $\mu_f$. The action $T_f$ is called a *principal algebraic action*.

When $\Gamma = \mathbb{Z}^d$, we identify $\mathbb{Z}\Gamma$ with the ring $\mathbb{Z}[u_1^{\pm 1}, \ldots, u_d^{\pm 1}]$ of Laurent polynomials. This way we may view $f \in \mathbb{Z}\Gamma$ as a polynomial function $f : \mathbb{C}^d \to \mathbb{C}$. With this identification, Lind–Schmidt–Ward proved in [**LSW90**] that for non-zero $f$,

$$h(T_f) = h_{\mu_f}(T_f) = \int_{\mathbb{T}^d} \log|f(e^{2\pi i\theta})|\, d\theta, \tag{1}$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

where $\mathbb{T}^d = \{(z_1, \ldots, z_d) \in \mathbb{C}^d : |z_i| = 1 \ \forall i\}$ denotes the $d$-dimensional torus. This extends earlier work of Yuzvinskii in the case $d = 1$ [**Juz65a**, **Juz65b**]. The right-hand side of the equation above is the log-Mahler measure of $f$.

Deninger noted that there is a generalization of the Mahler measure to non-abelian $\Gamma$ known as the Fuglede–Kadison determinant [**Den06**] and conjectured that, for amenable $\Gamma$, the entropy of $\Gamma \curvearrowright X_f$ equals the log of the Fuglede–Kadison determinant. Special cases were confirmed in [**Den06**, **DS07**, **Li12**] before the general amenable group case was handled in [**LT14**]. The case of expansive principal algebraic actions of residually finite groups was handled in [**Bow11a**] and extended in [**BL12**] to some non-expansive actions. Then in a stunning breakthrough Hayes obtained the most general result for sofic groups.

THEOREM 3.7. [**Hay16b**, Theorem 1.1] *Let $\Gamma$ be a sofic group with sofic approximation $\Sigma$. Let $f \in \mathbb{Z}\Gamma$ and let $\det^+(f) = \exp(\int_{(0,\infty)} \log(t)\, d\zeta_{|f|}(t))$ denote the positive Fuglede–Kadison determinant of $f$ where $\zeta_{|f|}$ denotes the spectral measure of $|f| = (f^* f)^{1/2}$. Then the following holds.*

(1) *$h_\Sigma(\Gamma \curvearrowright X_f) < +\infty$ if and only if $f$ is injective as a convolution operator on $\ell^2(\Gamma)$.*
(2) *If $f$ is injective as an convolution operator on $\ell^2(\Gamma)$, then*

$$h_\Sigma(\Gamma \curvearrowright X_f) = h_{\Sigma,\mu_f}(\Gamma \curvearrowright X_f) = \log \det^+(f).$$

*Remark 16.* The paper [**Hay16b**] also handles the more general case in which $f$ is a finite-dimensional matrix over $\mathbb{Z}\Gamma$, $I_f \subset \mathbb{Z}(\Gamma)^{\oplus n}$ is its range and $X_f$ is the Pontryagin dual of $\mathbb{Z}(\Gamma)^{\oplus n}/I_f$.

*Question 5.* Does the Rokhlin entropy of a principal algebraic action equal the logarithm of the Fuglede–Kadison determinant?

3.4.4. *Yuzvinskii's addition formula.* Let $G$ be a compact metrizable group and $\Gamma \curvearrowright G$ an action by continuous automorphisms. Suppose that $N \lhd G$ is a closed $\Gamma$-invariant normal subgroup. We will compare the entropy of $\Gamma \curvearrowright G$ with that of the restricted action $\Gamma \curvearrowright N$ and the induced action $\Gamma \curvearrowright G/N$. More precisely, we say the *addition formula* holds for $(\Gamma \curvearrowright G, N)$ if the entropy of $\Gamma \curvearrowright G$ equals the sum of the entropy of $\Gamma \curvearrowright N$ with the entropy of $\Gamma \curvearrowright G/N$. In general, this might depend on which entropy is intended.

In [**Juz65b**] Yuzvinskii proved the addition formula for $G = \mathbb{Z}$ with respect to both topological and measure entropy. The proof has been extended to more general kinds of skew-products [**Tho71**], to $\mathbb{Z}^d$ [**LSW90**], various other amenable groups [**MB09**, **Mil08**] and to arbitrary amenable groups [**Li12**] (and independently in unpublished work of Lind-Schmidt). Using this, Li–Thom obtained a general procedure for computing the entropy of algebraic actions of amenable groups on compact abelian groups [**LT14**] under very mild conditions. The formula shows that entropy can be viewed as $L^2$-torsion.

However it fails for non-amenable groups. Indeed the Ornstein–Weiss example (§1.3) is algebraic. Recall that $\mathbb{F}_2 = \langle a, b\rangle$ acts on the compact abelian group $G = (\mathbb{Z}/2)^{\mathbb{F}_2}$ with invariant normal subgroup $N \cong \mathbb{Z}/2$ consisting of the constants and quotient $G/N \cong G \times G$. With respect to any sofic approximation, the entropy of $\mathbb{F}_2 \curvearrowright G$ is $\log(2)$, the entropy of $\mathbb{F}_2 \curvearrowright N$ is 0 or $-\infty$ and the entropy of $\mathbb{F}_2 \curvearrowright G/N$ is $\log(4)$.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

Rokhlin entropy behaves similarly. Recent work of Bartholdi [**BK17**] shows that for any non-amenable $\Gamma$ and field $K$ there are $n \in \mathbb{N}$ and injective $K\Gamma$-module homomorphism $K\Gamma^{\oplus(n+1)} \to \mathbb{K}\Gamma^{\oplus n}$. Consider the case when $K$ is finite. Then the action $\Gamma \curvearrowright \widehat{K\Gamma^{\oplus n}}$ has entropy $n \log |K|$ while the $\Gamma$-action on the quotient group $\Gamma \curvearrowright \widehat{K\Gamma^{\oplus(n+1)}}$ has entropy $(n+1)\log|K|$. Thus the addition formula fails†.

In a different direction, Gaboriau and Seward consider the following construction: let $\Gamma$ be a finitely generated group, $\mathbb{K}$ a finite field, $\Gamma \curvearrowright \mathbb{K}^\Gamma$ the shift action, $\mathbb{K} \leq \mathbb{K}^\Gamma$ the constants. Then the following theorem holds.

THEOREM 3.8. [**GS15**] *For any sofic approximation* $\Sigma$*,*

$$(1+\beta^1_{(2)}(\Gamma))\log|\mathbb{K}| \leq h_\Sigma(\Gamma \curvearrowright \mathbb{K}^\Gamma/\mathbb{K}) = h_{\Sigma,\mu}(\Gamma \curvearrowright \mathbb{K}^\Gamma/\mathbb{K}) \leq \mathcal{C}_{\sup}(\Gamma)\log|\mathbb{K}|,$$

*where* $\mu$ *denotes Haar probability measure on* $\mathbb{K}^\Gamma/\mathbb{K}$*,* $\beta^1_{(2)}(\Gamma)$ *is the first* $L^2$*-Betti number of* $\Gamma$ *and* $\mathcal{C}_{\sup}(\Gamma)$ *is the* sup *cost of* $\Gamma$*.*

Conjecturally, $1+\beta^1_{(2)}(\Gamma) = \mathcal{C}_{\sup}(\Gamma)$. In spite of these negative results, the following remains open.

*Problem 8.* Does the $f$-invariant satisfy an addition formula?

In [**Bow10c**] a positive answer is claimed. However, there is a serious flaw. It was partially corrected in [**BG14**], which proved that indeed the $f$-invariant does satisfy an addition formula whenever $X$ is totally disconnected and there exists a special kind of generating partition. When $X$ is a connected finite-dimensional Lie group, the $f$-invariant of $\Gamma \curvearrowright X$ is minus infinity and an addition formula also holds for degenerate reasons.

*Problem 9.* Is there any formula or algorithm for computing the sofic entropy of $\Gamma \curvearrowright X_I$ where $I \subset \mathbb{Z}\Gamma$ is a finitely generated ideal and $X_I = \widehat{\mathbb{Z}\Gamma/I}$ is the Pontryagin dual? Such a formula is known when $\Gamma$ is amenable [**LT14**] but unknown for non-amenable groups.

3.4.5. *Further results.* In recent stunning work, Hayes has shown that for any algebraic action $\Gamma \curvearrowright X$ of a sofic group, there is a closed $\Gamma$-invariant normal subgroup $Y \leq X$ such that the outer Pinsker algebra is the sigma-algebra of $Y$-invariant Borel subsets (whenever the action satisfies the mild condition of admitting an lde-convergent sequence of model measures) [**Hay16c**]. Outer Pinsker algebras are defined in §11 below. The proof uses the product formula for outer Pinsker algebras (Theorem 11.1).

It follows that if $\Gamma \curvearrowright X/Y_0$ has positive outer entropy for every closed normal $\Gamma$-invariant subgroup $Y_0 \leq X$ with $Y_0 \neq X$, then $\Gamma \curvearrowright X$ has completely positive outer entropy. In [**Hay17a**] Hayes uses local sofic entropy theory (developed in [**KL13a**]) to show that if $f \in \mathbb{Z}\Gamma$ is invertible in the group von Neumann algebra $L\Gamma$ but not in $\mathbb{Z}\Gamma$ then every $k$-tuple of points in $X_f$ is a $\Sigma$-IE-$k$-tuple. This implies positive sofic entropy. So the Fuglede–Kadison determinant of $f$ is $> 1$ (answering a question of Deninger) and the $\Sigma$-entropy of $\Gamma \curvearrowright X_f$ is positive. Moreover, in [**Hay16c**] it is shown how this implies completely positive outer entropy (assuming the same mild condition as above). In [**Hay17**] it is shown that completely positive outer entropy implies the Koopman representation embeds into the countable sum of left-regular representations.

† Thanks to an anonymous reviewer for pointing this out.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*Question 6.* If $f \in \mathbb{Z}\Gamma$ is invertible in the group von Neumann algebra $L\Gamma$ but not in $\mathbb{Z}\Gamma$ is the action $\Gamma \curvearrowright (X_f, \mu_f)$ orbit-equivalent to a Bernoulli shift? Is it measurably conjugate to a Bernoulli shift? The latter is open even in the special case of amenable $\Gamma$. It is known to be true when $\Gamma = \mathbb{Z}^d$ [**Sch95**]. (More precisely, one should consider $\Gamma/N \curvearrowright (X_f, \mu_f)$ where $N$ is the kernel of $\Gamma \curvearrowright X_f$. This kernel is finite and therefore trivial if $\Gamma$ is torsion-free.)

Finally, Hayes has used Austin's lde-$\Sigma$-entropy to prove that the topological $\Sigma$-entropy of $\Gamma \curvearrowright X$ agrees with the measure $\Sigma$-entropy whenever the action satisfies the above-mentioned mild condition [**Hay16a**]. The case in which $\Gamma$ is amenable was handled earlier in [**Den06**, Theorem 3].

*Question 7.* Suppose $\Gamma$ is sofic, $\Sigma$ is a sofic approximation to $\Gamma$ and $f \in \mathbb{Z}\Gamma$ is such that the action $\Gamma \curvearrowright (X_f, \mu_f)$ has completely positive $\Sigma$-entropy (abbreviated CPE$^{\Sigma}$). Is the Haar measure on $X_f$ the unique measure of maximal $\Sigma$-entropy? By [**CL15b**, Theorem 8.6], if $\Gamma$ is amenable, then the Haar measure is the unique measure of maximal entropy if and only if $\Gamma \curvearrowright (X_f, \mu_f)$ is CPE.

3.5. *Gaussian actions.* Associated to any orthogonal representation $\rho : \Gamma \to \mathcal{O}(\mathcal{H})$ on a real Hilbert space $\mathcal{H}$ is a *Gaussian action* $\Gamma \curvearrowright (X_\rho, \mu_\rho)$. The details of this construction can be found in [**Kec10**] for example. For intuition, if $\mathcal{H}$ is finite-dimensional then the Gaussian action is the action of $\Gamma$ on $\mathcal{H}$ with respect to the standard Gaussian measure. Hayes computed the entropy of Gaussian actions in [**Hay17b**].

THEOREM 3.9. *The representation $\rho$ decomposes as $\rho = \rho_1 \oplus \rho_2$ where $\rho_1$ is singular with respect to the left-regular representation (so no non-trivial subrepresentation of $\rho_1$ embeds into the left regular representation) and $\rho_2$ embeds into the countable power of the left-regular representation. Moreover,*

$$h_{\Sigma,\mu_\rho}(\Gamma \curvearrowright X_\rho) = \begin{cases} -\infty & \text{if } h_{\Sigma,\mu_{\rho_1}}(\Gamma \curvearrowright X_{\rho_1}) = -\infty, \\ 0 & \text{if } \rho_2 = 0 \text{ and } h_{\Sigma,\mu_{\rho_1}}(\Gamma \curvearrowright X_{\rho_1}) \neq -\infty, \\ +\infty & \text{if } \rho_2 \neq 0 \text{ and } h_{\Sigma,\mu_{\rho_1}}(\Gamma \curvearrowright X_{\rho_1}) \neq -\infty. \end{cases}$$

The proof uses a Polish model for the action (this is a continuous action $\Gamma \curvearrowright Y$ where $Y$ is a completely metrizable separable space with a measure $\nu$ that is measurably conjugate to the original action). It also uses a weak form of Sinai's factor theorem [**Hay18**]: if the Koopman representation of $\Gamma \curvearrowright (X, \mu)$ is singular to the left regular representation then the sofic entropy is non-positive. To see the connection, let $\sigma : \Gamma \to \mathrm{Sym}(V)$ be given. Then there is an induced map from $\Gamma$ into the unitary group of $\mathbb{C}^V$. If $\sigma$ is regarded as a sofic approximation to $\Gamma$, then this map approximates the left regular representation.

3.6. *Distal actions.*

*Definition 10.* An action $\Gamma \curvearrowright^T (X, d)$ by homeomorphisms on a compact metric space is *distal* if $\inf_{g \in \Gamma} d(T^g x, T^g y) > 0$ for every pair of distinct points $x, y \in X$. Note that profinite, compact and equicontinuous actions are distal [**CZ15**, Lemma 4.1]. A pmp action $\Gamma \curvearrowright^T (X, \mu)$ is *measure distal* if it is measurably conjugate to a distal action.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

THEOREM 3.10. [**KL13a**, **Hay17**] *Measure distal actions have zero Rokhlin entropy. Therefore they have non-positive sofic entropy. Also, distal actions have non-positive topological sofic entropy with respect to every sofic approximation.*

*Remarks on the proof.* The paper [**KL13a**] shows that if the topological sofic entropy of $\Gamma \curvearrowright^T X$ is positive (with respect to some sofic approximation) then the action is *Li–Yorke chaotic*. This condition means there is an uncountable subset $Z \subset X$ such that every non-diagonal pair $(x, y) \in Z \times Z$ satisfies

$$\limsup_{G \ni s \to \infty} d(T^s x, T^s y) > 0, \quad \liminf_{G \ni s \to \infty} d(T^s x, T^s y) = 0.$$

It follows that distal actions have non-positive topological sofic entropy. □

The special case of topological $\mathbb{Z}$-actions was handled earlier by Keynes [**Key70**]. Recall that naive entropy is an upper bound for sofic entropy. This motivates the following question.

*Question 8.* Do distal actions have zero topological naive entropy? The measure version of this question is also open.

Partial progress has been made by Burton.

THEOREM 3.11. [**Bur17**] *If $\Gamma$ contains an infinite cyclic subgroup then every distal action has zero topological naive entropy (and therefore non-positive topological sofic entropy).*

3.7. *Smooth actions.* The purpose of this section is to show that, under mild conditions, the entropy of an action of a large group by diffeomorphisms on a smooth manifold is non-positive. The starting point is a bound on the topological entropy of a single Lipschitz map. So let $(X, d)$ be a compact metric space. The *ball dimension* of $(X, d)$ is

$$\dim_{\text{ball}}(X, d) = \limsup_{\epsilon \searrow 0} \frac{\log S_\epsilon(X, d)}{|\log \epsilon|},$$

where $S_\epsilon(X, d)$ is the minimum cardinality of an $\epsilon$-spanning subset of $(X, d)$. Given any map $T : X \to X$, the *Lipschitz contant* $\text{Lip}(T)$ is

$$\text{Lip}(T) := \sup_{x \neq y} \frac{d(Tx, Ty)}{d(x, y)}.$$

LEMMA 3.12. *Let $(X, d)$ be a compact metric space and $T : X \to X$ a continuous map. If $\dim_{\text{ball}}(X, d) < \infty$ and $\text{Lip}(T) < \infty$ then*

$$h_{\text{top}}(T) \leq \dim_{\text{ball}}(X, d) \max(0, \log \text{Lip}(T)),$$

*where $h_{\text{top}}(T)$ is the topological entropy of $T$. In particular, if $X$ is a smooth Riemannian manifold and $T$ is a diffeomorphism then $h_{\text{top}}(T) < \infty$.*

*Proof.* This is [**KH95**, Theorem 3.2.9]. □

The next result shows that if the induced action of an infinite index amenable subgroup has finite entropy then the action of the group has non-positive entropy.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

LEMMA 3.13. *Let $\Gamma$ be a group with an infinite index amenable subgroup $\Lambda \leq \Gamma$. Let $\Gamma \curvearrowright^T X$ be a continuous action on a compact metrizable space. Also let $\mu$ be an invariant probability measure on $X$ that is ergodic with respect to $T$. Let $T_\Lambda$ denote the restriction of the action to $\Lambda$. Then*

$$h_{\text{top}}(T_\Lambda) < \infty \Rightarrow h_\Sigma(T) \leq 0 = h^{\text{naive}}(T),$$
$$h_\mu(T_\Lambda) < \infty \Rightarrow h_{\Sigma,\mu}(T) \leq 0 = h_\mu^{\text{naive}}(T)$$

*for any sofic approximation $\Sigma$ to $\Gamma$.*

*Proof.* The proofs in the topological and measure settings are similar, so we will just give the proof in the measure setting. Let $\mathcal{P}$ be any partition of $X$ with finite Shannon entropy. Let $F \subset \Gamma$ be any finite set of coset representatives of $\Gamma/\Lambda$. So if $f_1 \neq f_2$ with $f_1, f_2 \in F$ then $f_1\Lambda \cap f_2\Lambda = \emptyset$. Let $K \subset \Lambda$ be an arbitrary finite set. Then

$$\lim_{K \nearrow \Lambda} \frac{1}{|K|} H_\mu(\mathcal{P}^{FK}) \leq h_\mu(T_\Lambda) < \infty,$$

where the limit is along any Følner sequence for $\Lambda$. So

$$h_\mu^{\text{naive}}(T, \mathcal{P}) \leq \lim_{K \nearrow \Lambda} \frac{1}{|FK|} H_\mu(\mathcal{P}^{FK}) = \frac{1}{|F|} h_\mu(T_\Lambda).$$

Taking the supremum over all $\mathcal{P}$ proves $h_\mu^{\text{naive}}(T) \leq (1/|F|)h_\mu(T_\Lambda)$. Since $|F|$ is arbitrary, this implies $h_\mu^{\text{naive}}(T) = 0$. The rest follows from Propositions 2.14 and 2.11. □

THEOREM 3.14. *Suppose $\Gamma$ is a countable group with an infinite cyclic subgroup of infinite index. If $\Gamma \curvearrowright X$ is a continuous action by Lipschitz maps on a compact metric space $(X, d)$ that has finite ball dimension then $\Gamma \curvearrowright X$ has non-positive sofic entropy with respect to every sofic approximation. In particular, smooth actions on manifolds have non-positive sofic entropy.*

*Proof.* This follows immediately from the previous two lemmas. □

There is a similar result for the $f$-invariant [**BG14**, Lemma 3.5].

3.8. *Non-free actions.* Given an action $\Gamma \curvearrowright^T (X, \mu)$ and $x \in X$, let $\text{Stab}_T(x) = \{g \in \Gamma : T^g x = x\}$ be the stabilizer. An action is *non-free* if there is a positive measure set of $x$ such that $\text{Stab}_T(x)$ is non-trivial. There are two results concerning the entropy of non-free actions.

THEOREM 3.15. *Suppose $\Gamma \curvearrowright^T (X, \mu)$ is an ergodic $\Gamma$-action with positive sofic entropy with respect to some sofic approximation. Then $\text{Stab}_T(x)$ is finite for a.e. $x$.*

This result is [**Mey16**, Theorem 2.3]. The special case of amenable groups is handled in a remark in the last section of [**Wei03**].

THEOREM 3.16. [**Sew16a**, Theorem 1.1] *Let $\mathbb{F} \curvearrowright^T (X, \mu)$ be a pmp action of a finitely generated free group. Suppose the action has a finite-entropy generating partition, $f_\mu(T) \neq -\infty$ and $\mathbb{F} \curvearrowright^S (Y, \nu)$ is a factor action. Then for $\nu$-a.e. $y \in Y$ either:*

- $\text{Stab}_S(y)$ *is trivial or*
- $\text{Stab}_S(y)$ *has finite-index in $\Gamma$ and $y$ is an atom (i.e. $\nu(\{y\}) > 0$).*

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

An anonymous reviewer pointed out that there are ergodic actions $\Gamma \curvearrowright^T (X, \mu)$ with positive Rokhlin entropy such that $\mathrm{Stab}_T(x)$ is infinite for a.e. $x$. For example, let $\mathbb{Z} \curvearrowright^S (X, \mu)$ be any ergodic action with positive Rokhlin entropy. Thinking of $\mathbb{Z}$ as a quotient of $\mathbb{Z}^2$, we obtain an ergodic action $\mathbb{Z}^2 \curvearrowright^T (X, \mu)$. Moreover, the Rokhlin entropy of this action is the same as the Rokhlin entropy of $\mathbb{Z} \curvearrowright^S (X, \mu)$ since the two actions have the same generating partitions.

*Question 9.* If $\Gamma \curvearrowright^T (X, \mu)$ is ergodic and has positive naive entropy (topological or measure) then is $\mathrm{Stab}_T(x)$ finite for a.e. $x$?

## 4. *Perturbations*

4.1. *Perturbing the sofic approximation.* When $\Gamma$ is amenable the sofic entropy agrees with classical entropy and therefore does not depend on the choice of sofic approximation. This is not true in the non-amenable case even when the action is trivial (§3.1). This subsection provides another explicit counterexample, which works for both topological and measure entropy. However, the example is not entirely satisfying because it leaves open a major problem.

*Question 10.* Suppose $\Gamma \curvearrowright X$ is a continuous action on a compact metrizable space by homeomorphisms. Let $\Sigma_1, \Sigma_2$ be two sofic approximations of $\Gamma$. Suppose $h_{\Sigma_i}(\Gamma \curvearrowright X)$ is not minus infinity for $i = 1, 2$. Is it true that $h_{\Sigma_1}(\Gamma \curvearrowright X) = h_{\Sigma_2}(\Gamma \curvearrowright X)$? The measure entropy version of this question is also open. In fact, it is open even in the special case of the Ising model on the free group (§3.3). It is possible that if the measure sofic entropy of an action is not minus infinity then it must equal the Rokhlin entropy.

THEOREM 4.1. *There exists a sofic group $\Gamma$ with sofic approximations $\Sigma_1, \Sigma_2$ and a continuous action $\Gamma \curvearrowright X$ on a compact metric space such that*

$$h_{\Sigma_1}(\Gamma \curvearrowright X) \neq h_{\Sigma_2}(\Gamma \curvearrowright X).$$

*Moreover, there exists an invariant probability measure $\mu$ on $X$ such that*

$$h_{\Sigma_1,\mu}(\Gamma \curvearrowright X) \neq h_{\Sigma_2,\mu}(\Gamma \curvearrowright X).$$

*Proof.* Let $\mathbb{F}_2 = \langle a, b \rangle$ be the rank-two free group and let $\mathbb{F}_2 \curvearrowright \mathbb{Z}/2$ be the action in which each generator in $\{a, b\}$ acts non-trivially. Recall that if $\sigma : \mathbb{F}_2 \to \mathrm{Sym}(V)$ is any map then the associated graph $G_\sigma$ has vertex set $V$ and edges $\{v, \sigma(a)v\}, \{v, \sigma(b)v\}$ for $v \in V$.

Let $\Sigma_1 = \{\sigma_{1,n}\}_{n\in\mathbb{N}}$ be a sofic approximation whose associated graphs are far from bipartite. To be precise, we require the existence of an $\epsilon > 0$ such that if $\mathcal{P} = \{P_1, P_2\}$ is a partition of $V_n$ with

$$|P_1| \geq (1 - \epsilon)|V_n|/2, \quad |P_2| \geq (1 - \epsilon)|V_n|/2$$

then the number of edges $\{v, w\}$ of the associated graph $\Gamma_{\sigma_{1,n}}$ with either $v, w \in P_1$ or $v, w \in P_2$ is at least $\epsilon|V_n|$. For example, we could choose $\sigma_{1,n} : \mathbb{F}_2 \to \mathrm{Sym}(V_n)$ uniformly at random (among all homomorphisms from $\mathbb{F}_2$ to $\mathrm{Sym}(V_n)$). With probability 1 the resulting sofic approximation satisfies the above property. This can be proven using the $f$-invariant for example or directly with combinatorial estimates.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

The strong non-bipartiteness of the graphs of $\Sigma_1$ immediately implies $h_{\Sigma_1}(\Gamma \curvearrowright X) = -\infty$. On the other hand, if $\Sigma_2$ is a sofic approximation whose associated graphs are bipartite, then $h_{\Sigma_2}(\Gamma \curvearrowright X) \geq 0$ since the bipartition of the graphs yields two microstates for the action. In fact, $h_{\Sigma_2}(\Gamma \curvearrowright X) = 0$ since any microstate must be close to one of the two bipartitions of the graphs. If $\mu$ is the unique invariant probability on $\mathbb{Z}/2$ then $h_{\Sigma_i,\mu}(\Gamma \curvearrowright X) = h_{\Sigma_i}(\Gamma \curvearrowright X)$ for $i = 1, 2$, which proves the last statement. □

*Problem 10.* Does the sofic entropy of either a topological action or a measure-preserving action vary upper semi-continuously with respect to the edit distance on sofic approximations (§2.2.4)? This seems likely but it has not been worked out. It is unknown whether entropy varies continuously (excluding the value negative infinity).

4.2. *Subgroups.* It is well known that if $T$ is an automorphism of $(X, \mu)$ and $n \in \mathbb{Z} - \{0\}$ then

$$h_\mu(T^n) = |n| h_\mu(T).$$

More generally, if $\Gamma$ is an amenable group, $\Gamma \curvearrowright (X, \mu)$ is a pmp action and $\Gamma' \leq \Gamma$ has finite index then

$$h_\mu(\Gamma' \curvearrowright X) = [\Gamma : \Gamma'] h_\mu(\Gamma \curvearrowright X).$$

See [**Dan01**, Corollary 3.5]. This is called the *subgroup formula*.

In the case of the $f$-invariant, a similar formula holds.

THEOREM 4.2. [**Sew14b**, Theorem 1.1, Corollary 1.2] *Let $\Gamma$ be a finitely generated free group, $\Lambda \leq \Gamma$ and $\Gamma \curvearrowright (X, \mu)$ be a pmp action with a finite-entropy generating partition. If $\Lambda$ has finite index in $\Gamma$ then the induced action $\Lambda \curvearrowright (X, \mu)$ also has a finite-entropy generating partition and*

$$f_\mu(\Lambda \curvearrowright X) = [\Gamma : \Lambda] f_\mu(\Gamma \curvearrowright X).$$

*If $[\Gamma : \Lambda] = +\infty$, there are infinitely many finite index subgroups $\Gamma'$ with $\Lambda < \Gamma' < \Gamma$ and there is a finite-entropy generating partition for the action $\Lambda \curvearrowright (X, \mu)$ then $f_\mu(\Gamma \curvearrowright X) \leq 0$.*

For Rokhlin and sofic entropy we have the following theorem.

THEOREM 4.3. [**Sew16b**, Theorem 1.7] *Let $\Gamma$ be a sofic group with sofic approximation $\Sigma$. Suppose that for all finite-index normal subgroups $N \lhd \Gamma$, the $\Sigma$-entropy of $\Gamma \curvearrowright (\Gamma/N, u_{\Gamma/N})$ is not minus infinity. Then for every finite index subgroup $\Lambda \leq \Gamma$ and aperiodic pmp action $\Gamma \curvearrowright^T (X, \mu)$,*

$$[\Gamma : \Lambda] h_\Sigma(T) \leq h^{\mathrm{Rok}}(T \restriction \Lambda) \leq [\Gamma : \Lambda] h^{\mathrm{Rok}}(T),$$

*where $T \restriction \Lambda$ is the restriction of the action to $\Lambda$.*

THEOREM 4.4. *If $\Lambda \leq \Gamma$ and $\Gamma \curvearrowright X$ is a continuous action on a compact metrizable space then*

$$h_\Sigma(\Gamma \curvearrowright X) \leq h_{\Sigma \restriction \Lambda}(\Lambda \curvearrowright X),$$

*where $\Sigma \restriction \Lambda$ denotes the restriction of $\Sigma$ to $\Lambda$. A similar inequality holds in the measure-entropy case.*

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*Proof.* This follows immediately upon realizing that for any continuous pseudometric $\rho$ on $X$, any finite $F \subset \Gamma$, $\delta > 0$ and $\sigma : \Gamma \to \mathrm{Sym}(d)$,

$$\mathrm{Map}(T, \rho, F, \delta, \sigma) \subset \mathrm{Map}(T \restriction \Lambda, \rho, F, \delta, \sigma \restriction \Lambda).$$

The argument for the measure-entropy case is similar. □

*Question 11.* Are the bounds in Theorems 4.3 and 4.4 tight?

Next we present a counterexample.

THEOREM 4.5. *There exists a finite-index subgroup $\Lambda \leq \Gamma$ of a sofic group $\Gamma$, a sofic approximation $\Sigma$ to $\Gamma$ and a pmp action $\Gamma \curvearrowright^T (X, \mu)$ such that*

$$h_{\Sigma,\mu}(T) = -\infty, \quad h_{\Sigma\restriction\Lambda,\mu}(T \restriction \Lambda) = 0.$$

*Proof.* Let $\Lambda = \langle a, b \rangle$ be the rank-two free group, $\Gamma = \Lambda \times \mathbb{Z}/2$. We identify $\Lambda$ with the subgroup $\Lambda \times \{0\}$ of $\Gamma$. Let $(X, \mu) = (\mathbb{Z}/2, u_2)$ and let $T$ be the trivial action of $\Gamma$ on $\mathbb{Z}/2$. Let $\Sigma' = \{\sigma'_i : \mathbb{F} \to \mathrm{Sym}(V_i)\}_{i=1}^{\infty}$ be a sofic approximation to $\mathbb{F} = \Lambda$ by expanders as in §3.1.1. Define $\sigma_i : \mathbb{F} \times \mathbb{Z}/2 \to \mathrm{Sym}(V_i \times \mathbb{Z}/2)$ by

$$\sigma_i(f, x)(v, y) = (\sigma'_i(f)v, x + y).$$

Then $\Sigma = \{\sigma_i\}_{i=1}^{\infty}$ is a sofic approximation to $\Gamma = \mathbb{F} \times \mathbb{Z}/2$.

Because $\Sigma'$ is by expanders, $\Sigma$ is also by expanders. So Proposition 3.1 implies $h_{\Sigma,u_2}(T) = -\infty$. However, the restriction $\Sigma \restriction \mathbb{F}$ is the disjoint union of two copies of $\Sigma'$. So the function

$$\phi : V_i \times \mathbb{Z}/2 \to \mathbb{Z}/2, \quad \phi(v, x) = x$$

is a microstate for the trivial action $\Lambda \curvearrowright \mathbb{Z}/2$. This shows that $h_{\Sigma\restriction\Lambda,u_2}(T \restriction \Lambda) \geq 0$. The upper bound can be derived directly or via naive entropy. □

### 4.3. *Co-induction.*

*Definition 11.* Let $\Lambda \leq \Gamma$ be countable discrete groups. Let $X$ be a compact metrizable space and $\Lambda \curvearrowright X$ an action by homeomorphisms. Let

$$Y = \{f : \Gamma \to X : f(gh) = h^{-1} f(g) \ \forall g \in \Gamma, h \in \Lambda\} \subset X^{\Gamma}$$

and give $Y$ the subspace topology (and $X^{\Gamma}$ the product topology). So $Y$ is a compact metrizable space and $\Gamma \curvearrowright Y$ by

$$(gf)(x) = f(g^{-1}x) \quad \text{for all } g, x \in \Gamma, \ f \in Y.$$

The action $\Gamma \curvearrowright Y$ is the *action of $\Lambda \curvearrowright X$ co-induced to $\Gamma$*.

For example, if $\Lambda$ is the trivial subgroup, then the co-induced action is a Bernoulli shift over $\Gamma$.

THEOREM 4.6. [**Hay16b**, Proposition 5.29] *Let $\Lambda \leq \Gamma$ and $\Sigma$ be a sofic approximation to $\Gamma$. Let $\Lambda \curvearrowright X$ be an action on a compact metrizable space by homeomorphisms and $\Gamma \curvearrowright Y$ the co-induced action. Then $h_\Sigma(\Gamma \curvearrowright Y) = h_{\Sigma\restriction\Lambda}(\Lambda \curvearrowright X)$.*

*Question 12.* Is there a measure entropy version of Theorem 4.6 or a Rokhlin entropy version? Results have been obtained by Austin for a variant of sofic entropy [**Aus16a**].

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

4.4. *Perturbing the partition.*

*Definition 12.* (The space of partitions) Let $(X, \mu)$ be a standard probability space. Given countable partitions $\mathcal{P}$, $\mathcal{Q}$ of $X$ the *relative Shannon entropy* of $\mathcal{P}$ given $\mathcal{Q}$ is

$$H_\mu(\mathcal{P}|\mathcal{Q}) := H_\mu(\mathcal{P} \vee \mathcal{Q}) - H_\mu(\mathcal{Q}).$$

The *Rokhlin distance* between $\mathcal{P}$ and $\mathcal{Q}$ is

$$d(\mathcal{P}, \mathcal{Q}) := H_\mu(\mathcal{P}|\mathcal{Q}) + H_\mu(\mathcal{Q}|\mathcal{P}).$$

This is a metric on the space of (mod 0 equivalence classes of) partitions of $(X, \mu)$ with finite Shannon entropy.

*Notation 1.* Let $\Gamma \curvearrowright^T (X, \mu)$ be a pmp action. Given a measurable partition $\mathcal{P}$ of $X$, the Mackey realization theorem implies the existence of a factor $\Gamma \curvearrowright^{T_\mathcal{P}} (X_\mathcal{P}, \mu_\mathcal{P})$ of $T$ such that if $\Phi : X \to X_\mathcal{P}$ is the factor map then $\Phi^{-1}(\mathcal{B}_{X_\mathcal{P}})$ is the sigma-algebra $\bigvee_{g\in\Gamma} T^g\mathcal{P}$ (up to measure zero).

PROPOSITION 4.7. *Let $\Sigma$ be a sofic approximation to a countable group $\Gamma$ and $\Gamma \curvearrowright^T (X, \mu)$ a pmp action. Then both the sofic entropy $h_{\Sigma,\mu_\mathcal{P}}(T_\mathcal{P})$ and the Rokhlin entropy $h^{\mathrm{Rok}}(T_\mathcal{P})$ vary upper semi-continuously in $\mathcal{P}$ with respect to the Rokhlin metric. If $\Gamma$ is amenable then the entropy varies continuously in $\mathcal{P}$. In the non-amenable case, neither entropy is continuous in general.*

*Proof.* The first statement is contained in [**Bow10b**, Corollary 6.3] (sofic entropy) and [**Sew15b**] (Rokhlin entropy). The amenable case is handled in [**MO85**, Proposition 4.3.13]. For the last statement, we perturb the Ornstein–Weiss example as follows. Let $\Gamma = \mathbb{F}_2 = \langle a, b\rangle$ be the rank-two free group, $X = (\mathbb{Z}/2)^\Gamma$ be the full two-shift and $\mu = u_2^\Gamma$ be the product measure on $X$. So $T$ is the usual shift action $(T^g)x(f) = x(g^{-1}f)$. Let $\phi : X \to \mathbb{Z}/2 \times \mathbb{Z}/2$ be the observable

$$\phi(x) = (x(1_\Gamma) + x(a), x(1_\Gamma) + x(b))$$

and $\mathcal{P} = \{\phi^{-1}(i, j) : i, j \in \mathbb{Z}/2\}$ be the corresponding partition. For every $n > 0$, choose a subset $A_n \subset \{x \in X : x(1_\Gamma) = 0\}$ such that $0 < \mu(A_n) < 1/n$. Finally, let

$$\mathcal{P}_n = \mathcal{P} \vee \{A_n, X\backslash A_n\}.$$

The Ornstein–Weiss example (§1.3) shows that $\Gamma \curvearrowright (X_\mathcal{P}, \mu_\mathcal{P})$ is isomorphic to the full four-shift. In particular,

$$h_{\Sigma,\mu_\mathcal{P}}(T_\mathcal{P}) = \log 4 = h^{\mathrm{Rok}}(T_\mathcal{P}).$$

Moreover, the factor map from $(X, \mu)$ to $(X_\mathcal{P}, \mu_\mathcal{P})$ is two-one. Indeed, if $\bar{1} \in (\mathbb{Z}/2)^\Gamma$ denotes the constant function then for any $x \in X$, $x$ and $\bar{1} + x$ have the same image in $(X_\mathcal{P}, \mu_\mathcal{P})$.

We claim that $\mathcal{P}_n$ is generating for all $n$. It suffices to show that for a.e. $x \in X$ there exists $g \in \Gamma$ such that $T^g x$ and $T^g(\bar{1} + x)$ lie in different parts of $\mathcal{P}_n$. Since the action is ergodic, there exists $g \in \Gamma$ such that $T^g x \in A_n$. This implies $(T^g x)(1_\Gamma) = 1$ and therefore $(T^g(x + \bar{1}))(1_\Gamma) = 0$. In particular, $T^g(\bar{1} + x)$ cannot be in $A_n$. So $\mathcal{P}_n$ is generating as claimed.

Because $\mathcal{P}_n$ is generating,

$$h_{\Sigma,\mu_{\mathcal{P}_n}}(T_{\mathcal{P}_n}) = \log 2 = h^{\mathrm{Rok}}(T_{\mathcal{P}_n})$$

for all $n$. Since $\mu(A_n) \to 0$ as $n \to \infty$, $\mathcal{P}_n \to \mathcal{P}$ in the Rokhlin metric. □

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

4.5. *Perturbing the measure.* Let $K$ be a finite set, $\Gamma \curvearrowright^T K^\Gamma$ the shift action and $\mathrm{Prob}_\Gamma(K^\Gamma)$ the space of all $T(\Gamma)$-invariant Borel probability measures on $K^\Gamma$ with the weak* topology.

PROPOSITION 4.8. *Let $\Sigma$ be a sofic approximation to $\Gamma$. Then both the $\Sigma$-entropy and the Rokhlin entropy of the action $\Gamma \curvearrowright^T (K^\Gamma, \mu)$ vary upper semi-continuously in $\mu \in \mathrm{Prob}_\Gamma(K^\Gamma)$ with respect to the weak* topology. In general, these are not continuous (even when $\Gamma = \mathbb{Z}$).*

*Proof remarks.* The Rokhlin entropy case is handled in [**Sew15b**]. For sofic entropy, a more general statement is proven in [**CZ15**]: whenever $\Gamma$ acts expansively on a compact metric space $X$ then $\Sigma$-entropy varies upper semi-continuously with respect to the measure $\mu$. In fact, this is proven for a weak form of expansitivity called asymptotic $h$-expansitivity.

In [**Bow09**] it is shown that if $\mathbb{F}$ is a finite rank free group (including $\mathbb{Z}$) then the set of shift-invariant measures $\mu$ that have finite support are dense in $\mathrm{Prob}_\mathbb{F}(K^\mathbb{F})$. If a shift-invariant measure has finite support then it has zero naive entropy and therefore non-positive sofic and Rokhlin entropy (by Propositions 2.15 and 2.11). So entropy is not a continuous function of $\mu \in \mathrm{Prob}_\mathbb{F}(K^\mathbb{F})$ (even when $\mathbb{F} = \mathbb{Z}$). □

*Remark 17.* In §7.4, we discuss a stronger topology on $\mathrm{Prob}_\Gamma(K^\Gamma)$ called the $d$-bar topology. If $\Gamma$ is amenable then entropy varies continuously in the $d$-bar topology but there are explicit counterexamples when $\Gamma$ is a rank-two free group.

4.6. *Orbit-equivalence.* Two actions $\Gamma \curvearrowright (X, \mu)$, $\Lambda \curvearrowright (Y, \nu)$ are *orbit-equivalent* (abbreviated OE) if there exists a measure space isomorphism $\Phi : X \to Y$ such that $\Phi(\Gamma x) = \Lambda \Phi(x)$ for a.e. $x$. By work of Dye and Ornstein–Weiss [**Dye63**, **Dye59**, **OW80**], all ergodic essentially free pmp actions of infinite amenable groups are OE. In particular, entropy is not an OE-invariant. However, every non-amenable group admits uncountably many non-OE actions [**Eps08**, **IKT09**]. Moreover, there are many groups for which Bernoulli shifts are OE-rigid in the sense that OE implies measure-conjugacy. These include ICC groups with property (T) [**Pop06b**] and direct products of non-amenable groups with infinite groups that have no non-trivial finite normal subgroups [**Pop08**]. So for these actions, entropy is automatically an OE-invariant.

This leads to the general question.

*Question 13.* Under what conditions on a group action can one conclude that entropy (whether sofic/Rokhlin/naive) is an OE-invariant?

For example in [**Kid08**] it is proven that $\Gamma$ is a non-exceptional mapping class group then any pmp action in which all finite-index subgroups acts ergodically is OE-rigid. It follows that entropy is an OE-invariant for such actions. In §4.6.1 below, we show there is a property of actions called weak compactness, which is an OE-invariant and implies zero entropy whenever the group is non-amenable.

4.6.1. *Weakly compact actions.* (The ideas of this section have been gracefully provided by Hayes.)

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

If $(X, \mu), (Y, \nu)$ are standard probability spaces and $f : X \to \mathbb{C}, g : Y \to \mathbb{C}$ are measurable then $f \otimes g : X \times Y \to \mathbb{C}$ is defined by $(f \otimes g)(x, y) = f(x)g(y)$. If $\Gamma \curvearrowright (X, \mu)$ is a pmp action then $\kappa : \Gamma \to \mathcal{U}(L^2(\mu))$ (where $\mathcal{U}(\cdot)$ is the unitary group) denotes the Koopman representation given by

$$(\kappa_g \xi)(x) = \xi(g^{-1}x) \quad \text{for all } \xi \in L^2(\mu),\ g \in \Gamma,\ x \in X.$$

The map $\kappa_g \otimes \kappa_g : \Gamma \to \mathcal{U}(L^2(\mu \times \mu))$ is the Koopman representation of $\Gamma$ on $X \times X$. Define $\kappa_g(\xi)$, $\kappa_g \otimes \kappa_g(\zeta)$ by the same formula for $\xi \in L^1(\mu)$, $\zeta \in L^1(\mu \times \mu)$.

*Definition 13.* [**OP10**] A pmp action $\Gamma \curvearrowright (X, \mu)$ is *weakly compact* if there is a sequence $\xi_1, \xi_2, \ldots \in L^2(\mu \times \mu)$ such that each $\xi_n \geq 0$,

- $\|\xi_n - (v \otimes \overline{v})\xi_n\|_2 \to 0$ as $n \to \infty$ for all $v : X \to S^1$ measurable;
- $\|\xi_n - (\kappa_g \otimes \kappa_g)(\xi_n)\|_2 \to 0$ as $n \to \infty$ for all $g \in \Gamma$;
- $\langle (f \otimes 1)\xi_n, \xi_n \rangle = \int_X f \, d\mu = \langle (1 \otimes f)\xi_n, \xi_n \rangle$, for all $f \in L^\infty(X, \mu)$.

It is shown in [**OP10**, Proposition 3.2] that compact actions are weakly compact. Moreover, weak compactness is an OE-invariant (see the remarks after [**OP10**, Proposition 3.4]). It is shown below that all weakly compact actions of non-amenable groups have zero Rokhlin entropy. Moreover, because weak compactness is preserved under factors, no factor of a weakly compact action has positive entropy.

Set $\zeta_n = \xi_n^2$ to see that weak compactness is equivalent to saying that there exists $\zeta_n \in L^1(\mu \times \mu)$ so that $\zeta_n \geq 0$ and:

- $\|\zeta_n - (v \otimes \overline{v})\zeta_n\|_1 \to 0$ as $n \to \infty$ for all measurable $v : X \to S^1 \subset \mathbb{C}$;
- $\|\zeta_n - (\kappa_g \otimes \kappa_g)(\zeta_n)\|_1 \to 0$ as $n \to \infty$ for all $g \in \Gamma$;
- $\int f(x)\zeta_n(x, y) \, d\mu \times \mu(x, y) = \int_X f \, d\mu = \int f(y)\zeta_n(x, y) \, d\mu \times \mu(x, y)$ for all $f \in L^\infty(\mu)$.

PROPOSITION 4.9. *All factors of weakly compact actions are weakly compact.*

*Proof.* Let $\Gamma \curvearrowright (Y, \nu)$ be a factor of $\Gamma \curvearrowright (X, \mu)$ and $\mathbb{E}_{Y \times Y}(f)$ be the conditional expectation of $f \in L^1(\mu \times \mu)$ onto $L^1(\nu \times \nu)$ (which we identify with the obvious subspace of $L^1(\mu \times \mu)$). Let $\zeta_n \in L^1(\mu \times \mu)$ be as in the remarks after Definition 13 and set $\eta_n = \mathbb{E}_{Y \times Y}(\zeta_n)$. It is straightforward to check that

$$\|\eta_n - (v \otimes \overline{v})\eta_n\|_1 \to 0 \quad \text{for all measurable } v : Y \to S^1 \subset \mathbb{C},$$
$$\|\eta_n - (\kappa_g \otimes \kappa_g)(\eta_n)\|_1 \to 0 \quad \text{for all } g \in \Gamma,$$

$$\int f(x)\eta_n(x, y) \, d\nu \times \nu(x, y) = \int_Y f \, d\nu = \int f(y)\eta_n(x, y) \, d\nu \times \nu(x, y) \quad \text{for all } f \in L^\infty(Y, \mu).$$

Thus $\{\eta_n\}$ witnesses that $\Gamma \curvearrowright (Y, \nu)$ is weakly compact. □

PROPOSITION 4.10. *If $\Gamma$ is non-amenable then no non-trivial Bernoulli action of $\Gamma$ is weakly compact.*

*Proof.* Weak compactness of $\Gamma \curvearrowright (X, \mu)$ tautologically implies the product action $\Gamma \curvearrowright (X \times X, \mu \times \mu)$ does not have spectral gap. However, if $\Gamma \curvearrowright (X, \mu)$ is Bernoulli

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

then so is the product action. It is well known that Bernoulli actions of non-amenable groups have spectral gap. □

COROLLARY 4.11. *If* $\Gamma$ *is non-amenable and* $\Gamma \curvearrowright (X, \mu)$ *is essentially free and weakly compact then it has zero Rokhlin entropy. Moreover, every action OE to this action has zero Rokhlin entropy.*

*Proof.* If $\Gamma \curvearrowright (X, \mu)$ has positive Rokhlin entropy then it factors onto a Bernoulli shift by Seward's generalization of Sinai's theorem (Theorem 7.4). The corollary now follows from Propositions 4.9 and 4.10. □

*Question 14.* Do all weakly compact actions of non-amenable groups have zero naive entropy?

4.6.2. *Orbit equivalence and relative entropy.* Rudolph and Weiss showed that while entropy is not an OE-invariant, entropy relative to the orbit-change sigma-algebra is [**RW00**]. They developed this tool to prove that CPE actions of amenable groups are uniformly mixing. It has since been used to show that CPE actions of amenable groups have countable Lebesgue spectrum [**DG02**] and to give an alternative development of Ornstein theory [**Dan01**, **DP02**]. We give the precise statement next.

Suppose essentially free actions $\Gamma \curvearrowright (X, \mu)$ and $\Lambda \curvearrowright (Y, \nu)$ are orbit-equivalent via an isomorphism $\Phi : X \to Y$. Then we can define cocycles $\alpha : \Gamma \times X \to \Lambda$ and $\beta : \Lambda \times Y \to \Gamma$ by

$$\alpha(g, x)\Phi(x) = \Phi(gx), \quad \beta(h, y)\Phi^{-1}(y) = \Phi^{-1}(hy).$$

These satisfy the cocycle equations

$$\alpha(gh, x) = \alpha(g, hx)\alpha(h, x), \quad \beta(gh, y) = \beta(g, hy)\beta(h, y).$$

Let $\mathcal{F}_X \subset \mathcal{B}_X$ be the smallest sigma-subalgebra such that $\alpha(g, \cdot) : X \to \Lambda$ is $\mathcal{F}_X$-measurable for all $g \in \Gamma$. If $\mathcal{F}_Y \subset \mathcal{B}_Y$ is defined similarly, then $\Phi$ maps $\mathcal{F}_X$ to $\mathcal{F}_Y$. These are the *orbit-change* sigma-algebras.

THEOREM 4.12. (Rudolph–Weiss theorem [**RW00**]) *Let* $\Gamma$, $\Lambda$ *be amenable groups and* $\Gamma \curvearrowright (X, \mu)$, $\Lambda \curvearrowright (Y, \nu)$ *essentially free pmp actions as above. Then*

$$h_\mu(\Gamma \curvearrowright X | \mathcal{F}_X) = h_\nu(\Lambda \curvearrowright Y | \mathcal{F}_Y).$$

This theorem has been generalized to Rokhlin entropy and arbitrary countable groups [**Sew19**]. It plays a major role in Seward's extension of Krieger's theorem. It also inspired sofic entropy theory for actions of groupoids [**Bow14**].

*Question 15.* Is there an analog of the Rudolph–Weiss theorem for arbitrary sofic groups if one takes for entropy the supremum over all sofic entropies? Perhaps one should use Hayes' definition of relative sofic entropy [**Hay16c**]?

4.6.3. *Integrable orbit-equivalence.* With notation as in §4.6.2, suppose that $\Gamma$ and $\Lambda$ are endowed with word lengths $\|\cdot\|_\Gamma$, $\|\cdot\|_\Lambda$. Then the orbit-equivalence is said to be

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*integrable* if for every $g \in \Gamma$ and $h \in \Lambda$,

$$\int \|\alpha(g, x)\|_\Lambda \, d\mu(x) < \infty, \quad \int \|\beta(h, y)\|_\Gamma \, d\nu(y) < \infty.$$

Austin proved in [**Aus16b**] that if $\Gamma$ and $\Lambda$ are amenable then entropy is invariant under integrable orbit-equivalence (abbreviated IOE).

*Question 16.* Is either Rokhlin entropy, $\Sigma$-entropy, naive entropy or the $f$-invariant IOE-invariant? Is the supremum of $\Sigma$-entropies over all $\Sigma$ an IOE-invariant?

## 5. *Factors and extensions*

5.1. *A variant of the Ornstein–Weiss example.* The Ornstein–Weiss example (§1.3) is an entropy-increasing finite-to-one factor map of a Bernoulli shift action of the free group onto another Bernoulli shift. It was generalized by Gaboriau–Seward via an algebraic construction that applies to all groups (§3.4.4). Their entropy bounds show that the map is entropy-increasing whenever $\Gamma$ has positive first $L^2$-Betti number. Next we present an example which is even more extreme: a zero entropy action with a finite-to-one factor map onto a Bernoulli shift.

THEOREM 5.1. *Let $\mathbb{F} = \langle a, b \rangle$. Then there exists an ergodic essentially free pmp action $\mathbb{F} \curvearrowright^T (X, \mu)$ such that:*

(1) $h_{\Sigma,\mu}(T) = f_\mu(T) = h^{\mathrm{Rok}}(T) = 0$ *for every sofic approximation $\Sigma$;*
(2) $\mathbb{F} \curvearrowright^T (X, \mu)$ *admits a two-one factor map onto a Bernoulli shift.*
(3) *Moreover the action is algebraic and the factor map is a continuous homomorphism.*

*Proof.* Let $\Lambda = \langle a \rangle$ denote the subgroup of $\mathbb{F}$ generated by $a$. We regard $(\mathbb{Z}/2)^{\mathbb{F}}$ as a compact abelian group under pointwise addition. Let $X \leq (\mathbb{Z}/2)^{\mathbb{F}}$ denote the subgroup consisting of all $x \in (\mathbb{Z}/2)^{\mathbb{F}}$ such that $x_{ga^n} = x_{ga^m}$ for every $g \in \mathbb{F}$ and $n, m \in \mathbb{Z}$. Let $(T^g)_{g\in\Gamma}$ be the shift action on $X$. So $(T^g)x(f) = x(g^{-1}f)$. This action is by group-automorphisms and so preserves the Haar measure.

The action $\mathbb{F} \curvearrowright^T X$ is measurably conjugate to the trivial action of $\Lambda$ on $\mathbb{Z}/2$ co-induced to $\mathbb{F}$. Because co-induction preserves topological entropy (Theorem 4.6) the action $\mathbb{F} \curvearrowright^T X$ has zero topological sofic entropy with respect to every sofic approximation. By the variational principle (Theorem 8.1), the measure sofic entropy is also zero. The time-zero partition is a Markov partition on $X$ (when viewed as a subset of $(\mathbb{Z}/2)^{\mathbb{F}}$). So its $f$-invariant can be computed directly and shown to equal zero.

To see that the Rokhlin entropy is zero, define $\phi_n : X \to \{0, 1\}$ by $\phi_n(x) = 0$ if $x(b^i) = 0$ for all $|i| \leq n$ and $\phi_n(x) = 1$ otherwise. We claim that $\phi_n$ is generating in the sense that the smallest $\mathbb{F}$-invariant sigma-algebra containing $\phi_n^{-1}(0)$ is the sigma-algebra of all Borel sets (modulo measure zero).

If $x \in X$ is Haar random then for any $g \in \Gamma$, $\{x(gb^n) : n \in \mathbb{Z}\}$ are i.i.d. random variables uniformly distributed on $\{0, 1\}$. It follows that for a.e. $x \in X$ if $g \in \Gamma$ is such that $x(g) = 0$ then there exists $m \in \mathbb{Z}$ such that $\phi_n(T^{a^m g^{-1}}x) = 0$ (i.e. $x(ga^{-m}b^i) = 0$ for all $|i| \leq n$).

Now suppose $x, y \in X$ satisfy the condition above and $x \neq y$. It suffices to show there exists $g \in \mathbb{F}$ such that $\phi_n(T^g x) \neq \phi_n(T^g y)$. Since $x \neq y$, there exists $h \in \Gamma$ such that $x(h) \neq y(h)$. Without loss of generality, $x(h) = 0$, $y(h) = 1$. By the condition above there

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

exists $m \in \mathbb{Z}$ such that $\phi_n(T^{a^m h^{-1}} x) = 0$. However $\phi_n(T^{a^m h^{-1}} y) = 1$ since $y$ is constant on $\Lambda$-cosets. Setting $g = a^m h^{-1}$ shows $\phi_n(T^g x) \neq \phi_n(T^g y)$ and so $\phi_n$ is generating.

Since the measure of $\phi_n^{-1}(0)$ tends to zero as $n \to \infty$, the Shannon entropy of $\{\phi_n^{-1}(0), \phi_n^{-1}(1)\}$ also tends to zero as $n \to \infty$. This shows $h^{\mathrm{Rok}}(\mathbb{F} \curvearrowright (X, \mu)) = 0$.

Let $\Phi : X \to (\mathbb{Z}/2)^{\mathbb{F}}$ be the map $\Phi(x)_g = x_g + x_{gb}$. This is a surjective homomorphism with kernel $N = \{\mathbf{0}, \mathbf{1}\}$. It is $\mathbb{F}$-equivariant. So it is a two-one factor map onto the Bernoulli shift with entropy $\log(2)$. □

*Question 17.* Is it true that all proper factors of the system described above have positive entropy? If so, then this action has zero sofic entropy but is 'almost' CPE.

*Question 18.* In [**Rud78**] Rudolph showed that every finite-to-one extension of a Bernoulli shift over $\mathbb{Z}$ is either Bernoulli or has a non-trivial factor map onto a finite action. By Theorem 5.1, Rudolph's theorem does not extend to non-abelian free groups. Is there some other classification of finite-to-one extensions of Bernoulli shifts over a non-abelian free group?

5.2. *Bernoulli factors of Bernoulli shifts.* This section sketches a proof of the following.

THEOREM 5.2. [**Bow17**] *Let $\Gamma$ be a non-amenable countable group. Then all Bernoulli shifts of $\Gamma$ factor onto each other.*

The first step is to prove the theorem when $\Gamma = \mathbb{F}_2 = \langle a, b \rangle$ is the rank-two free group (details are in [**Bow11b**]). Let $\Phi : (\mathbb{Z}/2)^{\mathbb{F}_2} \to (\mathbb{Z}/2 \times \mathbb{Z}/2)^{\mathbb{F}_2}$ denote the Ornstein–Weiss map. Now consider the map $\Psi : (\mathbb{Z}/2 \times \mathbb{Z}/2)^{\mathbb{F}_2} \to (\mathbb{Z}/2 \times \mathbb{Z}/2 \times \mathbb{Z}/2)^{\mathbb{F}_2}$ given by $\Psi(x, y) = (x, \Phi(y))$. By composing $\Psi$ with $\Phi$, we see that $\mathbb{F}_2 \curvearrowright (\mathbb{Z}/2)^{\mathbb{F}_2}$ factors onto $\mathbb{F}_2 \curvearrowright (\mathbb{Z}/2 \times \mathbb{Z}/2 \times \mathbb{Z}/2)^{\mathbb{F}_2}$. This construction can be iterated to show that $(\mathbb{Z}/2)^{\mathbb{F}_2}$ factors onto $((\mathbb{Z}/2)^{\mathbb{N}})^{\mathbb{F}_2}$. Since the latter has a diffuse base space, it factors onto all Bernoulli shifts over $\mathbb{F}_2$. Therefore $\mathbb{F}_2 \curvearrowright (\mathbb{Z}/2)^{\mathbb{F}_2}$ factors onto all Bernoulli shifts.

A co-induction argument using Sinai's factor theorem for $\mathbb{Z}$ shows that whenever $(K, \kappa)$ is a probability space with $H(K, \kappa) \geq \log 2$, the corresponding Bernoulli shift $\mathbb{F}_2 \curvearrowright (K, \kappa)^{\mathbb{F}_2}$ factors onto $\mathbb{F}_2 \curvearrowright (\mathbb{Z}/2, u_2)^{\mathbb{F}_2}$ and therefore factors onto all Bernoulli shifts (the factor map for the $\mathbb{F}_2$-action is constructed by applying the factor map for $\mathbb{Z}$-actions to each $\langle a \rangle$ coset). Another co-induction argument (due to Stepin [**Ste75**]) applied to Ornstein's isomorphism theorem shows that whenever two probability spaces have the same Shannon entropy then their corresponding Bernoulli shifts over $\mathbb{F}_2$ are isomorphic.

It now suffices to show that for any $\epsilon > 0$ there is a probability space $(K, \kappa)$ with Shannon entropy $< \epsilon$ such that the corresponding Bernoulli shift $\mathbb{F}_2 \curvearrowright (K, \kappa)^{\mathbb{F}_2}$ factors onto a Bernoulli shift with base entropy $\geq \log 2$. This is possible by a variant of the Ornstein–Weiss map. Consider the space $K = \mathbb{Z}/2 \cup \{*\}$ with a probability measure $\kappa_1$ satisfying $\kappa_1(\{0\}) = \kappa_1(\{1\})$. Such a measure can be found to have Shannon entropy $< \epsilon$. Let $L = \mathbb{Z}/2 \times \mathbb{Z}/2 \cup \{*\}$ with the probability measure $\lambda_1 \in \mathrm{Prob}(L)$ satisfying $\lambda_1(\{(i, j)\}) = \kappa_1(\{1\})/2$ for all $i, j \in \mathbb{Z}/2$.

The factor map $\Phi : K^{\mathbb{F}_2} \to L^{\mathbb{F}_2}$ is defined by $\Phi(x)(g) = \phi(g^{-1}x)$ where $\phi : K^{\mathbb{F}_2} \to L$ is defined by $\phi(x) = *$ if $x(1_{\mathbb{F}_2}) = *$ and otherwise

$$\phi(x) = (x(1_{\mathbb{F}_2}) + x(a^n), x(1_{\mathbb{F}_2}) + x(b^m)),$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

where $n, m \in \mathbb{N}$ are the smallest natural numbers such that $x(a^n) \neq *$, $x(b^m) \neq *$. It can be shown $\Phi$ pushes the product measure $\kappa_1^{\mathbb{F}_2}$ forward to the measure $\lambda_1^{\mathbb{F}_2}$. Therefore, entropy increases. Note

$$H(L, \lambda_1) = H(K, \kappa_1) + (1 - \kappa_1(\{*\})) \log(2) > H(K, \kappa_1).$$

If $H(L, \lambda_1) \geq \log 2$ then we are done. Otherwise, there exists an isomorphism from $\mathbb{F}_2 \curvearrowright (L, \lambda_1)^{\mathbb{F}_2}$ to a Bernoulli shift of the form $\mathbb{F}_2 \curvearrowright (K, \kappa_2)^{\mathbb{F}_2}$ where $\kappa_2$ is a probability measure satisfying $\kappa_2(\{0\}) = \kappa_2(\{1\})$. So we can find a new factor map from $\mathbb{F}_2 \curvearrowright (K, \kappa_2)^{\mathbb{F}_2}$ to $\mathbb{F}_2 \curvearrowright (L, \lambda_2)^{\mathbb{F}_2}$ where $H(L, \lambda_2) > H(K, \kappa_2)$. By composing, we obtain a factor map from the original action $\mathbb{F}_2 \curvearrowright (K, \kappa_1)^{\mathbb{F}_2}$ onto $\mathbb{F}_2 \curvearrowright (L, \lambda_2)^{\mathbb{F}_2}$. After a finite number of similar steps, $\mathbb{F}_2 \curvearrowright (K, \kappa_1)^{\mathbb{F}_2}$ factors onto a Bernoulli shift with base space entropy $\geq \log(2)$. This is because after the $i$th step the entropy of the base increases by $(1 - \kappa_i(\{*\})) \log(2)$ and $\kappa_i(\{*\})$ is decreasing to zero. This concludes the case $\Gamma = \mathbb{F}_2$.

The general case proceeds by a measurable co-induction argument. The key new ingredient is a generalization of the Gaboriau–Lyons theorem [**GL09**] to arbitrary Bernoulli shifts. That theorem states that if $\Gamma$ is an arbitrary non-amenable group then *there exists* a probability space $(K, \kappa)$ and there exists an ergodic essentially free pmp action $\mathbb{F}_2 \curvearrowright (K, \kappa)^\Gamma$ whose orbits are contained in the orbits of the Bernoulli action $\Gamma \curvearrowright (K, \kappa)^\Gamma$. The generalization proves the same statement with 'there exists a probability space $(K, \kappa)$' replaced by 'for every non-trivial probability space $(K, \kappa)$'. Details will appear in [**Bow17**]. This generalizes previous work of Ball [**Bal05**] who proved for every non-amenable group $\Gamma$ there exists some Bernoulli shift with finite base entropy that factors onto all Bernoulli shifts over $\Gamma$.

5.3. *Zero entropy extensions.* This section sketches a proof of the following.

THEOREM 5.3. [**Bow16**] *Let $\Gamma$ be a non-amenable countable group and $\Gamma \curvearrowright (X, \mu)$ a free ergodic action. Then there exists a free ergodic action $\Gamma \curvearrowright (\tilde{X}, \tilde{\mu})$ that factors onto $\Gamma \curvearrowright (X, \mu)$ and has zero Rokhlin entropy.*

*Remark 18.* Seward's generalization of Krieger's generator theorem, (Theorem 7.3) implies $\Gamma \curvearrowright (\tilde{X}, \tilde{\mu})$ admits a generating partition with two parts. This improves an earlier result of Seward [**Sew14a**] which proved, under the same hypotheses as Theorem 5.3, the existence of an extension $\Gamma \curvearrowright (\tilde{X}, \tilde{\mu})$ that admits a generating partition with at most $n$ parts where $n = n(\Gamma)$ depends only on $\Gamma$.

The next lemma is the key step. The proof given here is simpler than the one in [**Bow16**] (which was written before Theorem 5.2 was known).

LEMMA 5.4. *Let $\Gamma$ be any countable non-amenable group. There exists a pmp action $\Gamma \curvearrowright^T (Z, \zeta)$ satisfying:*

- *$T$ is an inverse limit of Bernoulli shifts;*
- *$h^{\mathrm{Rok}}(T) = 0$;*
- *$\Gamma \curvearrowright^T (Z, \zeta)$ factors onto all Bernoulli shifts over $\Gamma$.*

*In particular, if $\Gamma$ is sofic then $\Gamma \curvearrowright^T (Z, \zeta)$ is not isomorphic to a Bernoulli shift.*

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*Proof.* Let $(K_n, \kappa_n)$ be a sequence of probability spaces with Shannon entropy bounds $0 < H(K_n, \kappa_n) < 2^{-n}$. By Theorem 5.2, there exist factor maps $\Phi_i : K_i^\Gamma \to K_{i-1}^\Gamma$ for $i \geq 2$. Let $\Gamma \curvearrowright^T (Z, \zeta)$ denote the inverse limit of this system. Since $T$ factors onto all the Bernoulli shifts $\Gamma \curvearrowright (K_n, \kappa_n)^\Gamma$, Theorem 5.2 implies it factors onto all Bernoulli shifts. So it suffices to show $h^{\text{Rok}}(T) = 0$. This follows from [**Sew15b**, Corollary 3.9]. Alternatively, let $\alpha_n$ be a generating partition of $K_n^\Gamma$ with $H(\alpha_n) < 2^{-n}$. By pulling back, we may consider $\alpha_n$ as a partition of $Z$. Let $\beta_m = \bigvee_{n=m}^\infty \alpha_n$. Then $H(\beta_m) < 2^{-m+1}$ and $\beta_m$ is generating for $T$. So $h^{\text{Rok}}(T) = 0$. □

*Proof of Theorem 5.3.* By Seward's generalization of Sinai's factor theorem (Theorem 7.4), there exists a Bernoulli factor $\Gamma \curvearrowright (Y, \nu)$ of $\Gamma \curvearrowright (X, \mu)$ such that

$$h^{\text{Rok}}(\Gamma \curvearrowright (X, \mu) | \mathcal{B}_Y) = 0,$$

where $\mathcal{B}_Y$ denotes the pullback of the Borel sigma-algebra of $Y$.

Let $\Gamma \curvearrowright (Z, \zeta)$ be as in Lemma 5.4. Fix a factor map of $\Gamma \curvearrowright (Z, \zeta)$ onto $\Gamma \curvearrowright (Y, \nu)$. Let $\Gamma \curvearrowright (\tilde{X}, \tilde{\mu})$ be the independent joining of $\Gamma \curvearrowright (Z, \zeta)$ and $\Gamma \curvearrowright (X, \mu)$ over $\Gamma \curvearrowright (Y, \nu)$.

It suffices to show $h^{\text{Rok}}(\Gamma \curvearrowright (\tilde{X}, \tilde{\mu})) = 0$. To see this, let $\epsilon > 0$, $\alpha$ be a generating partition of $Z$ with $H(\alpha) < \epsilon$ and $\beta$ be a partition of $X$ with $H(\beta) < \epsilon$ such that $\sigma - \text{alg}_\Gamma(\beta \cup \mathcal{B}_Y) = \mathcal{B}_X$ (up to measure zero). By pulling back, $\alpha$ and $\beta$ may be thought of as partitions on $\tilde{X}$. Clearly $\alpha \vee \beta$ is generating for the action $\Gamma \curvearrowright (\tilde{X}, \tilde{\mu})$ and $H(\alpha \vee \beta) < 2\epsilon$. Since $\epsilon > 0$ is arbitrary, this implies the claim. □

*Remark 19.* The paper [**Bow16**] uses Theorem 5.3 to prove that the generic measure-preserving action of $\Gamma$ on a fixed probability space $(X, \mu)$ has zero Rokhlin entropy. The special case when $\Gamma$ is amenable was handled earlier by Rudolph (see the Subclaim after [**FW04**, Claim 19]).

5.4. *Finite-to-one factors.* A factor map $\phi : X \to Y$ is finite-to-one if for a.e. $y \in Y$, $\phi^{-1}(y)$ is finite. It is well known that, for $\mathbb{Z}$-actions, finite-to-one factor maps preserve entropy. This fact readily extends to actions of amenable groups. However it does not hold for non-amenable groups. The Ornstein–Weiss map is one counterexample §1.3. For another, suppose $\Gamma$ has a sofic approximation $\Sigma$ by expanders (as in §3.1.1). Then the trivial action on a two point space has entropy $-\infty$ (with respect to $\Sigma$). However the trivial action on a one-point space has entropy zero. Nonetheless there are some positive results.

THEOREM 5.5. [**Bow10c**] *If $\mathbb{F}_r$ denotes the rank $r$ free group and $\mathbb{F}_r \curvearrowright (Y, \nu)$ is an $n$-to-1 factor of $\mathbb{F}_r \curvearrowright (X, \mu)$ then*

$$f_\nu(\mathbb{F}_r \curvearrowright Y) = (r-1)\log(n) + f_\mu(\mathbb{F}_r \curvearrowright X).$$

*Proof remarks.* The proof is almost immediate from the Abramov–Rokhlin formula for the $f$-invariant (§10.1). □

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*Remark 20.* Theorem 5.5 gives new examples of entropy-increasing factor maps. For example, if $\mathbb{F}_r \curvearrowright^T (X, \mu)$ is any pmp action and $c : \mathbb{F}_r \times X \to G$ a cocycle (where $G$ is a finite group) then one can form the skew-product action $\mathbb{F}_r \curvearrowright^{S_c} X \times G$ defined by

$$S_c^g(x, h) = (T^g x, c(g, x)h).$$

Since the factor map $X \times G \to X$ is $|G|$-to-1, Theorem 5.5 implies

$$f_\mu(T) = (r - 1)\log(|G|) + f_{\mu \times u_G}(S_c).$$

If $T$ is Bernoulli and $G = \mathbb{Z}/2$ say, are there simple conditions on the cocycle $c$ such that $S_c$ is also Bernoulli? Note the Ornstein–Weiss example is of this form.

PROPOSITION 5.6. *Assuming ergodicity, Rokhlin entropy does not decrease under a finite-to-one factor map.*

*Proof.* Let $\Gamma \curvearrowright (Y, \nu)$, $\Gamma \curvearrowright (Z, \zeta)$ be ergodic pmp actions and $\pi : Y \to Z$ a finite-to-one factor map. Let $\mathcal{P}$ be a generating partition for $\Gamma \curvearrowright (Z, \zeta)$. Because $\pi$ is finite-to-one there exists a measurable partition $\{Y_1, \ldots, Y_n\}$ of $Y$ such that $\pi$ restricted to $Y_i$ is injective for all $i$. For $\epsilon > 0$, let $Y_i' \subset Y_i$ be a subset with $0 < \nu(Y_i') < \epsilon$ and let $\mathcal{Q} = \{Y_1', Y_2', \ldots, Y_n', Y \setminus \cup_i Y_i'\}$ be the coarsest partition containing every $Y_i'$. Finally, let

$$\mathcal{R} = \mathcal{Q} \vee \pi^{-1}(\mathcal{P}).$$

We claim that $\mathcal{R}$ is generating for the action $\Gamma \curvearrowright Y$. It suffices to show that for a.e. pair of distinct elements $x, y \in Y$ there exists $g \in \Gamma$ such that $gx, gy$ are in different parts of $\mathcal{R}$. So let $x, y \in Y$ be distinct. By ergodicity, we may assume $\Gamma x$ intersects $Y_i'$ for some $i$.

If $\pi(x) \neq \pi(y)$ then, because $\mathcal{P}$ is generating there exists $g \in \Gamma$ such that $g\pi(x), g\pi(y)$ are in different parts of $\mathcal{P}$ and therefore $gx, gy$ are separated by $\mathcal{R}$.

Now assume $\pi(x) = \pi(y)$. Let $g \in \Gamma$ be such that $gx \in Y_i'$. Because $\pi$ restricted to $Y_i'$ is injective and $\pi(gx) = \pi(gy)$ it follows that $gy \notin Y_i'$. So $\mathcal{R}$ separates $gx$ and $gy$ and we are done.

Because $\mathcal{R}$ is generating,

$$h^{\mathrm{Rok}}(\Gamma \curvearrowright (Y, \nu)) \leq H_\nu(\mathcal{R}) \leq H_\zeta(\mathcal{P}) + H_\nu(\mathcal{Q}).$$

The partition $\mathcal{Q}$ depends on $\epsilon > 0$ and $H_\nu(\mathcal{Q}) \searrow 0$ as $\epsilon \searrow 0$. So it follows that $h^{\mathrm{Rok}}(\Gamma \curvearrowright (Y, \nu)) \leq H_\zeta(\mathcal{P})$. Since $\mathcal{P}$ is an arbitrary generating partition for $\Gamma \curvearrowright (Z, \zeta)$, the proposition follows. □

*Remark 21.* In work-in-progress by Alpeev–Seward, the Rokhlin entropy of an action is the convex integral of the Rokhlin entropies of its ergodic components. So the previous remains true without the ergodicity assumption.

*Question 19.* Is there an upper bound for the (sofic or Rokhlin) entropy of a finite-to-one factor in terms of the entropy of the source?

PROPOSITION 5.7. *Sofic entropy does not decrease under a finite-to-one factor map. More precisely, suppose $\Gamma \curvearrowright (Y, \nu)$, $\Gamma \curvearrowright (Z, \zeta)$ are pmp actions and there is a finite-to-one factor map $\pi : Y \to Z$. Then $h_{\Sigma,\zeta}(\Gamma \curvearrowright Z) \geq h_{\Sigma,\nu}(\Gamma \curvearrowright Y)$. The variational principle (Theorem 8.1) implies a similar result for topological sofic entropy.*

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

More generally, whenever $\Gamma \curvearrowright (Y, \nu)$ is a compact extension of $\Gamma \curvearrowright (Z, \zeta)$ then $h_{\Sigma,\zeta}(\Gamma \curvearrowright Z) \geq h_{\Sigma,\nu}(\Gamma \curvearrowright Y)$. This follows from [**Hay17**, Theorem 1.1] although it might not be obvious. Hayes graciously provided the following explanation.

For each $g \in \Gamma$ and measurable function $f : Z \to \mathbb{C}$, define a new measurable function $\alpha_g(f) : Z \to \mathbb{C}$ by

$$\alpha_g(f)(h) = f(g^{-1}h).$$

Let $L^\infty(Z) \rtimes_{\text{alg}} \Gamma$ be the set of all formal sums $\sum_{g\in\Gamma} f_g u_g$ where $f_g \in L^\infty(Z)$ and $f_g = 0$ for all but finitely many $g$. The set $L^\infty(Z) \rtimes_{\text{alg}} \Gamma$ has a $*$-algebra structure: addition is defined in the obvious manner, multiplication and the $*$-operation are defined by

$$\left(\sum_g f_g u_g\right)\left(\sum_g k_g u_g\right) = \sum_g \left(\sum_h f_h \alpha_h(k_{h^{-1}g})\right) u_g$$

$$\left(\sum_g f_g u_g\right)^* = \sum_g \overline{\alpha_g(f_{g^{-1}})} u_g.$$

We need to consider representations of $L^\infty(Z) \rtimes_{\text{alg}} \Gamma$. A natural one is $\lambda_Z : L^\infty(Z) \rtimes_{\text{alg}} \Gamma \to B(L^2(Z \times \Gamma, \zeta \times \eta))$, (where $\eta$ is the counting measure and $B(\cdot)$ denotes the algebra of bounded operators) defined by

$$(\lambda_Z(f)\xi)(z, g) = f(z)\xi(z, g) \quad \text{for } f \in L^\infty(Z), \xi \in L^2(Z \times \Gamma), z \in Z, g \in \Gamma$$
$$(\lambda_Z(u_h)\xi)(z, g) = \xi(h^{-1}z, h^{-1}g) \quad \text{for } \xi \in L^2(Z \times \Gamma), z \in Z, h, g \in \Gamma.$$

Given an extension $\Gamma \curvearrowright (Y, \nu) \to \Gamma \curvearrowright (Z, \zeta)$ with factor map $\pi : Y \to Z$, there is a natural representation $\rho : L^\infty(Z) \rtimes_{\text{alg}} \Gamma \to B(L^2(Y, \nu))$ given by

$$(\rho(f)\xi)(y) = f(\pi(y))\xi(y) \quad \text{for } f \in L^\infty(Z), \xi \in L^2(Y) \tag{2}$$

$$(\rho(u_g)\xi)(y) = \xi(g^{-1}y) \quad \text{for } g \in \Gamma, \xi \in L^2(Y). \tag{3}$$

*Definition 14.* Suppose that for $i = 1, 2$, $\rho_i : L^\infty(Z) \rtimes_{\text{alg}} \Gamma \to B(\mathcal{H}_i)$ is a representation. A bounded linear map $T : \mathcal{H}_1 \to \mathcal{H}_2$ is $L^\infty(Z) \rtimes_{\text{alg}} \Gamma$-*modular* if

$$T(\rho_1(\phi)\xi) = \rho_2(\phi)T(\xi) \quad \text{for all } \phi \in L^\infty(Z) \rtimes_{\text{alg}} \Gamma, \xi \in \mathcal{H}_1.$$

The representations $\rho_1$, $\rho_2$ are *mutually singular* if every $L^\infty(Z) \rtimes_{\text{alg}} \Gamma$-modular map $T : \mathcal{H}_1 \to \mathcal{H}_2$ equals zero. An exercise shows this definition is symmetric in $\rho_1$, $\rho_2$.

We will prove that if $\Gamma \curvearrowright (Y, \nu) \to \Gamma \curvearrowright (Z, \zeta)$ is a compact extension, then $\rho$ as defined above is mutually singular with respect to $\lambda_Z$. To simplify the proof, we introduce a few definitions. Recall that if $(X, d)$ is a metric space, $A, B \subseteq X$, and $\varepsilon > 0$ then we write $A \subseteq_\varepsilon B$ if for every $a \in A$, there is a $b \in B$ with $d(a, b) < \varepsilon$.

*Definition 15.* Let $\mathcal{H}$ be a Hilbert space and $\rho : L^\infty(Z) \rtimes_{\text{alg}} \Gamma \to B(\mathcal{H})$ a $*$-representation. We say that $\rho$ is:

- *compact over* $L^\infty(Z)$ if for every $\xi \in \mathcal{H}$ and every $\varepsilon > 0$, there are $\eta_1, \ldots, \eta_k \in \mathcal{H}$ so that

$$\rho(\Gamma)\xi \subseteq_{\varepsilon,\|\cdot\|} \left\{\sum_{j=1}^k \rho(f_j)\eta_j : f_j \in L^\infty(Z), \|f_j\|_\infty \leq 1\right\}.$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

- *mixing* if for every $\xi, \eta \in \mathcal{H}$

$$\lim_{g\to\infty} \sup_{f\in L^\infty(Z),\|f\|_\infty\le 1} |\langle \rho(f)\rho(u_g)\xi, \eta\rangle| = 0.$$

The idea for each of these definitions is that we are replacing the usual complex scalars with $L^\infty(Z)$. So an element in $L^\infty(Z)$ of norm at most one should be thought of as a replacement for a complex number of size at most 1. Proposition 5.7 now follows from [**Hay17**, Theorem 1.1] and the next result.

PROPOSITION 5.8. *Let $\Gamma$ be a countable discrete group, $(Z, \zeta)$ a probability space and $\Gamma \curvearrowright (Z, \zeta)$ a measure-preserving action.*

(1) *Let $\mathcal{H}_j$, $j = 1, 2$ be Hilbert spaces and let $\rho_j : L^\infty(Z) \rtimes_{\text{alg}} \Gamma \to B(\mathcal{H}_j)$, $j = 1, 2$ be two $*$-representations. If $\rho_1$ is mixing and $\rho_2$ is compact over $L^\infty(Z)$, then $\rho_1$, $\rho_2$ are mutually singular.*

(2) *The representation $\lambda_Z$ is mixing.*

(3) *If $\Gamma \curvearrowright (Y, \nu) \to \Gamma \curvearrowright (Z, \zeta)$ is a compact extension and $\rho : L^\infty(Z) \rtimes_{\text{alg}} \Gamma \to B(L^2(Y))$ is defined by (2), (3), then $\rho$ is compact over $L^\infty(Z)$.*

*Proof.* (1) Let $T \in B(\mathcal{H}_2, \mathcal{H}_1)$ be $L^\infty(Z) \rtimes_{\text{alg}} \Gamma$-modular. This means that $T(\rho_2(\phi)\xi) = \rho_1(\phi)T(\xi)$ for all $\phi \in L^\infty(Z) \rtimes_{\text{alg}} \Gamma$, $\xi \in \mathcal{H}_2$. Let $\xi \in \mathcal{H}_2$ and $\varepsilon > 0$. Let $\eta_1, \ldots, \eta_k \in \mathcal{H}$ be as in the definition of compact over $L^\infty(Z)$. Given $g \in \Gamma$, there exist $f_{g,1}, \ldots, f_{g,k} \in L^\infty(Z)$ with $\|f_{g,j}\|_\infty \le 1$ for $j = 1, \ldots, k$ so that

$$\left\| \rho_2(u_g)\xi - \sum_{j=1}^k \rho_2(f_{g,j})\eta_j \right\| < \varepsilon.$$

We have, for any $g \in \Gamma$ :

$$\begin{aligned}
\|T(\xi)\|^2 = |\langle T(\xi), T(\xi)\rangle| &= |\langle \rho_1(u_{g^{-1}})T(\rho_2(u_g)\xi), T(\xi)\rangle| \\
&\le \varepsilon\|T\|\|T(\xi)\| + \left|\sum_{j=1}^k \langle \rho_1(u_{g^{-1}})\rho_1(f_{g,j})T(\eta_j), T(\xi)\rangle\right| \\
&= \varepsilon\|T\|\|T(\xi)\| + \left|\sum_{j=1}^k \langle \rho_1(\alpha_{g^{-1}}(f_{g,j}))\rho_1(u_{g^{-1}})T(\eta_j), T(\xi)\rangle\right| \\
&\le \varepsilon\|T\|\|T(\xi)\| + \sum_{j=1}^k \sup_{f\in L^\infty(Z):\|f\|_\infty\le 1} |\langle \rho_1(f)\rho_1(u_{g^{-1}})T(\eta_j), T(\xi)\rangle|.
\end{aligned}$$

Letting $g \to \infty$ we find that

$$\|T(\xi)\|^2 \le \varepsilon\|T\|\|T(\xi)\|.$$

Letting $\varepsilon \to 0$ proves that $T(\xi) = 0$ and, as $\xi$ was arbitrary, that $T = 0$.

(2) Given $\xi \in L^2(Z)$, $\eta \in \ell^2(\Gamma)$ we define $\xi \otimes \eta \in L^2(Z \times \Gamma)$ by $(\xi \otimes \eta)(z, g) = \xi(z)\eta(g)$. Let

$$D = \text{span}\{\xi \otimes \delta_g : \xi \in L^2(Z), g \in \Gamma\}.$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

By the fact that $\lambda_Z|_{L^\infty(Z)}$ is contractive and the density of $D$, it is enough to show that for every $\xi, \eta \in D$ we have

$$\lim_{g\to\infty} \sup_{f\in L^\infty(Z):\|f\|_\infty\le 1} |\langle \lambda_Z(f)\lambda_Z(u_g)\xi, \eta\rangle| = 0.$$

Let $\xi, \eta \in D$ and write $\xi = \sum_h \xi_h \otimes \delta_h$, $\eta = \sum_h \eta_h \otimes \delta_h$. Let $E$ be a finite subset of $\Gamma$ so that $\xi_h = 0$ and $\eta_h = 0$ if $h \in \Gamma\backslash E$. It is then straightforward to see that if $g \in \Gamma\backslash EE^{-1}$, then

$$\langle \lambda_Z(f)\lambda_Z(u_g)\xi, \eta\rangle = 0$$

for any $f \in L^\infty(Z)$. This proves (2).

(3) $L^2(Y|Z)$ is the set of $\xi \in L^2(Y)$ with $\mathbb{E}_Z(|\xi|^2) \in L^\infty(Z)$. It is a Banach space under the norm

$$\|\xi\|_{L^2(Y|Z)} = \sqrt{\|\mathbb{E}_Z(|\xi|^2)\|_\infty}.$$

See [**KL16**, Ch. 3] for more detail.

Given $f \in L^\infty(Y)$ and $\varepsilon > 0$, compactness of the extension $\Gamma \curvearrowright (Y, \nu) \to \Gamma \curvearrowright (Z, \zeta)$ implies the existence (see [**KL16**, Definition 3.8]) of vectors $\zeta_1, \dots, \zeta_k \in L^2(Y|Z)$, a subset $Y_0 \subseteq Y$ with $\nu(Y_0) > 1 - \varepsilon$ and $Z_0 \subset Z$ with $\zeta(Z_0) > 1 - \epsilon$ so that

$$\chi_{Z_0}\rho(\Gamma)(\chi_{Y_0} f) \subseteq_{\varepsilon,\|\cdot\|_{L^2(Y|Z)}} \left\{\sum_{j=1}^k \rho(k_j)\zeta_j : k_j \in L^\infty(Z),\ \|k_j\|_\infty \le 1,\ j = 1, \dots, k\right\}.$$

Since

$$\|\rho(u_g)(f - \chi_{Y_0} f)\|_2 = \|\chi_{Y\setminus Y_0} f\|_2 \le \sqrt{\varepsilon}\|f\|_\infty$$

and $\|\cdot\|_2 \le \|\cdot\|_{L^2(Y|Z)}$ we have

$$\rho(\Gamma) f \subseteq_{\varepsilon+\sqrt{\varepsilon}\|f\|_\infty,\|\cdot\|_2} \left\{\sum_{j=1}^k \rho(k_j)\zeta_j : k_j \in L^\infty(Z),\ \|k_j\|_\infty \le 1,\ j = 1, \dots, k\right\}.$$

Since $\varepsilon > 0$ is arbitrary and $L^\infty(Y)$ is dense in $L^2(Y)$, this proves part (3). □

## 6. *Combinations*

6.1. *Ergodic decomposition.* For any pmp action $\Gamma \curvearrowright (X, \mu)$ there exists a map $x \mapsto \nu_x$ from $X$ to $\mathrm{Prob}_\Gamma(X)$, the space of $\Gamma$-invariant ergodic probability measures on $X$, such that

$$\mu = \int \nu_x \, d\mu(x).$$

See [**GS00b**] for example. If $\Gamma$ is amenable then it is well known that

$$h_\mu(\Gamma \curvearrowright X) = \int h_{\nu_x}(\Gamma \curvearrowright X)\, d\mu(x).$$

This is known as the *ergodic decomposition formula* [**MO85**]. In general such a formula cannot hold for non-amenable $\Gamma$ with an arbitrary sofic approximation. For example, trivial actions can have entropy minus infinity (§3.1). Nonetheless there are some positive results for uniformly diffuse sofic approximations and for the $f$-invariant as explained next.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*Definition 16.* Let $\Sigma = \{\sigma_n\}_{n\in\mathbb{N}}$ be an arbitrary sofic approximation to $\Gamma$. Also let $\{p_n\}_{n\in\mathbb{N}}$ be a sequence of positive integers with $\lim_{n\to\infty} p_n = +\infty$. For each $n$, define $\sigma_n^{\oplus p_n} : \Gamma \to \mathrm{Sym}([d_n] \times [p_n])$ by

$$\sigma_n^{\oplus p_n}(g)(j, k) = (\sigma_n(g)j, k).$$

Then $\Sigma' := \{\sigma_n^{\oplus p_n}\}_{n\in\mathbb{N}}$ is a *uniformly diffuse* sofic approximation. This is stronger than being diffuse (which was considered in §3.1.2).

*Exercise 11.* Let $\Gamma \curvearrowright (X, \mu)$, $\Sigma$, $\Sigma'$ be as above. Then

$$h_{\Sigma',\mu}(\Gamma \curvearrowright X) = \int h_{\Sigma',\nu_x}(\Gamma \curvearrowright X)\, d\mu(x) = \int h_{\Sigma,\nu_x}(\Gamma \curvearrowright X)\, d\mu(x).$$

In §6.2.5 below, a variant of sofic entropy, called average-local sofic entropy is introduced. It is almost immediate that it satisfies the ergodic decomposition formula. It also agrees with $h_{\Sigma',\mu}(\Gamma \curvearrowright X)$, which proves the exercise.

The next result shows that the $f$-invariant satisfies the ergodic decomposition formula with a correction term.

THEOREM 6.1. [**Sew16a**, Theorem 1.4] *Let $\mathbb{F} \curvearrowright (X, \mu)$ be a pmp action of a rank $r$ free group. Assume $\mathbb{F} \curvearrowright (X, \mu)$ admits a finite-entropy generating partition. Then $\mathbb{F} \curvearrowright (X, \nu_x)$ admits a finite-entropy generating partition for $\mu$-a.e. $x$ and*

$$f_\mu(\mathbb{F} \curvearrowright X) = \int f_{\nu_x}(\mathbb{F} \curvearrowright X)\, d\mu(x) - (r-1)H(\tau),$$

*where $\tau \in \mathrm{Prob}(\mathrm{Prob}(X))$ is the law of $\nu_x$ (so $\tau = \int \delta_{\nu_x}\, d\mu(x)$).*

*Proof sketch.* The statement is directly verified for Markov chains. The general case can be obtained from approximating by Markov chains. □

In work-in-progress, Alpeev and Seward have proven that Rokhlin entropy satisfies the ergodic decomposition formula.

6.2. *Direct products.* It is well known that if $\Gamma$ is amenable then topological and measure entropy are additive under direct products:

$$h_{\mathrm{top}}(T \times S) = h_{\mathrm{top}}(T) + h_{\mathrm{top}}(S),$$
$$h_{\mu\times\nu}(T \times S) = h_\mu(T) + h_\nu(S),$$

where $\Gamma \curvearrowright^T (X, \mu)$, $\Gamma \curvearrowright^S (Y, \nu)$ and $T \times S$ is the action on $X \times Y$ defined by

$$(T \times S)^g(x, y) = (T^g x, S^g y) \quad \text{for all } g \in \Gamma,\ x \in X,\ y \in Y.$$

Here is a brief summary of this section: topological sofic entropy is also additive under direct products. However, measure sofic entropy is only subadditive and there are explicit counterexamples to additivity. The $f$-invariant is additive under the restriction that all actions involved have finite generating partitions. However, it is unknown whether Rokhlin entropy is additive. There are several variants of sofic entropy introduced by Austin, one of which is additive under direct products. In fact, these variants 'explain' how additivity can fail.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

6.2.1. *Direct products and topological entropy.*

THEOREM 6.2. (Topological entropy product formula) *Let* $\Gamma$ *be a sofic group with a sofic approximation* $\Sigma$ *and fix a non-principal ultrafilter* $\mathcal{U}$ *on* $\mathbb{N}$. *For continuous actions* $\Gamma \curvearrowright X$, $\Gamma \curvearrowright Y$ *on compact metrizable spaces define* $(\Sigma, \mathcal{U})$*-entropy by*

$$h_{\Sigma,\mathcal{U}}(\Gamma \curvearrowright X) = \sup_{\epsilon>0} \inf_{F\in\Gamma} \inf_{\delta>0} \lim_{i\to\mathcal{U}} |V_i|^{-1} \log(N_\epsilon(\mathrm{Map}(T, \rho, F, \delta, \sigma_i), \rho_\infty))$$

*for any continuous generating pseudometric* $\rho$ *on* $X$. *That is, the definition of sofic entropy is modified by replacing* $\limsup_{i\to\infty}$ *with the ultralimit along* $\mathcal{U}$. *Then*

$$h_{\Sigma,\mathcal{U}}(\Gamma \curvearrowright X \times Y) = h_{\Sigma,\mathcal{U}}(\Gamma \curvearrowright X) + h_{\Sigma,\mathcal{U}}(\Gamma \curvearrowright Y)$$

*where, by convention,* $+\infty + (-\infty) = -\infty$.

The proof is a straightforward exercise. The crux of the argument is that if $\phi : V \to X$, $\psi : V \to Y$ are good microstates then $\phi \times \psi : V \to X \times Y$ is also a good microstate.

6.2.2. *Subadditivity.*

THEOREM 6.3. (Subadditive product formula) *Let* $\Gamma$ *be a sofic group with a sofic approximation* $\Sigma$ *and pmp actions* $\Gamma \curvearrowright (X, \mu)$, $\Gamma \curvearrowright (Y, \nu)$. *Then*

$$h_{\Sigma,\mu\times\nu}(\Gamma \curvearrowright X \times Y) \le h_{\Sigma,\mu}(\Gamma \curvearrowright X) + h_{\Sigma,\nu}(\Gamma \curvearrowright Y).$$

*If one of these actions is a Bernoulli shift then equality holds.*

The proof is a direct exercise. The key observation is that if $\psi : V \to X \times Y$ is a good microstate for $\mu \times \nu$ then the projections $\psi_X : V \to X$ and $\psi_Y : V \to Y$ are good microstates for the marginals. The special case of actions with finite-entropy generating partitions was handled in [**Bow10b**]. The general case can be proven similarly using the partition definition of sofic entropy (§2.4.2).

6.2.3. *Direct products and the* $f$*-invariant.* It is a brief exercise to show that the $f$-invariant is additive under direct products in the following sense: if $\mathbb{F}_r \curvearrowright^T (X, \mu)$, $\mathbb{F}_r \curvearrowright^S (Y, \nu)$ are both pmp actions with finite-entropy generating partitions then

$$f_{\mu\times\nu}(T \times S) = f_\mu(T) + f_\nu(S).$$

*Question 20.* Does this formula extend to actions that do not have finite-entropy generating partitions? In this case the $f$-invariant is defined via a random sofic approximation as in §2.5.2. Examples below show there is reason to be cautious.

In the next two examples, we consider sofic entropy with respect to the random sofic approximation $\mathbb{P}$ defined in §2.5.2. Recall that $\mathbb{P}$-entropy extends the $f$-invariant to actions that need not admit finite generating partitions.

*Example 5.* (The infinite entropy Bernoulli shift and the trivial action) The $\mathbb{P}$-entropy of the Bernoulli shift $\mathbb{F}_r \curvearrowright ([0, 1], \mathrm{Leb})^{\mathbb{F}_r}$ is $+\infty$. The $\mathbb{P}$-entropy of the trivial action of $\mathbb{F}_r$ on $([0, 1], \mathrm{Leb})$ is $-\infty$ (for example this follows from Theorem 5.5). The direct product of these two actions has $\mathbb{P}$-entropy $-\infty$. The reason is that the number of $\epsilon$-separated

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

microstates $\phi : [n] \to [0, 1]^{\mathbb{F}_r} \times [0, 1]$ for this action is roughly bounded by $\exp(n/\epsilon)$. On the other hand, with high probability a random sofic approximation $\sigma : \mathbb{F}_r \to \mathrm{Sym}(n)$ admits no microstates at all.

*Example 6.* (Countable direct products) The $\mathbb{P}$-entropy is not additive under countable direct products. To see this recall that the trivial action $\mathbb{F}_2 \curvearrowright (\mathbb{Z}/2, u_2)$ has $f$-invariant $-\log(2)$ while the Bernoulli shift $\mathbb{F}_2 \curvearrowright (\mathbb{Z}/2, u_2)^{\mathbb{F}_2}$ has $f$-invariant $\log(2)$. So the direct product action

$$\mathbb{F}_2 \curvearrowright (\mathbb{Z}/2 \times \mathbb{Z}/2^{\mathbb{F}_2}, u_2 \times u_2^{\mathbb{F}_2})$$

has $f$-invariant 0. The infinite direct power of this action is measurably conjugate to the direct product of the trivial action $\mathbb{F}_r \curvearrowright ([0, 1], \mathrm{Leb})$ with the Bernoulli shift $\mathbb{F}_r \curvearrowright ([0, 1], \mathrm{Leb})^{\mathbb{F}_r}$. By the previous example, the $\mathbb{P}$-entropy of this action is $-\infty$.

6.2.4. *A counterexample to additivity.*

THEOREM 6.4. *There exists a sofic group* $\Gamma$ *with a sofic approximation* $\Lambda$ *and a pmp action* $\Gamma \curvearrowright^T (X, \mu)$ *such that*

$$h_{\Lambda, \mu \times \mu}(T \times T) = -\infty \neq 0 = 2h_{\Lambda, \mu}(T).$$

*Proof.* Let $\mathbb{F}_2 = \langle a, b \rangle$ be a rank-two free group and $\Sigma = \{\sigma_i : \mathbb{F}_2 \to \mathrm{Sym}(V_i)\}$ be a sofic approximation to $\mathbb{F}_2$ by expanders as in §3.1.1. For each $i$, define $\sigma_i^{\oplus 2} : \mathbb{F}_2 \to \mathrm{Sym}(V_i \times \{0, 1\})$ by

$$\sigma_i^{\oplus 2}(g)(v, i) = (\sigma(g)v, i)$$

(as in Definition 16) and let $\Lambda = \Sigma^{(2)} = \{\sigma_i^{\oplus 2}\}$.

Let $T$ be the trivial action of $\mathbb{F}_2$ on $\{0, 1\}$ (so $T^g x = x$ for all $g \in \mathbb{F}_2$, $x \in \{0, 1\}$). By Proposition 3.1, $h_{\Sigma, \mu}(T) = -\infty$ where $\mu$ is the uniform probability measure. An exercise shows that $h_{\Sigma^{(2)}, \mu}(T) = 0$. An argument similar to Proposition 3.1 shows $h_{\Sigma^{(2)}, \mu \times \mu}(T \times T) = -\infty$. □

6.2.5. *Variants of sofic entropy and direct products.* In [**Aus16a**] Austin introduced two variants of sofic entropy, one of which is additive under direct products. Both variants replace the number of microstates with covering numbers of measures on model spaces. The idea to use measures on model spaces goes back to [**Bow11a**] where yet another variant of sofic entropy was introduced.

The starting point is to assume our action has the form $\Gamma \curvearrowright (\mathcal{X}^\Gamma, \mu)$ where $(\mathcal{X}, d_\mathcal{X})$ is a compact metric space and $\mu$ is a shift-invariant measure. This special form does not lose generality: any pmp action has a topological model and therefore we can assume it has the form $\Gamma \curvearrowright (\mathcal{X}, \nu)$ where $(\mathcal{X}, d_\mathcal{X})$ is as above. We can equivariantly embed $\mathcal{X}$ into $\mathcal{X}^\Gamma$ via $x \mapsto (g^{-1}x)_{g \in \Gamma}$. This pushes $\nu$ forward to a shift-invariant measure on $\mathcal{X}^\Gamma$.

Let $V$ be a finite set. A point in the product space $\mathcal{X}^V$ is represented as $\mathbf{x} = (x_v)_{v \in V}$. This space is endowed with the normalized Hamming metric

$$d_\mathcal{X}^{(V)}(\mathbf{x}, \mathbf{y}) = |V|^{-1} \sum_{v \in V} d_\mathcal{X}(x_v, y_v).$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

Given $\mathbf{x} \in \mathcal{X}^V$ and a map $\sigma : \Gamma \to \mathrm{Sym}(V)$, the *pullback name of* $\mathbf{x}$ *at* $v$ is

$$\Pi_v^\sigma(\mathbf{x}) := (x_{\sigma(g)^{-1}v})_{g\in\Gamma} \in \mathcal{X}^\Gamma.$$

This defines a map $\Pi_v^\sigma : \mathcal{X}^V \to \mathcal{X}^\Gamma$. The *empirical distribution* of $\mathbf{x}$ is

$$P_{\mathbf{x}}^\sigma = |V|^{-1} \sum_{v\in V} \delta_{\Pi_v^\sigma(\mathbf{x})} \in \mathrm{Prob}(\mathcal{X}^\Gamma).$$

If $\sigma$ is a homomorphism then $P_{\mathbf{x}}^\sigma$ is shift-invariant. In general if $\mathcal{O} \subset \mathrm{Prob}(\mathcal{X}^\Gamma)$ is any weak* open neighborhood of the subspace $\mathrm{Prob}_\Gamma(\mathcal{X}^\Gamma)$ of shift-invariant probability measures, $\Sigma = \{\sigma_n\}$ is a sofic approximation and $\mathbf{x}_n \in \mathcal{X}^{V_n}$ then $P_{\mathbf{x}_n}^{\sigma_n} \in \mathcal{O}$ for all sufficiently large $n$.

If $\mathcal{O}$ is any weak* open neighborhood of $\mu$ in $\mathrm{Prob}(\mathcal{X}^\Gamma)$, then we let $\Omega(\mathcal{O}, \sigma_n)$ denote the set of all microstates $\mathbf{x} \in \mathcal{X}^{V_n}$ such that $P_{\mathbf{x}}^{\sigma_n} \in \mathcal{O}$. The metric space $(\Omega(\mathcal{O}, \sigma_n), d_{\mathcal{X}}^{(V_n)})$ is a *model space*. It plays a role here that is very similar to the role played by $\mathrm{Map}(\cdots)$ in §2.4.1.

In order to avoid taking limsups and liminfs, we will work with ultralimits. So let $\mathcal{U}$ be a non-principal ultrafilter on $\mathbb{N}$. A sequence $\{\mu_n\}$ of probability measures on $\mathcal{X}^{V_n}$ is said to converge to $\mu$ with respect to $(\Sigma, \mathcal{U})$:

- *locally on average* if $\lim_{n\to\mathcal{U}} \int P_{\mathbf{x}}^{\sigma_n}\, d\mu_n(\mathbf{x}) = \mu$;
- *locally* if for every weak* open neighborhood $\mathcal{O}$ of $\mu$ it is the case that

$$\lim_{n\to\mathcal{U}} |V_n|^{-1} \#\{v \in V_n : (\Pi_v^{\sigma_n})_* \mu_n \in \mathcal{O}\} = 1;$$

- *locally and empirically* if it converges locally and for every weak* open neighborhood $\mathcal{O}$ of $\mu$

$$\lim_{n\to\mathcal{U}} \mu_n(\Omega(\mathcal{O}, \sigma_n)) = 1;$$

- *locally and doubly empirically* if $\mu_n \times \mu_n$ locally and empirically converges to $\mu \times \mu$.

The notions above are listed in order of increasing strength. In [**Aus16a**], local and empirical convergence is called quenched convergence. The new terminology has been chosen to avoid a conflict with common statistical physics language.

Every notion of convergence above corresponds to a variant of sofic entropy in which the number of microstates in the usual formula for entropy is replaced with a covering number. To explain, suppose $(\mathcal{Y}, d_{\mathcal{Y}})$ is a metric space and $\delta > 0$. Then the *δ-covering number* of $(Y, d_{\mathcal{Y}})$, denoted $\mathrm{cov}_\delta(\mathcal{Y}, d_{\mathcal{Y}})$, is the minimum cardinality of a subset $S \subset \mathcal{Y}$ whose $\delta$-neighborhood is all of $\mathcal{Y}$. If $\mu$ is a probability measure on $\mathcal{Y}$ then $\mathrm{cov}_{\epsilon,\delta}(\mu, d_{\mathcal{Y}})$ is the infimum of $\mathrm{cov}_\delta(\mathcal{Z}, d_{\mathcal{Y}} \restriction \mathcal{Z})$ over all subsets $\mathcal{Z} \subset \mathcal{Y}$ with measure $\mu(\mathcal{Z}) \geq 1 - \epsilon$.

The *local-on-average* sofic entropy of $\Gamma \curvearrowright (\mathcal{X}^\Gamma, \mu)$ (with respect to $(\Sigma, \mathcal{U})$) is defined by

$$h_{\Sigma,\mathcal{U}}^{\text{loc-avg}}(\Gamma \curvearrowright (\mathcal{X}^\Gamma, \mu)) = \sup\Big\{ \sup_{\epsilon,\delta>0} \lim_{n\to\mathcal{U}} |V_n|^{-1} \log \mathrm{cov}_{\epsilon,\delta}(\mu_n, d_{\mathcal{X}}^{(V_n)}) \Big\},$$

where the first supremum is over all sequences $\{\mu_n\}$ that locally-on-average converge to $\mu$ with respect to $(\Sigma, \mathcal{U})$. The definitions of *local sofic entropy, local and empirical sofic entropy* and *local and doubly empirical sofic entropy* (denoted $h_{\Sigma,\mathcal{U}}^{\text{loc}}(\cdot)$, $h_{\Sigma,\mathcal{U}}^{\text{le}}(\cdot)$, $h_{\Sigma,\mathcal{U}}^{\text{lde}}(\cdot)$) are similar. Local and empirical is abbreviated to le- and local and doubly empirical to lde-.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

PROPOSITION 6.5. *The four notions of entropy defined above are measure-conjugacy invariants.*

*Remarks on the proof.* The fact that le- and lde-sofic entropy are measure-conjugacy invariants is proven in [**Aus16a**]. The proof there generalizes to local-on-average and local sofic entropy. The main idea is that any factor map gives rise to a sequence of 'almost Lipschitz' maps between model spaces. These maps essentially push-forward a sequence $\{\mu_n\}$ converging to $\mu$ to a new sequence $\{\nu_n\}$ converging to $\nu$ in such a way that the type of convergence is preserved. □

Any sequence $\{\mu_n\}$ of probability measures $\mu_n$ on $\Omega(\mathcal{O}_n, \sigma_n)$ converges locally-on-average to $\mu$ whenever $\{\mathcal{O}_n\}$ is a sequence of weak* open sets that decrease to $\mu$. Therefore sofic entropy lower bounds $h^{\text{loc-avg}}_{\Sigma,\mathcal{U}}(\cdot)$. On the other hand, the increasing strength of the notions of convergence imply

$$h^{\text{loc-avg}}_{\Sigma,\mathcal{U}}(\cdot) \geq h^{\text{loc}}_{\Sigma,\mathcal{U}}(\cdot) \geq h^{\text{le}}_{\Sigma,\mathcal{U}}(\cdot) \geq h^{\text{lde}}_{\Sigma,\mathcal{U}}(\cdot).$$

PROPOSITION 6.6. *Let $\Gamma \curvearrowright^T (X, \mu)$ be a pmp action.*

- *If $T$ is ergodic then $h_{\Sigma,\mathcal{U},\mu}(T) = h^{\text{loc-avg}}_{\Sigma,\mathcal{U},\mu}(T)$.*
- *If $T$ is ergodic then any sequence $\{\mu_n\}$ that locally converges to $\mu$ must also le-converge to $\mu$. So $h^{\text{loc}}_{\Sigma,\mathcal{U},\mu}(T) = h^{le}_{\Sigma,\mathcal{U},\mu}(T)$.*
- *If $T$ is weakly mixing then any sequence $\{\mu_n\}$ that locally converges to $\mu$ must also lde-converge to $\mu$. So $h^{\text{loc}}_{\Sigma,\mathcal{U},\mu}(T) = h^{le}_{\Sigma,\mathcal{U},\mu}(T) = h^{lde}_{\Sigma,\mathcal{U},\mu}(T)$.*

*Proof.* The first statement is proven in [**Bow11a**] (under unnecessarily restrictive simplifying hypotheses). In fact the proof shows that if $\mu_n$ locally-on-average converges to $\mu$ then for any weak* open neighborhood $\mathcal{O}$ of $\mu$, $\mu_n(\Omega(\mathcal{O}, \sigma_n)) \to 1$ as $n \to \mathcal{U}$. This also explains the second statement. The third statement follows from the second since weak mixing implies that $\mu \times \mu$ is ergodic and it can be shown directly that if $\mu_n \to \mu$ locally then $\mu_n \times \mu_n \to \mu \times \mu$ locally. The second and third statements are proven in [**Aus16a**]. □

THEOREM 6.7. [**Aus16a**, Theorems B and C] *Let $\Gamma \curvearrowright^T (\mathcal{X}^\Gamma, \mu)$, $\Gamma \curvearrowright^S (\mathcal{Y}^\Gamma, \nu)$ be pmp actions. As above, let $\Sigma$ be a sofic approximation to $\Gamma$ and $\mathcal{U}$ a non-principal ultrafilter on $\mathbb{N}$. Then*

$$h_{\Sigma,\mathcal{U},\mu\times\nu}(T \times S) \geq h^{\text{lde}}_{\Sigma,\mathcal{U},\mu}(T) + h_{\Sigma,\mathcal{U},\nu}(S),$$
$$h^{\text{lde}}_{\Sigma,\mathcal{U},\mu\times\nu}(T \times S) = h^{\text{lde}}_{\Sigma,\mathcal{U},\mu}(T) + h^{\text{lde}}_{\Sigma,\mathcal{U},\nu}(S).$$

*Remarks on the proof.* The statements of Theorems B and C in [**Aus16a**] differ from the above. Instead of using ultrafilters one quantifies over all subsequences. The proof of the version above can be derived from the proof in [**Aus16a**] with only minimal changes. The main idea is that if $\{\mu_n\}$ is a sequence of measures on model spaces that locally and doubly empirically converges to $\mu$ then for any weak* open neighborhood $\mathcal{N}$ of $\mu \times \nu$ in $\text{Prob}(\mathcal{X}^\Gamma \times \mathcal{Y}^\Gamma)$ there is a weak* open neighborhood $\mathcal{O}$ of $\nu$ in $\text{Prob}(\mathcal{Y}^\Gamma)$ such that

$$\inf_{y \in \Omega(\mathcal{O},\sigma_n)} \mu_n(\{\mathbf{x} \in \mathcal{X}^{V_n} : (\mathbf{x}, \mathbf{y}) \in \Omega(\mathcal{N}, \sigma_n)\}) \to 1$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

as $n \to \mathcal{U}$. This is reminiscent of the well-known fact that if $\mu \times \mu$ is ergodic and $\nu$ is ergodic then $\mu \times \nu$ is ergodic. □

Finally, in order to further justify the notion of lde-entropy we have the following.

THEOREM 6.8. [**Aus16a**, Theorem D] *Given a pmp action* $\Gamma \curvearrowright^T (X, \mu)$*, sofic approximation* $\Sigma$ *and a non-principal ultrafilter* $\mathcal{U}$ *on* $\mathbb{N}$*, the power-stabilized* $(\Sigma, \mathcal{U})$*-entropy is defined by*

$$h^{\mathrm{ps}}_{\Sigma,\mathcal{U},\mu}(T) = \lim_{n\to\infty} \frac{1}{n} h_{\Sigma,\mathcal{U},\mu^{\times n}}(T^{\times n}),$$

*where* $\Gamma \curvearrowright^{T^{\times n}} (X, \mu)^{\times n}$ *is the n-fold Cartesian power of T. The limit exists by subadditivity. Then*

$$h^{\mathrm{ps}}_{\Sigma,\mathcal{U},\mu}(T) \geq h^{\mathrm{lde}}_{\Sigma,\mathcal{U},\mu}(T)$$

*and equality holds if* $\Gamma \curvearrowright^T (X, \mu)$ *admits a finite-entropy generating partition.*

One further justification is as follows.

THEOREM 6.9. [**Aus16a**, Corollary D′] *Given a pmp action* $\Gamma \curvearrowright^T (\mathcal{X}^\Gamma, \mu)$ *with a finite-entropy generating partition, sofic approximation* $\Sigma$ *and a non-principal ultrafilter* $\mathcal{U}$ *on* $\mathbb{N}$*,*

$$h^{\mathrm{lde}}_{\Sigma,\mathcal{U},\mu}(T) = \inf_S h_{\Sigma,\mathcal{U},\mu\times\nu}(T \times S) - h_{\Sigma,\mathcal{U},\nu}(S),$$

*where the infimum is over all pmp actions* $\Gamma \curvearrowright^S (\mathcal{Y}^\Gamma, \nu)$.

## 7. *Ornstein theory*

In 1970, Ornstein introduced a powerful set of tools, known collectively as the 'Ornstein machine', for proving that a given automorphism is measurably conjugate to a Bernoulli shift [**Orn70a**, **Orn70c**, **Orn70b**]. This machine also unifies the proofs of the following major results.

- (Sinai's factor theorem): every ergodic automorphism $T$ with positive entropy factors onto every Bernoulli shift $\mathbb{Z} \curvearrowright (K, \kappa)^{\mathbb{Z}}$ with $H(K, \kappa) \leq h_\mu(T)$ [**Sin64**];
- (Krieger's generator theorem): every ergodic automorphism $T$ admits a generating partition $\mathcal{P}$ with $|\mathcal{P}| < 1 + \exp(h_\mu(T))$ [**Kri70**];
- (Ornstein's isomorphism theorem): Bernoulli shifts are classified up to measure-conjugacy by entropy.

See [**Gla03**, **Dow11**, **Pet89**, **Rud90**] for modern treatments.

These results were generalized by Ornstein and Weiss [**OW80**] to arbitrary countable amenable groups via quasi-tiling machinery. Alternatively, this generalization can be made via orbit-equivalence theory [**Dan01**, **DP02**].

In recent work, all three major results have been partially generalized to all countable groups. These generalizations are discussed next, followed by a section on the $\bar{d}$-metric that plays a crucial role in Ornstein theory.

### 7.1. *The isomorphism theorem.*

Ornstein's isomorphism theorem has recently been extended to all countable groups. The final piece was put in by Seward in work that is still in progress.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

THEOREM 7.1. *Let $\Gamma$ be a countably infinite group. Let $(K, \kappa)$, $(L, \lambda)$ be two probability spaces with the same Shannon entropy. Then the corresponding Bernoulli shifts $\Gamma \curvearrowright (K, \kappa)^\Gamma$, $\Gamma \curvearrowright (L, \lambda)^\Gamma$ are measurably conjugate.*

*Proof sketch.* Following Stepin, we say that a group $\Gamma$ is *Ornstein* if whenever $(K, \kappa)$, $(L, \lambda)$ are any two probability spaces with the same Shannon entropy then the Bernoulli shifts $\Gamma \curvearrowright (K, \kappa)^\Gamma$, $\Gamma \curvearrowright (L, \lambda)^\Gamma$ are measurably conjugate. So our goal is to prove that all countably infinite groups are Ornstein.

Stepin showed that if $\Gamma$ contains an Ornstein subgroup then $\Gamma$ must be Ornstein itself [**Ste75**]. This is because if $H \leq \Gamma$ is an Ornstein subgroup then an isomorphism from $\Gamma \curvearrowright (K, \kappa)^\Gamma$ to $\Gamma \curvearrowright (L, \lambda)^\Gamma$ can be built out of an isomorphism from $H \curvearrowright (K, \kappa)^H$ to $H \curvearrowright (L, \lambda)^H$ coset-by-coset. In other words, it is a co-induction argument (the details are spelled out in [**Bow11b**]). By Ornstein–Weiss [**OW80**] all infinite amenable groups are Ornstein. So any group that contains an infinite amenable subgroup must be Ornstein.

On the other hand, Ol'shankii proved the existence of countable non-amenable groups that contain no proper infinite subgroups [**Ol′91**]. Stepin's trick cannot be directly applied to such groups. Nonetheless there is a measurable version of Stepin's trick that works and is applied in [**Bow12a**].

Suppose there is a non-trivial probability space $(M, \mu)$ such that both $(K, \kappa)$ and $(L, \lambda)$ factor onto $(M, \mu)$. Non-trivial means that $\mu(\{x\}) < 1$ for every $x \in M$. Then the Bernoulli shifts $\Gamma \curvearrowright (K, \kappa)^\Gamma$ and $\Gamma \curvearrowright (L, \lambda)^\Gamma$ both factor onto $\Gamma \curvearrowright (M, \mu)^\Gamma$. By Zimmer [**Zim84**], there exists an ergodic automorphism $T$ of $(M, \mu)^\Gamma$ whose orbits are contained in the $\Gamma$-orbits. We lift $T$ to automorphisms $\tilde{T}_K$, $\tilde{T}_L$ of $(K, \kappa)^\Gamma$, $(L, \lambda)^\Gamma$ respectively. Using Thouvenot's relative version of Ornstein's isomorphism theorem [**Tho75**] we see that $\tilde{T}_K$ and $\tilde{T}_L$ are isomorphic via an isomorphism compatible with $T$. This isomorphism is used in a manner similar to Stepin's trick to build an isomorphism from $\Gamma \curvearrowright (K, \kappa)^\Gamma$ to $\Gamma \curvearrowright (L, \lambda)^\Gamma$ (see [**Bow12a**] for details).

Next suppose that neither $(K, \kappa)$ nor $(L, \lambda)$ is a two-atom space. For example, this means that for any $k_0, k_1 \in K$, $\kappa(\{k_0, k_1\}) < 1$. Then an elementary argument shows the existence of a third probability space $(N, \nu)$ such that

$$H(K, \kappa) = H(L, \lambda) = H(N, \nu)$$

and $(K, \kappa)$ has a non-trivial common factor with $(N, \nu)$, which has a non-trivial common factor with $(L, \lambda)$. So the previous result shows that $\Gamma \curvearrowright (K, \kappa)^\Gamma$ is isomorphic to $\Gamma \curvearrowright (L, \lambda)^\Gamma$.

The final piece to the puzzle is to handle the case when $K$ is a two-atom space. This is handled in work-in-progress by Seward. The main idea is to find a common factor between $\Gamma \curvearrowright (K, \kappa)^\Gamma$ and $\Gamma \curvearrowright (L, \lambda)^\Gamma$ for a specific choice of $(L, \lambda)$ and then apply arguments similar to the above. □

By contrast, it is trivial to check that no finite group is Ornstein. So among countable groups, the Ornstein property characterizes the infinite groups.

7.1.1. *Non-Bernoulli factors of Bernoulli shifts.* In the special case $\Gamma = \mathbb{Z}$, many natural actions are known to be isomorphic to Bernoulli shifts. This includes mixing

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

Markov chains, inverse limits of Bernoulli shifts, factors of Bernoulli shifts, algebraic actions with completely positive entropy, hyperbolic toral automorphisms, the time-one map of geodesic flow on a negatively curved surface of finite volume. All of these results were obtained using Ornstein theory. By contrast, when $\Gamma = \mathbb{F}_2$ we gave in §3.3 Example 3 an example of a mixing Markov chain that is not Bernoulli and in §5.4 an example of an inverse limit of Bernoulli shifts that is not Bernoulli. Next we give an example, due to Popa–Sasyk of a factor of a Bernoulli shift that is non-Bernoulli. Moreover, this factor is algebraic and has completely positive sofic entropy by the main result of [**Ker14**].

THEOREM 7.2. [**PS07**, **Pop06a**, **Aus16c**] *Let* $\Gamma$ *be a countably infinite group with an infinite normal subgroup* $H$ *such that* $(\Gamma, H)$ *has relative property (T). Then* $\Gamma$ *admits a Bernoulli action with a non-trivial factor that is non-Bernoulli. In fact, the factor action is not even orbit-equivalent to a Bernoulli shift.*

*Proof outline.* The idea is to compute the cohomology of the actions taking values in the circle $\mathbb{R}/\mathbb{Z}$. To be precise, let $(X, \mu)$ be a standard measure space with an action $\Gamma \curvearrowright X$. A *1-cocycle* is a map $c : \Gamma \times X \to \mathbb{R}/\mathbb{Z}$ satisfying

$$c(g_2 g_1, x) = c(g_2, g_1 x) + c(g_1, x)$$

for $g_1, g_2 \in \Gamma$ and a.e. $x \in X$. A cocycle $c$ is a *coboundary* if there is a map $\phi : X \to \mathbb{R}/\mathbb{Z}$ such that

$$c(g, x) = \phi(gx) - \phi(x).$$

Two cocycles are *cohomologous* if their difference is a coboundary.

The first cohomology group is $H^1(\Gamma \curvearrowright (X, \mu)) = Z(\Gamma \curvearrowright (X, \mu))/B(\Gamma \curvearrowright (X, \mu))$ where $Z(\cdot)$ is the additive group of 1-cocycles and $B(\cdot)$ is the additive group of coboundaries.

In [**PS07**], it is shown that for any Bernoulli shift over $\Gamma$, every cocycle is cohomologous to a homomorphism $\Gamma \to \mathbb{R}/\mathbb{Z}$. Therefore the cohomology group is isomorphic to $\mathrm{Hom}(\Gamma, \mathbb{R}/\mathbb{Z})$. This has been vastly generalized by Popa's cocycle superrigidity theorem [**Pop07**].

Now let $K$ be a compact abelian group. $K$ is embedded into $K^\Gamma$ by $k \mapsto$ the constant function $(g \mapsto k)$. This is a closed $\Gamma$-invariant subgroup. So $\Gamma \curvearrowright K^\Gamma/K$ is an algebraic action (called an Ornstein–Weiss factor in [**GS15**] where bounds on its entropy are obtained). In [**Pop06a**], it is shown that the cohomology group of $\Gamma \curvearrowright K^\Gamma/K$ (with respect to Haar measure on $K^\Gamma/K$) is isomorphic to $\mathrm{Hom}(\Gamma, \mathbb{R}/\mathbb{Z}) \times \mathrm{Hom}(K, \mathbb{R}/\mathbb{Z})$. So if $\mathrm{Hom}(\Gamma, \mathbb{R}/\mathbb{Z}) \times \mathrm{Hom}(K, \mathbb{R}/\mathbb{Z})$ is not isomorphic to $\mathrm{Hom}(\Gamma, \mathbb{R}/\mathbb{Z})$, then this action cannot be orbit-equivalent to a Bernoulli shift, let alone measurably conjugate to one.

The group $\mathrm{Hom}(\Gamma, \mathbb{R}/\mathbb{Z})$ is compact while $\mathrm{Hom}(K, \mathbb{R}/\mathbb{Z})$ is countable. As $K$ varies over all compact abelian groups, $\mathrm{Hom}(K, \mathbb{R}/\mathbb{Z})$ varies over all countable abelian groups. Since there are uncountably many countable abelian groups, there are uncountably factors of Bernoulli shifts over $\Gamma$ that are not Bernoulli.

The proof in [**Aus16c**] differs from the above. It assumes $\Gamma$ is sofic and shows that the model spaces of Bernoulli shifts have an asymptotic connectivity property that the model spaces of $\Gamma \curvearrowright K^\Gamma/K$ lacks. □

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*Question 21.* Does there exist a non-Bernoulli factor of a Bernoulli shift over the free group?

*Question 22.* If a factor of a Bernoulli shift is orbit-equivalent to a Bernoulli shift, must it be Bernoulli? This is true if $\Gamma$ is an ICC property (T) group by Popa's cocycle superrigidity theorem [**Pop07**].

*Question 23.* Let $K$ be a finite set and $\mathrm{Prob}_{\mathbb{Z}}(K^{\mathbb{Z}})$ denote the space of shift-invariant Borel probability measures on $K^{\mathbb{Z}}$ with the weak* topology. Let $0 < c < \log|K|$ and consider the subset $X_c \subset \mathrm{Prob}_{\mathbb{Z}}(K^{\mathbb{Z}})$ of all measures with entropy rate $\geq c$. By upper semicontinuity, $X_c$ is closed. In unpublished work, Rudolph proved the subset $X_c' \subset X_c$ of all measures $\mu \in X_c$ such that the shift action $\mathbb{Z} \curvearrowright (K^{\mathbb{Z}}, \mu)$ is Bernoulli and has entropy $= c$ is a dense $G_\delta$ subset of $X_c$. Density is a consequence of Rokhlin's lemma. The statement that $X_c'$ is $G_\delta$ can be derived from the fact that a process is Bernoulli if and only if it is finitely determined. It uses the full strength of Ornstein theory. Is there an analogous result for any or every non-amenable group?

7.2. *Krieger's generator theorem.* Seward generalized Krieger's generator theorem to all countable groups using Rokhlin entropy.

THEOREM 7.3. [**Sew19**] *Let $\Gamma \curvearrowright^T (X, \mu)$ be a countably infinite group acting ergodically, but not necessarily freely, by measure-preserving bijections on a non-atomic standard probability space $(X, \mu)$. If $\bar{p} = (p_i)$ is any finite or countable probability vector with*

$$h^{\mathrm{Rok}}(T) < H(\bar{p}) = -\sum_i p_i \log p_i,$$

*then there is a generating partition $\alpha = \{A_i\}$ with $\mu(A_i) = p_i$ for all $i$.*

The proof works almost entirely within the pseudogroup of the orbit equivalence relation of the action. It also uses previous (very accessible) work of Seward [**Sew15a**] to show that there exists a finite generating partition of the action whenever it has finite Rokhlin entropy. There is also a relative version of [**Sew19**, Theorem 7.3] and a non-ergodic version is being written [**AS16**].

7.3. *Sinai's theorem.* Seward recently generalized Sinai's factor theorem.

THEOREM 7.4. [**Sew18**] *Let $\Gamma$ be an arbitrary countable group and $\Gamma \curvearrowright^\alpha (X, \mu)$ an ergodic essentially free pmp action. Let $(K, \kappa)$ be a probability space and suppose*

$$0 < H(K, \kappa) \leq h^{\mathrm{Rok}}(\alpha).$$

*Then the action $\Gamma \curvearrowright^\alpha lpha(X, \mu)$ factors onto the Bernoulli shift $\Gamma \curvearrowright (K, \kappa)^\Gamma$.*

*Remarks on the proof.* This is only a light sketch of the proof of this deep result in the special case in which $H(K, \kappa) < h^{\mathrm{Rok}}(\alpha)$. By [**STD16**] there exists an essentially free action $\Gamma \curvearrowright^\beta (Y, \nu)$ with

$$h^{\mathrm{Rok}}(\beta) < h^{\mathrm{Rok}}(\alpha) - H(K, \kappa)$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

such that $\Gamma \curvearrowright^{\alpha} (X, \mu)$ factors onto $\Gamma \curvearrowright^{\beta} (Y, \nu)$. Let $\mathcal{B}_Y$ denote the Borel sigma-algebra of $Y$, which we consider to be a sub-sigma-algebra of $\mathcal{B}_X$ via this factor map.

There exists a Borel map $f : Y \to [0, 1]$ and an aperiodic automorphism $T \in \mathrm{Aut}(Y, \nu)$ such that $f_*\nu$ is the Lebesgue measure and each orbit of $T$ is the intersection of an orbit of $\Gamma$ with the preimage of a point under $f$. Let $\tilde{f} : X \to [0, 1]$ and $\tilde{T} \in \mathrm{Aut}(X, \mu)$ denote the lifts of $\tilde{f}$ and $\tilde{T}$ respectively. Note that every $\tilde{T}$-orbit is the intersection of an orbit of $\Gamma$ with the preimage of a point under $\tilde{f}$.

Given a countable partition $\xi$ of $X$ and $S \subset [0, 1]$ let

$$\xi_S = \xi \restriction^X \tilde{f}^{-1}(S)$$

denote its quasi-restriction (this is the smallest partition of $X$ containing all sets of the form $Z \cap \tilde{f}^{-1}(S)$ for $Z \in \xi$). If $\mathcal{F}$ is a sigma-algebra of $X$, we define $\mathcal{F}_S$ similarly. Also let

$$\mathcal{P}_\xi = \mathcal{B}_Y \vee \bigvee_{t \in [0,1]} (\sigma - \mathrm{alg}_\Gamma(\xi_{[0,t)}) \restriction^X \tilde{f}^{-1}([t, 1])).$$

This is called the *external past* of $\xi$. It is $\tilde{T}$-invariant and importantly

$$h^{\mathrm{Rok}}(\alpha \restriction \sigma - \mathrm{alg}_\Gamma(\xi) | \mathcal{B}_Y) \leq h_\mu(\tilde{T}, \xi | \mathcal{P}_\xi),$$

where $\alpha \restriction \sigma - \mathrm{alg}_\Gamma(\xi)$ denotes the factor of the action $\alpha$ generated by the partition $\xi$. Therefore

$$H(K, \kappa) \leq h_\mu(\tilde{T}, \xi | \mathcal{P}_\xi).$$

So we can apply the relative Sinai factor theorem (due to Thouvenot [**Tho75**]) to $\tilde{T}$ relative to $\mathcal{P}_\xi$ to obtain a Bernoulli factor for $\tilde{T}$ that is independent of $\mathcal{P}_\xi$. This Bernoulli factor can be 'put together' to obtain a Bernoulli factor for the $\Gamma$-action. □

As spectacular as the result above is; it might not be the 'best possible'. As explained in §5.1 there exist actions with zero Rokhlin entropy that factor onto Bernoulli shifts. This leads us to the following.

*Question 24.* Suppose $\Gamma \curvearrowright (X, \mu)$ is an ergodic pmp action and $(K, \kappa)$ is a probability space whose Shannon entropy lower bounds the naive entropy of $\Gamma \curvearrowright (X, \mu)$. Then does $\Gamma \curvearrowright (X, \mu)$ factor onto the Bernoulli shift $\Gamma \curvearrowright (K, \kappa)^\Gamma$? By Sinai's theorem, if $\Gamma$ is amenable then the answer is 'yes'. When $\Gamma$ is non-amenable then the naive entropy of $\Gamma \curvearrowright (X, \mu)$ is in $\{0, \infty\}$. When the naive entropy is zero, the action cannot factor onto any Bernoulli shift since naive entropy is monotone under factor maps.

7.3.1. *Spectral theory implications.* Given a measure space $(X, \mu)$, let $L^2_0(X, \mu) \subset L^2(X, \mu)$ denote the orthogonal complement of the constant functions. Given a pmp action $\Gamma \curvearrowright^T (X, \mu)$ there is a corresponding homomorphism of $\Gamma$ into the unitary group of $L^2_0(X, \mu)$ given by

$$\kappa : \Gamma \to L^2_0(X, \mu), \quad \kappa_g f = f \circ g^{-1}.$$

This is called the *Koopman representation*. It is well known that the Koopman representation of a Bernoulli shift is isomorphic to the countable sum of left regular representations, which means that it has a *countable Lebesgue spectrum*. So it follows

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

from Sinai's factor theorem that the Koopman representation of a positive-entropy action of an amenable group $\Gamma$ necessarily contains a subrepresentation isomorphic to a countable sum of left-regular representations. A more difficult result to obtain is that any action with completely positive entropy has a countable Lebesgue spectrum. This was first proven for $\Gamma = \mathbb{Z}$ using the Rokhlin–Sinai theorem and then extended to all amenable groups by [**DG02**] using orbit-equivalence techniques.

We now have versions of these results for arbitrary groups.

THEOREM 7.5. [**Sew18**] *Suppose $\Gamma \curvearrowright (X, \mu)$ is an essentially free ergodic pmp action. Let $\mathcal{H} \subset L^2_0(X, \mu)$ be a $\Gamma$-invariant closed subspace and $\mathcal{F} \subset \mathcal{B}_X$ be the smallest $\Gamma$-invariant sigma-algebra such that all functions in $\mathcal{H}$ are $\mathcal{F}$-measurable. If $\mathcal{H}$ has no non-zero subrepresentation that embeds into the left regular representation $\Gamma \curvearrowright \ell^2(\Gamma)$, then the factor corresponding to $\mathcal{F}$ has zero Rokhlin entropy.*

The sofic entropy version of this theorem was obtained previously by Hayes [**Hay18**] with a completely different proof relying on von Neumann algebra machinery.

COROLLARY 7.6. [**Sew18**] *Suppose $\Gamma \curvearrowright (X, \mu)$ is an essentially free ergodic action with completely positive outer Rokhlin entropy as defined in §11.2. Then the Koopman representation $\Gamma \curvearrowright L^2_0(X, \mu)$ is unitarily isomorphic to the countable sum of left regular representations.*

The sofic entropy version of the above corollary is stated as Theorem 11.5 below.

7.3.2. *Markov chains factor onto Bernoulli shifts.* We do not know whether every mixing Markov chain over a free group has positive Rokhlin entropy. So Theorem 7.4 cannot be applied. However, the existence of a Bernoulli factor for these systems can be obtained directly.

THEOREM 7.7. *Every mixing Markov chain over a non-abelian free group factors onto a Bernoulli shift.*

*Proof.* Let $\mathbb{F} = \langle S \rangle$ be a non-abelian free group and $\mathbf{X} = (X_g)_{g \in \mathbb{F}}$ be a mixing Markov chain taking values in a finite or countable state space $K$. Let $\mu$ be the law of $\mathbf{X}$ so that $\mathbb{F} \curvearrowright (K^{\mathbb{F}}, \mu)$ is a pmp action. Without loss of generality, we assume that the law of $X_e$ is fully supported on $K$.

Let $a, b \in S$ be distinct elements. Let $\Lambda \leq \mathbb{F}$ be the cyclic subgroup generated by $a$. The law of $(X_g)_{g \in \Lambda}$ conditioned on $X_e = k$ is non-atomic (for any $k \in K$). Denote this law by $\mu_k \in \mathrm{Prob}(K^\Lambda)$. So there exists a measure-space isomorphism

$$\phi_k : (K^\Lambda, \mu_k) \to (\mathbb{T}, \lambda),$$

where the latter denotes the circle with Haar measure. Define $\phi : K^\Lambda \to \mathbb{T}$ by $\phi(x) = \phi_k(x)$ where $k = x_e$.

Define processes $\mathbf{Y} = (Y_g)_{g \in \mathbb{F}}$ and $\mathbf{Z} = (Z_g)_{g \in \mathbb{F}}$ by

$$Y_g = \phi(a^n \mapsto X_{ga^n}), \quad Z_g = Y_g + Y_{gb}.$$

Since $\mathbf{Z}$ is a factor of $\mathbf{X}$, it suffices to show that it is i.i.d. Before doing that, let us consider some general properties of the variables $Y_g$ and $Z_g$.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

Suppose that $g \in \Gamma$ and $U \subset \Gamma \backslash \{g\}$ have the property that every path from a vertex in $U\Lambda$ to a vertex in $g\Lambda$ in the Cayley graph $\mathrm{Cay}(\mathbb{F}, S)$ necessarily passes through $g$. In this case, $Y_g$ is independent of $\{Y_h\}_{h \in U}$. This is because the Markov property implies $\{X_{ha^n} : h \in U, n \in \mathbb{Z}\}$ and $\{X_{ga^n} : n \in \mathbb{Z}\}$ are independent conditioned on $X_g$. Since the law of $Y_g$ conditioned on $X_g$ is Haar measure, $Y_g$ is independent of $X_g$. So $Y_g$ is independent of $(Y_h)_{h \in U}$.

In general, if $A$, $B$, $C$ are random variables satisfying: (1) $A$, $B$ are independent, (2) $B$, $C$ take values in the circle and (3) $B$ is distributed according to Haar measure, then $A$ and $B + C$ are independent. It follows that if $U \subset \Gamma$ and $g \in \Gamma$ are such that either $Y_g$ or $Y_{gb}$ is independent of $\{Z_h\}_{h \in U}$ then $Z_g = Y_g + Y_{gb}$ is independent of $\{Z_h\}_{h \in U}$. In particular, this holds if either every path from a vertex in $(U \cup Ub)\Lambda$ to a vertex in $g\Lambda$ passes through $g$ or if every path from a vertex in $(U \cup Ub)\Lambda$ to a vertex in $gb\Lambda$ passes through $gb$.

Now let $W \subset \mathbb{F}$ be a finite set such that the induced subgraph of $W$ in the Cayley graph $\mathrm{Cay}(\mathbb{F}, S)$ is connected. To finish the proof, it suffices to show that $(Z_g)_{g \in W}$ is i.i.d.

Let $f, g \in W$ be such that there exists $s \in S \cup S^{-1}$ with $fs = g$ and the induced subgraph of $U := W \backslash \{g\}$ is connected. By induction, we may assume that the variables $(Z_g)_{g \in U}$ are i.i.d.

Because the induced subgraph of $U$ is connected, the induced subgraph of $(U \cup Ub)\Lambda$ is also connected. Moreover, $(U \cup Ub)\Lambda$ contains $g$ if and only if $s \in \{a^{-1}, a, b\}$. In this case, $gb \notin (U \cup Ub)\Lambda$ and every path from a vertex in $(U \cup Ub)\Lambda$ to a vertex in $gb\Lambda$ necessarily passes through $gb$. On the other hand, if $(U \cup Ub)\Lambda$ does not contain $g$ then every path from a vertex in $(U \cup Ub)\Lambda$ to a vertex in $g\Lambda$ necessarily passes through $g$. Therefore, $Z_g$ is independent of $\{Z_h\}_{h \in U}$ and $(Z_g)_{g \in W}$ is i.i.d. □

7.4. *The $\bar{d}$ metric.* Let $(K, d_K)$ be a compact metric space. The metric $d_K$ induces a new metric, denoted $\bar{d}$, on the space of $\Gamma$-invariant probability measures on $K^\Gamma$. Intuitively, this new metric measures how closely two measures $\mu$, $\nu$ can be 'joined'. More precisely, recall that a *joining* between measures $\mu, \nu \in \mathrm{Prob}_\Gamma(K^\Gamma)$ is a $\Gamma$-invariant Borel probability measure $\lambda$ on the product space $K^\Gamma \times K^\Gamma$ whose marginals are $\mu$ and $\nu$. The $\bar{d}$-distance between $\mu$ and $\nu$ is

$$\bar{d}(\mu, \nu) = \inf_\lambda \int d_K(x_e, y_e)\, d\lambda(x, y),$$

where the infimum is over all joinings $\lambda$ of $\mu$ and $\nu$. If $K$ is finite then it is usually assumed that $d_K$ is the trivial metric $d_K(k_1, k_2) = 1$ if $k_1 \neq k_2$. This metric plays a key role in most developments of Ornstein theory.

THEOREM 7.8. *If $\Gamma$ is amenable and $K$ is finite then:*

(1) *entropy function $\mu \mapsto h_\mu(\Gamma \curvearrowright K^\Gamma)$ is continuous in the topology induced by $\bar{d}$;*
(2) *the set of Bernoulli measures in $\mathrm{Prob}_\Gamma(K^\Gamma)$ is $\bar{d}$-closed.*

*Proof.* The first statement is an exercise in [**Rud90**] (for $\Gamma = \mathbb{Z}$). The last statement is contained in [**ST75**] (again for $\Gamma = \mathbb{Z}$) although it also follows from the characterization of Bernoulli shifts as finitely determined processes (see e.g. [**Rud90**] for details). □

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

Here we will show that both statements above fail for at least some non-amenable groups. This is interesting because the first statement is a key ingredient in Ornstein theory and the last is a consequence. The next theorem is due to Austin. It improves on an earlier example due to myself and Seward.

THEOREM 7.9. *If* $\Gamma$ *contains a non-abelian free group then there exists a sequence* $\{\mu_n\}_{n=1}^{\infty}$ *of* $\Gamma$*-invariant measures* $\mu_n \in \mathrm{Prob}_\Gamma(K^\Gamma)$ *for some finite* $K$ *such that:*

- $\Gamma \curvearrowright (K^\Gamma, \mu_n)$ *is isomorphic to the Bernoulli shift* $\Gamma \curvearrowright ((\mathbb{Z}/2), u_2)^\Gamma$ *for all* $n$*;*
- *the sequence* $\{\mu_n\}$ *converges in the* $\bar{d}$*-metric to a measure* $\mu_\infty$ *(as* $n \to \infty$*) such that* $\Gamma \curvearrowright (K^\Gamma, \mu_\infty)$ *is isomorphic to the Bernoulli shift* $\Gamma \curvearrowright ((\mathbb{Z}/2 \times \mathbb{Z}/2), u_2 \times u_2)^\Gamma$.

*In particular, neither sofic entropy, the* $f$*-invariant nor Rokhlin entropy is continuous in the* $\bar{d}$*-metric.*

*Proof.* Let $K = (\mathbb{Z}/2)^4$. Let $\beta_n : (\mathbb{Z}/2)^\Gamma \to \{0, 1\}$ be any sequence of measurable functions such that

$$\lim_{n\to\infty} u_2^\Gamma(\beta_n^{-1}(1)) = 0$$

and $u_2^\Gamma(\beta_n^{-1}(1)) > 0$ for all $n$.

Let $a, b \in \Gamma$ generate a rank-two free subgroup. Define factor maps $\alpha_n : (\mathbb{Z}/2)^\Gamma \to K^\Gamma$ by

$$\alpha_n(x)_g = \begin{cases} (x_g + x_{ga}, x_g + x_{gb}, 0, 0) & \text{if } \beta_n(g^{-1}x) = 0, \\ (x_g + x_{ga}, x_g + x_{gb}, x_g, 1) & \text{if } \beta_n(g^{-1}x) = 1. \end{cases}$$

Let $\mu_n = (\alpha_n)_* u_2^\Gamma$. To verify the conclusions, let $\pi : K^\Gamma \to (\mathbb{Z}/2 \times \mathbb{Z}/2)^\Gamma$ be the projection map onto the first two coordinates. Observe that $\pi\alpha_n$ is the Ornstein–Weiss factor map. Since this map is two-one (when restricted to the free subgroup generated by $a, b$) and $u_2^\Gamma(\beta_n^{-1}(1)) > 0$ it follows that $\alpha_n$ is an isomorphism onto its image.

Define $\alpha_\infty : (\mathbb{Z}/2)^\Gamma \to K^\Gamma$ by $\alpha_\infty(x)_g = (x_g + x_{ga}, x_g + x_{gb}, 0, 0)$. The $\bar{d}$-limit of $\mu_n$ is the measure $\mu_\infty := \alpha_{\infty *} u_2^\Gamma$. The Ornstein–Weiss example shows that $\Gamma \curvearrowright (K^\Gamma, \mu_\infty)$ is isomorphic to the Bernoulli shift over the base space with entropy $\log(4)$. □

*Question 25.* Is the $f$-invariant finitely observable in the sense of [**OW07**]? The theorem above suggests the answer may be 'no'.

THEOREM 7.10. *Let* $\Gamma$ *be an infinite property (T) group. Then there exists a finite set* $K$ *and a sequence* $\{\mu_n\}$ *of* $\Gamma$*-invariant measures on* $K^\Gamma$ *such that:*

- $\Gamma \curvearrowright (K^\Gamma, \mu_n)$ *is isomorphic to a Bernoulli shift for all* $n$*;*
- *the sequence* $\{\mu_n\}_n$ *converges in the* $\bar{d}$*-topology to a measure* $\mu_\infty$ *(as* $n \to \infty$*) such that* $\Gamma \curvearrowright (K^\Gamma, \mu_\infty)$ *is isomorphic to a non-Bernoulli factor of a Bernoulli shift.*

*Proof.* This example is similar to the previous one. To be precise, let $S \subset \Gamma$ be a finite generating set and $K = (\mathbb{Z}/2)^S \times \mathbb{Z}/2 \times \mathbb{Z}/2$. Let $\beta_n : (\mathbb{Z}/2)^\Gamma \to \{0, 1\}$ be any sequence of measurable functions such that

$$\lim_{n\to\infty} u_2^\Gamma(\beta_n^{-1}(1)) = 0$$

and $u_2^\Gamma(\beta_n^{-1}(1)) > 0$ for all $n$.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

Define factor maps $\alpha_n : (\mathbb{Z}/2)^\Gamma \to K^\Gamma$ by

$$\alpha_n(x)_g = \begin{cases} ((x_g + x_{gs})_{s\in S}, 0, 0) & \text{if } \beta_n(g^{-1}x) = 0, \\ ((x_g + x_{gs})_{s\in S}, x_g, 1) & \text{if } \beta_n(g^{-1}x) = 1. \end{cases}$$

Let $\mu_n = (\alpha_n)_* u_2^\Gamma$. To verify the conclusions, let $\pi : K^\Gamma \to ((\mathbb{Z}/2)^S)^\Gamma$ be the projection map. Observe that $\pi\alpha_n$ is the Ornstein–Weiss factor map. Since this map is two-one and $u_2^\Gamma(\beta_n^{-1}(1)) > 0$, it follows that $\alpha_n$ is an isomorphism onto its image.

Define $\alpha_\infty : (\mathbb{Z}/2)^\Gamma \to K^\Gamma$ by $\alpha_\infty(x)_g = ((x_g + x_{gs})_{s\in S}, 0, 0)$. The $\bar{d}$-limit of $\mu_n$ is the measure $\mu_\infty := \alpha_{\infty *} u_2^\Gamma$ and $\Gamma \curvearrowright (K^\Gamma, \mu_\infty)$ is isomorphic to the Ornstein–Weiss factor $\Gamma \curvearrowright (\mathbb{Z}/2)^\Gamma/(\mathbb{Z}/2)$.

By the proof of Theorem 7.2, the cohomology group of any Bernoulli shift over $\Gamma$ with values in the circle $\mathbb{R}/\mathbb{Z}$ is $\mathrm{Hom}(\Gamma, \mathbb{R}/\mathbb{Z})$ while the cohomology group of the Ornstein–Weiss factor $\Gamma \curvearrowright (\mathbb{Z}/2)^\Gamma/(\mathbb{Z}/2)$ is $\mathrm{Hom}(\Gamma, \mathbb{R}/\mathbb{Z}) \times \mathbb{Z}/2$. Since $\Gamma$ has property (T), $\mathrm{Hom}(\Gamma, \mathbb{R}/\mathbb{Z})$ is finite. Therefore, these cohomology groups are non-isomorphic and so $\Gamma \curvearrowright (K^\Gamma, \mu_\infty)$ is not orbit-equivalent to a Bernoulli shift and so cannot be measurably conjugate to one. □

*Remark 22.* Corollary 5.4 shows that for any non-amenable group there exists an inverse limit of factors of Bernoulli shifts that has zero Rokhlin entropy. This inverse limit can be realized as a $\bar{d}$-limit of factors of Bernoulli shifts (with $K = [0, 1]$ for example). This provides another example of a non-Bernoulli $\bar{d}$-limit of Bernoulli shifts although in this case, $K$ is infinite. Is it possible to modify this example using finite $K$?

## 8. *The variational principle*

THEOREM 8.1. (The variational principle) *Let $\Gamma \curvearrowright X$ be a continuous action on a compact metrizable space. Let $\Sigma$ be a sofic approximation of $\Gamma$. Then*

$$h_\Sigma(\Gamma \curvearrowright X) = \sup_\mu h_{\Sigma,\mu}(\Gamma \curvearrowright X),$$

*where the supremum is over all $\Gamma$-invariant Borel probability measures $\mu$ on $X$. In particular, if there does not exist a $\Gamma$-invariant probability measure on $X$ then $h_\Sigma(\Gamma \curvearrowright X) = -\infty$.*

*Remark 23.* For $\mathbb{Z}$-actions, this theorem was obtained over several papers [**Goo69**, **Goo72**, **Din70**, **Goo71**]. The proof that now appears in most textbooks is due to Misiurewicz [**Mis76**]. A number of other variational principles in entropy theory are provided in [**Dow11**]. The case of general sofic groups is [**KL11b**, Theorem 6.1]. There are versions of this result for sofic pressure [**Chu13**], sofic groupoids [**Bow14**] and a local version (with respect to a finite open cover) in [**Zha12**].

*Proof sketch.* The inequality

$$h_\Sigma(\Gamma \curvearrowright X) \geq \sup_\mu h_{\Sigma,\mu}(\Gamma \curvearrowright X)$$

is immediate from the pseudometric definition of sofic entropy (§2.4.1). The opposite inequality is achieved in the following way. Fix a finite partition $\mathcal{P}$ of $\mathrm{Prob}(X)$, the space of

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

probability measures on $X$. Fix a scale $\epsilon > 0$. Each microstate $\phi : V_n \to X$ has an empirical measure

$$P_\phi^{\sigma_n} = |V_n|^{-1} \sum_{v \in V_n} \delta_{\phi(v)} \in \mathrm{Prob}(X).$$

By pulling back the partition $\mathcal{P}$, we obtain a finite partition $\tilde{\mathcal{P}}$ on the space of topological microstates. It follows that the exponential growth rate of the maximum cardinality of an $\epsilon$-separated subset of topological microstates is approximated by the same growth rate only restricted to microstates whose empirical measures lie in a fixed part of the partition. By refining this partition and taking Benjamini–Schramm limits we can build an invariant measure whose sofic entropy is bounded below by the exponential rate of growth of the maximum cardinality of an $\epsilon$-separated subset of the topological microstate space. Sending $\epsilon \searrow 0$ finishes the proof. □

8.1. *Measures of maximal entropy: existence.* The variational principle naturally leads to two problems: under what conditions does there exist a measure of maximal entropy and if one exists is it unique? If sofic entropy is upper semi-continuous as a function on $\mathrm{Prob}_\Gamma(X)$ with respect to the weak* topology then compactness of $\mathrm{Prob}_\Gamma(X)$ implies existence. Upper semi-continuity is discussed in §4.5 and in greater depth in [**CZ15**]. For example, it holds whenever $X$ is a closed $\Gamma$-invariant subpace of $\mathcal{A}^\Gamma$ for some finite alphabet $\mathcal{A}$.

8.2. *Measures of maximal entropy: uniqueness.* It appears that there are no general results concerning uniqueness of measures of maximal entropy outside of $\Gamma = \mathbb{Z}^d$. However, there is a counterexample to a natural conjecture in the case of $\Gamma = \mathbb{F}_2$ and subshifts of finite type which we will go over next.

*Definition 17.* (Subshifts of finite type) Let $\Gamma$ be a countable group, $K$ a finite set and $\Gamma \curvearrowright K^\Gamma$ the shift action: $(gx)(f) = x(g^{-1}f)$ for $g,\ f \in \Gamma,\ x \in K^\Gamma$. Let $F \Subset \Gamma$ be a finite set and $\Omega \subset K^F$ be a collection of maps from $F$ to $K$. Let $X$ be the set of all $x \in K^\Gamma$ such that for every $g \in \Gamma$, $gx$ restricted to $F$ is in $\Omega$. Then $X$ is the *subshift of finite type* determined by $\Omega$. Because subshifts of finite type are expansive, they admit measures of maximal entropy.

It is a well known fact that if $X$ is a subshift of finite type (over $\mathbb{Z}$), which is topologically transitive, then $X$ admits a *unique* measure of maximal entropy. Indeed more is true—for any continuous potential $\phi : X \to \mathbb{R}$, there is a unique equilibrium measure on $X$ [**Bow08**]. By contrast, Burton *et al* proved in [**BS94**] that $\mathbb{Z}^2$ admits strongly irreducible subshifts of finite type with more than one measure of maximal entropy. See also [**BS95**] for further constructions and general criteria for uniqueness in the $\mathbb{Z}^d$-setting.

THEOREM 8.2. *There exists a topologically transitive subshift of finite type over the free group $\mathbb{F}_2 = \langle a, b \rangle$ that admits more than one measure of maximal $f$-invariant.*

*Proof.* Let $n \geq 4$ and define $X \subset [n]^{\mathbb{F}_2}$ by an element $x \in X$ if and only if for every $g \in \mathbb{F}_2$ and $s \in \{a, b\}$ there is an $\epsilon \in \{0, 1\}$ such that $x(gs) = x(g) + \epsilon \mod n$. This is a topologically transitive subshift of finite type.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

Let $T = (T^g)_{g\in\mathbb{F}_2}$ denote the shift action $T^g x(f) = x(g^{-1}f)$ on $X$. Because the $f$-invariant is an infimum of continuous functions (namely, $f_\mu(T) = \inf_{W\Subset\mathbb{F}_2} F_\mu(T, \mathcal{P}^W)$ where $\mathcal{P}$ is the canonical partition on $[n]^{\mathbb{F}_2}$) it is upper semi-continuous in $\mu$ with respect to the weak* topology. Therefore, a measure of maximal $f$-invariant exists. Let $\mu$ be such a measure.

Let $\nu$ be the Markov approximation to $\mu$. To be precise, $\nu$ is the law of a Markov process $\mathbf{X} = (X_g)_{g\in\mathbb{F}_2}$ with values in $[n]$ satisfying the following: if $\mathbf{Y} = (Y_g)_{g\in\mathbb{F}_2}$ is a stationary process with law $\mu$ then the law of the pair $(X_e, X_s)$ equals the law of $(Y_e, Y_s)$ for every $s \in \{a, b, a^{-1}, b^{-1}\}$. By Theorem 3.5,

$$f_\mu(T) \leq F_\mu(T, \mathcal{P}) = F_\nu(T, \mathcal{P}) = f_\nu(T).$$

Moreover, equality holds if and only if $\mu = \nu$. So $\mu$ is a Markov measure.

Let $A : [n]^{\mathbb{F}_2} \to [n]^{\mathbb{F}_2}$ be the +1 map: $A(x)_g \equiv x_g + 1 \mod n$. This map commutes with $\mathbb{F}_2$-action and preserves $X$. Therefore $A_*\mu$ is also a measure of maximal $f$-invariant. Let us assume to obtain a contradiction that $\mu$ is the unique measure of maximal $f$-invariant. Then $A_*\mu = \mu$ which implies the existence of a parameter $0 \leq \alpha_s \leq 1/n$ (for $s \in \{a, b\}$) such that

$$\mu(\{x \in X : x_e = i\}) = 1/n,$$
$$\mu(\{x \in X : x_e = i, x_s = i\}) = \alpha_s,$$
$$\mu(\{x \in X : x_e = i, x_s = i + 1\}) = 1/n - \alpha_s.$$

Recall

$$f_\mu(T) = -3H_\mu(\mathcal{P}) + \sum_{s\in S} H_\mu(\mathcal{P} \vee T^s\mathcal{P}).$$

In our case, $H_\mu(\mathcal{P}) = \log(n)$ and $H_\mu(\mathcal{P} \vee T^s\mathcal{P})$ is uniquely maximized by $\alpha_s = 1/2n$ in which case $H_\mu(\mathcal{P} \vee T^s\mathcal{P}) = \log(2n)$. Therefore,

$$f_\mu(T) = -3\log(n) + 2\log(2n) = \log(4) - \log(n).$$

Because $n \geq 4$, this is non-positive. However, $X$ admits fixed points; for example the element $x_g = 0$ for all $g \in \mathbb{F}_2$. The Dirac measure concentrated on a fixed point is also an invariant probability measure and its $f$-invariant is zero. This contradiction proves the theorem. □

*Problem 11.* The example above is not completely satisfying because it uses the $f$-invariant instead of sofic entropy. It would be interesting to find a sofic example or prove that one does not exist.

*Remark 24.* The example above exploits the fact that $F_\mu(\mathcal{P})$ is not concave in $\mu$ and therefore $f_\mu(\mathcal{P})$ is also not concave. By contrast, Shannon entropy $H_\mu(\mathcal{P})$ is concave in $\mu$ (this is important in proving uniqueness of measures of maximal entropy for topologically transitive subshifts of finite type over the integers).

*Question 26.* Hoffman constructed a subshift of finite type over $\mathbb{Z}^2$ that has a measure of maximal entropy that has completely positive entropy but is non-Bernoulli [**Hof11**]. Does an analogous example exist for the free group?

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

9. *Gibbs measures, pressure and equilibrium states*

9.1. *Motivation: finite graphs.* To motivate the notion of sofic pressure, we begin by recalling Gibbs measures in the setting of finite graphs. Let $\mathcal{X}$ be a finite set of 'spins' and $G = (V, E)$ a finite graph. A *spin configuration* is a function $\phi : V \to \mathcal{X}$. A *vertex potential* is a function $\psi_{\text{vert}} : \mathcal{X} \to \mathbb{R}$ and an *edge potential* is a symmetric function $\psi_{\text{edge}} : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$. The *energy* of a spin configuration $\phi$ (with respect to $\psi_{\text{vert}}$ and $\psi_{\text{edge}}$) is

$$E(\phi) := \sum_{v \in V} \psi_{\text{vert}}(\phi(v)) + \sum_{e=\{v,w\} \in E} \psi_{\text{edge}}(\phi(v), \phi(w)).$$

Consider the following problem: given $E_0 \in \mathbb{R}$, find a probability measure $\mu$ on $\mathcal{X}^V$ such that the $\mu$-average energy is $E_0$:

$$\sum_{\phi} E(\phi)\mu(\{\phi\}) = E_0$$

and so that $\mu$ maximizes entropy over all measures satisfying the above. By way of Lagrange multipliers, it can be shown that the unique measure solving this problem has the form

$$\mu(\{\phi\}) = Z^{-1} \exp(-\beta E(\phi))$$

for some constants $Z$, $\beta$. Moreover,

$$Z = Z(G) = \sum_{\phi} \exp(-\beta E(\phi))$$

is called the *partition function*. It is of central importance to understand the exponential growth rate of $Z(G)$ as the graph $G = (V, E)$ Benjamini–Schramm converges to a fixed graph of interest (for example, the Cayley graph of $\mathbb{Z}^d$ or $\mathbb{F}_r$). See [**DM10a**, **Geo11**, **Kel98**].

In the next two sections we define the sofic pressure of a given topological action $\Gamma \curvearrowright X$ together with a potential function $\psi : X \to \mathbb{R}$ and sofic approximation $\Sigma$ as the exponential growth rate of an analogous partition function.

9.2. *Pressure.* Let $(X, \rho)$ be a compact metric space, $\Gamma \curvearrowright^T X$ an action by homeomorphisms, $\Psi : X \to \mathbb{R}$ a continuous function which we will call a *potential function* and $\Sigma = \{\sigma_n : \Gamma \to \text{Sym}(V_n)\}$ a sofic approximation to $\Gamma$. Given $\sigma : \Gamma \to \text{Sym}(V)$ and a microstate $\phi : V \to X$ define its energy by

$$E(\phi) = \sum_{v \in V} \Psi(\phi(v)).$$

Now suppose $\mathcal{Z} \subset X^V$ is a collection of microstates. Then define the associated partition function by

$$Z(\mathcal{Z}) = \sum_{\phi \in \mathcal{Z}} \exp(E(\phi)).$$

Given $F \Subset \Gamma$, $\delta, \epsilon > 0$, let

$$Z_\epsilon(\Psi, \Gamma \curvearrowright X, \rho, F, \delta, \sigma) = \sup Z(\mathcal{Z}),$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

where the sup is over all $(\rho_\infty, \epsilon)$-separated subsets of $\mathrm{Map}(T, \rho, F, \delta, \sigma)$. The $\Sigma$-*pressure* of $(\Gamma \curvearrowright X, \Psi)$ is

$$P_\Sigma(\Gamma \curvearrowright X, \Psi) := \sup_{\epsilon>0} \inf_{F \Subset \Gamma} \inf_{\delta>0} \limsup_{n\to\infty} |V_n|^{-1} \log Z_\epsilon(\Psi, \Gamma \curvearrowright X, \rho, F, \delta, \sigma_n).$$

This definition was introduced in [**Chu13**] where it is also proven to be independent of the choice of metric $\rho$ (one can even use a generating pseudometric and $\rho_\infty$ can be replaced by $\rho_2$). For example, when $\Psi = 0$ the pressure is the same as the sofic entropy.

If $\Gamma$ is amenable then there is another definition of pressure given in terms of a Følner sequence. In [**Chu13**], Chung proves that this definition agrees with the above. There is also a variational principle for pressure generalizing the one for entropy.

THEOREM 9.1. (The variational principle [**Chu13**]) *Let $\Gamma \curvearrowright^T X$ be a continuous action on a compact metrizable space and $\Psi : X \to \mathbb{R}$ be continuous. Let $\Sigma$ be a sofic approximation of $\Gamma$. Then*

$$P_\Sigma(T, \Psi) = \sup_\mu h_{\Sigma,\mu}(T) + \int \Psi \, d\mu,$$

*where the supremum is over all $T(\Gamma)$-invariant Borel probability measures on $X$.*

9.3. *Pressure in symbolic systems.* Here we specialize to the following set-up: let $\mathcal{X}$ be a finite set and $\Gamma \curvearrowright \mathcal{X}^\Gamma$ be the action by shifting: $(gx)(f) = x(g^{-1}f)$. Let $\Psi : \mathcal{X}^\Gamma \to \mathbb{R}$ be a continuous potential function. In this case, the sofic pressure admits a more intuitive formulation. Given $\sigma : \Gamma \to \mathrm{Sym}(V)$, $v \in V$ and $\phi : V \to \mathcal{X}$, let $\Pi_v^\sigma(\phi) \in \mathcal{X}^\Gamma$ be the *pullback name* of $\phi$:

$$\Pi_v^\sigma(\phi)_g = \phi(\sigma(g)^{-1}v).$$

The *energy* of $\phi$ is

$$E(\phi) = \sum_{v\in V} \Psi(\Pi_v^\sigma(\phi))$$

and the partition function associated to $\sigma$, $\Psi$ is

$$Z(\sigma, \Psi) = \sum_{\phi \in \mathcal{X}^V} \exp(E(\phi)).$$

Finally, the $\Sigma$-*pressure* of $(\Gamma \curvearrowright \mathcal{X}^\Gamma, \Psi)$ is

$$P_\Sigma(\Gamma \curvearrowright \mathcal{X}^\Gamma, \Psi) := \limsup_{n\to\infty} |V_n|^{-1} \log Z(\sigma_n, \Psi).$$

It is straightforward to show that this definition agrees with the previous definition.

9.4. *Equilibrium states.* A measure $\mu$ is called an *equilibrium state* for $(\Gamma \curvearrowright X, \Psi)$ if it realizes the supremum in the variational principle:

$$P_\Sigma(\Gamma \curvearrowright X, \Psi) = h_{\Sigma,\mu}(\Gamma \curvearrowright X) + \int \Psi \, d\mu.$$

For example, if the action $\Gamma \curvearrowright X$ is expansive then entropy is upper semi-continuous in the measure $\mu$ and therefore there exists an equilibrium state. In the special case of $\Gamma = \mathbb{Z}$, if $X$ is a subshift of finite type (over $\mathbb{Z}$), which is topologically transitive, then there is a unique equilibrium measure on $X$ [**Bow08**].

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*Example 7.* Suppose $\mathcal{X}$ is a finite set and $\Psi : \mathcal{X}^\Gamma \to \mathbb{R}$ depends only on the time 0 coordinate:

$$\Psi(x) = \Psi_0(x_e)$$

for some function $\Psi_0 : \mathcal{X} \to \mathbb{R}$. Because $\Gamma \curvearrowright \mathcal{X}^\Gamma$ is expansive, there exists an equilibrium state $\mu$ for the pair $(\Gamma \curvearrowright \mathcal{X}^\Gamma, \Psi)$. Let $\kappa$ be the one-dimensional marginal on $\mathcal{X}$:

$$\kappa(\{k_0\}) := \mu(\{x \in \mathcal{X}^\Gamma : x_e = k_0\}) \quad \text{for all } k_0 \in \mathcal{X}.$$

We claim that $\kappa^\Gamma = \mu$. Because $\Psi$ depends only on the time 0 coordinate, $\kappa^\Gamma(\Psi) = \mu(\Psi)$. So it suffices to show that $\kappa^\Gamma$ uniquely maximizes entropy over all invariant measures with the given 1-dimensional marginal. This follows from Seward's theorem 2.15.

So we have shown that every equilibrium measure is a product measure. In fact, there is a unique equilibrium measure given by

$$\kappa(\{k_0\}) = \frac{\exp(\Psi_0(k_0))}{\sum_{k \in \mathcal{X}} \exp(\Psi_0(k))}.$$

This is proven in [**Chu13**]. Alternatively, it follows from Lagrange multipliers. We have now answered [**Chu13**, Question 5.4] by demonstrating the uniqueness of the equilibrium measure.

9.5. *The Ising model.* The Ising model is a well-studied model of magnetism in statistical mechanics. In the notation of §9.1, it amounts to choosing constants $B, \beta \in \mathbb{R}$ and setting $\mathcal{X} = \{-1, +1\}$, $\Psi_{\text{vert}}(k) = Bk$, $\Psi_{\text{edge}}(k, l) = \beta kl$ so that for any spin configuration $\phi : V \to \{-1, +1\}$,

$$E(\phi) := B \sum_{v \in V} \phi(v) + \beta \sum_{e=\{v,w\} \in E} \phi(v)\phi(w)$$

and

$$Z = Z(G) = \sum_{\phi : V \to \{-1,+1\}} \exp(-E(\phi))$$

(we have employed a small change of variables). If $G_n = (V_n, E_n)$ is a sequence of finite graphs then the *asymptotic free energy density* (or free entropy or pressure) is the limit

$$\lim_{n \to \infty} |V_n|^{-1} \log Z(G_n)$$

if it exists. In the special case in which $\{G_n\}$ Benjamini–Schramm converges to a locally finite tree, an exact formula for the above limit is computed in [**DM10b**]. Here we will see that asymptotic free energy density can be re-interpreted as sofic pressure whenever the sequence $\{G_n\}$ arises from a sofic approximation.

So let $\Gamma$ be a group, $\Sigma = \{\sigma_n : \Gamma \to \text{Sym}(V_n)\}$ a sofic approximation to $\Gamma$ and $S \subset \Gamma$ a finite symmetric generating set. Let $G_n = (V_n, E_n)$ be the graph with edges $(v, \sigma_n(s)v)$ for $s \in S$, $v \in V$. Because $\Sigma$ is a sofic approximation, $\{G_n\}$ Benjamini–Schramm converges to the Cayley graph of $(\Gamma, S)$.

Consider the potential function $\Psi : \{-1, 1\}^\Gamma \to \mathbb{R}$

$$\Psi(x) = Bx(e) + \beta \sum_{s \in S} x(e)x(s).$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

By §9.3,

$$P_\Sigma(\Gamma \curvearrowright \{-1, +1\}^\Gamma, \Psi) = \limsup_{n\to\infty} n^{-1} \log Z(G_n). \tag{4}$$

In the special case in which $\Gamma = \mathbb{F}_r$ is a rank $r > 1$ free group and $S$ is a free generating set, it follows from the analysis in [**DM10b**] that the sofic pressure of $(\Gamma \curvearrowright \{-1, +1\}^\Gamma, \Psi)$ does not depend on the choice of sofic approximation and an explicit formula is known.

9.6. *Gibbs measures.* Let $\mathcal{X}$ be a finite set. A potential function $\Psi : \mathcal{X}^\Gamma \to \mathbb{R}$ has *finite range* if there is a finite subset $J \subset \Gamma$ such that $\Psi(x)$ depends only on the restriction of $x$ to $J$. In this setting, a *Gibbs measure* is any Borel probability measure $\mu$ on $\mathcal{X}^\Gamma$ satisfying the following. Let $Y = (Y_g)_{g\in\Gamma}$ be a $\Gamma$-indexed process with law $\mu$. Then $\mu$ is a Gibbs measure if for any fixed $x \in \mathcal{X}^\Gamma$ and finite $\Lambda \Subset \Gamma$, the law of $(Y_g)_{g\in\Lambda}$ conditioned on $Y_f = x(f)$ for all $f \in \Gamma\backslash\Lambda$ is given by, for any $z \in \mathcal{X}^\Gamma$ with $z(g) = x(g)$ for all $g \in \Gamma\backslash\Lambda$,

$$\mu(Y = z | Y_g = x(g) \quad \text{for all } g \in \Gamma\backslash\Lambda) = Z^{-1} \exp\left(\sum_{g\in\Lambda} \Psi(g^{-1}z)\right),$$

where

$$Z = \sum_{z\in\mathcal{X}^\Gamma : z(g)=x(g) \text{ for all } g\notin\Lambda} \exp\left(\sum_{g\in\Lambda} \Psi(g^{-1}z)\right).$$

It is well known, in the case of $\Gamma = \mathbb{Z}^d$ that every equilibrium measure for $(\Gamma \curvearrowright \mathcal{X}^\Gamma, \Psi)$ is a Gibbs measure [**Geo11**]. It appears that this question has not been explored in the context of sofic entropy. However, Alpeev proved in [**Alp15**] that when $\Gamma$ is sofic, Gibbs measures exist. Moreover, if there is a unique Gibbs measure for $(\Gamma \curvearrowright \mathcal{X}^\Gamma, \beta\Psi)$ and all $\beta \in [0, 1]$ then the modified sofic entropies for these measures do not depend on the choice of sofic approximation. Modified sofic entropy refers to the sofic entropy defined in [**Aus16a**] via measures on model spaces.

10. *Relative entropy*

Suppose $\Gamma$ is an amenable group, $\Gamma \curvearrowright (X, \mu)$ a pmp action, $\mathcal{P}$ a partition of $X$ and $\mathcal{F}$ a $\Gamma$-invariant sigma-sub-algebra of the sigma-algebra of Borel sets. Then the *entropy rate of the process* $\Gamma \curvearrowright (X, \mu, \mathcal{P})$ *relative to* $\mathcal{F}$ is

$$h_\mu(\Gamma \curvearrowright X, \mathcal{P}|\mathcal{F}) = \lim_{n\to\infty} |F_n|^{-1} H_\mu(\mathcal{P}^{F_n}|\mathcal{F}),$$

where $\{F_n\}$ is a Følner sequence in $\Gamma$. The relative entropy of the action with respect to $\mathcal{F}$ is

$$h_\mu(\Gamma \curvearrowright X|\mathcal{F}) = \sup_{\mathcal{P}} h_\mu(\Gamma \curvearrowright X, \mathcal{P}|\mathcal{F})$$

where the sup is over all finite partitions $\mathcal{P}$ of $X$. This is a measure-conjugacy invariant in the sense that if $\Gamma \curvearrowright (Y, \nu)$ is another pmp action and $\phi : X \to Y$ is a measure-conjugacy, then

$$h_\mu(\Gamma \curvearrowright X|\mathcal{F}) = h_\nu(\Gamma \curvearrowright Y|\phi(\mathcal{F})).$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

10.1. *The Abramov–Rokhlin formula.*

THEOREM 10.1. (Abramov–Rokhlin formula) *If* $\Gamma$ *is amenable and* $\Gamma \curvearrowright (Y, \nu)$ *is a factor of* $\Gamma \curvearrowright (X, \mu)$ *then*

$$h_\mu(\Gamma \curvearrowright X) = h_\nu(\Gamma \curvearrowright Y) + h_\mu(\Gamma \curvearrowright X|\mathcal{B}_Y),$$

*where* $\mathcal{B}_Y \subset \mathcal{B}_X$ *is the pullback sigma-sub-algebra.*

*Remarks on the proof.* In the case $\Gamma = \mathbb{Z}$, this result was obtained in [**AR62**]. The general amenable case, due to Ward–Zhang, makes heavy use of the Ornstein–Weiss quasi-tiling machinery [**WZ92**]. Another proof, due to Danilenko, uses orbit-equivalence theory [**Dan01**]. A new short proof appears in [**KL16**, §9.7] □

The Ornstein–Weiss example shows this formula does not extend to Rokhlin entropy. However, it does extend to the $f$-invariant. To describe this let $\mathbb{F}_r = \langle s_1, \ldots, s_r \rangle$ denote the rank $r$ free group. Suppose $\mathbb{F}_r \curvearrowright^T (X, \mu)$ is a pmp action, $\mathcal{F} \subset \mathcal{B}_X$ an $\mathbb{F}_r$-invariant sub-sigma-algebra and $\mathcal{P}$ finite-entropy partitions of $X$. Define

$$F_\mu(T, \mathcal{P}|\mathcal{F}) := H_\mu(\mathcal{P}|\mathcal{F}) + \sum_{s \in S} (H_\mu(\mathcal{P} \vee T^s\mathcal{P}|\mathcal{F}) - 2H_\mu(\mathcal{P}|\mathcal{F})),$$

$$f_\mu(T, \mathcal{P}|\mathcal{F}) := \inf_{W \Subset \Gamma} F(T, \mathcal{P}^W|\mathcal{F}).$$

In [**Bow10c**] it is shown if $\mathcal{P}$, $\mathcal{P}'$ are generating partitions then $f_\mu(T, \mathcal{P}|\mathcal{F}) = f_\mu(T, \mathcal{P}'|\mathcal{F})$. So it makes sense to define

$$f_\mu(T|\mathcal{F}) = f_\mu(T, \mathcal{P}|\mathcal{F})$$

for any finite-entropy generating partition $\mathcal{P}$.

THEOREM 10.2. [**Bow10c**] *With notation as above, if* $\mathcal{Q}$ *is a finite-entropy partition of* $X$ *contained in the sigma-algebra* $\bigvee_{g \in \mathbb{F}_r} T^g\mathcal{P} = \sigma\text{-alg}_{\mathbb{F}_r}(\mathcal{P})$, *then*

$$f_\mu(T, \mathcal{P}) = f_\mu(T, \mathcal{Q}) + f_\mu(T, \mathcal{P}|\sigma\text{-alg}_{\mathbb{F}_r}(\mathcal{Q})).$$

*So if* $\mathcal{P}$ *is generating and* $\mathbb{F}_r \curvearrowright^S (Y, \nu)$ *is a Mackey realization of the sigma-algebra corresponding to* $\sigma\text{-alg}_{\mathbb{F}_r}(\mathcal{Q})$ *then*

$$f_\mu(T) = f_\nu(S) + f_\mu(T|\sigma\text{-alg}_{\mathbb{F}_r}(\mathcal{Q})).$$

*Remarks on the proof.* The proof is obtained from the alternative formulation of the $f$-invariant in §2.5.1 and the classical Abramov–Rokhlin formula for $\mathbb{Z}$-actions. □

*Remark 25.* The Abramov–Rokhlin formula is used to prove the entropy formula for finite-to-one factor maps (Theorem 5.5) and can be used to prove the ergodic decomposition formula (Theorem 6.1). It is also used in the proof of (special cases of) Yuzvinskii's formula §3.4.4. The Ornstein–Weiss map gives an example where the relative $f$-invariant is negative.

*Remark 26.* Hayes has recently defined a notion of relative sofic entropy [**Hay16c**].

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

11. *Outer/extension entropy*

In classical entropy theory, one considers the entropy of an action with respect to a factor. In the new non-amenable theory, we also have to consider the entropy of a factor relative to the source! This gives a non-trivial concept that has been called *extension entropy*, *outer entropy* and *entropy in the presence*. It is the exponential rate of growth of the number of microstates of the target action that lift to microstates of the extension. The ideas originated in Kerr's partition definition of sofic entropy (§2.4.2) and were developed in [**LL16**, **Hay17**, **Hay16c**, **Sew15b**, **Sew16b**].

11.1. *Outer sofic entropy.* Let $\Gamma \curvearrowright^T X$ and $\Gamma \curvearrowright^S Y$ be continuous actions on compact metrizable spaces and suppose $\Phi : X \to Y$ is a continuous factor map. The *outer* $\Sigma$-*entropy* of $\Phi$ is defined by

$$h_\Sigma(\Phi) = \sup_{\epsilon>0} \inf_{F\Subset\Gamma} \inf_{\delta>0} \limsup_{i\to\infty} |V_i|^{-1} \log(N_\epsilon(\Phi^{V_i}(\mathrm{Map}(T, \rho_X, F, \delta, \sigma_i), \rho_{Y,\infty}))),$$

where $\rho_X$, $\rho_Y$ are generating continuous pseudometrics on $X$, $Y$, $\mathrm{Map}(\cdots)$ is as defined in §2.3, $N_\epsilon(\cdot, \rho_{Y,\infty})$ denotes the maximum cardinality of an $\epsilon$-separated subset and $\Phi^{V_i} : X^{V_i} \to Y^{V_i}$ is the map

$$\Phi^{V_i}(\mathbf{x})_v = \Phi(x_v)$$

for $\mathbf{x} = (x_v)_{v\in V_i} \in X^{V_i}$. It can be shown that this definition does not depend on the choice of generating pseudometrics $\rho_X$, $\rho_Y$.

If $\mu$ is a $\Gamma$-invariant measure on $X$, $\nu$ is a $\Gamma$-invariant measure on $Y$ and $\Phi_*\mu = \nu$ then define the *outer* $\Sigma$-*measure-entropy* of $\Phi$ by

$$h_{\Sigma,\mu}(\Phi) = \sup_{\epsilon>0} \inf_{\mathcal{O}} \inf_{F\Subset\Gamma} \inf_{\delta>0} \limsup_{i\to\infty} |V_i|^{-1} \log(N_\epsilon(\Phi^{V_i}(\mathrm{Map}(T, \rho_X, \mathcal{O}, F, \delta, \sigma_i), \rho_{Y,\infty}))),$$

where $\mathcal{O}$ varies over all weak* open neighborhoods of $\mu$ in $\mathrm{Prob}(X)$. There is an equivalent definition based on partitions:

$$h_{\Sigma,\mu}(\Phi) = \sup_{\mathcal{Q}} \inf_{\mathcal{P}} \inf_{1_\Gamma\in F\Subset\Gamma} \inf_{\delta>0} \limsup_{i\to\infty} \frac{1}{|V_i|} \log |\mathrm{Hom}_\mu(\mathcal{P}, F, \delta, \sigma_i)|_{\mathcal{Q}},$$

where the supremum is over all finite partitions $\mathcal{Q}$ that are measurable with respect to $\Phi^{-1}(\mathcal{B}_Y)$ and the infimum is over all finite partitions $\mathcal{P} \geq \mathcal{Q}$.

This equivalence shows that outer sofic entropy is a measure-conjugacy invariant in the following sense. Suppose that $\Gamma \curvearrowright^{T'} (X', \mu')$ and $\Gamma \curvearrowright^{S'} (Y', \nu')$ are pmp actions, $\Phi' : (X', \mu') \to (Y', \nu')$ is a factor map and there are measure-conjugacies $\tau_1 : X \to X'$, $\tau_2 : Y \to Y'$ that make the diagram commute:

$$\begin{array}{ccc} X & \xrightarrow{\tau_1} & X' \\ \downarrow{\scriptstyle \Phi} & & \downarrow{\scriptstyle \Phi'} \\ Y & \xrightarrow{\tau_2} & Y' \end{array}$$

then $h_{\Sigma,\mu}(\Phi) = h_{\Sigma,\mu'}(\Phi')$. For details justifying these claims see [**Hay17**, Theorem 1.20]. There is also a relative notion of outer sofic entropy developed in [**Hay16c**].

The definition implies the outer sofic entropy is bounded from above by the sofic entropies of both the target and source. If $\Gamma$ is amenable then every microstate for the

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

factor action $\Gamma \curvearrowright Y$ lifts to a microstate for the source $\Gamma \curvearrowright X$. For this reason, outer sofic entropy agrees with the classical entropy of the target whenever $\Gamma$ is amenable [**Hay17**, Appendix A]. If $\Gamma$ is non-amenable then this no longer holds. Consider the Ornstein–Weiss map

$$\mathbb{F}_2 \curvearrowright (\mathbb{Z}/2, u_2)^{\mathbb{F}_2} \to^{\Phi} \mathbb{F}_2 \curvearrowright (\mathbb{Z}/2 \times \mathbb{Z}/2, u_2 \times u_2)^{\mathbb{F}_2}.$$

Because it is finite-to-one, it can be shown that its outer sofic entropy agrees with the sofic entropy of the *source*, which is $\log(2)$ instead of the entropy of the target which is $\log(4)$.

*Question 27.* Is there a variational principle connecting outer topological sofic entropy with outer measure sofic entropy?

As the Ornstein–Weiss example shows, sofic entropy is not necessarily monotone under factor maps. However, outer sofic entropy is: if $X \to^{\Phi} Y \to^{\Psi} Z$ are factor maps then

$$h_\Sigma(\Phi) \geq h_\Sigma(\Psi \circ \Phi).$$

This is because any microstate for $\Gamma \curvearrowright Z$ which lifts to a microstate for $\Gamma \curvearrowright X$ via $\Psi \circ \Phi$ necessarily lifts to a microstate for $\Gamma \curvearrowright Y$ via $\Psi$.

11.1.1. *Outer sofic Pinsker algebras.* Let $\Gamma \curvearrowright (X, \mu)$ be a pmp action and $\Sigma$ a sofic approximation to $\Gamma$. The $\Sigma$-Pinsker algebra is the sigma-sub-algebra of $\mathcal{B}_X$ generated by all factors with zero $\Sigma$-entropy. Because of subadditivity, it has zero $\Sigma$-entropy itself. However, it does not have good monotonicity properties because there exist actions with zero $\Sigma$-entropy that factor onto actions with positive $\Sigma$-entropy (Theorem 5.1). A better alternative is the outer $\Sigma$-Pinsker algebra.

The *outer* $\Sigma$*-Pinsker algebra* of the action $\Gamma \curvearrowright (X, \mu)$ is the largest sigma-sub-algebra $\Pi^\Sigma(\mu) \subset \mathcal{B}_X$ such that the corresponding factor has outer $\Sigma$-entropy zero. Because outer $\Sigma$-entropy is monotone, any invariant sub-sigma-algebra of $\Pi^\Sigma(\mu)$ also has zero outer $\Sigma$-entropy.

It is an important classical result that the Pinsker algebra of a direct product of transformations is the direct product of the Pinsker algebras. This follows from the Rokhlin–Sinai theorem that the Pinsker algebra is the one-sided tail sigma-algebra of any generating partition. The case of general amenable groups is handled in [**GTW00**] using Sinai's factor theorem and joinings arguments. The case of sofic groups is new.

THEOREM 11.1. [**Hay16c**] *Suppose that* $\Gamma \curvearrowright (X, \mu)$, $\Gamma \curvearrowright (Y, \nu)$ *are pmp actions,* $\Sigma$ *is a sofic approximation to* $\Gamma$ *and there exist model measures* $\{\mu_n\}$, $\{\nu_n\}$ *that lde-converge to* $\mu$, $\nu$ *respectively in the sense of §6.2.5. Then* $\Pi^\Sigma(\mu \times \nu) = \Pi^\Sigma(\mu) \vee \Pi^\Sigma(\nu)$.

11.2. *Outer Rokhlin entropy.* Let $\Gamma \curvearrowright^T (X, \mu)$, $\Gamma \curvearrowright^S (Y, \nu)$ be pmp actions and suppose there is a factor map $\Phi : X \to Y$. The *outer Rokhlin entropy* of the $\Phi$ is

$$h^{\mathrm{Rok}}(\Phi) = \inf H_\mu(\mathcal{P}),$$

where the infimum is over all measurable partitions $\mathcal{P}$ of $X$ such that $\Phi^{-1}(\mathcal{B}_Y) \subset \sigma\text{-alg}_\Gamma(\mathcal{P})$ and $\mathcal{B}_Y$ is the Borel sigma-algebra of $Y$.

This outer Rokhlin entropy is bounded above by the Rokhlin entropy of the factor and by the Rokhlin entropy of the source. In the special case in which $\Gamma$ is amenable, the

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

outer Rokhlin entropy equals the Rokhlin entropy of the target because of monotonicity of entropy under factor maps.

As in the case of outer sofic entropy, the Ornstein–Weiss factor map has outer Rokhlin entropy log(2). The Ornstein–Weiss example shows Rokhlin entropy is not necessarily monotone under factor maps. However, outer Rokhlin entropy is: if $X \to^{\Phi} Y \to^{\Psi} Z$ are factor maps then

$$h^{\mathrm{Rok}}(\Phi) \geq h^{\mathrm{Rok}}(\Psi \circ \Phi).$$

This is because any partition $\mathcal{P}$ of $X$ satisfying $\Phi^{-1}(\mathcal{B}_Y) \subset \sigma\text{-alg}_\Gamma(\mathcal{P})$, also satisfies $\Phi^{-1}\Psi^{-1}(\mathcal{B}_Z) \subset \sigma\text{-alg}_\Gamma(\mathcal{P})$.

11.2.1. *Outer Rokhlin Pinsker algebras.* The *outer Rokhlin Pinsker algebra* of the action $\Gamma \curvearrowright (X, \mu)$ is the largest sigma-sub-algebra $\Pi^{\mathrm{Rok}}(\mu) \subset \mathcal{B}_X$ such that the corresponding factor has outer Rokhlin entropy zero. Because outer Rokhlin entropy is monotone, any invariant sub-sigma-algebra of $\Pi^{\mathrm{Rok}}(\mu)$ also has zero outer Rokhlin entropy. The action is said to have *completely positive outer Rokhlin entropy* if $\Pi^{\mathrm{Rok}}(\mu)$ is trivial.

THEOREM 11.2. [**Sew16b**] *If $\Gamma \curvearrowright (X, \mu)$ and $\Gamma \curvearrowright (Y, \nu)$ are essentially free pmp actions that are weakly contained in all essentially free pmp actions of $\Gamma$ then*

$$\Pi^{\mathrm{Rok}}(\mu \times \nu) = \Pi^{\mathrm{Rok}}(\mu) \vee \Pi^{\mathrm{Rok}}(\nu).$$

*Remark 27.* The notions of outer Rokhlin entropy and the theorem above admit relative versions. See [**Sew16b**] for details.

11.3. *Completely positive outer entropy.* An action has *completely positive outer $\Sigma$-entropy*, denoted $\mathrm{CPE}^{\Sigma}$, if every non-trivial factor has positive outer $\Sigma$-entropy. Because outer $\Sigma$-entropy is bounded from above by $\Sigma$-entropy, this implies that every non-trivial factor has positive $\Sigma$-entropy.

It follows from the Rokhlin–Sinai theorem that Bernoulli shifts over the integers have completely positive entropy. The case of amenable groups follows from the fact that Bernoulli shifts are uniformly mixing and uniform mixing implies CPE. This is the easy half part of [**RW00**], which proves that CPE is equivalent to uniformly mixing. The main result of [**Ker14**] is that if $\Gamma$ is any sofic group, then every non-trivial factor of a Bernoulli shift over $\Gamma$ has positive $\Sigma$-entropy (for every $\Sigma$). This is obtained via a positive lower bound on the local entropy that is uniform over all good enough sofic approximations. The proof shows more: that Bernoulli shifts over $\Gamma$ are $\mathrm{CPE}^{\Sigma}$ for every $\Sigma$. Similarly, every non-trivial factor of a Bernoulli shift over a free group has positive $f$-invariant.

In [**Hay16c**], Theorem 11.1 is used to show that a large class of algebraic actions are $\mathrm{CPE}^{\Sigma}$ (see §3.4 for details).

11.4. *Uniformly mixing.*

*Definition 18.* A sequence $\{F_i\}_{i\in\mathbb{N}}$ of finite subsets of $\Gamma$ is said to *spread out* if for every $g \in \Gamma\backslash\{1_\Gamma\}$, $g \notin F_i F_i^{-1}$ for all but finitely many $i$.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*Definition 19.* A pmp action $\Gamma \curvearrowright (X, \mu)$ is *uniformly mixing* if for every finite partition $\mathcal{P}$ of $X$ and any sequence $F_i \subset \Gamma$ of finite subsets that spread out,

$$\lim_{i\to\infty} \frac{1}{|F_i|} H_\mu\left(\bigvee_{g\in F_i} g^{-1}\mathcal{P}\right) = H_\mu(\mathcal{P}).$$

The main result of [**RW00**] is that if $\Gamma$ is amenable group then CPE implies uniform mixing. The converse was proven earlier (see [**GS00a**] or [**DGRS08**, Theorem 4.2]). In [**AB16**] a new concept, called *uniform model mixing*, that is adapted to a sofic approximation, is shown to imply completely positive le-$\Sigma$-entropy (le-$\Sigma$-entropy is defined in §6.2.5). This is used to prove that if $\Gamma$ is sofic and contains an infinite cyclic subgroup, then there exist uncountably many completely positive le-$\Sigma$-entropy actions that are not factors of Bernoulli shifts.

The goal of this section is to show that uniform mixing does not imply completely positive sofic entropy (Corollary 11.7). In fact, all mixing Markov chains over a free group are uniformly mixing (Theorem 11.3). However, with a spectral criterion due to Hayes (Theorem 11.5) we show that the Ising model with small transition probability does not have completely positive sofic entropy with respect to any sofic approximation.

11.4.1. *Markov chains.*

THEOREM 11.3. *If $\Gamma = \mathbb{F}_r$ is the free group then all mixing Markov chains with finite state space over $\Gamma$ are uniformly mixing.*

We need a lemma first. Recall that a *tree* is a simply connected graph and a *leaf* of a tree is a vertex with degree one.

LEMMA 11.4. *Let $T$ denote a finite tree with at least two vertices. For each leaf $v$ of $T$, let $T_v \subset T$ denote the smallest subtree that contains every leaf of $T$ except for $v$. Suppose that for some $n \geq 0$ every pair of distinct leaves of $T$ are at a distance $\geq n$ apart (in the path metric). Then there exists a leaf $v$ such that the distance between $v$ and $T_v$ is at least $\lceil n/2 \rceil$.*

*Proof.* † Let $u, v$ be leaves of $T$ that are a maximum distance apart. Notice that $T_v$ is the union of the paths from $u$ to $w$, where $w$ ranges over all leaves of $T$ except for $v$. Fix a leaf $w \neq v$. Let $z$ be the point on the path from $w$ to $u$ that is closest to $v$. Then $z$ must also lie on the path from $v$ to $u$. Because $u$ and $v$ are at maximum distance apart, $d(v, u) \geq d(w, u)$. Since $d(v, u) = d(v, z) + d(z, u)$ and $d(w, u) = d(w, z) + d(z, u)$, $d(v, z) \geq d(w, z)$. Therefore,

$$n \leq d(v, w) = d(v, z) + d(z, w) \leq 2d(v, z).$$

This holds for all leaves $w \neq v$ and therefore, $n \leq 2d(v, T_v)$. □

*Proof of Theorem 11.3.* Let $d(\cdot, \cdot)$ denote the word metric in $\Gamma$ with respect to a free generating set. Let $\Gamma \curvearrowright (K^\Gamma, \mu)$ be a mixing Markov chain with state space $K$. Let $\mathcal{P}$ be the time 0 partition: $\mathcal{P} = \{P_k : k \in K\}$ where

$$P_k = \{x \in K^\Gamma : x_e = k\}.$$

† This proof, which is shorter than my original, was gracefully provided by an anonymous reviewer.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

Because $\mathcal{P}$ is a Markov partition of a mixing Markov chain for every $\epsilon > 0$ there exists $n = n(\epsilon) \in \mathbb{N}$ such that if $g, h \in \Gamma$ satisfy $d(g, h) \geq n$ then $H_\mu(g^{-1}\mathcal{P}|h^{-1}\mathcal{P}) \geq H_\mu(\mathcal{P}) - \epsilon$.

Suppose $F \subset \Gamma$ is a finite subset such that the word distance between any two distinct elements $g, h \in F$ is at least $2n$. Recall that $\mathcal{P}^F = \bigvee_{f \in F} f^{-1}\mathcal{P}$. By the previous lemma, there exists a $w \in F$ such that if $T$ denotes the smallest subtree containing $F - \{w\}$ then $d(T, w) \geq n$. Let $g \in T$ minimize the distance to $w$. The Markov property implies

$$\begin{aligned} H_\mu(\mathcal{P}^F) &= H_\mu(w^{-1}\mathcal{P}|\mathcal{P}^{F-\{w\}}) + H_\mu(\mathcal{P}^{F-\{w\}}) \geq H_\mu(w^{-1}\mathcal{P}|\mathcal{P}^T) + H_\mu(\mathcal{P}^{F-\{w\}}) \\ &= H_\mu(w^{-1}\mathcal{P}|g^{-1}\mathcal{P}) + H_\mu(\mathcal{P}^{F-\{w\}}) \geq H_\mu(\mathcal{P}) + H_\mu(\mathcal{P}^{F-\{w\}}) - \epsilon. \end{aligned}$$

So by induction on $|F|$ we obtain, $H_\mu(\mathcal{P}^F) \geq |F|(H_\mu(\mathcal{P}) - \epsilon)$. This implies uniform mixing with respect to $\mathcal{P}$.

Now let $\mathcal{Q}$ be a partition that is contained in $\mathcal{P}$. Note that if $g, h \in \Gamma$ satisfy $d(g, h) \geq n$ then $H_\mu(g^{-1}\mathcal{Q}|h^{-1}\mathcal{P}) \geq H_\mu(\mathcal{Q}) - \epsilon$. Therefore, if $F$, $T$, $w$, $g$ are as above then

$$\begin{aligned} H_\mu(\mathcal{Q}^F) &= H_\mu(w^{-1}\mathcal{Q}|\mathcal{Q}^{F-\{w\}}) + H_\mu(\mathcal{Q}^{F-\{w\}}) \geq H_\mu(w^{-1}\mathcal{Q}|\mathcal{Q}^T) + H_\mu(\mathcal{Q}^{F-\{w\}}) \\ &= H_\mu(w^{-1}\mathcal{Q}|g^{-1}\mathcal{Q}) + H_\mu(\mathcal{Q}^{F-\{w\}}) \geq H_\mu(\mathcal{Q}) + H_\mu(\mathcal{Q}^{F-\{w\}}) - \epsilon. \end{aligned}$$

Uniform mixing with respect to $\mathcal{Q}$ now follows from induction on $|F|$.

Observe that for any finite $K \subset \Gamma$ with the property that the induced subgraph of the Cayley graph is connected, the partition $\mathcal{P}^K$ is Markov. So the previous paragraphs imply the following: for any $\mathcal{Q} \leq \mathcal{P}^K$, the action $\Gamma \curvearrowright (X, \mu)$ is uniformly mixing with respect to $\mathcal{Q}$.

Now let $\mathcal{R}$ be an arbitrary partition. Let $\delta > 0$. Because the partition $\mathcal{P}$ is generating, there exists a partition $\mathcal{Q} \leq \mathcal{P}^K$ (for some finite $K \subset \Gamma$) such that

$$d(\mathcal{Q}, \mathcal{R}) := H_\mu(\mathcal{Q}|\mathcal{R}) + H_\mu(\mathcal{R}|\mathcal{Q}) < \delta,$$

where $d(\cdot, \cdot)$ denotes the Rokhlin distance on partitions (as defined in §4.4).

It follows that $d(\mathcal{Q}^F, \mathcal{R}^F) < \delta|F|$ for any finite $F$. Therefore if $\{F_i\}$ spreads out then because the action is uniformly mixing with respect to $\mathcal{Q}$,

$$\liminf_i \frac{1}{|F_i|} H_\mu(\mathcal{R}^{F_i}) \geq \liminf_i \frac{1}{|F_i|} H_\mu(\mathcal{Q}^{F_i}) - \delta = H_\mu(\mathcal{Q}) - \delta \geq H_\mu(\mathcal{R}) - 2\delta.$$

Because $\delta > 0$ is arbitrary, this implies the theorem. □

11.4.2. *A spectral criterion for CPE.*

THEOREM 11.5. [**Hay18**, Corollary 1.4] *Let $\Gamma \curvearrowright (X, \mu)$ be ergodic and let $\kappa^0 : \Gamma \to U(L^2_0(X, \mu))$ be the Koopman representation on the orthogonal complement of the constants. If $\Gamma \curvearrowright (X, \mu)$ is CPE with respect to some sofic approximation then $\kappa^0$ embeds into the countable sum of the left regular representation of $\Gamma$.*

*Remark 28.* The Rokhlin entropy version of the above theorem is Corollary 7.6.

COROLLARY 11.6. *Let $\mathbb{F}_r = \langle s_1, \ldots, s_r \rangle$ denote the rank $r \geq 2$ free group. Let $\Delta \in \mathbb{C}\Gamma$ denote the 'discrete Lapacian'*

$$\Delta = \frac{1}{2r} \sum_{i=1}^{r} s_i + s_i^{-1}.$$

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

*There is some $\epsilon_r > 0$ such that if $\Gamma \curvearrowright (X, \mu)$ is CPE with respect to some sofic approximation then*

$$\|\kappa_0(\Delta)\| \leq 1 - \epsilon_r.$$

*Proof.* The operator norm of $\Delta$, considered as an operator on $\ell^2(\Gamma)$ is $1 - \epsilon_r$ for some $\epsilon_r > 0$ by Kesten's criterion [**BdlHV08**, Theorem G.4.4]. Therefore, the operator norm of $\Delta$, considered as an operator on $\ell^2(\Gamma)^{\oplus\infty}$ is also $1 - \epsilon_r$. So this corollary follows from Theorem 11.5. □

COROLLARY 11.7. *Let $\mathbb{F}_2 = \langle s_1, \ldots, s_r \rangle$ denote the rank r free group. If $\epsilon > 0$ is sufficiently small then the Ising model (see Example 3 §3.3) is not CPE with respect to any sofic approximation.*

*Proof.* Let $\mu_\epsilon \in \text{Prob}(\{0, 1\}^{\mathbb{F}_r})$ denote the law of the Markov chain $\mathbf{X} = (X_g)_{g \in \mathbb{F}_r}$ with state space $\{-1, 1\}$ satisfying

$$P(X_e = -1) = P(X_e = 1) = 1/2$$
$$P(X_e = k | X_s = k) = 1 - \epsilon, \quad P(X_e \neq k | X_s = k) = \epsilon$$

for all $s \in S := \{s_1, \ldots, s_r\} \cup \{s_1^{-1}, \ldots, s_r^{-1}\}$. Then $\mathbb{E}[X_e] = 0$, $\mathbb{E}[X_e^2] = 1$ and $\mathbb{E}[\Delta(X_e)X_e] = 1 - 2\epsilon$. This last computation requires some explanation: we consider $X_e \in L_0^2(\mu_\epsilon)$. Then

$$\Delta(X_e) = |S|^{-1} \sum_{s \in S} X_e \circ \kappa_0(s) = |S|^{-1} \sum_{s \in S} X_s.$$

Because $\langle X_e, X_s \rangle = (1 - \epsilon) - \epsilon = 1 - 2\epsilon$ for all $s \in S$, it follows that $\langle \Delta(X_e), X_e \rangle = \mathbb{E}[\Delta(X_e)X_e] = 1 - 2\epsilon$. Thus $\|\kappa_0(\Delta)\| \geq 1 - 2\epsilon$. The previous corollary now implies this one. □

*Acknowledgements.* I am most grateful for discussions with Tim Austin, Peter Burton, Ben Hayes, David Kerr, Hanfeng Li, Sorin Popa, Brandon Seward, Jean-Paul Thouvenot, Robin Tucker-Drob and Benjy Weiss. Many of these researchers have contributed examples which appear in this article, some for the first time. Supported in part by NSF grant DMS-0968762, NSF CAREER Award DMS-0954606 and BSF grant 2008274.

## REFERENCES

[**AB16**] T. Austin and P. Burton. Uniform mixing and completely positive sofic entropy. *J. Anal. Math.*, to appear.

[**AKM65**] R. L. Adler, A. G. Konheim and M. H. McAndrew. Topological entropy. *Trans. Amer. Math. Soc.* **114** (1965), 309–319.

[**Alp15**] A. Alpeev. The entropy of Gibbs measures on sofic groups. *Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI)* **436** (2015), 34–48 (Teoriya Predstavlenii, Dinamicheskie Sistemy, Kombinatornye Metody. XXV).

[**AR62**] L. M. Abramov and V. A. Rohlin. Entropy of a skew product of mappings with invariant measure. *Vestnik Leningrad. Univ.* **17**(7) (1962), 5–13.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

[**AS16**] A. Alpeev and B. Seward. Krieger's finite generator theorem for ergodic actions of countable groups III. *Preprint*, 2016.
[**Aus16a**] T. Austin. Additivity properties of sofic entropy and measures on model spaces. *Forum Math. Sigma* **4** (2016), e25, 79.
[**Aus16b**] T. Austin. Behaviour of entropy under bounded and integrable orbit equivalence. *Geom. Funct. Anal.* **26**(6) (2016), 1483–1525.
[**Aus16c**] T. Austin. The geometry of model spaces for probability-preserving actions of sofic groups. *Anal. Geom. Metr. Spaces* **4** (2016), Art. 6.
[**AW13**] M. Abért and B. Weiss. Bernoulli actions are weakly contained in any free action. *Ergod. Th. & Dynam. Sys.* **33**(2) (2013), 323–333.
[**Ax68**] J. Ax. The elementary theory of finite fields. *Ann. of Math. (2)* **88** (1968), 239–271.
[**Bal05**] K. Ball. Factors of independent and identically distributed processes with non-amenable group actions. *Ergod. Th. & Dynam. Sys.* **25**(3) (2005), 711–730.
[**BdlHV08**] B. Bekka, P. de la Harpe and A. Valette. *Kazhdan's Property (T)*. Cambridge University Press, New York, 2008.
[**BG14**] L. Bowen and Y. Gutman. A Juzvinskii addition theorem for finitely generated free group actions. *Ergod. Th. & Dynam. Sys.* **34**(1) (2014), 95–109.
[**BK17**] L. Bartholdi and D. Kielak. Amenability of groups is characterized by Myhill's Theorem. *J. Eur. Math. Soc.*, to appear.
[**BL12**] L. Bowen and H. Li. Harmonic models and spanning forests of residually finite groups. *J. Funct. Anal.* **263**(7) (2012), 1769–1808.
[**BM97**] M. Burger and S. Mozes. Finitely presented simple groups and products of trees. *C. R. Acad. Sci. Paris Sér. I Math.* **324**(7) (1997), 747–752.
[**BM00**] M. Burger and S. Mozes. Lattices in product of trees. *Publ. Math. Inst. Hautes Études Sci.* (92) (2000), 151–194, 2001.
[**Bow08**] R. Bowen. *Equilibrium States and the Ergodic Theory of Anosov Diffeomorphisms (Lecture Notes in Mathematics, 470)* revised edn.. Springer, Berlin, 2008. With a preface by David Ruelle, Edited by Jean-René Chazottes.
[**Bow09**] L. Bowen. Free groups in lattices. *Geom. Topol.* **13**(5) (2009), 3021–3054.
[**Bow10a**] L. Bowen. The ergodic theory of free group actions: entropy and the $f$-invariant. *Groups Geom. Dyn.* **4**(3) (2010), 419–432.
[**Bow10b**] L. Bowen. Measure conjugacy invariants for actions of countable sofic groups. *J. Amer. Math. Soc.* **23**(1) (2010), 217–245.
[**Bow10c**] L. Bowen. Non-abelian free group actions: Markov processes, the Abramov–Rohlin formula and Yuzvinskii's formula. *Ergod. Th. & Dynam. Sys.* **30**(6) (2010), 1629–1663.
[**Bow10d**] L. P. Bowen. A measure-conjugacy invariant for free group actions. *Ann. of Math. (2)* **171**(2) (2010), 1387–1400.
[**Bow11a**] L. Bowen. Entropy for expansive algebraic actions of residually finite groups. *Ergod. Th. & Dynam. Sys.* **31**(3) (2011), 703–718.
[**Bow11b**] L. Bowen. Weak isomorphisms between Bernoulli shifts. *Israel J. Math.* **183** (2011), 93–102.
[**Bow12a**] L. Bowen. Every countably infinite group is almost Ornstein. *Dynamical Systems and Group Actions (Contemporary Mathematics, 567)*. American Mathematical Society, Providence, RI, 2012, pp. 67–78.
[**Bow12b**] L. Bowen. Sofic entropy and amenable groups. *Ergod. Th. & Dynam. Sys.* **32**(2) (2012), 427–466.
[**Bow14**] L. Bowen. Entropy theory for sofic groupoids I: the foundations. *J. Anal. Math.* **124** (2014), 149–233.
[**Bow16**] L. Bowen. Zero entropy is generic. *Entropy* **18**(6) (2016), Paper No. 220, 20.
[**Bow17**] L. Bowen. Finitary random interlacements and the Gaboriau–Lyons problem. *Preprint*, 2017, arXiv:1707.09573.
[**BS94**] R. Burton and J. E. Steif. Non-uniqueness of measures of maximal entropy for subshifts of finite type. *Ergod. Th. & Dynam. Sys.* **14**(2) (1994), 213–235.
[**BS95**] R. Burton and J. E. Steif. New results on measures of maximal entropy. *Israel J. Math.* **89**(1–3) (1995), 275–300.
[**BS01**] I. Benjamini and O. Schramm. Recurrence of distributional limits of finite planar graphs. *Electron. J. Probab.* **6**(23) (2001), 13 (electronic).
[**Bur17**] P. Burton. Naive entropy of dynamical systems. *Israel J. Math.* **219**(2) (2017), 637–659.
[**CHR14**] L. Ciobanu, D. F. Holt and S. Rees. Sofic groups: graph products and graphs of groups. *Pacific J. Math.* **271**(1) (2014), 53–64.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

[**Chu13**] N.-P. Chung. Topological pressure and the variational principle for actions of sofic groups. *Ergod. Th. & Dynam. Sys.* **33**(5) (2013), 1363–1390.
[**CL15a**] V. Capraro and M. Lupini. *Introduction to Sofic and Hyperlinear Groups and Connes' Embedding Conjecture (Lecture Notes in Mathematics, 2136)*. Springer, Cham, 2015. With an appendix by Vladimir Pestov.
[**CL15b**] N.-P. Chung and H. Li. Homoclinic groups, IE groups, and expansive algebraic actions. *Invent. Math.* **199**(3) (2015), 805–858.
[**Cor11**] Y. Cornulier. A sofic group away from amenable groups. *Math. Ann.* **2** (2011), 269–275.
[**CZ15**] N.-P. Chung and G. Zhang. Weak expansiveness for actions of sofic groups. *J. Funct. Anal.* **268**(11) (2015), 3534–3565.
[**Dan01**] A. I. Danilenko. Entropy theory from the orbital point of view. *Monatsh. Math.* **134**(2) (2001), 121–141.
[**Den06**] C. Deninger. Fuglede–Kadison determinants and entropy for actions of discrete amenable groups. *J. Amer. Math. Soc.* **19**(3) (2006), 737–758 (electronic).
[**DG02**] A. H. Dooley and V. Ya. Golodets. The spectrum of completely positive entropy actions of countable amenable groups. *J. Funct. Anal.* **196**(1) (2002), 1–18.
[**DGRS08**] A. H. Dooley, V. Ya. Golodets, D. J. Rudolph and S. D. Sinel'shchikov. Non-Bernoulli systems with completely positive entropy. *Ergod. Th. & Dynam. Sys.* **28**(1) (2008), 87–124.
[**Din70**] E. I. Dinaburg. A correlation between topological entropy and metric entropy. *Dokl. Akad. Nauk SSSR* **190** (1970), 19–22.
[**DKP14**] K. Dykema, D. Kerr and M. Pichot. Sofic dimension for discrete measured groupoids. *Trans. Amer. Math. Soc.* **366**(2) (2014), 707–748.
[**DM10a**] A. Dembo and A. Montanari. Gibbs measures and phase transitions on sparse random graphs. *Braz. J. Probab. Stat.* **24**(2) (2010), 137–211.
[**DM10b**] A. Dembo and A. Montanari. Ising models on locally tree-like graphs. *Ann. Appl. Probab.* **20**(2) (2010), 565–592.
[**Dow11**] T. Downarowicz. *Entropy in Dynamical Systems (New Mathematical Monographs, 18)*. Cambridge University Press, Cambridge, 2011.
[**DP02**] A. I. Danilenko and K. K. Park. Generators and Bernoullian factors for amenable actions and cocycles on their orbits. *Ergod. Th. & Dynam. Sys.* **22**(6) (2002), 1715–1745.
[**DS07**] C. Deninger and K. Schmidt. Expansive algebraic actions of discrete residually finite amenable groups and their entropy. *Ergod. Th. & Dynam. Sys.* **27**(3) (2007), 769–786.
[**Dye59**] H. A. Dye. On groups of measure preserving transformation. I. *Amer. J. Math.* **81** (1959), 119–159.
[**Dye63**] H. A. Dye. On groups of measure preserving transformations. II. *Amer. J. Math.* **85** (1963), 551–576.
[**Eps08**] I. Epstein. Orbit inequivalent actions of non-amenable groups. *Preprint*, 2008, arXiv:0707.4215.
[**ES04**] G. Elek and E. Szabó. Sofic groups and direct finiteness. *J. Algebra* **280**(2) (2004), 426–434.
[**ES05**] G. Elek and E. Szabó. Hyperlinearity, essentially free actions and $L^2$-invariants. The sofic property. *Math. Ann.* **332**(2) (2005), 421–441.
[**ES06**] G. Elek and E. Szabó. On sofic groups. *J. Group Theory* **9**(2) (2006), 161–171.
[**ES11**] G. Elek and E. Szabó. Sofic representations of amenable groups. *Proc. Amer. Math. Soc.* **139**(12) (2011), 4285–4291.
[**ES12**] G. Elek and B. Szegedy. A measure-theoretic approach to the theory of dense hypergraphs. *Adv. Math.* **231**(3–4) (2012), 1731–1772.
[**FO70**] N. A. Friedman and D. S. Ornstein. On isomorphism of weak Bernoulli transformations. *Adv. Math.* **5**(1970) (1970), 365–394.
[**FW04**] M. Foreman and B. Weiss. An anti-classification theorem for ergodic measure preserving transformations. *J. Eur. Math. Soc. (JEMS)* **6**(3) (2004), 277–292.
[**Gab17**] D. Gaboriau. Entropie sofique. *Astérisque* **390** (2017), 101–138, Exp. No. 1108, Séminaire Bourbaki 2015/2016. Exposés 1104–1119.
[**Geo11**] H.-O. Georgii. *Gibbs Measures and Phase Transitions (de Gruyter Studies in Mathematics, 9)*, 2nd edn. Walter de Gruyter & Co, Berlin, 2011.
[**GL09**] D. Gaboriau and R. Lyons. A measurable-group-theoretic solution to von Neumann's problem. *Invent. Math.* **177**(3) (2009), 533–540.
[**Gla03**] E. Glasner. *Ergodic Theory Via Joinings (Mathematical Surveys and Monographs, 101)*. American Mathematical Society, Providence, RI, 2003.
[**Goo69**] L. Wayne Goodwyn. Topological entropy bounds measure-theoretic entropy. *Proc. Amer. Math. Soc.* **23** (1969), 679–688.
[**Goo71**] T. N. T. Goodman. Relating topological entropy and measure entropy. *Bull. Lond. Math. Soc.* **3** (1971), 176–180.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

[**Goo72**] L. Wayne Goodwyn. Comparing topological entropy with measure-theoretic entropy. *Amer. J. Math.* **94** (1972), 366–388.
[**Gro99**] M. Gromov. Endomorphisms of symbolic algebraic varieties. *J. Eur. Math. Soc. (JEMS)* **1**(2) (1999), 109–197.
[**GS00a**] V. Ya. Golodets and S. D. Sinel'shchikov. Complete positivity of entropy and non-Bernoullicity for transformation groups. *Colloq. Math.* **84/85**(part 2) (2000), 421–429. Dedicated to the memory of Anzelm Iwanik.
[**GS00b**] G. Greschonig and K. Schmidt. Ergodic decomposition of quasi-invariant probability measures. *Colloq. Math.* **84/85**(part 2) (2000), 495–514. Dedicated to the memory of Anzelm Iwanik.
[**GS15**] D. Gaboriau and B. Seward. Cost, $\ell^2$-Betti numbers, and the sofic entropy of some algebraic actions. *J. Anal. Math. Preprint*, 2015, arXiv:1509.02482.
[**GTW00**] E. Glasner, J.-P. Thouvenot and B. Weiss. Entropy theory without a past. *Ergod. Th. & Dynam. Sys.* **20**(5) (2000), 1355–1370.
[**Hay17**] B. Hayes. Mixing and spectral gap relative to Pinsker factors for sofic groups. *Proc. 2014 Maui and 2015 Qinhuangdao Conferences in Honour of Vaughan F. R. Jones' 60th Birthday (Proc. Centre Math. Appl. Austral. Nat. Univ., 46)*. Australian National University, Canberra, 2017, pp. 193–221.
[**Hay16a**] B. Hayes. Local and doubly empirical convergence and the entropy of algebraic actions of sofic groups. *Ergod. Th. & Dynam. Sys.*, to appear. *Preprint*, 2016, arXiv:1603.06450.
[**Hay16b**] B. Hayes. Fuglede–Kadison determinants and sofic entropy. *Geom. Funct. Anal.* **26**(2) (2016), 520–606.
[**Hay16c**] B. Hayes. Relative entropy and the Pinsker product formula for sofic groups. *Preprint*, 2016, arXiv: 1605.01747.
[**Hay17a**] B. Hayes. Independence tuples and Deninger's problem. *Groups Geom. Dyn.* **11**(1) (2017), 245–289.
[**Hay17b**] B. Hayes. Sofic entropy of Gaussian actions. *Ergod. Th. & Dynam. Sys.* **37**(7) (2017), 2187–2222.
[**Hay18**] B. Hayes. Polish models and sofic entropy. *J. Inst. Math. Jussieu* **17**(2) (2018), 241–275.
[**HLS14**] H. Hatami, L. Lovász and B. Szegedy. Limits of locally-globally convergent graph sequences. *Geom. Funct. Anal.* **24**(1) (2014), 269–296.
[**Hof11**] C. Hoffman. Subshifts of finite type which have completely positive entropy. *Discrete Contin. Dyn. Syst.* **29**(4) (2011), 1497–1516.
[**HS16**] B. Hayes and A. Sale. The wreath product of two sofic groups is sofic. *Preprint*, 2016, arXiv:1601. 03286.
[**HS18**] B. Hayes and A. W. Sale. Metric approximations of wreath products. *Ann. Inst. Fourier (Grenoble)* **68**(1) (2018), 423–455.
[**IKT09**] A. Ioana, A. S. Kechris and T. Tsankov. Subequivalence relations and positive-definite functions. *Groups Geom. Dyn.* **3**(4) (2009), 579–625.
[**Juz65a**] S. A. Juzvinskiĭ. Metric properties of automorphisms of locally compact commutative groups. *Sibirsk. Mat. Ž.* **6** (1965), 244–247.
[**Juz65b**] S. A. Juzvinskiĭ. Metric properties of the endomorphisms of compact groups. *Izv. Akad. Nauk SSSR Ser. Mat.* **29** (1965), 1295–1328.
[**Kec10**] A. S. Kechris. *Global Aspects of Ergodic Group Actions (Mathematical Surveys and Monographs, 160)*. American Mathematical Society, Providence, RI, 2010.
[**Kel98**] G. Keller. *Equilibrium States in Ergodic Theory (London Mathematical Society Student Texts, 42)*. Cambridge University Press, Cambridge, 1998.
[**Ker13**] D. Kerr. Sofic measure entropy via finite partitions. *Groups Geom. Dyn.* **7**(3) (2013), 617–632.
[**Ker14**] D. Kerr. Bernoulli actions of sofic groups have completely positive entropy. *Israel J. Math.* **202**(1) (2014), 461–474.
[**Key70**] H. B. Keynes. Lifting of topological entropy. *Proc. Amer. Math. Soc.* **24** (1970), 440–445.
[**KH95**] A. Katok and B. Hasselblatt. *Introduction to the Modern Theory of Dynamical Systems (Encyclopedia of Mathematics and its Applications, 54)*. Cambridge University Press, Cambridge, 1995, with a supplementary chapter by Katok and Leonardo Mendoza.
[**Kid08**] Y. Kida. Orbit equivalence rigidity for ergodic actions of the mapping class group. *Geom. Dedicata* **131** (2008), 99–109.
[**Kie75**] J. C. Kieffer. A generalized Shannon–McMillan theorem for the action of an amenable group on a probability space. *Ann. Probab.* **3**(6) (1975), 1031–1037.
[**KL11a**] D. Kerr and H. Li. Bernoulli actions and infinite entropy. *Groups Geom. Dyn.* **5**(3) (2011), 663–672.
[**KL11b**] D. Kerr and H. Li. Entropy and the variational principle for actions of sofic groups. *Invent. Math.* **186**(3) (2011), 501–558.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

[**KL13a**] D. Kerr and H. Li. Combinatorial independence and sofic entropy. *Commun. Math. Stat.* **1**(2) (2013), 213–257.
[**KL13b**] D. Kerr and H. Li. Soficity, amenability, and dynamical entropy. *Amer. J. Math.* **135**(3) (2013), 721–761.
[**KL16**] D. Kerr and H. Li. *Ergodic Theory (Springer Monographs in Mathematics)*. Springer, Cham, 2016, Independence and dichotomies.
[**Kol58**] A. N. Kolmogorov. A new metric invariant of transient dynamical systems and automorphisms in Lebesgue spaces. *Dokl. Akad. Nauk SSSR (N.S.)* **119** (1958), 861–864.
[**Kol59**] A. N. Kolmogorov. Entropy per unit time as a metric invariant of automorphisms. *Dokl. Akad. Nauk SSSR* **124** (1959), 754–755.
[**Kri70**] W. Krieger. On entropy and generators of measure-preserving transformations. *Trans. Amer. Math. Soc.* **149** (1970), 453–464.
[**Kun16**] G. Kun. On sofic approximations of property (T) groups. *Preprint*, 2016, arXiv:1606.04471.
[**Li12**] H. Li. Compact group automorphisms, addition formulas and Fuglede–Kadison determinants. *Ann. of Math. (2)* **176**(1) (2012), 303–347.
[**LL16**] H. Li and B. Liang. Sofic mean length. *Adv. Math.*, to appear.
[**LSW90**] D. Lind, K. Schmidt and T. Ward. Mahler measure and entropy for commuting automorphisms of compact groups. *Invent. Math.* **101**(3) (1990), 593–629.
[**LT14**] H. Li and A. Thom. Entropy, determinants, and $L^2$-torsion. *J. Amer. Math. Soc.* **27**(1) (2014), 239–292.
[**Mal40**] A. Malcev. On isomorphic matrix representations of infinite groups. *Rec. Math. [Mat. Sbornik] N.S.* **8**(50) (1940), 405–422.
[**MB09**] R. Miles and M. Björklund. Entropy range problems and actions of locally normal groups. *Discrete Contin. Dyn. Syst.* **25**(3) (2009), 981–989.
[**Mey16**] T. Meyerovitch. Positive sofic entropy implies finite stabilizer. *Entropy* **18**(7) (2016), 14, Paper No. 263.
[**Mil08**] R. Miles. The entropy of algebraic actions of countable torsion-free abelian groups. *Fund. Math.* **201**(3) (2008), 261–282.
[**Mis76**] M. Misiurewicz. A short proof of the variational principle for a $\mathbf{Z}_+^N$ action on a compact space. *Astérisque* **40** (1976), 147–157.
[**MO85**] J. Moulin Ollagnier. *Ergodic Theory and Statistical Mechanics (Lecture Notes in Mathematics, 1115)*. Springer, Berlin, 1985.
[**Ol′91**] A. Yu. Ol'shanskiĭ. *Geometry of Defining Relations in Groups (Mathematics and its Applications (Soviet Series), 70)*. Kluwer Academic Publishers Group, Dordrecht, 1991, translated from the 1989 Russian original by Yu. A. Bakhturin.
[**OP10**] N. Ozawa and S. Popa. On a class of $\mathrm{II}_1$ factors with at most one Cartan subalgebra. *Ann. of Math. (2)* **172**(1) (2010), 713–749.
[**Orn70a**] D. Ornstein. Bernoulli shifts with the same entropy are isomorphic. *Adv. Math.* **4**(1970) (1970), 337–352.
[**Orn70b**] D. Ornstein. Factors of Bernoulli shifts are Bernoulli shifts. *Adv. Math.* **5**(1970) (1970), 349–364.
[**Orn70c**] D. Ornstein. Two Bernoulli shifts with infinite entropy are isomorphic. *Adv. Math.* **5**(1970) (1970), 339–348.
[**OW80**] D. S. Ornstein and B. Weiss. Ergodic theory of amenable group actions. I. The Rohlin lemma. *Bull. Amer. Math. Soc. (N.S.)* **2**(1) (1980), 161–164.
[**OW87**] D. S. Ornstein and B. Weiss. Entropy and isomorphism theorems for actions of amenable groups. *J. Anal. Math.* **48** (1987), 1–141.
[**OW07**] D. Ornstein and B. Weiss. Entropy is the only finitely observable invariant. *J. Mod. Dyn.* **1**(1) (2007), 93–105.
[**Pău11**] L. Păunescu. On sofic actions and equivalence relations. *J. Funct. Anal.* **261**(9) (2011), 2461–2485.
[**Pes08**] V. G. Pestov. Hyperlinear and sofic groups: a brief guide. *Bull. Symbolic Logic* **14**(4) (2008), 449–480.
[**Pet89**] K. Petersen. *Ergodic Theory (Cambridge Studies in Advanced Mathematics, 2)*. Cambridge University Press, Cambridge, 1989, corrected reprint of the 1983 original.
[**PK09**] V. G. Pestov and A. Kwiatkowska. An introduction to hyperlinear and sofic groups. *Appalachian Set Theory (London Mathematical Society Lecture Notes Series)*, to appear. *Preprint*, 2009, arXiv: 0911.4266v2.
[**Pop06a**] S. Popa. Some computations of 1-cohomology groups and construction of non-orbit-equivalent actions. *J. Inst. Math. Jussieu* **5**(2) (2006), 309–332.
[**Pop06b**] S. Popa. Strong rigidity of $\mathrm{II}_1$ factors arising from malleable actions of $w$-rigid groups. II. *Invent. Math.* **165**(2) (2006), 409–451.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press

[**Pop07**] S. Popa. Cocycle and orbit equivalence superrigidity for malleable actions of $w$-rigid groups. *Invent. Math.* **170**(2) (2007), 243–295.
[**Pop08**] S. Popa. On the superrigidity of malleable actions with spectral gap. *J. Amer. Math. Soc.* **21**(4) (2008), 981–1000.
[**PS07**] S. Popa and R. Sasyk. On the cohomology of Bernoulli actions. *Ergod. Th. & Dynam. Sys.* **27**(1) (2007), 241–251.
[**Roh67**] V. A. Rohlin. Lectures on the entropy theory of transformations with invariant measure. *Uspekhi Mat. Nauk* **22**(5 (137)) (1967), 3–56.
[**Rud78**] D. J. Rudolph. If a finite extension of a Bernoulli shift has no finite rotation factors, it is Bernoulli. *Israel J. Math.* **30**(3) (1978), 193–206.
[**Rud90**] D. J. Rudolph. *Fundamentals of Measurable Dynamics (Oxford Science Publications)*. The Clarendon Press, Oxford University Press, New York, 1990, Ergodic theory on Lebesgue spaces.
[**RW00**] D. J. Rudolph and B. Weiss. Entropy and mixing for amenable group actions. *Ann. of Math. (2)* **151**(3) (2000), 1119–1150.
[**Sch95**] K. Schmidt. *Dynamical Systems of Algebraic Origin (Modern Birkhäuser Classics)*. Birkhäuser/Springer Basel AG, Basel, 1995, 2011 reprint of the 1995 original.
[**Sew14a**] B. Seward. Every action of a nonamenable group is the factor of a small action. *J. Mod. Dyn.* **8**(2) (2014), 251–270.
[**Sew14b**] B. Seward. A subgroup formula for $f$-invariant entropy. *Ergod. Th. & Dynam. Sys.* **34**(1) (2014), 263–298.
[**Sew15a**] B. Seward. Ergodic actions of countable groups and finite generating partitions. *Groups Geom. Dyn.* **9**(3) (2015), 793–810.
[**Sew15b**] B. Seward. Krieger's finite generator theorem for ergodic actions of countable groups II. *J. Mod. Dynam.*, to appear.
[**Sew16a**] B. Seward. Finite entropy actions of free groups, rigidity of stabilizers, and a Howe–Moore type phenomenon. *J. Anal. Math.* **129** (2016), 309–340.
[**Sew16b**] B. Seward. Weak containment and Rokhlin entropy. *Preprint*, 2016, arXiv:1602.06680.
[**Sew18**] B. Seward. Positive entropy actions of countable groups factor onto Bernoulli shifts. *J. Amer. Math. Soc.*, to appear.
[**Sew19**] B. Seward. Krieger's finite generator theorem for actions of countable groups I. *Invent. Math.* **215**(1) (2019), 265–310.
[**Sin64**] Ja. G. Sinaĭ. On a weak isomorphism of transformations with invariant measure. *Mat. Sb. (N.S.)* **63**(105) (1964), 23–42.
[**ST75**] P. Shields and J.-P. Thouvenot. Entropy zero $\times$ Bernoulli processes are closed in the $\bar{d}$-metric. *Ann. Probab.* **3**(4) (1975), 732–736.
[**STD16**] B. Seward and R. D. Tucker-Drob. Borel structurability on the 2-shift of a countable group. *Ann. Pure Appl. Logic* **167**(1) (2016), 1–21.
[**Ste75**] A. M. Stepin. Bernoulli shifts on groups. *Dokl. Akad. Nauk SSSR* **223**(2) (1975), 300–302.
[**Tho71**] R. K. Thomas. The addition theorem for the entropy of transformations of $G$-spaces. *Trans. Amer. Math. Soc.* **160** (1971), 119–130.
[**Tho75**] J.-P. Thouvenot. Quelques propriétés des systèmes dynamiques qui se décomposent en un produit de deux systèmes dont l'un est un schéma de Bernoulli. *Israel J. Math* **21**(2–3) (1975), 177–207. Conference on Ergodic Theory and Topological Dynamics (Kibbutz, Lavi, 1974).
[**Tho10**] A. Thom. Examples of hyperlinear groups without factorization property. *Groups Geom. Dyn.* **4**(1) (2010), 195–208.
[**Wei00**] B. Weiss. Sofic groups and dynamical systems. *Sankhyā Ser. A* **62**(3) (2000), 350–359. Conference on Ergodic Theory and Harmonic Analysis (Mumbai, 1999).
[**Wei03**] B. Weiss. Actions of amenable groups. *Topics in Dynamics and Ergodic Theory (London Mathematical Society Lecture Note Series, 310)*. Cambridge University Press, Cambridge, 2003, pp. 226–262.
[**Wei15**] B. Weiss. Entropy and actions of sofic groups. *Discrete Contin. Dyn. Syst. Ser. B* **20**(10) (2015), 3375–3383.
[**WZ92**] T. Ward and Q. Zhang. The Abramov–Rokhlin entropy addition formula for amenable group actions. *Monatsh. Math.* **114**(3–4) (1992), 317–329.
[**Zha12**] G. Zhang. Local variational principle concerning entropy of sofic group action. *J. Funct. Anal.* **262**(4) (2012), 1954–1985.
[**Zim84**] R. J. Zimmer. *Ergodic Theory and Semisimple Groups (Monographs in Mathematics, 81)*. Birkhäuser, Basel, 1984.

https://doi.org/10.1017/etds.2019.18 Published online by Cambridge University Press