# UNIFORM MIXING AND COMPLETELY POSITIVE SOFIC ENTROPY

*By*

TIM AUSTIN AND PETER BURTON

**Abstract.** Let $G$ be a countable discrete sofic group. We define a concept of uniform mixing for measure-preserving $G$-actions and show that it implies completely positive sofic entropy. When $G$ contains an element of infinite order, we use this to produce an uncountable family of pairwise nonisomorphic $G$-actions with completely positive sofic entropy. None of our examples is a factor of a Bernoulli shift.

## 1 Introduction

Let $G$ be a countable discrete sofic group, $(X, \mu)$ a standard probability space and $T : G \curvearrowright X$ a measurable $G$-action preserving $\mu$. In [2], Lewis Bowen defined the sofic entropy of $(X, \mu, T)$ relative to a sofic approximation under the hypothesis that the action admits a finite generating partition. The definition was extended to general $(X, \mu, T)$ by Kerr and Li in [10] and Kerr gave a more elementary approach in [8]. In [3] Bowen showed that when $G$ is amenable, sofic entropy relative to any sofic approximation agrees with the standard Kolmogorov–Sinai entropy. Despite some notable successes such as the proof in [2] that Bernoulli shifts with distinct base-entropies are nonisomorphic, many aspects of the theory of sofic entropy are still relatively undeveloped.

Rather than work with abstract measure-preserving $G$-actions, we will use the formalism of $G$-processes. If $G$ is a countable group and $A$ is a standard Borel space, we will endow $A^G$ with the right-shift action given by $(g \cdot a)(h) = a(hg)$ for $g, h \in G$ and $a \in A^G$. A $G$-process over $A$ is a Borel probability measure $\mu$ on $A^G$ which is invariant under this action. Any measure-preserving action of $G$ on a standard probability space is measure-theoretically isomorphic to a $G$-process over some standard Borel space $A$. We will assume the state space $A$ is finite, which corresponds to the case of measure-preserving actions which admit a finite generating partition. Note that by results of Seward from [12] and [13], the last



JOURNAL D'ANALYSE MATHÉMATIQUE, Vol. 138 (2019)
DOI 10.1007/s11854-019-0033-7

condition is equivalent to an action admitting a countable generating partition with finite Shannon entropy.

In [1], the first author introduced a modified invariant called model-measure sofic entropy which is a lower bound for Bowen's sofic entropy. Let

$$\Sigma = (\sigma_n : G \to \mathrm{Sym}(V_n))$$

be a sofic approximation to $G$. Model-measure sofic entropy is constructed in terms of sequences $(\mu_n)_{n=1}^{\infty}$ where $\mu_n$ is a probability measure on $A^{V_n}$. If these measures replicate the process $(A^G, \mu)$ in an appropriate sense, then we say that $(\mu_n)_{n=1}^{\infty}$ locally and empirically converges to $\mu$. We refer the reader to [1] for the precise definitions. We have substituted the phrase 'local and empirical convergence' for the phrase 'quenched convergence' which appeared in [1]. This has been done to avoid confusion with an alternative use of the word 'quenched' in the physics literature. A process is said to have completely positive model-measure sofic entropy if every nontrivial factor has positive model-measure sofic entropy. The goal of this paper is to prove the following theorem, which generalizes the main theorem of [5].

**Theorem 1.1.** *Let $G$ be a countable sofic group containing an element of infinite order. Then there exists an uncountable family of pairwise nonisomorphic $G$-processes each of which has completely positive model-measure sofic entropy (and hence completely positive sofic entropy) with respect to any sofic approximation to $G$. None of these processes is a factor of a Bernoulli shift.*

In order to prove Theorem 1.1 we introduce a concept of uniform mixing for sequences of model-measures. This uniform model-mixing will be defined formally in Section 3. It implies completely positive model-measure sofic entropy.

**Theorem 1.2.** *Let $G$ be a countable sofic group and let $(A^G, \mu)$ be a $G$-process with finite state space $A$. Suppose that for some sofic approximation $\Sigma$ to $G$, there is a uniformly model-mixing sequence $(\mu_n)_{n=1}^{\infty}$ which locally and empirically converges to $\mu$ over $\Sigma$. Then $(A^G, \mu)$ has completely positive lower model-measure sofic entropy with respect to $\Sigma$.*

As in [5], the examples we exhibit to establish Theorem 1.1 are produced via a coinduction method for lifting $H$-processes to $G$-processes when $H \leq G$. If $(A^H, \nu)$ is an $H$-process, then we can construct a corresponding $G$-process $(A^G, \mu)$ as follows. Let $T$ be a transversal for the right cosets of $H$ in $G$. Identify $G$ as a set with $H \times T$ and thereby identify $A^G$ with $(A^H)^T$. Set $\mu = \nu^T$. We call $(A^G, \mu)$ the coinduced process and denote it by $\mathrm{CInd}_H^G(\nu)$. (See page 72 of [7] for more details on this construction.) When $H \cong \mathbb{Z}$ this procedure preserves uniform mixing.

**Theorem 1.3.** *Let $G$ be a countable sofic group and let $(A^{\mathbb{Z}}, \nu)$ be a uniformly mixing $\mathbb{Z}$-process with finite state space $A$. Let $H \leq G$ be a subgroup isomorphic to $\mathbb{Z}$ and identify $A^{\mathbb{Z}}$ with $A^H$. Then for any sofic approximation $\Sigma$ to $G$, there is a uniformly model-mixing sequence of measures which locally and empirically converges to $\mathrm{CInd}_H^G(\nu)$ over $\Sigma$.*

We remark that it is easy to see that if $(A^G, \mu)$ is a Bernoulli shift (that is to say, $\mu$ is a product measure), then there is a uniformly model-mixing sequence which locally and empirically converges to $\mu$. Indeed, if $\mu = \eta^G$ for a measure $\eta$ on $A$, then the measures $\eta^{V_n}$ on $A^{V_n}$ are uniformly model-mixing and locally and empirically converge to $\mu$. Thus Theorem 1.2 shows that Bernoulli shifts with finite state space have completely positive sofic entropy, giving another proof of this case of the main theorem from [9]. We believe that completely positive sofic entropy for general Bernoulli shifts can be deduced along the same lines, requiring only a few additional estimates, but do not pursue the details here.

**Acknowledgements.** The first author's research was partially supported by the Simons Collaboration on Algorithms and Geometry. The second author's research was partially supported by NSF grants DMS-0968710 and DMS-1464475.

## 2 Preliminaries

**2.1 Notation.** The notation we use closely follows that in [1]; we refer the reader to that reference for further discussion. Let $A$ be a finite set. For any pair of sets $W \subseteq S$ we let $\pi_W : A^S \to A^W$ be projection onto the $W$-coordinates (thus our notation leaves the larger set $S$ implicit). Let $G$ be a countable group and let $(A^G, \mu)$ be a $G$-process. For $F \subseteq G$ we will write $\mu_F = \pi_{F*}\mu \in \mathrm{Prob}(A^F)$ for the $F$-marginal of $\mu$.

Let $B$ be another finite set and let $\phi : A^G \to B$ be a measurable function. If $F \subseteq G$, we will say that $\phi$ is $F$-local if it factors through $\pi_F$. We will say $\phi$ is local if it is $F$-local for some finite $F$. Let $\phi^G : A^G \to B^G$ be given by $\phi^G(a)(g) = \phi(g \cdot a)$ and note that $\phi^G$ is equivariant between the right-shift on $A^G$ and the right-shift on $B^G$.

Let $V$ be a finite set and let $\sigma$ be a map from $G$ to $\mathrm{Sym}(V)$. For $g \in G$ and $v \in V$ we write $\sigma^g \cdot v$ instead of $\sigma(g)(v)$. For $F \subseteq G$ and $S \subseteq V$ we define

$$\sigma^F(S) = \{\sigma^g \cdot s : g \in F, s \in S\}.$$

For $v \in V$ we write $\sigma^F(v)$ for $\sigma^F(\{v\})$. We write $\Pi^\sigma_{v,F}$ for the map from $A^V$ to $A^F$ given by $\Pi^\sigma_{v,F}(\overline{a})(g) = \overline{a}(\sigma^g \cdot v)$ for $\overline{a} \in A^V$ and $g \in F$. We write $\Pi^\sigma_v$ for $\Pi^\sigma_{v,G}$.

With $\phi : A^G \to B$ as before, we write $\phi^\sigma$ for the map from $A^V$ to $B^V$ given by $\phi^\sigma(\overline{a})(v) = \phi(\Pi_v^\sigma(\overline{a}))$.

If $D$ is a finite set and $\eta$ is a probability measure on $D$, then $\mathrm{H}(\eta)$ denotes the Shannon entropy of $\eta$, and for $\epsilon > 0$ we define

$$\mathrm{cov}_\epsilon(\eta) = \min\{|F| : F \subseteq D \text{ is such that } \eta(F) > 1-\epsilon\}.$$

If $\phi : D \to E$ is a map to another finite set, then we may write $\mathrm{H}_\mu(\phi)$ in place of $\mathrm{H}(\phi_*\mu)$. For $p \in [0,1]$ we let $\mathrm{H}(p) = -p\log p - (1-p)\log(1-p)$.

We use the $o(\cdot)$ and $\lesssim$ asymptotic notations with respect to the limit $n \to \infty$. Given two functions $f$ and $g$ on $\mathbb{N}$, the notation $f \lesssim g$ means that there is a positive constant $C$ such that $f(n) \leq Cg(n)$ for all $n$.

## 2.2 An information theoretic estimate

**Lemma 2.1.** *Let $A$ be a finite set and let $(V_n)_{n=1}^\infty$ be a sequence of finite sets such that $|V_n|$ increases to infinity. Let $\mu_n$ be a probability measure on $A^{V_n}$. We have*

$$\liminf_{n\to\infty} \frac{\mathrm{H}(\mu_n)}{|V_n|} \leq \sup_{\epsilon>0} \liminf_{n\to\infty} \frac{1}{|V_n|} \log \mathrm{cov}_\epsilon(\mu_n).$$

**Proof.** Let $\mu$ be a probability measure on a finite set $F$ and let $E \subseteq F$. By conditioning on the partition $\{E, F\setminus E\}$ and then recalling that entropy is maximized by uniform distributions we obtain

$$\begin{aligned} \mathrm{H}(\mu) &= \mu(E)\cdot \mathrm{H}(\mu(\cdot\,|\,E)) + \mu(F\setminus E)\cdot \mathrm{H}(\mu(\cdot\,|\,F\setminus E)) + \mathrm{H}(\mu(E)) \\ &\leq \mu(E)\cdot\log(|E|) + (1-\mu(E))\cdot\log(|F\setminus E|) + \mathrm{H}(\mu(E)). \end{aligned} \tag{2.1}$$

Now let $\mu_n$ and $V_n$ be as in the statement of the lemma. Let $\epsilon > 0$ and let $S_n \subseteq A^{V_n}$ be a sequence of sets with $\mu_n(S_n) > 1-\epsilon$ and $|S_n| = \mathrm{cov}_\epsilon(\mu_n)$. By applying (2.1) with $F = A^{V_n}$ and $E = S_n$ we have

$$\begin{aligned} &\liminf_{n\to\infty}\frac{\mathrm{H}(\mu_n)}{|V_n|} \\ &\quad\leq \liminf_{n\to\infty}\frac{1}{|V_n|}(\mu(S_n)\cdot\log(|S_n|) + (1-\mu(S_n))\cdot\log(|A^{V_n}\setminus S_n|) + \mathrm{H}(\mu(S_n))) \\ &\quad\leq \liminf_{n\to\infty}\frac{1}{|V_n|}(\log(|S_n|) + \epsilon\cdot\log(|A^{V_n}|) + \mathrm{H}(\epsilon)) \\ &\quad\leq \Big(\liminf_{n\to\infty}\frac{1}{|V_n|}\log\mathrm{cov}_\epsilon(\mu_n)\Big) + \epsilon\cdot\log(|A|). \end{aligned}$$

Now let $\epsilon$ tend to zero to obtain the lemma. □

## 3 Metrics on sofic approximations and uniform model-mixing

Let us fix a proper right-invariant metric $\rho$ on $G$: for instance, if $G$ is finitely generated, then $\rho$ can be a word metric, and, more generally, we may let $w : G \to [0, \infty)$ be any proper weight function and define $\rho$ to be the resulting weighted word metric. Again let $V$ be a finite set and let $\sigma$ be a map from $G$ to $\mathrm{Sym}(V)$. Let $H_\sigma$ be the graph on $V$ with an edge from $v$ to $w$ if and only if $\sigma^g \cdot v = w$ or $\sigma^g \cdot w = v$ for some $g \in G$. Define a weight function $W$ on the edges of $H_\sigma$ by setting

$$W(v, w) = \min\{\rho(g, 1_G) : \sigma^g \cdot v = w \text{ or } \sigma^g \cdot w = v\}.$$

If $v$ and $w$ are in the same connected component of $H_\sigma$ let $\rho_\sigma$ be the $W$-weighted graph distance between $v$ and $w$, that is

$$\rho_\sigma(v, w) = \min\left\{\sum_{i=0}^{k-1} W(p_i, p_{i+1}) : (v = p_0, p_1, \ldots, p_{k-1}, p_k = w) \text{ is an } H_\sigma\text{-path from } v \text{ to } w\right\}.$$

Having defined $\rho_\sigma$ on the connected components of $H_\sigma$, choose some number $M$ much larger than the $\rho_\sigma$-distance between any two points in the same connected component. Set $\rho_\sigma(v, w) = M$ for any pair $v, w$ of vertices in distinct connected components of $H_\sigma$. Note that if $(\sigma_n : G \to \mathrm{Sym}(V_n))$ is a sofic approximation to $G$, then for any fixed $r < \infty$, once $n$ is large enough the map $g \mapsto \sigma_n^g \cdot v$ restricts to an isometry from $B_\rho(1_G, r)$ to $B_{\rho_{\sigma_n}}(v, r)$ for most $v \in V_n$.

In the sequel the sofic approximation will be fixed, and we will abbreviate $\rho_{\sigma_n}$ to $\rho_n$. We can now state the main definition of this paper.

**Definition 3.1.** Let $(V_n)_{n=1}^\infty$ be a sequence of finite sets with $|V_n| \to \infty$ and for each $n$ let $\sigma_n$ be a map from $G$ to $\mathrm{Sym}(V_n)$. Let $A$ be a finite set. For each $n \in \mathbb{N}$ let $\mu_n$ be a probability measure on $A^{V_n}$. We say the sequence $(\mu_n)_{n=1}^\infty$ is **uniformly model-mixing** if the following holds. For every finite $F \subseteq G$ and every $\epsilon > 0$ there is some $r < \infty$ and a sequence of subsets $W_n \subseteq V_n$ such that

$$|W_n| = (1 - o(1))|V_n|,$$

and if $S \subseteq W_n$ is $r$-separated according the metric $\rho_n$ then

$$\mathrm{H}(\pi_{\sigma_n^F(S)*}\mu_n) \geq |S| \cdot (\mathrm{H}(\mu_F) - \epsilon).$$

This definition is motivated by Weiss' notion of uniform mixing from the special case when $G$ is amenable: see [14] and also Section 4 of [5]. Let us quickly recall that notion in the setting of a $G$-process $(A^G, \mu)$. First, if $K \subseteq G$ is finite and $S \subseteq G$ is another subset, then $S$ is **$K$-spread** if any distinct elements $s_1, s_2 \in S$ satisfy $s_1 s_2^{-1} \notin K$. The process $(A^G, \mu)$ is **uniformly mixing** if, for any finite-valued measurable function $\phi : A^G \to B$ and any $\epsilon > 0$, there exists a finite subset $K \subseteq G$ with the following property: if $S \subseteq G$ is another finite subset which is $K$-spread, then

$$\mathrm{H}((\phi_*^G \mu)_S) \geq |S| \cdot (\mathrm{H}_\mu(\phi) - \epsilon).$$

Beware that we have reversed the order of multiplying $s_1$ and $s_2^{-1}$ in the definition of '$K$-spread' compared with [5]. This is because we work in terms of observables such as $\phi$ rather than finite partitions of $A^G$, and shifting an observable by the action of $g \in G$ corresponds to shifting the partition that it generates by $g^{-1}$.

The principal result of [11] is that completely positive entropy implies uniform mixing. The reverse implication also holds: see [6] or Theorem 4.2 in [5]. Thus, uniform mixing is an equivalent characterization of completely positive entropy.

The definition of uniform mixing may be rephrased in terms of our proper metric $\rho$ on $G$ as follows. The process $(A^G, \mu)$ is uniformly mixing if and only if, for any finite-valued measurable function $\phi : A^G \to B$ and any $\epsilon > 0$, there exists an $r < \infty$ with the following property: if $S \subseteq G$ is $r$-separated according to $\rho$, then

$$\mathrm{H}((\phi_*^G \mu)_S) \geq |S| \cdot (\mathrm{H}_\mu(\phi) - \epsilon).$$

This is equivalent to the previous definition because a subset $S \subseteq G$ is $r$-separated according to $\rho$ if and only if it is $B_\rho(1_G, r)$-spread. The balls $B_\rho(1_G, r)$ are finite, because $\rho$ is proper, and any other finite subset $K \subseteq G$ is contained in $B_\rho(1_G, r)$ for all sufficiently large $r$.

This is the point of view on uniform mixing which motivates Definition 3.1. We use the right-invariant metric $\rho$ rather than the general definition of '$K$-spread' sets because it is more convenient later.

Definition 3.1 is directly compatible with uniform mixing in the following sense. If $G$ is amenable and $(F_n)_{n=1}^\infty$ is a Følner sequence for $G$, then the sets $F_n$ may be regarded as a sofic approximation to $G$: an element $g \in G$ acts on $F_n$ by translation wherever this stays inside $F_n$ and arbitrarily at points which are too close to the boundary of $F_n$. If $(A^G, \mu)$ is an ergodic $G$-process, then it follows easily that the sequence of marginals $\mu_{F_n}$ locally and empirically converge to $\mu$ over this Følner-set sofic approximation. If $(A^G, \mu)$ is uniformly mixing, then this sequence of marginals is clearly uniformly model-mixing in the sense of Definition 3.1.

On the other hand, suppose that $(A^G, \mu)$ admits a sofic approximation and a locally and empirically convergent sequence of measures over that sofic approximation which is uniformly model-mixing. Then our Theorem 1.2 shows that $(A^G, \mu)$ has completely positive sofic entropy. If $G$ is amenable, then sofic entropy always agrees with Kolmogorov–Sinai entropy [3], and this implies that $(A^G, \mu)$ has completely positive entropy and hence is uniformly mixing, by the result of [11].

Thus if $G$ is amenable, then completely positive entropy and uniform mixing are both equivalent to the existence of a sofic approximation and a locally and empirically convergent sequence of measures over it which is uniformly model-mixing. If these conditions hold, then we expect that one can actually find a locally and empirically convergent and uniformly model-mixing sequence of measures over *any* sofic approximation to $G$. This should follow using a similar kind of decomposition of the sofic approximants into Følner sets as in Bowen's proof in [3]. However, we have not explored this argument in detail.

Definition 3.1 applies to a shift-system with a finite state space. It can be transferred to an abstract measure-preserving $G$-action on $(X, \mu)$ by fixing a choice of finite measurable partition of $X$. However, in order to study actions which do not admit a finite generating partition, it might be worth looking for an extension of Definition 3.1 to $G$-processes with arbitrary compact metric state spaces, similarly to the setting in [1]. We also do not pursue this generalization here.

## 4 Proof of Theorem 1.2

We will use basic facts about the Shannon entropy of observables (i.e., random variables with finite range), for which we refer the reader to Chapter 2 of [4]. Let $\Sigma = (\sigma_n : G \to \mathrm{Sym}(V_n))$, $(A^G, \mu)$ and $(\mu_n)_{n=1}^{\infty}$ be as in the statement of Theorem 1.2. The following is the 'finitary' model-measure analog of Lemma 5.1 in [5].

**Lemma 4.1.** *Let $F \subseteq G$ be finite. Let $B$ be a finite set and let $\phi : A^G \to B$ be an $F$-local observable. Let $S_n \subseteq V_n$ be a sequence of sets such that $|S_n| \gtrsim |V_n|$. Then we have*

$$\mathrm{H}(\mu_F) - \frac{1}{|S_n|}\mathrm{H}(\pi_{\sigma_n^F(S_n)*}\mu_n) \geq \mathrm{H}_\mu(\phi) - \frac{1}{|S_n|}\mathrm{H}(\pi_{S_n*}\phi_*^{\sigma_n}\mu_n) - o(1).$$

**Proof of Lemma 4.1.** Let $\theta : A^F \to B$ be a map with $\theta \circ \pi_F = \phi$. Fix $n \in \mathbb{N}$ and $S \subseteq V_n$. Let $\alpha = \pi_{\sigma_n^F(S)} : A^{V_n} \to A^{\sigma_n^F(S)}$ and let $\beta = \pi_S \circ \phi^{\sigma_n} : A^{V_n} \to B^S$. For $s \in S$ let $\alpha_s = \Pi_{s,F}^{\sigma_n} : A^{V_n} \to A^F$ and let $\beta_s = \theta \circ \Pi_{s,F}^{\sigma_n} : A^{V_n} \to B$. Then we have $\alpha = (\alpha_s)_{s\in S}$ and $\beta = (\beta_s)_{s\in S}$. Enumerate $S = (s_k)_{k=1}^m$ and write $\alpha_{s_k} = \alpha_k$. All entropies in the following display are computed with respect to $\mu_n$. We have

$$\begin{aligned}
\mathrm{H}(\alpha) &= \mathrm{H}(\alpha_1,\dots,\alpha_m) = \mathrm{H}(\alpha_1) + \sum_{k=1}^{m-1} \mathrm{H}(\alpha_{k+1}|\alpha_1,\dots,\alpha_k)\\
&= \mathrm{H}(\alpha_1,\beta_1) + \sum_{k=1}^{m-1} \mathrm{H}(\alpha_{k+1},\beta_{k+1}|\alpha_1,\dots,\alpha_k)\\
&= \mathrm{H}(\beta_1) + \mathrm{H}(\alpha_1|\beta_1) + \sum_{k=1}^{m-1} \mathrm{H}(\beta_{k+1}|\alpha_1,\dots,\alpha_k) + \sum_{k=1}^{m-1} \mathrm{H}(\alpha_{k+1}|\beta_{k+1},\alpha_1,\dots,\alpha_k)\\
&\le \mathrm{H}(\beta_1) + \sum_{k=1}^{m-1} \mathrm{H}(\beta_{k+1}|\beta_1,\dots,\beta_k) + \sum_{k=1}^{m} \mathrm{H}(\alpha_k|\beta_k)\\
&= \mathrm{H}(\beta) + \sum_{k=1}^{m} \mathrm{H}(\alpha_k|\beta_k).
\end{aligned}$$

Let $\imath$ be the identity map on $A^F$. Then

$$\begin{aligned}
&|S|\cdot \mathrm{H}(\mu_F) - \mathrm{H}(\pi_{\sigma_n^F(S)*}\mu_n)\\
&\qquad = |S|\cdot \mathrm{H}_{\mu_F}(\imath) - \mathrm{H}_{\mu_n}(\alpha)\\
&\qquad \ge |S|\cdot \mathrm{H}_{\mu_F}(\theta) + |S|\cdot \mathrm{H}_{\mu_F}(\imath|\theta) - \mathrm{H}_{\mu_n}(\beta) - \sum_{s\in S} \mathrm{H}_{\mu_n}(\alpha_s|\beta_s) \qquad (4.1)\\
&\qquad = |S|\cdot \mathrm{H}_{\mu}(\phi) - \mathrm{H}\big(\pi_{S*}\phi_*^{\sigma_n}\mu_n\big) + |S|\cdot \mathrm{H}_{\mu_F}(\imath|\theta) - \sum_{s\in S} \mathrm{H}_{\mu_n}(\alpha_s|\beta_s).
\end{aligned}$$

Now allowing $n$ to vary, let $S_n \subseteq V_n$ be a sequence of sets such that $|S_n| \gtrsim |V_n|$. Write $\nu_n = \pi_{\sigma_n^F(S_n)*}\mu_n$. Let $s \in S_n$ be such that the obvious map from $F$ to $\sigma_n^F(s)$ is injective. Then the function $\overline{a} \mapsto \Pi_{s,F}^{\sigma_n}(\overline{a})$ provides an identification of $A^{\sigma_n^F(s)}$ with $A^F$. This identification sends $\alpha_s$ to $\imath$ and $\beta_s$ to $\theta$. When $n$ is large, the $\sigma_n^F(s)$ marginal of $\mu_n$ will resemble $\mu_F$ for most $s \in S_n$. Since $\alpha_s$ and $\beta_s$ are $\pi_{\sigma_n^F(s)}$ measurable, this implies that $\mathrm{H}_{\mu_F}(\imath|\theta) \approx \mathrm{H}_{\nu_n}(\alpha_s|\beta_s)$ for most $s$. More precisely, we can find a sequence of sets $C_n \subseteq S_n$ with

$$|C_n| = (1-o(1))|S_n|$$

such that

$$\max_{s\in C_n} |\mathrm{H}_{\mu_F}(\imath|\theta) - \mathrm{H}_{\nu_n}(\alpha_s|\beta_s)| = o(1).$$

Thus

$$\begin{aligned}
&\left| |S_n|\cdot \mathrm{H}_{\mu_F}(\imath|\theta) - \sum_{s\in S_n} \mathrm{H}_{\nu_n}(\alpha_s|\beta_s)\right|\\
&\qquad \le \sum_{s\in C_n} |\mathrm{H}_{\mu_F}(\imath|\theta) - \mathrm{H}_{\nu_n}(\alpha_s|\beta_s)| + \sum_{s\in S_n\setminus C_n} |\mathrm{H}_{\mu_F}(\imath|\theta) - \mathrm{H}_{\nu_n}(\alpha_s|\beta_s)| = o(|S_n|).
\end{aligned}$$

The lemma then follows from (4.1) and the above. □

Recall that for a measure space $(X, \mu)$ and two observables $\alpha$ and $\beta$ on $X$, the Rokhlin distance between $\alpha$ and $\beta$ is defined by

$$d_\mu^{\mathrm{Rok}}(\alpha, \beta) = \mathrm{H}_\mu(\alpha|\beta) + \mathrm{H}_\mu(\beta|\alpha).$$

This is a pseudometric on the space of observables on $X$. An easy computation shows that if $\alpha_1, \dots, \alpha_n$ and $\beta_1, \dots, \beta_n$ are two families of observables on $X$, then

$$d_\mu^{\mathrm{Rok}}((\alpha_1, \dots, \alpha_n), (\beta_1, \dots, \beta_n)) \le \sum_{k=1}^{n} d_\mu^{\mathrm{Rok}}(\alpha_k, \beta_k).$$

**Lemma 4.2.** *Let $\phi, \psi : A^G \to B$ be two local observables. Let $S_n \subseteq V_n$ be a sequence of sets with $|S_n| \gtrsim |V_n|$. Then we have*

$$\frac{1}{|S_n|}|\mathrm{H}(\pi_{S_n *}\phi_*^{\sigma_n}\mu_n) - \mathrm{H}(\pi_{S_n *}\psi_*^{\sigma_n}\mu_n)| \le d_\mu^{\mathrm{Rok}}(\phi, \psi) + o(1).$$

**Proof.** Let $\alpha_n = \pi_{S_n} \circ \phi^{\sigma_n} : A^{V_n} \to B^{S_n}$ and let $\beta_n = \pi_{S_n} \circ \psi^{\sigma_n} : A^{V_n} \to B^{S_n}$. Let $F$ be a finite subset of $G$ such that both $\phi$ and $\psi$ are $F$-local. Let $\theta : A^F \to B$ be a map such that $\theta \circ \pi_F = \phi$ and let $\kappa : A^F \to B$ be a map such that $\kappa \circ \pi_F = \psi$. For $s \in S_n$ let $\alpha_{n,s} = \theta \circ \Pi_{s,F}^{\sigma_n} : A^{V_n} \to B$ so that $\alpha_n = (\alpha_{n,s})_{s \in S_n}$. Also let $\beta_{n,s} = \kappa \circ \Pi_{s,F}^{\sigma_n} : A^{V_n} \to B$. Then we have

$$\begin{aligned} &\frac{1}{|S_n|}|\mathrm{H}(\pi_{S_n *}\phi_*^{\sigma_n}\mu_n) - \mathrm{H}(\pi_{S_n *}\psi_*^{\sigma_n}\mu_n)| \\ &= \frac{1}{|S_n|}|\mathrm{H}_{\mu_n}(\alpha_n) - \mathrm{H}_{\mu_n}(\beta_n)| \le \frac{1}{|S_n|} \cdot d_{\mu_n}^{\mathrm{Rok}}(\alpha_n, \beta_n) \\ &= \frac{1}{|S_n|} \cdot d_{\mu_n}^{\mathrm{Rok}}((\alpha_{n,s})_{s \in S_n}, (\beta_{n,s})_{s \in S_n}) \le \frac{1}{|S_n|} \sum_{s \in S_n} d_{\mu_n}^{\mathrm{Rok}}(\alpha_{n,s}, \beta_{n,s}). \end{aligned} \tag{4.2}$$

If the map $g \mapsto \sigma_n^g \cdot s$ is injective on $F$, we can identify $A^{\sigma_n^F(s)}$ with $A^F$ and thereby identify $\alpha_{n,s}$ with $\theta$ and $\beta_{n,s}$ with $\kappa$. Note that

$$d_{\mu_F}^{\mathrm{Rok}}(\theta, \kappa) = d_\mu^{\mathrm{Rok}}(\phi, \psi).$$

It follows that for any $\epsilon > 0$ we can find a weak star neighborhood $\mathcal{O}$ of $\mu$ such that if $s \in S_n$ is such that $(\Pi_s^{\sigma_n})_*\mu_n \in \mathcal{O}$, then

$$|d_{\mu_n}^{\mathrm{Rok}}(\alpha_{n,s}, \beta_{n,s}) - d_\mu^{\mathrm{Rok}}(\phi, \psi)| < \epsilon.$$

Thus, since $\mu_n$ locally and empirically converges to $\mu$, there are sets $C_n \subseteq S_n$ with $|C_n| = (1 - o(1))|S_n|$ such that

$$\max_{s \in C_n} |d_{\mu_n}^{\mathrm{Rok}}(\alpha_{n,s}, \beta_{n,s}) - d_\mu^{\mathrm{Rok}}(\phi, \psi)| = o(1). \tag{4.3}$$

The lemma now follows from (4.2) and (4.3). □

**Corollary 4.1.** *Let* $(\phi_m : A^G \to B)_{m=1}^{\infty}$ *be a sequence of local observables and let* $\phi : A^G \to B$ *be a local observable. Let* $S_n \subseteq V_n$ *be a sequence of sets with* $|S_n| \gtrsim |V_n|$. *Then if* $(m_n)_{n=1}^{\infty}$ *increases to infinity at a slow enough rate we have*

$$\frac{1}{|S_n|}|\mathrm{H}(\pi_{S_n*}\phi_*^{\sigma_n}\mu_n) - \mathrm{H}(\pi_{S_n*}\phi_{m_n*}^{\sigma_n}\mu_n)| \leq d_\mu^{\mathrm{Rok}}(\phi, \phi_{m_n}) + o(1).$$

**Proof of Theorem 1.2.** Let $B$ be a finite set and let $\psi : A^G \to B$ be an observable with $\mathrm{H}_\mu(\psi) > 0$. Let $(\phi_m)_{m=1}^{\infty}$ be an AL approximating sequence for $\psi$ rel $\mu$ (see Definition 4.4 in [1]). Then the sequence $\phi_m$ converges to $\psi$ in $d_\mu^{\mathrm{Rok}}$. In particular, $\phi_m$ is a Cauchy sequence and so we can find $M \in \mathbb{N}$ so that for all $m \geq M$ we have

$$d_\mu^{\mathrm{Rok}}(\phi_m, \phi_M) \leq \frac{\mathrm{H}_\mu(\psi)}{8}. \tag{4.4}$$

We will also assume $M$ is large enough that

$$\mathrm{H}_\mu(\phi_M) \geq \frac{\mathrm{H}_\mu(\psi)}{2}. \tag{4.5}$$

Let $F$ be a finite subset of $G$ such that $\phi_M$ is $F$-local. Then Definition 3.1 provides an $r < \infty$ and a sequence of subsets $W_n \subseteq V_n$ such that $|W_n| = (1 - o(1))|V_n|$ and if $S \subseteq W_n$ is $r$-separated then

$$\mathrm{H}(\mu_F) - \frac{1}{|S|}\mathrm{H}(\pi_{\sigma_n^F(S)*}\mu_n) \leq \frac{\mathrm{H}_\mu(\phi_M)}{2}. \tag{4.6}$$

Let $K = |B_\rho(1_G, r)|$. Since $\sigma_n$ is a sofic approximation there are sets $W_n' \subseteq V_n$ with $|W_n'| = (1 - o(1))|V_n|$ such that if $w \in W_n'$, then the $\rho_n$ ball of radius $r$ around $w$ has cardinality at most $K$. Write $Y_n = W_n \cap W_n'$ and note that we have $|Y_n| = (1-o(1))|V_n|$. For each $n$ let $S_n$ be an $r$-separated subset of $Y_n$ with maximal cardinality. Then $Y_n \subseteq \bigcup_{s \in S_n} B_{\rho_n}(s, r)$ so that

$$|S_n| \geq \frac{|Y_n|}{K} = (1 - o(1))\frac{|V_n|}{K}. \tag{4.7}$$

By Lemma 4.1 and (4.6) we have

$$\mathrm{H}_\mu(\phi_M) - \frac{1}{|S_n|}\mathrm{H}(\pi_{S_n*}\phi_{M*}^{\sigma_n}\mu_n) - o(1) \leq \frac{\mathrm{H}_\mu(\phi_M)}{2}$$

so that from (4.5) we have

$$\frac{\mathrm{H}_\mu(\psi)}{4} - o(1) \leq \frac{1}{|S_n|}\mathrm{H}(\pi_{S_n*}\phi_{M*}^{\sigma_n}\mu_n). \tag{4.8}$$

By Proposition 5.15 in [1], if $(m_n)_{n=1}^{\infty}$ increases to infinity at a slow enough rate then $(\phi_{m_n}^{\sigma_n})_*\mu_n$ will locally and empirically converge to $\psi_*^G\mu$. Since $A$ is finite, by the same argument as for Proposition 8.1 in [1] we have

$$\begin{aligned} \underline{\mathrm{h}}_{\Sigma}^{\mathrm{q}}(\psi_*^G\mu) &\geq \sup_{\epsilon>0}\liminf_{n\to\infty}\frac{1}{|V_n|}\log\mathrm{cov}_\epsilon((\phi_{m_n}^{\sigma_n})_*\mu_n) \\ &\geq \liminf_{n\to\infty}\frac{1}{|V_n|}\mathrm{H}((\phi_{m_n}^{\sigma_n})_*\mu_n) \end{aligned} \tag{4.9}$$

where the second inequality follows from Lemma 2.1. We also assume that $(m_n)_{n=1}^{\infty}$ increases slowly enough for Corollary 4.1 to hold. By (4.4) we have

$$\left|\frac{1}{|S_n|}\mathrm{H}(\pi_{S_n *}\phi_{M*}^{\sigma_n}\mu_n) - \frac{1}{|S_n|}\mathrm{H}(\pi_{S_n *}(\phi_{m_n}^{\sigma_n})_*\mu_n)\right| \leq \frac{\mathrm{H}_\mu(\psi)}{8} + o(1).$$

Combining this with (4.8) we see that

$$\frac{1}{|S_n|}\mathrm{H}(\pi_{S_n *}(\phi_{m_n}^{\sigma_n})_*\mu_n) \geq \frac{\mathrm{H}_\mu(\psi)}{8} - o(1).$$

By the above and (4.7) we have that for all sufficiently large $n$,

$$\mathrm{H}((\phi_{m_n}^{\sigma_n})_*\mu_n) \geq \frac{\mathrm{H}_\mu(\psi)}{8K+1}|V_n|. \tag{4.10}$$

Theorem 1.2 now follows from (4.9) and (4.10). □

## 5 Proof of Theorem 1.3

Let $(A^{\mathbb{Z}}, \nu)$ be a uniformly mixing $\mathbb{Z}$-process, and for each positive integer $l$ let $\nu_l$ be the marginal of $\nu$ on $A^l$. Let $\Sigma = (\sigma_n : G \to \mathrm{Sym}(V_n))$ be an arbitrary sofic approximation to $G$. Let $h \in G$ have infinite order and write $H = \langle h\rangle \cong \mathbb{Z}$. We construct a measure $\mu_n$ on $A^{V_n}$ for each $n \in \mathbb{N}$. We will later show that the sequence $(\mu_n)_{n=1}^{\infty}$ is uniformly model-mixing and locally and empirically converges to $\mu$ over $\Sigma$.

We first construct a measure $\mu_n^l$ on $A^{V_n}$ for each pair $(n, l)$ with $l$ much smaller than $n$. For a given $n$, the single permutation $\sigma_n^h$ partitions $V_n$ into a disjoint union of cycles. Since $h$ has infinite order and $\Sigma$ is a sofic approximation, once $n$ is large most points will be in very long cycles. In particular, we assume that most points are in cycles with length much larger than $l$. Partition the cycles into disjoint paths so that as many of the paths have length $l$ as possible, and let $\mathcal{P}_n^l = (P_{n,1}^l, \ldots, P_{n,k_n}^l)$ be the collection of all length-$l$ paths that result (so $\mathcal{P}_n^l$ is not a partition of the whole of $V_n$, but covers most of it). Fix any element $\overline{a}_0 \in A^{V_n}$ and define a random element $\overline{a} \in A^{V_n}$ by choosing each restriction $\overline{a}\restriction_{P_{n,i}^l}$ independently with the distribution of $\nu_l$ and extending to the rest of $V_n$ according to $\overline{a}_0$. Let $\mu_n^l$ be the law of this $\overline{a}$.

Now let $(l_n)_{n=1}^{\infty}$ increase to infinity at a slow enough rate that the following two conditions are satisfied:

(a) The number of points of $V_n$ that lie in some member of the family $\mathcal{P}_n^{l_n}$ is $(1 - o(1))|V_n|$.

(b) Whenever $g, g' \in G$ lie in distinct right cosets of $H$, so that $g^{-1}h^p g' \neq 1_G$ for all $p \in \mathbb{Z}$, we have

$$|\{v \in V_n : (\sigma_n^g)^{-1}(\sigma_n^h)^p \sigma_n^{g'} \cdot v = v \text{ for some } p \in \{-l_n, \dots, l_n\}\}| = o(|V_n|)$$

.

Set $\mu_n = \mu_n^{l_n}$. We separate the proof that $(\mu_n)_{n=1}^{\infty}$ has the required properties into two lemmas.

**Lemma 5.1.** *$(\mu_n)_{n=1}^{\infty}$ locally and empirically converges to $\mu$ over $\Sigma$.*

**Proof of Lemma 5.1.** Since $(A^G, \mu)$ is ergodic, by Corollary 5.6 in [1] it suffices to show that $\mu_n$ locally weak star converges to $\mu$. For a set $I \subseteq \mathbb{Z}$ write $h^I = \{h^i : i \in I\}$. Fix a finite set $F \subseteq G$. By enlarging $F$ if necessary we can assume there is an interval $I \subseteq \mathbb{Z}$ such that $F = \bigcup_{k=1}^m h^I t_k$ for $t_1, \dots, t_m$ in some transversal for the right cosets of $H$ in $G$. For each $g \in F$ let $j_g$ be a fixed element of $A$. Let $B \subseteq A^G$ be defined by

$$B = \{a \in A^G : a(g) = j_g \text{ for all } g \in F\}$$

and let $\epsilon > 0$. Then sets such as

$$\mathcal{O} = \{\eta \in \mathrm{Prob}(A^G) : \eta(B) \approx_\epsilon \mu(B)\}$$

form a subbasis of neighborhoods around $\mu$. It therefore suffices to show that when $n$ is large we have $(\Pi_v^{\sigma_n})_* \mu_n \in \mathcal{O}$ with high probability in the choice of $v \in V_n$.

For $k \in \{1, \dots, m\}$ let

$$B_k = \{x \in A^{\mathbb{Z}} : x(i) = j_{h^i t_k} \text{ for all } i \in I\}.$$

Note that $\mu$ is defined in such a way that $\mu(B) = \prod_{i=1}^k \nu(B_k)$. Now, let $W_n$ be the set of all points $v \in V_n$ such that the following conditions holds.

(i) The map $g \mapsto \sigma_n^g \cdot v$ is injective on $F$.

(ii) $\sigma_n^{h^i t_k} \cdot v = (\sigma_n^h)^i \sigma_n^{t_k} \cdot v$ for all $i \in I$ and $k \in \{1, \dots, m\}$.

(iii) For all pairs $g, g' \in F$, $\sigma_n^g \cdot v$ is in the same path as $\sigma_n^{g'} \cdot v$ if and only if $g$ and $g'$ lie in the same right coset of $H$. In particular, each of the images $\sigma_n^g \cdot v$ for $g \in F$ is contained in some member of $\mathcal{P}_n^{l_n}$.

We claim that $|W_n| = (1-o(1))|V_n|$. Clearly Conditions (i) and (ii) are satisfied with high probability in $v$, and so is the last part of Condition (iii), by Condition (a) in the choice of $(l_n)_{n=1}^{\infty}$.

Fix $g, g' \in F$ and suppose that $g$ and $g'$ are in the same coset of $H$, so that we have $g = h^i t_k$ and $g' = h^{i'} t_k$ for some $k \in \{1, \dots, m\}$ and $i, i' \in I$. If $v$ satisfies Condition (ii), then we have

$$(\sigma_n^h)^{i'-i}\sigma_n^g \cdot v = (\sigma_n^h)^{i'-i}(\sigma_n^h)^i\sigma_n^{t_k} \cdot v = (\sigma_n^h)^{i'}\sigma_n^{t_k} \cdot v = \sigma_n^{g'} \cdot v$$

so that $\sigma_n^g \cdot v$ and $\sigma_n^{g'} \cdot v$ will lie in the same path assuming that $\sigma_n^{t_k} \cdot v$ is not one of the first or last $|I|$ elements of its path. Note that for any $v \in V_n$ we have

$$|\{w : \sigma_n^{t_k} \cdot w = v \quad \text{for some } k \in \{1, \dots, m\}\}| \le m.$$

It follows that the number of points $v \in V_n$ such that $\sigma_n^{t_k} \cdot v$ is one of the first or last $|I|$ elements of a path is at most $2mp_n|I| + o(|V_n|)$ where $p_n$ is the total number of paths in $V_n$. By Condition (a) in the choice of $(l_n)_{n=1}^{\infty}$, most of $V_n$ is covered by paths whose lengths increase to infinity. Since also $p_n = o(V_n)$, it follows that $\sigma_n^g \cdot v$ lies in the same path as $\sigma_n^{g'} \cdot v$ with high probability in $v$.

On the other hand, suppose that $g$ and $g'$ are in distinct cosets of $H$. Assume that $\sigma_n^g \cdot v$ and $\sigma_n^{g'} \cdot v$ are in the same path. Then there is $p \in \{-l_n, \dots, l_n\}$ with $\sigma_n^g \cdot v = (\sigma_n^h)^p \sigma_n^{g'} \cdot v$, and hence $(\sigma_n^g)^{-1}(\sigma_n^h)^p\sigma_n^{g'} \cdot v = v$. By Condition (b) in the choice of $(l_n)_{n=1}^{\infty}$ there are only $o(|V_n|)$ vertices $v$ for which this holds. Thus we have established the claim.

Now let $v \in W_n$. We have

$$(\Pi_v^{\sigma_n})_*\mu_n(B) = \mu_n(\{\overline{a} \in A^{V_n} : \overline{a}(\sigma_n^g \cdot v) = j_g \text{ for all } g \in F\}).$$

For each $k \in \{1, \dots, m\}$ the set $\{(\sigma_n^h)^i\sigma_n^{t_k} \cdot v : i \in I\}$ is contained in a single path. Since the marginal of $\mu_n$ on each path is $\nu_{l_n}$, the probability that

$$\overline{a}((\sigma_n^h)^i\sigma_n^{t_k} \cdot v) = j_{h^i t_k}$$

for all $i \in I$ is equal to $\nu_{l_n}(B_k) = \nu(B_k)$. On the other hand, the marginals of $\mu_n$ on distinct paths are independent, so the probability that $\overline{a}(\sigma_n^g \cdot v) = j_g$ for all $g \in F$ is actually equal to $\prod_{i=1}^k \nu(B_k)$. □

If $(A^{\mathbb{Z}}, \nu)$ is weakly mixing, then so is the co-induced $G$-action. In particular, this certainly holds if $(A^{\mathbb{Z}}, \nu)$ is uniformly mixing. Therefore we may immediately promote Lemma 5.1 to the fact that $(\mu_n)_{n=1}^{\infty}$ locally and doubly empirically converges to $\mu$ over $\Sigma$, by Lemma 5.15 of [1]. In fact, we suspect that local and double empirical convergence holds here whenever $(A^{\mathbb{Z}}, \nu)$ is ergodic.

**Lemma 5.2.** $(\mu_n)_{n=1}^\infty$ *is uniformly model-mixing.*

**Proof of Lemma 5.2.** Let $F \subseteq G$ be finite and let $\epsilon > 0$. Again decompose $F = \bigcup_{k=1}^m h^I t_k$ for some interval $I \subseteq \mathbb{Z}$ and elements $t_k \in T$. Note that the restriction of the metric $\rho$ to $H$ is a proper right invariant metric on $H \cong \mathbb{Z}$, even though it might be different from the usual metric on $\mathbb{Z}$. Thus since $\nu$ is uniformly mixing, we can find some $r_0 < \infty$ such that if $(I_j)_{j=1}^q$ is a family of intervals in $\mathbb{Z}$ which are each of length $|I|$ and are pairwise at distance at least $r_0$, then writing $K = \bigcup_{j=1}^q I_j$ we have

$$\mathrm{H}(\nu_K) \geq q \cdot \left(\mathrm{H}(\nu_I) - \frac{\epsilon}{m}\right). \tag{5.1}$$

Let $r < \infty$ be large enough that for all $g, g' \in G$, if $\rho(g, g') \geq r$, then

$$\rho(fg, f'g') \geq r_0 \quad \text{for all } f, f' \in F.$$

Such a choice of $r$ is possible since by right-invariance of $\rho$ we have

$$\rho(fg, g) = \rho(f, 1_G) \quad \text{and} \quad \rho(f'g', g') = \rho(f', 1_G).$$

Let $W_n$ be as in the proof of Lemma 5.1 and recall that $|W_n| = (1 - o(1))|V_n|$. Let $S \subseteq W_n$ be $r$-separated according to $\rho_n$.

Fix a path $P \in \mathcal{P}_n^{l_n}$ and let $S_P$ be the set of points $v \in S$ such that $\sigma_n^{t_{k(v)}} \cdot v \in P$ for some $k(v) \in \{1, \ldots, m\}$. Since $S \subseteq W_n$, Condition (iii) from the previous proof implies that

$$\sigma_n^F(S) \cap P = \bigcup_{v \in S_P} \{(\sigma_n^h)^i \sigma_n^{t_{k(v)}} \cdot v : i \in I\}.$$

Each of the sets in the latter union is an interval of length $|I|$ in $P$ and by our choice of $r$ these are pairwise at distance $r_0$ in $\rho_n$ restricted to $P$. Since the marginal of $\mu_n$ on $P$ is equal to $\nu_{n_l}$, (5.1) implies that

$$\mathrm{H}(\pi_{(\sigma_n^F(S) \cap P)*}\mu_n) \geq |S_P| \cdot \left(\mathrm{H}(\nu_I) - \frac{\epsilon}{m}\right).$$

Since the marginals of $\mu_n$ on distinct paths are independent, this implies that

$$\mathrm{H}(\pi_{\sigma_n^F(S)*}\mu_n) \geq \left(\sum_{P \in \mathcal{P}_n^{l_n}} |S_P|\right) \cdot \left(\mathrm{H}(\nu_I) - \frac{\epsilon}{m}\right). \tag{5.2}$$

By Condition (iii) in the definition of $W_n$, each $v \in S$ appears in $S_P$ for exactly $m$ paths $P$. Therefore

$$\sum_{P \in \mathcal{P}_n^{l_n}} |S_P| = m \cdot |S|. \tag{5.3}$$

Now $\mathrm{H}(\mu_F) = m \cdot \mathrm{H}(\nu_I)$, so from (5.2) and (5.3) we have

$$\mathrm{H}(\pi_{\sigma_n^F(S)*}\mu_n) \geq |S| \cdot (\mathrm{H}(\mu_F) - \epsilon)$$

as required. □

**Proof of Theorem 1.3.** Theorem 1.3 now follows from Theorem 1.2 and Lemmas 5.1 and 5.2. □

## 6 Proof of Theorem 1.1

**Proof of Theorem 1.1.** This part of the argument is essentially the same as the corresponding part of [5]. Consider the family of uniformly mixing $\mathbb{Z}$-processes $\{(4^{\mathbb{Z}}, \nu_\omega) : \omega \in 2^{\mathbb{N}}\}$ constructed in Section 6 of [5]. Fix an isomorphic copy $H$ of $\mathbb{Z}$ in $G$ and let $\mu_\omega = \mathrm{CInd}_H^G(\nu_\omega)$. By Theorems 1.2 and 1.3 the process $(4^G, \mu_\omega)$ has completely positive model-measure sofic entropy. Note that the restriction of the $G$-action to $H$ is a permuted power of the original $\mathbb{Z}$-process in the sense of Definition 6.5 from [5]. Thus by Proposition 6.6 in that reference, the processes $\{(4^G, \mu_\omega) : \omega \in 2^{\mathbb{N}}\}$ are pairwise nonisomorphic.

Suppose toward a contradiction that for some $\omega$, $(4^G, \mu_\omega)$ is a factor of a Bernoulli shift $(Z^G, \eta^G)$ over some standard probability space $(Z, \eta)$. Let $\psi : Z^G \to 4^G$ be an equivariant measurable map with $\psi_*\eta^G = \mu_\omega$. Note that the restricted right-shift action $H \curvearrowright (Z^G, \eta^G)$ is still isomorphic to a Bernoulli shift and $\psi$ is still a factor map from this process to the restricted action $H \curvearrowright (4^G, \mu_\omega)$. Thus the latter $\mathbb{Z}$-process is isomorphic to a Bernoulli shift and so is its factor $(4^{\mathbb{Z}}, \nu_\omega)$. This contradicts Corollary 6.4 in [5]. □

### References


[1] T. Austin. *Additivity properties of sofic entropy and measures on model spaces*, Forum Math. Sigma **4** (2016), e25.

[2] L. Bowen, *Measure conjugacy invariants for actions of countable sofic groups*, J. Amer. Math. Soc. **23** (2010), 217–245.

[3] L. Bowen, *Sofic entropy and amenable groups*, Ergodic Theory Dynam. Systems **32** (2012), 427–466.

[4] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Hoboken, NJ, 2006.

[5] A. Dooley, V. Golodets, D. Rudolph and S. Sinel'shchikov, *Non-Bernoulli systems with completely positive entropy*, Ergodic Theory Dynam. Systems **28** (2007), 87–124.

[6] V. Ya. Golodets and S. D. Sinel'shchikov, *Complete positivity of entropy and non-Bernoullicity for transformation groups*, Colloq. Math. **84/85** (2000), 421–429.

[7] A. S. Kechris, *Global Aspects of Ergodic Group Actions*, American Mathematical Society, Providence, RI, 2010.

[8] D. Kerr, *Sofic measure entropy via finite partitions*, Groups Geom. Dyn. **7** (2013), 617–632.

[9] D. Kerr, *Bernoulli actions of sofic groups have completely positive entropy*, Israel J. Math. **202** (2014), 461–474.

[10] D. Kerr and H. Li, *Entropy and the variational principle for actions of sofic groups*, Invent. Math. **186** (2011), 501–558.

[11] D. Rudolph and B. Weiss, *Entropy and mixing for amenable group actions*, Ann. of Math. (2) **151** (2000), 1119–1150.

[12] B. Seward, *Krieger's finite generator theorem for ergodic actions of countable groups I*, Invent. Math. **215** (2019), 265–310.

[13] B. Seward. *Krieger's finite generator theorem for ergodic actions of countable groups II*, arXiv: 1501.03367 [math.DS].

[14] B. Weiss, *Actions of amenable groups*, in *Topics in Dynamics and Ergodic Theory*, Cambridge University Press, Cambridge, 2003, pp. 226–262.

*Tim Austin*
Courant Institute of Mathematical Sciences
New York University
New York, NY 10012, USA
email: tim@cims.nyu.edu

*Peter Burton*
Department of Mathematics
California Institute of Technology
Pasadena, CA 91125, USA
email: pjburton@caltech.edu

(Received May 13, 2016 and in revised form November 1, 2016)