Ergod. Th. & Dynam. Sys. (2012), **32**, 427–466 doi:10.1017/S0143385711000253
© *Cambridge University Press, 2011*

Sofic entropy and amenable groups

LEWIS BOWEN

Texas A&M University, Mailstop 3368, College Station, TX, 77843-3368, USA
(e-mail: lpbowen@math.tamu.edu)

(*Received* 24 *August* 2010 *and accepted in revised form* 14 *March* 2011)

In memory of Dan Rudolph

Abstract. In previous work, the author introduced a measure-conjugacy invariant for sofic group actions called sofic entropy. Here, it is proven that the sofic entropy of an amenable group action equals its classical entropy. The proof uses a new measure-conjugacy invariant called upper-sofic entropy and a theorem of Rudolph and Weiss for the entropy of orbit-equivalent actions relative to the orbit change σ-algebra.

Contents

1 Introduction 427
1.1 Sofic groups . . . 428
1.2 Sofic entropy . . . 428
1.3 Overview . . . 429
2 Random sofic approximations, approximate processes and entropy 430
3 Amenable groups 432
4 The case of $\mathbb{Z}$ 433
5 Relative entropy 438
6 Orbit-equivalence and entropy 454
6.1 Lifting factors . . . 455
6.2 Lifting orbit-equivalences . . . 456
Acknowledgements 466
References 466

1. *Introduction*

The paper [**Bo10a**] introduced a family of measure-conjugacy invariants referred to as sofic entropy for actions of sofic groups. This entropy is inspired by the classical Kolmogorov–Sinai entropy and shares many of its features. The main goal of this paper is to show that the sofic entropy of an amenable group action equals its classical entropy. An alternative approach based on operator algebras is being developed by Kerr and Li [**KL1, KL2**]. The reader is encouraged to review [**Bo10a**] for more background.

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

1.1. *Sofic groups.* To begin, let us recall the definition of a sofic group.

Definition 1. (Sofic groups) Let G be a countable group. For any integer $m_i > 0$, let $[m_i] = \{1, \ldots, m_i\}$ and $\mathrm{Sym}(m_i)$ denote the symmetric group on $[m_i]$. Let $\Sigma = \{\sigma_i\}_{i=1}^{\infty}$ be a sequence of maps $\sigma_i : G \to \mathrm{Sym}(m_i)$ which are not assumed to homomorphisms. Then Σ is a *sofic approximation* to G if for every $g, h \in G$,

$$\lim_{i\to\infty} \frac{1}{m_i} \#\{p \in [m_i] : \sigma(g)\sigma(h)p = \sigma(gh)p\} = 1$$

and for every $g \neq h \in G$,

$$\lim_{i\to\infty} \frac{1}{m_i} \#\{p \in [m_i] : \sigma(g)p \neq \sigma(h)p\} = 1.$$

To avoid trivialities, we also assume $\lim_{i\to\infty} m_i = +\infty$, which is necessarily true if G is infinite. G is *sofic* if there exists a sofic approximation to G.

Example 1. If G is residually finite then there exists a decreasing sequence $\{N_i\}_{i=1}^{\infty}$ of finite-index normal subgroups of G with $\bigcap_i N_i = \{e\}$. Let $\sigma_i : G \to \mathrm{Sym}(G/N_i)$ be the canonical homomorphism given by the action of G on G/N_i. Then $\{\sigma_i\}_{i=1}^{\infty}$ is a sofic approximation to G.

Example 2. If G is amenable then there exists an increasing sequence $\{F_i\}_{i=1}^{\infty}$ of finite subsets of G such that $\bigcup_i F_i = G$ and for every finite $K \subset G$,

$$\lim_{i\to\infty} \frac{|KF_i \,\Delta\, F_i|}{|F_i|} = 1.$$

Let $\sigma_i : G \to \mathrm{Sym}(F_i)$ be any map such that if $f \in F_i$, $g \in G$ and $gf \in F_i$ then $\sigma_i(g)f = gf$. Then $\{\sigma_i\}_{i=1}^{\infty}$ is a sofic approximation to G.

Sofic groups were defined implicitly by Gromov in [**Gro99**] and explicitly by Weiss in [**We00**]. Since finitely generated linear groups (i.e., subgroups of $\mathrm{GL}_n(F)$ where F is a field) are residually finite (by [**Ma40**]), they are sofic. It is easy to check that a countable group is sofic if and only if all of its finitely generated subgroups are sofic. Thus, all countable linear groups are sofic. It is unknown whether every countable group is sofic but an unresolved case is that of the universal Burnside group on a finite set of generators. Pestov has written a beautiful up-to-date survey [**Pe08**] on sofic groups and their siblings, hyperlinear groups.

1.2. *Sofic entropy.* Let G be a countable discrete group. In this paper, an *action* of G is a triple (T, X, μ), where (X, μ) is a standard probability space and $T = (T_g)_{g\in G}$ is a collection of measure-preserving transformations $T_g : X \to X$ such that $T_g T_h = T_{gh}$ for all $g, h \in G$. The notation $G \curvearrowright^{T} (X, \mu)$ means (T, X, μ) is an action. Also, $G \curvearrowright (X, \mu)$ means that G acts on (X, μ) by measure-preserving transformations and the product of $g \in G$ with $x \in X$ is denoted by gx.

A *process* over G is a quadruple $\overline{X} := (T, X, \mu, \phi)$, where (T, X, μ) is a G-action and $\phi : X \to A$ is a measurable map into a finite or countable set A. ϕ is called an *observable* and A is the *range* of the process. We will implicitly assume that the range of every process

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

considered in this paper is finite. The next few paragraphs define the entropy rate of ϕ with respect to a sofic approximation Σ for G in the special case in which A is finite.

Suppose that $\sigma : G \to \mathrm{Sym}(m)$ is a map and $\psi : \{1, \ldots, m\} \to A$ is a function. In order to compare ψ with ϕ, let $W \subset G$ be finite (W is for *window*). Recall that A^W is the set of all functions from W to A. Let $\phi^W : X \to A^W$ be the map defined by $\phi^W(x)(w) := \phi(T_w x)$. Similarly, define $\psi_\sigma^W : \{1, \ldots, m\} \to A^W$ by $\psi_\sigma^W(p)(w) = \psi(\sigma(w)p)$. The measure μ pushes forward to a measure $\phi_*^W \mu$ on A^W. Similarly, if u is the uniform probability measure on $\{1, \ldots, m\}$ then $(\psi_\sigma^W)_* u$ is a measure of A^W. Let $d_W((\sigma, \psi), \phi)$ be the total variation distance between $\phi_*^W \mu$ and $(\psi_\sigma^W)_* u$. Explicitly,

$$d_W((\sigma, \psi), \phi) := \frac{1}{2} \|\phi_*^W \mu - (\psi_\sigma^W)_* u\|_1 = \frac{1}{2} \sum_{a \in A^W} |\phi_*^W \mu(\{a\}) - (\psi_\sigma^W)_* u(\{a\})|.$$

The *sofic entropy rate* of the process $\overline{X}$ with respect to a sofic approximation $\Sigma = \{\sigma_i\}_{i=1}^\infty$ to G (where $\sigma_i : G \to \mathrm{Sym}(m_i)$) is defined by

$$h(\Sigma, \overline{X}) := \inf_{W \subset G} \inf_{\epsilon > 0} \limsup_{i \to \infty} \frac{\log \#\{\psi : \{1, \ldots, m_i\} \to A : d_W((\sigma_i, \psi), \phi) < \epsilon\}}{m_i}. \quad (1)$$

The first infimum is over all finite subsets of G. The entropy $h(\Sigma, \overline{X})$ may alternatively be denoted by $h(\Sigma, \phi)$ or $h_\mu(\Sigma, \phi)$.

In order to obtain a measure-conjugacy invariant, consider a special class of observables as follows. The map ϕ is *generating* if the smallest G-invariant σ-algebra on X for which ϕ is measurable is the σ-algebra of all measurable sets up to sets of measure zero. The following is part of the main result of [**Bo10a**].

THEOREM 1.1. *Suppose $G \curvearrowright (X, \mu)$. If ϕ_1 and ϕ_2 are finite generating observables of X and Σ is a sofic approximation to G then $h(\Sigma, \phi_1) = h(\Sigma, \phi_2)$.*

Because of this result, the *entropy* of the action $G \curvearrowright^T (X, \mu)$ with respect to Σ is defined by $h(\Sigma, T) := h(\Sigma, \phi)$, where ϕ is any finite generating observable (if one exists).

In [**Bo10a**], an alternative but equivalent definition of entropy is given based on partitions instead of observables. Also, the entropy rate of an observable with countable range is defined under special conditions. That extension is not needed here. The main result of this paper is the following theorem.

THEOREM 1.2. *If G is infinite and amenable, $G \curvearrowright (X, \mu)$ and ϕ is a finite observable then for any sofic approximation Σ to G, $h(\Sigma, \phi)$ is the classical mean entropy rate of ϕ.*

The definition of classical mean entropy rate is reviewed in §3. By [**Ro88**], if the classical entropy of an ergodic, essentially free system $G \curvearrowright^T (X, \mu)$ is finite (and G is amenable) then there exists a finite generating observable for the action. So, the above theorem implies $h(\Sigma, T)$ is the classical entropy of the action in this case.

1.3. *Overview.* Section 2 discusses several variations on the definition of sofic entropy. These arise from allowing randomness in the sofic approximation and also in the approximations to the observable. Section 3 provides basic background on amenable groups and classical entropy theory. Section 4 proves the main Theorem 1.2 in the special case in which $G = \mathbb{Z}$. This uses the above variations on sofic entropy but is

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

otherwise elementary. Section 5 discusses relative entropy theory, both the classical case and the sofic case. Section 6 proves the main results in entropy/orbit-equivalence theory that allow us to conclude Theorem 1.2.

2. *Random sofic approximations, approximate processes and entropy*

It will be helpful to broaden the notion of sofic approximation to allow for 'random' sofic approximations, defined next.

Definition 2. (Random sofic approximation) Let G be a countable group. Let $\{m_i\}_{i=1}^{\infty}$ be a sequence of natural numbers, $\mathrm{Sym}(m_i)^G$ the set of maps from G to $\mathrm{Sym}(m_i)$ and u_{m_i} the uniform probability measure on $[m_i]$. Let κ_i be a probability measure on $\mathrm{Sym}(m_i)^G$. We say that the sequence $\mathbb{K} = \{\kappa_i\}_{i=1}^{\infty}$ is a *random sofic approximation* to G if for every $g, h \in G$,

$$\lim_{i\to\infty} \kappa_i \times u_{m_i}(\{(\sigma, p) \in \mathrm{Sym}(m_i)^G \times [m_i] : \sigma(g)\sigma(h)p = \sigma(gh)p\}) = 1$$

and for every $g \neq h \in G$,

$$\lim_{i\to\infty} \kappa_i \times u_{m_i}(\{(\sigma, p) \in \mathrm{Sym}(m_i)^G \times [m_i] : \sigma(g)p \neq \sigma(h)p\}) = 1.$$

Remark 1. This notion generalizes sofic approximations in the following sense. If $\Sigma = \{\sigma_i\}_{i=1}^{\infty}$ is a sofic approximation of a group G (where $\sigma_i \in \mathrm{Sym}(m_i)^G$) and δ_i is the probability measure on $\mathrm{Sym}(m_i)^G$ supported on σ_i then $\{\delta_i\}_{i=1}^{\infty}$ is a random sofic approximation of G.

Definition 3. (Sofic entropy) Let $\overline{X} := (T, X, \mu, \phi)$ be a process over a group G with random sofic approximation $\mathbb{K} = \{\kappa_i\}_{i=1}^{\infty}$. Suppose $\phi : X \to A$, where A is a finite set. For each $\sigma \in \mathrm{Sym}(m_i)^G$, finite $W \subset G$ and $\epsilon > 0$, let $N_i(\sigma, W, \epsilon)$ be the number of functions $\psi : [m_i] \to A$ such that $d_W((\sigma, \psi), \phi) < \epsilon$ (this is the notation used in §1.2). The *sofic entropy* of $\overline{X}$ with respect to $\mathbb{K}$ is

$$h(\mathbb{K}, \overline{X}) := \inf_{W\subset G} \inf_{\epsilon>0} \limsup_{i\to\infty} \frac{\int \log N_i(\sigma, W, \epsilon)\, d\kappa_i(\sigma)}{m_i}.$$

For example, if $\mathbb{K}'$ is any subsequence of $\mathbb{K}$ then $h(\mathbb{K}', \overline{X}) \leq h(\mathbb{K}, \overline{X})$.

The definition above generalizes the notion of sofic entropy by introducing randomness into the sofic approximation. It is also possible to introduce randomness into the observables. This leads to a new notion of entropy called upper-sofic entropy (which was introduced implicitly in [**Bo10b**]). To explain, it is necessary to have a notion of 'approximate process' which is motivated by the definition of a sofic group.

Definition 4. (Approximate process) An *approximate process* over G is a quadruple $\overline{X} = (T, X, \mu, \phi)$, where (X, μ) is a standard probability space, $T = (T_g)_{g\in G}$ is a set of measure-preserving Borel maps $T_g : X \to X$ and $\phi : X \to A$ is a Borel map to a finite or countable set A called the *range* of the process. An approximate process is a *process* if T defines an action, i.e., $T_{gh} = T_g T_h$ for all $g, h \in G$. The word 'approximate' is used to suggest that $\overline{X}$ is approximating some process. The definition by itself does not imply this but it is how these objects will be used.

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Definition 5. (Local statistics and distance between processes) Given a finite set $W \subset G$ and an approximate process $\overline{X} = (T, X, \mu, \phi)$, define $\phi_T^W : X \to A^W$ by

$$\phi_T^W(x) := [w \mapsto \phi(T_w x)].$$

Let $(\phi_T^W)_* \mu$ be the pushforward measure on A^W. This measure is called the *W-local statistics of* $\overline{X}$.

Given another approximate process $\overline{Y} = (S, Y, \nu, \psi)$ with range A, we define

$$d_W(\overline{X}, \overline{Y}) := \tfrac{1}{2} \|(\phi_T^W)_* \mu - (\psi_S^W)_* \nu\|_1.$$

Then $\overline{X}$ and $\overline{Y}$ are said to be equivalent if $d_W(\overline{X}, \overline{Y}) = 0$ for all finite $W \subset G$. Let $\mathcal{P}(G, A)$ be the set of all equivalence classes of approximate processes over G with range A. By abuse of the notation, we do not distinguish between an approximate process and its equivalence class. Let $\mathcal{P}(G, A)$ have the smallest topology such that for every finite $W \subset G$, the function d_W is continuous with respect to the product topology on $\mathcal{P}(G, A) \times \mathcal{P}(G, A)$.

Definition 6. (Approximate processes from random sofic approximations) Let $\mathbb{K} = \{\kappa_i\}_{i=1}^{\infty}$ be a random sofic approximation to G. Let η_i be a probability measure on $\mathrm{Sym}(m_i)^G \times A^{[m_i]}$, where A is a finite or countable set. Suppose that the projection of η_i to the first factor is κ_i. For each $g \in G$, define

$$\hat{\sigma}_i(g) : \mathrm{Sym}(m_i)^G \times A^{[m_i]} \times [m_i] \to \mathrm{Sym}(m_i)^G \times A^{[m_i]} \times [m_i],$$
$$\hat{\sigma}_i(g)(\sigma, \psi, p) = (\sigma, \psi, \sigma(g)p).$$

Define

$$\chi_i : \mathrm{Sym}(m_i)^G \times A^{[m_i]} \times [m_i] \to A, \quad \chi_i(\sigma, \psi, p) := \psi(p).$$

Define

$$\overline{X}_i := (\hat{\sigma}_i, \mathrm{Sym}(m_i)^G \times A^{[m_i]} \times [m_i], \eta_i \times u_{m_i}, \chi_i).$$

Then $\overline{X}_i$ is the approximate process *constructed from* η_i.

The sequence $\{\overline{X}_i\}_{i=1}^{\infty}$ is *adapted to the sofic approximation* $\mathbb{K} = \{\kappa_i\}_{i=1}^{\infty}$ if it arises from the above construction for some sequence of measures $\{\eta_i\}_{i=1}^{\infty}$.

Definition 7. If μ is a probability measure on a finite or countable set X then

$$H(\mu) := -\sum_{x \in X} \mu(\{x\}) \log(\mu(\{x\})).$$

By convention, $0 \log(0) = 0$.

Definition 8. (Upper-sofic entropy) Let η_i, κ_i, etc be as in Definition 6 and let $\eta_i = \int \nu_{i,\sigma} \, d\kappa_i(\sigma)$ be the decomposition over κ_i. So, $\nu_{i,\sigma}$ is a probability measure on the set $\{(\sigma, \xi) \in \mathrm{Sym}(m_i)^G \times A^{[m_i]}\}$. Then define

$$h(\overline{X}_i) := \frac{1}{m_i} \int H(\nu_{i,\sigma}) \, d\kappa_i(\sigma).$$

This definition depends implicitly on η_i (which might not be determined by the equivalence class of $\overline{X}_i$).

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

The *upper-sofic entropy* of a finite-range process $\overline{X} = (T, X, \mu, \phi)$ with respect to $\mathbb{K}$ is defined by

$$\bar{h}(\mathbb{K}, \overline{X}) := \sup \limsup_{j\to\infty} h(\overline{X}_j),$$

where the supremum is over all sequences $\{\overline{X}_j\}_{j=1}^{\infty}$ of approximate processes adapted to $\mathbb{K}'$ (where $\mathbb{K}'$ is a subsequence of $\mathbb{K}$) such that $\lim_{j\to\infty} \overline{X}_j = \overline{X}$. If no such subsequence exists then $\bar{h}(\mathbb{K}, \overline{X}) = -\infty$. The upper-sofic entropy $h(\Sigma, \overline{X})$ can alternatively be denoted by $\bar{h}(\mathbb{K}, \phi)$ or $\bar{h}_\mu(\mathbb{K}, \phi)$ if it is desirable to emphasize the dependence on μ or ϕ.

If each κ_i is supported on a singleton set $\{\sigma_i\} \subset \mathrm{Sym}(m_i)^G$ then let $\Sigma := \{\sigma_i\}_{i\in\mathbb{N}}$ and define $\bar{h}(\Sigma, \overline{X}) := \bar{h}(\mathbb{K}, \overline{X})$.

Using the methods of [**Bo10a**], it can be shown that upper-sofic entropy is an invariant: if ϕ, ψ are two generating observables with finite range then $\bar{h}(\mathbb{K}, \phi) = \bar{h}(\mathbb{K}, \psi)$ (but this is not needed here). Next, upper-sofic entropy is related to sofic entropy (which will be referred to as lower-sofic entropy so as to distinguish it).

Definition 9. (Strong convergence of approximate processes) As above, let $\{\overline{X}_i\}_{i=1}^{\infty}$ be a sequence of approximate processes constructed from measures $\{\eta_i\}_{i=1}^{\infty}$ on $\mathrm{Sym}(m_i)^G \times A^{[m_i]}$ as in Definition 6. Suppose that the limit $\lim_{i\to\infty} \overline{X}_i = \overline{X}$ is a process over G. The sequence $\{\overline{X}_i\}_{i=1}^{\infty}$ *converges to* $\overline{X}$ *strongly* (denoted $\lim_{i\to\infty} \overline{X}_i = \overline{X}$ strongly) if, for every finite $W \subset G$ and every $\epsilon > 0$,

$$\lim_{i\to\infty} \eta_i(\{(\sigma, \psi) \in \mathrm{Sym}(m_i)^G \times A^{[m_i]} : d_W((\sigma, \psi), \phi) < \epsilon\}) = 1.$$

PROPOSITION 2.1. *Let $\overline{X}$ be a process over G and let $\mathbb{K}$ be a random sofic approximation. Then*

$$h(\mathbb{K}, \overline{X}) = \sup \limsup_{j\to\infty} h(\overline{X}_j),$$

where the supremum is over all sequences $\{\overline{X}_j\}_{j=1}^{\infty}$ of approximate processes adapted to $\mathbb{K}'$ (where $\mathbb{K}'$ is a subsequence of $\mathbb{K}$) such that $\lim_{j\to\infty} \overline{X}_j = \overline{X}$ strongly.

Proof. The proof is an exercise in understanding the definitions. □

COROLLARY 2.2. *Let $\overline{X}$ be a process over G and let $\mathbb{K}$ be a random sofic approximation. Then*

$$h(\mathbb{K}, \overline{X}) \leq \bar{h}(\mathbb{K}, \overline{X}).$$

3. *Amenable groups*

Definition 10. Let G be a countable group, $F, K \subset G$ finite sets and $\epsilon > 0$. Then F is *left-(K, ϵ)-invariant* if

$$\frac{|KF \cap F|}{|F|} \geq 1 - \epsilon,$$

where $KF = \{kf \in G \mid k \in K,\ f \in F\}$. A *left-Følner sequence* of G is a sequence $\{F_n\}_{n=1}^{\infty}$ of finite subsets of G such that for all finite $K \subset G$ and all $\epsilon > 0$, there exists an N such that $n > N$ implies F_n is left-(K, ϵ)-invariant. G is *amenable* if there exists a left-Følner sequence for G.

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Definition 11. Let G be an amenable group with left-Følner sequence $\{F_n\}_{n=1}^{\infty}$. Let $\overline{X} = (T, X, \mu, \phi)$ be a process over G with range A. For a finite $W \subset G$, let $\phi^W : X \to A^W$ be the map $\phi^W(x)(w) = \phi(T_w x)$. The *classical entropy* of $\overline{X}$ is defined by

$$h(\overline{X}) := \lim_{n\to\infty} \frac{H(\phi_*^{F_n}\mu)}{|F_n|},$$

where $H(\cdot)$ is as in Definition 7. Some alternative notation for the entropy rate are $h(\overline{X}) = h(\phi) = h_\mu(\phi) = h_\mu(T, \phi) = h(T, \phi)$. The entropy rate does not depend on the choice of Følner sequence (for example, [**Ol85**]).

4. *The case of* $\mathbb{Z}$

The purpose of this section is to prove the following proposition.

PROPOSITION 4.1. *Let $\mathbb{K} = \{\kappa_i\}_{i=1}^{\infty}$ be a random sofic approximation of $\mathbb{Z}$. Let $\overline{X}$ be a process over $\mathbb{Z}$ with finite range. Then $\bar{h}(\mathbb{K}, \overline{X}) = h(\mathbb{K}, \overline{X}) = h(\overline{X})$, i.e., classical entropy, sofic entropy and upper-sofic entropy agree.*

In order to prove this, we will reduce to the case when the sofic approximation $\mathbb{K}$ is particularly simple (given by finite quotients of $\mathbb{Z}$). For this, we need to define what it means for two sofic approximations to be close.

Definition 12. Let $W \subset G$ be finite and let $\epsilon > 0$. Suppose that $\sigma : G \to \mathrm{Sym}(m)$ and $\sigma' : G \to \mathrm{Sym}(m')$ are two maps and there exist subsets $Q \subset [m]$, $Q' \subset [m']$ and a bijection $\beta : Q \to Q'$ such that:

(1) $\sigma'(w)\beta(q) = \beta(\sigma(w)q)$ for all $w \in W$ and $q \in Q$ with $\sigma(w)q \in Q$;
(2) $\sigma(w)\beta^{-1}(q') = \beta^{-1}(\sigma'(w)q')$ for all $w \in W$ and $q' \in Q'$ with $\sigma'(w)q' \in Q'$; and
(3) $|Q| \geq (1-\epsilon)m$, $|Q'| \geq (1-\epsilon)m'$.

Then σ and σ' are said to be (W, ϵ)-*close* to each other. We say two probability measures κ, κ' on $\mathrm{Sym}(m)^G$, $\mathrm{Sym}(m')^G$, respectively, are (W, ϵ)-*close* if there is a probability measure ϑ on $\mathrm{Sym}(m_i)^G \times \mathrm{Sym}(m_i')^G$ with marginals κ and κ' such that $\vartheta(\mathcal{G}(W, \epsilon)) \geq 1 - \epsilon$, where $\mathcal{G}(W, \epsilon)$ is the set of all $(\sigma, \sigma') \in \mathrm{Sym}(m)^G \times \mathrm{Sym}(m')^G$ that are (W, ϵ)-close to each other. Finally, we say that two random sofic approximations to G, $\mathbb{K} = \{\kappa_i\}_{i=1}^{\infty}$ and $\mathbb{L} = \{\lambda_i\}_{i=1}^{\infty}$, are *asymptotic* if for every finite $W \subset G$ and $\epsilon > 0$, κ_i is (W, ϵ)-close to λ_i for all sufficiently large i.

The following theorem is a special case of Theorem 5.5 proven in the next section.

THEOREM 4.2. *Let $\overline{X}$ be a process over a group G with random sofic approximations $\mathbb{K}$ and $\mathbb{L}$. If $\mathbb{K}$ and $\mathbb{L}$ are asymptotic then $\bar{h}(\mathbb{K}, \overline{X}) = \bar{h}(\mathbb{L}, \overline{X})$ and $h(\mathbb{K}, \overline{X}) = h(\mathbb{L}, \overline{X})$.*

This motivates the next result, which is a minor extension of [**ES10**, Theorem 2].

THEOREM 4.3. *Let G be an amenable group and $\mathbb{K} = \{\kappa_i\}_{i=1}^{\infty}$, $\mathbb{L} = \{\lambda_i\}_{i=1}^{\infty}$ be random sofic approximations to G. Suppose that for each i, $\kappa_i : G \to \mathrm{Sym}(m_i)$, $\lambda_i : G \to \mathrm{Sym}(m_i')$ and $\lim_{i\to\infty}(m_i/m_i') = 1$. Then $\mathbb{K}$ and $\mathbb{L}$ are asymptotic.*

To prove this, we need some terminology adapted from [**ES10**].

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Definition 13. Let G be a finitely generated group and $S = S^{-1} \subset G$ a finite generating set for G. Let $\sigma : G \to \mathrm{Sym}(m)$ be a map. Then σ is an *r-approximation* to (G, S) (for r a positive integer) if there exists a set $V \subset [m]$ such that $|V| \geq (1 - (1/r))m$ and for every $v \in V$ and every sequence $g_1, \ldots, g_r \in S \cup \{e\}$,

$$\sigma(g_1)\sigma(g_2) \cdots \sigma(g_r)v = \sigma(g_1 g_2 \cdots g_r)v$$

and if $g_1', \ldots, g_r' \in S \cup \{e\}$ are such that $g_1 \cdots g_r \neq g_1' \cdots g_r'$ then

$$\sigma(g_1 g_2 \cdots g_r)v \neq \sigma(g_1' g_2' \cdots g_r')v.$$

Definition 14. If $T \subset G$ and $\sigma : G \to \mathrm{Sym}(T)$ is a map then σ is a *copy* of T if $\sigma(g)t = gt$ for every $t \in T$ and $g \in G$ with $gt \in T$.

Definition 15. If $\sigma_i : G \to \mathrm{Sym}(m_i)$ $(i = 1, \ldots, n)$ are maps then $\sigma_1 + \cdots + \sigma_n$ is defined to be the map from G to $\mathrm{Sym}(m_1 + \cdots + m_n)$ defined by

$$(\sigma_1 + \cdots + \sigma_n)(g)p := \sigma_i(g)p,$$

where i is determined as folows: if $p \leq m_1$ then $i = 1$; otherwise i is such that $\sum_{j=1}^{i-1} m_j < p \leq \sum_{j=1}^{i} m_j$. Also, if $r \geq 1$ is an integer then $r\sigma_1 := \sigma_1 + \cdots + \sigma_1$ (r summands). So, if $\alpha = (\alpha_1, \ldots, \alpha_r)$ is a vector of positive integers, and $(\sigma_1, \ldots, \sigma_r)$ is a vector of maps, then $\alpha \cdot (\sigma_1, \ldots, \sigma_r)$ is defined to be the map $\alpha_1\sigma_1 + \cdots + \alpha_r\sigma_r$.

PROPOSITION 4.4. *Let G be a finitely generated amenable group with finite generating set $S = S^{-1}$. Let $\{F_i\}_{i=1}^{\infty}$ be a left-Følner sequence for G. Then, for each integer $r > 0$, there exists an integer $R(r) > 0$ (which also depends on $\{F_i\}_{i=1}^{\infty}$), a finite subsequence $\{T_1, \ldots, T_n\} \subset \{F_i\}_{i=1}^{\infty}$ and a vector $\alpha = (\alpha_1, \ldots, \alpha_n)$ of positive natural numbers such that the following holds. Every T_i satisfies $|ST_i \,\Delta\, T_i| \leq |T_i|/r$ and every map $\sigma : G \to \mathrm{Sym}(m)$ which is an $R(r)$-approximation to (G, S) is $(S, 1/r)$-close to an integer multiple of $\alpha \cdot (\sigma_1, \ldots, \sigma_n)$, where σ_i is any copy of T_i.*

Proof. This is [**ES10**, Proposition 2.8] (in different terminology). □

Proof of Theorem 4.3. Let $W \subset G$ be finite and $\epsilon > 0$. It suffices to show that κ_i is (W, ϵ)-close to λ_i for all sufficiently large i. Let $S = W \cup W^{-1}$. Without loss of generality, we may assume G is generated by S.

Given integers $m, r > 0$, let $X(m, r)$ be the set of all $(\sigma, p) \in \mathrm{Sym}(m)^G$ such that

$$\sigma(g_1)\sigma(g_2) \cdots \sigma(g_r)p = \sigma(g_1 \cdots g_r)p \quad \text{for all } g_1, g_2, \ldots, g_r \in S \cup \{e\}$$

and if $g_1', \ldots, g_r' \in S \cup \{e\}$ are such that $g_1 g_2 \cdots g_r \neq g_1' g_2' \cdots g_r'$ then

$$\sigma(g_1 g_2 \cdots g_r)p \neq \sigma(g_1' g_2' \cdots g_r')p.$$

For $\sigma \in \mathrm{Sym}(m)^G$, let $X_\sigma(m, r)$ be the set of all $p \in [m]$ with $(\sigma, p) \in X(m, r)$.

Now let $\mathbb{K} = \{\kappa_i\}_{i=1}^{\infty}$ be a random sofic approximation. By definition, there exists an N such that for all $i > N$, $\kappa_i \times u_{m_i}(X(m_i, r)) \geq 1 - (1/r^2)$, where u_{m_i} is the uniform

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

probability measure on $[m_i]$. Let σ be an element of $\mathrm{Sym}(m_i)^G$ chosen at random with law κ_i. By Markov's inequality,

$$\mathbb{P}(m_i - |X_\sigma(m_i, r)| > m_i/r) \leq \mathbb{E}[m_i - |X_\sigma(m_i, r)|](r/m_i) \leq (m_i/r^2)(r/m_i) = (1/r).$$

Therefore,

$$\kappa_i(\{\sigma \in \mathrm{Sym}(m_i)^G : |X_\sigma(m_i, r)| \geq m_i(1 - (1/r))\}) \geq 1 - (1/r).$$

Note that each $\sigma \in \mathrm{Sym}(m_i)^G$ with $|X_\sigma(m_i, r)| \geq m_i(1 - (1/r))$ is an r-approximation to (G, S).

By the previous proposition, there exist a finite sequence $\{T_1, \ldots, T_n\}$ of finite sets $T_j \subset G$ and a vector $\alpha = (\alpha_1, \ldots, \alpha_n)$ of positive natural numbers such that the following holds. Every T_j satisfies $|ST_j \,\Delta\, T_j| \leq |T_j|/r$ and for every map $\sigma : G \to \mathrm{Sym}(m_i)$ which is an r-approximation to (G, S), there is an integer $k_i(\sigma)$ such that σ is $(S, 1/r)$-close to $k_i(\sigma)\alpha \cdot (\sigma_1, \ldots, \sigma_n)$, where σ_j is a copy of T_j for $1 \leq j \leq n$.

We would like to choose the integers $k_i(\sigma)$ to be independent of σ (although they must depend on i). So, let k_i be the minimum of $k_i(\sigma)$ over all maps $\sigma : G \to \mathrm{Sym}(m_i)$ which are r-approximations to (G, S). We claim that each such σ is $(S, 3/r)$-close to $k_i\alpha \cdot (\sigma_1, \ldots, \sigma_n)$.

To prove the claim, first note that if $\sigma : G \to \mathrm{Sym}(m)$ and $\sigma' : G \to \mathrm{Sym}(m')$ are any two maps that are $(S, 1/r)$-close to each other then $1 - (1/r) \leq m/m' \leq 1/(1 - (1/r))$. So, let q be the integer such that $\alpha \cdot (\sigma_1, \ldots, \sigma_n)$ is a map from G into $\mathrm{Sym}(q)$. Then if σ is any map from G into $\mathrm{Sym}(m_i)$ which is an r-approximation to (G, S) then $1 - (1/r) \leq k_i(\sigma)q/m_i \leq 1/(1 - (1/r))$. Since this is true for every such σ, it follows that, for any such σ,

$$1 - 2/r \leq (1 - (1/r))^2 \leq \frac{k_i q}{k_i(\sigma)q}.$$

Clearly, $k_i(\sigma)\alpha \cdot (\sigma_1, \ldots, \sigma_n)$ and $k_i\alpha \cdot (\sigma_1, \ldots, \sigma_n)$ are (S, ϵ)-close for any ϵ with $1 - \epsilon \leq k_i q/k_i(\sigma)q$. So we can choose $\epsilon = 2/r$.

It follows readily from the definition of closeness that if σ is (S, ϵ)-close to σ', σ' is (S, ϵ')-close to σ'' and $\epsilon + \epsilon' \leq 1$ then σ is $(S, \epsilon + \epsilon')$-close to σ''. So, if $\sigma : G \to \mathrm{Sym}(m_i)$ is an r-approximation to (G, S) then it is $(S, 1/r)$-close to $k_i(\sigma)\alpha \cdot (\sigma_1, \ldots, \sigma_n)$ which is $(S, 2/r)$-close to $k_i\alpha \cdot (\sigma_1, \ldots, \sigma_n)$. Thus, σ is $(S, 3/r)$-close to $k_i\alpha \cdot (\sigma_1, \ldots, \sigma_n)$ (whenever $r \geq 3$).

Let $\tau_i = k_i\alpha \cdot (\sigma_1, \ldots, \sigma_n)$. From the claim, it follows that κ_i is $(S, 3/r)$-close to the probability measure supported on $\{\tau_i\}$.

By choosing N larger if necessary, we may assume that for $i > N$, there is (by a similar argument) an integer k' such that if $\tau_i' = k_i\alpha \cdot (\sigma_1, \ldots, \sigma_n)$ then λ_i is $(S, 3/r)$-close to the probability measure supported on $\{\tau_i'\}$. Note $1 - 3/r \leq k_i' q/m_i' \leq 1/(1 - 3/r)$. So,

$$(1 - 3/r)^2 \min(m_i/m_i', m_i'/m_i) \leq \frac{k_i' q}{k_i q} \leq \frac{\max(m_i/m_i', m_i'/m_i)}{(1 - 3/r)^2}.$$

This implies τ_i and τ_i' are $(S, 1 - (1 - 3/r)^2 \min(m_i/m_i', m_i'/m_i))$-close. So, κ_i and λ_i are $(S, (6/r) + 1 - (1 - 3/r)^2 \min(m_i/m_i', m_i'/m_i))$-close to each other. Since $\lim_{i\to\infty}(m_i/m_i') = 1$ and r can be made arbitrarily large, we have shown that for every

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

$\epsilon > 0$, there exists an N such that $i > N$ implies κ_i and λ_i are (S, ϵ)-close, which implies the theorem. □

LEMMA 4.5. *Let $\mathbb{K} = \{\kappa_i\}_{i=1}^{\infty}$ be a random sofic approximation of $\mathbb{Z}$. Let $\overline{X} = (T, X, \mu, \phi)$ be a process over $\mathbb{Z}$ with finite range A. Then $h(\mathbb{K}, \overline{X}) = \bar{h}(\mathbb{K}, \overline{X})$.*

Proof. By Theorems 4.3 and 4.2, it suffices to show $h(\Sigma, \phi) = \bar{h}(\Sigma, \phi)$, where $\Sigma = \{\sigma_i\}_{i=1}^{\infty}$ and $\sigma_i : \mathbb{Z} \to \mathrm{Sym}(m_i)$ is the homomorphism with $\sigma_i(1) = (1, 2, \ldots, m_i)$.

By Corollary 2.2, it suffices to show that $h(\Sigma, \phi) \geq \bar{h}(\Sigma, \phi)$. Let $\epsilon > 0$ and $W \subset G$ be finite. It suffices to show that

$$\liminf_{n\to\infty} \frac{\log(\#\{\psi : \{1, \ldots, m_n\} \to A : d_W((\sigma_n, \psi), \phi) < \epsilon\})}{m_n} \geq \bar{h}(\Sigma, \phi) - \epsilon.$$

The definition of $\bar{h}(\Sigma, \phi)$ implies there exists an i and a probability measure ν_i on $A^{[m_i]}$ such that if δ_i is the probability measure supported on the singleton $\{\sigma_i\}$, and $\overline{X}_i$ is the approximate process constructed from $\eta_i := \delta_i \times \nu_i$, then:

- $d_W(\overline{X}_i, \overline{X}) < \epsilon/2$;
- $h(\overline{X}_i) \geq \bar{h}(\Sigma, \phi) - \epsilon/2$; and
- there is an $N > 0$ such that $W \subset [-N, N]$ and $2N/m_i < \epsilon/2$.

After perturbing ν_i if necessary, we may assume that there is an integer $d > 0$ such that $d\nu_i$ is integral (i.e., $d\nu_i(\{\xi\}) \in \mathbb{Z}$ for all $\xi \in A^{[m_i]}$). We may also assume that ν_i is σ_i-invariant by replacing it with $(1/m_i)\sum_{j=1}^{m_i}(\sigma_i^j)_*\nu_i$ if necessary.

Let $n > i$ be a large number (to be specified later). Let $k = \lfloor m_n/dm_i \rfloor$. We will say that a function $\psi : [m_n] \to A$ is *good* if for every $\xi \in A^{[m_i]}$, the number of j with $0 \leq j < kd$ satisfying

$$\psi(jm_i + p) = \xi(p) \quad \text{for all } 1 \leq p \leq m_i$$

is exactly $kd\nu_i(\{\xi\})$. For such a ψ, let $\overline{X}_\psi$ be the approximate process constructed from $\delta_n \times \delta_\psi$, where δ_ψ is the probability measure concentrated on $\{\psi\} \subset A^{[m_n]}$.

In order to estimate $d_W(\overline{X}_\psi, \overline{X}_i)$, let u_{m_n} be the uniform probability measure on $[m_n]$ (so $\psi_*^W u_{m_n}$ is the W-local statistics of $\overline{X}_\psi$ by Definition 5). If u'_{m_n} is the uniform probability measure on

$$K = \{p \in [m_n] : p = jm_i + q \text{ for some } 0 \leq j < kd \text{ and some } 1 + N \leq q \leq m_i - N\}$$

then $\psi_*^W u'_{m_n}$ is the W-local statistics of $\overline{X}_i$. This uses the fact that ν_i is σ_i-invariant and $W \subset [-N, N]$. Since $|K| = kd(m_i - 2N)$, it follows that

$$d_W(\overline{X}_\psi, \overline{X}_i) \leq 1 - \frac{kd(m_i - 2N)}{m_n}.$$

Let N_0 be large enough so that if $n > N_0$ then $1 - (kd(m_i - 2N)/m_n) < \epsilon/2$. By choice of $\overline{X}_i$, this implies $d_W(\overline{X}_\psi, \overline{X}) < \epsilon$. We will now assume that $n > N_0$.

The number of good functions $\psi : [m_n] \to A$ is

$$|A|^{(m_n - kdm_i)}(kd)!\left(\prod_{\xi \in A^{[m_i]}} (kd\nu_i(\{\xi\}))!\right)^{-1}.$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Stirling's formula implies that

$$\lim_{m_n\to\infty} \frac{\log[|A|^{(m_n-kdm_i)}(kd)!(\prod_{\xi\in A^{[m_i]}}(kd\nu_i(\{\xi\}))!)^{-1}]}{kd} = H(\nu_i).$$

Therefore,

$$\liminf_{n\to\infty} \frac{\log(\#\{\psi : \{1,\dots,m_n\}\to A : d_W((\sigma_n,\psi),\phi)<\epsilon\})}{m_n} \geq \frac{H(\nu_i)}{m_i}$$
$$= h(\overline{X}_i) \geq \bar{h}(\Sigma,\phi)-\epsilon.$$

Because $\epsilon > 0$ is arbitrary, this implies the lemma. □

LEMMA 4.6. *Let $\mathbb{K} = \{\kappa_i\}_{i=1}^{\infty}$ be a random sofic approximation of $\mathbb{Z}$. Let $\overline{X} = (T, X, \mu, \phi)$ be a process over $\mathbb{Z}$. Then $h(\mathbb{K}, \overline{X}) \geq h(\overline{X})$.*

Proof. By Theorems 4.3 and 4.2 and the previous lemma it suffices to show that $\bar{h}(\Sigma,\phi) \geq h(\overline{X})$, where $\Sigma = \{\sigma_i\}_{i=1}^{\infty}$ and $\sigma_i : \mathbb{Z} \to \mathrm{Sym}(m_i)$ is the homomorphism with $\sigma_i(1) = (1, 2, \dots, m_i)$.

Let $\phi^{m_i} : X \to A^{[m_i]}$ be the map $\phi^{m_i}(x)(p) = \phi(T^p x)$ for all $x \in X$, $p \in [m_i]$. Let $\eta_i := \phi^{m_i}_* \mu$ be the pushforward measure on $A^{[m_i]}$. Let $\overline{X}_i$ be the approximate process constructed from $\delta_i \times \eta_i$, where δ_i is the probability measure concentrated on $\{\sigma_i\} \subset \mathrm{Sym}(m_i)^{\mathbb{Z}}$. The Følner property of the sequence of intervals $[m_i] \subset \mathbb{Z}$ implies that $\lim_{i\to\infty} \overline{X}_i = \overline{X}$. The definition of $h(\overline{X})$ implies $\lim_{i\to\infty} h(\overline{X}_i) = h(\overline{X})$. This implies the lemma. □

Proof of Proposition 4.1. By Theorems 4.3 and 4.2 and the previous lemma, it suffices to prove that if $\Sigma = \{\sigma_i\}_{i=1}^{\infty}$, where $\sigma_i : \mathbb{Z} \to \mathrm{Sym}(m_i)$ is the homomorphism with $\sigma_i(1) = (1, 2, \dots, m_i)$, then $\bar{h}(\Sigma, \overline{X}) \leq h(\overline{X})$. Let $\{\nu_i\}_{i=1}^{\infty}$ be a sequence of probability measures on $A^{[m_i]}$ such that if $\{\overline{X}_i\}_{i=1}^{\infty}$ is the sequence of approximate processes constructed from $\eta_i := \delta_i \times \nu_i$ (where δ_i is the probability measure concentrated on $\{\sigma_i\}_{i=1}^{\infty} \subset \mathrm{Sym}(m_i)^{\mathbb{Z}}$) then

$$\lim_{i\to\infty} \overline{X}_i = \overline{X} \quad \text{and} \quad \lim_{i\to\infty} h(\overline{X}_i) = \bar{h}(\Sigma, \overline{X}).$$

Let $\overline{X} = (T, X, \mu, \phi)$, where $\phi : X \to A$. Using a standard trick, $\overline{X}$ is equivalent to a process of the form $(\tau, A^{\mathbb{Z}}, \mu', \phi')$, where $\tau : A^{\mathbb{Z}} \to A^{\mathbb{Z}}$ is the shift map $\tau(y)(n) := y(n+1)$ and $\phi' : A^{\mathbb{Z}} \to A$ is the time-0 projection $\phi'(y) := y(0)$. To be precise, let $\phi^{\mathbb{Z}} : X \to A^{\mathbb{Z}}$ be the map $\phi^{\mathbb{Z}}(x)(n) := \phi(T^n x)$ for $x \in X$, $n \in \mathbb{Z}$. This map is equivariant. $\overline{X}$ is equivalent to $(\tau, A^{\mathbb{Z}}, \mu', \phi')$, where μ' is the pushforward measure $\phi^{\mathbb{Z}}_* \mu$. So, without loss of generality, we will assume that $X = A^{\mathbb{Z}}$, $T = \tau$ is the shift map and $\phi : A^{\mathbb{Z}} \to A$ is the time-0 projection.

Let $\pi_n : (A^{[m_n]})^{\mathbb{Z}} \to A^{\mathbb{Z}}$ be the map defined by

$$\pi_n[\Psi](im_n + j) := \Psi(i)(j) \quad \text{for all } \Psi \in (A^{[m_n]})^{\mathbb{Z}},\ i \in \mathbb{Z},\ j \in [m_n].$$

Let $\nu_n^{\mathbb{Z}}$ be the measure on $(A^{[m_n]})^{\mathbb{Z}}$ equal to the product of $\mathbb{Z}$-copies of ν_n. Let $\mu_n' = (\pi_n)_*(\nu_n^{\mathbb{Z}})$ be the pushforward measure on $A^{\mathbb{Z}}$. Let

$$\mu_n = \frac{1}{m_n}\sum_{j=1}^{m_n} \tau_*^j \mu_n'.$$

Note that μ_n is τ-invariant.

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

For an interval $[a, b] \subset \mathbb{Z}$, let $\phi^{[a,b]} : A^{\mathbb{Z}} \to A^{[a,b]}$ be the projection map. By concavity of entropy,

$$h_{\mu_n}(\phi) = \lim_{N\to\infty} \frac{H(\phi_*^{[-m_nN+1,m_nN]}\mu_n)}{2m_nN} \geq \lim_{N\to\infty} \frac{H(\phi_*^{[-m_nN+1,m_nN]}\mu_n')}{2m_nN}$$
$$= \lim_{N\to\infty} \frac{2NH(\nu_n)}{2m_nN} = \frac{H(\nu_n)}{m_n}.$$

Since $\lim_{i\to\infty} \overline{X}_i = \overline{X}$, it follows that $\lim_n \mu_n = \mu$ in the weak* topology on $M(A^{\mathbb{Z}})$, the space of all τ-invariant Borel probability measures on $A^{\mathbb{Z}}$. It is well known that the function $\lambda \in M(A^{\mathbb{Z}}) \mapsto h_\lambda(\phi)$ is upper-semi-continuous on $M(A^{\mathbb{Z}})$; for example, see [**Gl03**, Lemma 15.1, p. 270]. It follows that

$$\bar{h}(\Sigma, \phi) = \limsup_{n\to\infty} \frac{H(\nu_n)}{m_n} \leq \limsup_{n\to\infty} h_{\mu_n}(\phi) \leq h_\mu(\phi),$$

as required. □

5. *Relative entropy*

Definition 16. (Factors of approximate processes) Given an approximate process $\overline{X} = (T, X, \mu, \phi)$ with $\phi : X \to A$ and a function $\beta : A \to B$, let $\beta \circ \overline{X}$ be the approximate process

$$\beta \circ \overline{X} := (T, X, \mu, \beta \circ \phi).$$

If an approximate process $\overline{Y}$ is constructed from a measure η_i on $\mathrm{Sym}(m_i)^G \times A^{[m_i]}$, as in Definition 6, then $\beta \circ \overline{Y}$ has an alternative description as follows. Let

$$\tilde{\beta} : \mathrm{Sym}(m_i)^G \times A^{[m_i]} \to \mathrm{Sym}(m_i)^G \times B^{[m_i]}, \quad \tilde{\beta}(\sigma, \xi) = (\sigma, \beta \circ \xi).$$

Then $\beta \circ \overline{Y}$ is equivalent to the process $\overline{Z}$ constructed from the pushforward measure $\tilde{\beta}_*\eta_i$. (By equivalent, we mean that $d_W(\beta \circ \overline{Y}, \overline{Z}) = 0$ for every finite $W \subset G$ in the notation of Definition 4.)

The next lemma follows immediately from the definitions.

LEMMA 5.1. *If $\{\overline{X}_i\}_{i=1}^{\infty}$ is a sequence of approximate processes with range A, $\lim_{i\to\infty} \overline{X}_i = \overline{X}$ and $\beta : A \to B$ is a map then $\lim_{i\to\infty} \beta \circ \overline{X}_i = \beta \circ \overline{X}$. Moreover, if $\lim_{i\to\infty} \overline{X}_i = \overline{X}$ strongly then $\lim_{i\to\infty} \beta \circ \overline{X}_i = \beta \circ \overline{X}$ strongly.*

Definition 17. (Relative entropy) Let G be a countable amenable group acting by measure-preserving transformations on a standard probability space $(X, \mathcal{B}, \mu)$. Let $\phi : X \to A$ be a finite observable, Σ be a sofic approximation to G and $\mathcal{F} \subset \mathcal{B}$ be a G-invariant σ-algebra. Define

$$h_\mu(\phi|\mathcal{F}) := \inf_\psi h_\mu(\phi \vee \psi) - h_\mu(\psi),$$

where the infimum is over all finite-range $\mathcal{F}$-measurable observables $\psi : X \to B$ and $\phi \vee \psi : X \to A \times B$ is the observable $\phi \vee \psi(x) = (\phi(x), \psi(x))$. In case $\mathcal{F}$ is the G-invariant σ-algebra generated by an observable ψ then we write $h_\mu(\phi|\psi) = h_\mu(\phi|\mathcal{F})$.

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

In case $\overline{X} = (T, X, \mu, \phi)$ is a process over G and $\psi = \beta \circ \phi$ for some $\beta : A \to B$, we write $h(\overline{X}|\beta \circ \overline{X}) = h(\overline{X}) - h(\beta \circ \overline{X}) = h_\mu(\phi|\psi)$. The first equality holds from the Abramov–Rohlin formula [**WZ92**].

We can now define relative sofic entropy (in a special case).

Definition 18. (Relative sofic entropy) Let $\mathbb{K}$ be a random sofic approximation to G. Let $\overline{X}$ be a G-process with finite range A and $\beta : A \to B$ be a map. Define

$$\bar{h}(\mathbb{K}, \overline{X}|\beta \circ \overline{X}) = \sup \limsup_{j\to\infty} h(\overline{X}_j) - h(\beta \circ \overline{X}_j),$$

where the supremum is over all sequences $\{\overline{X}_j\}_{j=1}^\infty$ adapted to $\mathbb{K}'$ (where $\mathbb{K}'$ is a subsequence of $\mathbb{K}$) such that $\lim_{j\to\infty} \overline{X}_j = \overline{X}$. Similarly, let

$$h(\mathbb{K}, \overline{X}|\beta \circ \overline{X}) = \sup \limsup_{j\to\infty} h(\overline{X}_j) - h(\beta \circ \overline{X}_j),$$

where the supremum is over all sequences $\{\overline{X}_j\}_{j=1}^\infty$ adapted to $\mathbb{K}'$ (where $\mathbb{K}'$ is a subsequence of $\mathbb{K}$) such that $\lim_{j\to\infty} \overline{X}_j = \overline{X}$ strongly. If $\overline{X} = (T, X, \mu, \phi)$ and $\psi = \beta \circ \phi$ then an alternative notation for relative entropy is

$$\bar{h}(\mathbb{K}, \phi|\psi) := \bar{h}(\mathbb{K}, \overline{X}|\beta \circ \overline{X}), \quad h(\mathbb{K}, \phi|\psi) := h(\mathbb{K}, \overline{X}|\beta \circ \overline{X}).$$

We may also write $\bar{h}_\mu(\mathbb{K}, \phi|\psi)$ or $h_\mu(\mathbb{K}, \phi|\psi)$ if it is desirable to emphasize the dependence on the measure μ.

Before moving on, it is worthwhile to record some inequalities relating the entropy of direct products to the entropies of their direct factors. To be precise, if $\overline{X} = (T, X, \mu, \phi)$, $\overline{Y} = (S, Y, \nu, \psi)$ are two approximate processes over G then their direct product is the process $\overline{X} \times \overline{Y} := (T \times S, X \times Y, \mu \times \nu, \phi \times \psi)$, where

$$(T \times S)_g(x, y) := (T_g x, S_g y) \quad \text{for all } g \in G, (x, y) \in X \times Y.$$

Let $\pi_B : A \times B \to B$ and $\pi_A : A \times B \to A$ be the projection maps. To simplify, we let (for example)

$$\bar{h}(\mathbb{K}, \overline{X} \times \overline{Y}|\overline{Y}) := \bar{h}(\mathbb{K}, \overline{X} \times \overline{Y}|\pi_B \circ \overline{X} \times \overline{Y}).$$

LEMMA 5.2. *If $\overline{X}$, $\overline{Y}$ are two processes over G, as above, and $\mathbb{K}$ is a random sofic approximation then*

$$\bar{h}(\mathbb{K}, \overline{X} \times \overline{Y}) \le \bar{h}(\mathbb{K}, \overline{X}) + \bar{h}(\mathbb{K}, \overline{Y}),$$
$$\bar{h}(\mathbb{K}, \overline{X} \times \overline{Y}|\overline{Y}) \le \bar{h}(\mathbb{K}, \overline{X}).$$

Similar statements hold with lower-sofic entropy in place of upper-sofic entropy.

Proof. Let $\overline{X} = (T, X, \mu, \phi)$ and $\overline{Y} = (S, Y, \nu, \psi)$. Let $\{\overline{Z}_i\}_{i=1}^\infty$ be a sequence of approximate processes adapted to $\mathbb{K}$ so that $\lim_{i\to\infty} \overline{Z}_i = \overline{X} \times \overline{Y}$. Then $\{\pi_A \circ \overline{Z}_i\}_{i=1}^\infty$ converges to $\overline{X}$ and $\{\pi_B \circ \overline{Z}_i\}_{i=1}^\infty$ converges to $\overline{Y}$. Moreover,

$$h(\overline{Z}_i) \le h(\pi_A \circ \overline{Z}_i) + h(\pi_B \circ \overline{Z}_i).$$

This and Lemma 5.1 imply

$$\bar{h}(\mathbb{K}, \overline{X} \times \overline{Y}) \le \bar{h}(\mathbb{K}, \overline{X}) + \bar{h}(\mathbb{K}, \overline{Y}), \quad \bar{h}(\mathbb{K}, \overline{X} \times \overline{Y}|\overline{Y}) \le \bar{h}(\mathbb{K}, \overline{X}).$$

The proofs for lower-sofic entropy are similar. □

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

If G is non-amenable then there are examples showing that some of the inequalities of the lemma above can be strict. However, in the special case of Bernoulli actions, we have equality. To be precise, let (Ω, ω) be a standard probability space. G acts on the product space (Ω^G, ω^G) by $T_g y(f) = y(g^{-1}f)$ for all $y \in \Omega^G$, $g, f \in G$. Let $\phi : \Omega^G \to \Omega$ be the map $\phi(y) = y(e)$. The process $(T, \Omega^G, \omega^G, \phi)$ is the *Bernoulli process* over G with base (Ω, ω).

LEMMA 5.3. *If $\overline{X}$ is any finite-range process over G, $\overline{Y}$ is a Bernoulli process with base (Ω, ω) (where Ω is finite) and $\mathbb{K}$ is a random sofic approximation to G then*

$$\bar{h}(\mathbb{K}, \overline{X} \times \overline{Y}) = \bar{h}(\mathbb{K}, \overline{X}) + \bar{h}(\mathbb{K}, \overline{Y}),$$
$$\bar{h}(\mathbb{K}, \overline{X} \times \overline{Y} | \overline{Y}) = \bar{h}(\mathbb{K}, \overline{X}).$$

Similar statements hold with lower-sofic entropy in place of upper-sofic entropy.

Proof. The first statement above was proven in [**Bo10a**] for non-random sofic approximations and with lower-sofic entropy in place of upper-sofic entropy. We will handle here only the case of upper-sofic entropy as the other cases are similar.

Let $\{\overline{X}_j\}_{j=1}^{\infty}$ be a sequence of approximate processes adapted to $\mathbb{K}' = \{\kappa_j\}_{j=1}^{\infty}$ (where $\mathbb{K}'$ is a subsequence of $\mathbb{K}$) such that $\lim_{j\to\infty} \overline{X}_j = \overline{X}$ and $\lim_{j\to\infty} h(\overline{X}_j) = \bar{h}(\mathbb{K}, \overline{X})$. Let η_j be the probability measure on $\mathrm{Sym}(m_j)^G \times A^{[m_j]}$ from which $\overline{X}_j$ is constructed (as in Definition 6). Let $\tilde{\eta}_j = \eta_j \times \omega^{[m_j]}$ be a probability measure on $\mathrm{Sym}(m_j)^G \times A^{[m_j]} \times \Omega^{[m_j]} = \mathrm{Sym}(m_j)^G \times (A \times \Omega)^{[m_j]}$ and let $\overline{Z}_j$ be the approximate process constructed from $\tilde{\eta}_j$. Note $h(\overline{Z}_j) = h(\overline{X}_j) + H(\omega)$.

We claim that $\lim_{j\to\infty} \overline{Z}_j = \overline{X} \times \overline{Y}$. To see this, let $V \subset G$ be finite. Let $\mathcal{G}_j(V)$ be the set of all $(\sigma, p) \in \mathrm{Sym}(m_j)^G \times [m_j]$ such that $\sigma(g)\sigma(h)p = \sigma(gh)p$ for all $g, h \in V$ and for every $g \neq h$ with $g, h \in V$, $\sigma(g)p \neq \sigma(h)p$. We consider $\mathcal{G}_j(V) \times (A \times \Omega)^{[m_j]}$ as a subset of $\mathrm{Sym}(m_j)^G \times (A \times \Omega)^{[m_j]} \times [m_j]$ in the obvious way. Then $\lim_{j\to\infty} \tilde{\eta}_j \times u_{m_j}(\mathcal{G}_j(V) \times (A \times \Omega)^{[m_j]}) = 1$ because $\mathbb{K}$ is a sofic approximation.

In general, if τ is a measure on a set J and $J_0 \subset J$ then we write $\tau | J_0$ to denote τ restricted to J_0. Let $\chi_j : \mathrm{Sym}(m_j)^G \times (A \times \Omega)^{[m_j]} \times [m_j] \to A \times \Omega$ be the projection map $\chi_j(\sigma, \xi, p) = \xi(p)$. Then the measure $(\chi_j^V)_*(\tilde{\eta}_j \times u_{m_j} | \mathcal{G}_j(V) \times (A \times \Omega)^{[m_j]})$ splits as a product $\tau_j \times \omega^V$ for some measure τ_j on A^V. Let $\overline{X} = (T, X, \mu, \psi)$. Thus,

$$\begin{aligned}
\lim_{j\to\infty} d_V(\overline{Z}_j, \overline{X} \times \overline{Y}) &= \lim_{j\to\infty} \tfrac{1}{2} \| (\chi_j^V)_*(\tilde{\eta}_j \times u_{m_j}) - (\psi \times \phi)_*^V \mu \times \omega^G \|_1 \\
&= \lim_{j\to\infty} \tfrac{1}{2} \| (\chi_j^V)_*(\tilde{\eta}_j \times u_{m_j} | \mathcal{G}(V) \times (A \times \Omega)^{[m_j]}) \\
&\quad - (\psi \times \phi)_*^V \mu \times \omega^G \|_1 \\
&= \lim_{j\to\infty} \tfrac{1}{2} \| (\chi_j^V)_*(\eta_j \times u_{m_j} | \mathcal{G}(V) \times A^{[m_j]}) - \psi_*^V \mu \|_1 \\
&= \lim_{j\to\infty} d_V(\overline{X}_j, \overline{X}) = 0.
\end{aligned}$$

The first equality holds by definition of d_V. The second one holds because $\lim_{j\to\infty} \tilde{\eta}_j \times u_{m_j}(\mathcal{G}_j(V) \times (A \times \Omega)^{[m_j]}) = 1$. The third equality holds because

$$(\chi_j^V)_*(\tilde{\eta}_j \times u_{m_j} | \mathcal{G}(V) \times (A \times \Omega)^{[m_j]})$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

splits as a product $\tau_j \times \omega^V$ and $(\psi \times \phi)^V_* \mu \times \omega^G$ splits as the product $\psi^V_* \mu \times \omega^V$. The last equality holds because $\lim_{j\to\infty} \eta_j \times u_{m_j}(\mathcal{G}_j(V) \times A^{[m_j]}) = 1$ since $\mathbb{K}$ is a sofic approximation.

Since $\lim_{j\to\infty} \overline{Z}_j = \overline{X} \times \overline{Y}$, it follows that

$$\bar{h}(\mathbb{K}, \overline{X} \times \overline{Y}) \geq \bar{h}(\mathbb{K}, \overline{X}) + H(\omega), \quad \bar{h}(\mathbb{K}, \overline{X} \times \overline{Y} | \overline{Y}) \geq \bar{h}(\mathbb{K}, \overline{X}).$$

By the previous lemma, it now suffices to prove that $\bar{h}(\mathbb{K}, \overline{Y}) = H(\omega)$.

Applying the equation above to the case when $\overline{X}$ is trivial, we see that $\bar{h}(\mathbb{K}, \overline{Y}) \geq H(\omega)$. Suppose that $\{\overline{Y}_j\}_{j=1}^\infty$ is a sequence of approximate processes adapted to $\mathbb{K}' = \{\kappa_j\}_{j=1}^\infty$ (where $\mathbb{K}'$ is a subsequence of $\mathbb{K}$) such that $\lim_{j\to\infty} \overline{Y}_j = \overline{Y}$ and $\lim_{j\to\infty} h(\overline{Y}_j) = \bar{h}(\mathbb{K}, \overline{Y})$. Let η_j be the probability measure on $\mathrm{Sym}(m_j)^G \times \Omega^{[m_j]}$ from which $\overline{Y}_j$ is constructed. Let $\pi : \mathrm{Sym}(m_j)^G \times \Omega^{[m_j]} \times [m_j] \to \Omega$ be the map $\pi(\sigma, \xi, p) = \xi(p)$ and let $\omega_j = \pi_*(\eta_j \times u_{[m_j]})$ be the pushforward measure. Because $\lim_{j\to\infty} \overline{Y}_j = \overline{Y}$, ω_j converges to ω.

We claim that $h(\overline{Y}_j) \leq H(\omega_j)$. Let $\eta_{j,\sigma}$ be the fiber measure of η_j over $\sigma \in \mathrm{Sym}(m_j)^G$. This is a measure on $\Omega^{[m_j]}$. By abuse of notation we let $\pi : \Omega^{[m_j]} \times [m_j] \to \Omega$ denote the map $\pi(\xi, p) = \xi(p)$, and for $p \in [m_j]$ let $\pi_p : \Omega^{[m_j]} \to \Omega$ denote the map $\pi(\xi) = \xi(p)$.

By concavity of the entropy function,

$$\begin{aligned} \frac{1}{m_j} H(\eta_{j,\sigma}) &\leq \frac{1}{m_j} \sum_{p\in[m_j]} H((\pi_p)_*(\eta_{j,\sigma})) \\ &\leq H\left(\frac{1}{m_j} \sum_{p\in[m_j]} (\pi_p)_*(\eta_{j,\sigma})\right) = H(\pi_*(\eta_{j,\sigma} \times u_{[m_j]})). \end{aligned}$$

So, by concavity of entropy again,

$$\begin{aligned} h(\overline{Y}_j) &= \frac{1}{m_j} H(\eta_j | \kappa_j) = \frac{1}{m_j} \int H(\eta_{j,\sigma})\, d\kappa_j(\sigma) \\ &\leq \int H(\pi_*(\eta_{j,\sigma} \times u_{[m_j]}))\, d\kappa_j(\sigma) \leq H\left(\int \pi_*(\eta_{j,\sigma} \times u_{[m_j]})\, d\kappa_j(\sigma)\right) = H(\omega_j). \end{aligned}$$

So

$$\bar{h}(\mathbb{K}, \overline{Y}) = \lim_{j\to\infty} h(\overline{Y}_j) \leq \limsup_{j\to\infty} H(\omega_j) = H(\omega).$$

Since we have already shown that $\bar{h}(\mathbb{K}, \overline{Y}) \geq H(\omega)$, we now know that $\bar{h}(\mathbb{K}, \overline{Y}) = H(\omega)$. □

LEMMA 5.4. *If $\overline{X}$ is any finite-range process over G, $\overline{Y} = (T, \Omega^G, \omega^G, \phi)$ is a Bernoulli process with base (Ω, ω) (where Ω is finite), Ω' is a finite set, $\phi' : \Omega^G \to \Omega'$, $\beta : \Omega' \to \Omega$ are Borel maps such that $\phi = \beta \circ \phi'$, $\overline{Y}' = (T, \Omega^G, \omega^G, \phi')$ and $\mathbb{K}$ is a random sofic approximation to G then*

$$h(\mathbb{K}, \overline{X} \times \overline{Y}' | \overline{Y}') = h(\mathbb{K}, \overline{X}), \quad \bar{h}(\mathbb{K}, \overline{X} \times \overline{Y}' | \overline{Y}') = \bar{h}(\mathbb{K}, \overline{X}).$$

Proof. We will only prove the statement for upper-sofic entropy as the statement for lower-sofic entropy is similar. By Lemma 5.2, it suffices to prove $\bar{h}(\mathbb{K}, \overline{X} \times \overline{Y}' | \overline{Y}') \geq \bar{h}(\mathbb{K}, \overline{X})$.

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Let $\mathbb{K} = \{\kappa_i\}_{i=1}^{\infty}$, where κ_i is a measure on $\mathrm{Sym}(m_i)^G$. Let $F : \mathrm{Sym}(m_i)^G \to \mathrm{Sym}(m_i)^G$ be the map $F(\sigma)(g) = \sigma(g)$ if $g \neq e$ and $F(\sigma)(e)$ is the identity permutation. Let $\kappa_i' := F_*\kappa_i$. Observe that $\mathbb{K}' := \{\kappa_i'\}_{i=1}^{\infty}$ is asymptotic to $\mathbb{K}$. Therefore, by Theorem 5.5, we may assume, without loss of generality, that $\mathbb{K} = \mathbb{K}'$.

Define $\overline{X}_j$, η_j, κ_j (for $j = 1, 2, \ldots$) as in the proof of the previous lemma. For $V \subset G$ finite with $e \in V$, $\sigma \in \mathrm{Sym}(m_j)^G$ and $\zeta \in \Omega^{[m_j]}$, define $\zeta[\sigma, V] \in (\Omega^V)^{[m_j]}$ by

$$\zeta[\sigma, V](p) := [v \in V \mapsto \zeta(\sigma(v)p)].$$

Let $\tilde{\eta}_j^V$, η_j^V be the measures on $\mathrm{Sym}(m_j)^G \times A^{[m_j]} \times (\Omega^V)^{[m_j]}$ obtained by pushing $\eta_j \times \omega^{[m_j]}$ forward under the maps

$$(\sigma, \xi, \zeta) \mapsto (\sigma, \xi, \zeta[\sigma, V]), \quad (\sigma, \xi, \zeta) \mapsto (\sigma, \zeta[\sigma, V]),$$

respectively. Let $\overline{Y}_j^V$, $\overline{Z}_j^V$ be the approximate process constructed from η_j^V, $\tilde{\eta}_j^V$, respectively. Also, let $\overline{Y}_j = \overline{Y}_j^{\{e\}}$. As in the previous lemma, we obtain

$$\lim_{j\to\infty} \overline{Z}_j^V = \overline{X} \times \overline{Y}^V, \quad \lim_{j\to\infty} \overline{Y}_j^V = \overline{Y}^V,$$

where $\overline{Y}^V = (T, \Omega^G, \omega^G, \phi^V)$.

Suppose that $\beta' : \Omega^V \to \Omega'$ is a map such that $\beta\beta' : \Omega^V \to \Omega$ is the map at the identity element. Let id denote the identity map on A. Then,

$$\lim_{j\to\infty} \mathrm{id} \times \beta' \circ \overline{Z}_j^V = \overline{X} \times (\beta' \circ \overline{Y}^V)$$

and because the map $(\sigma, \xi, \zeta) \mapsto (\sigma, \xi, \beta'\zeta[\sigma, V])$ is injective, $h(\mathrm{id} \times \beta' \circ \overline{Z}_j^V) = h(\overline{Z}_j)$, which implies (by the previous lemma)

$$\lim_{j\to\infty} h(\mathrm{id} \times \beta' \circ \overline{Z}_j^V) = h(\mathbb{K}, \overline{X} \times \overline{Y}) = h(\mathbb{K}, \overline{X}) + H(\omega).$$

Also, the map $(\sigma, \zeta) \mapsto (\sigma, \beta'\zeta[\sigma, V])$ is injective. So, $h(\beta' \circ \overline{Y}_j^V) = h(\overline{Y}_j)$. By definition, $\overline{Y}_j$ is the approximate process constructed from $\kappa_j \times \omega^{[m_j]}$. So, $h(\overline{Y}_j) = H(\omega)$.

Because $\phi = \beta \circ \phi'$, there exist a sequence $\{V_k\}_{k=1}^{\infty}$ of finite subsets of G (with $e \in V_k$ for all k) and a sequence of maps $\beta_k' : \Omega^{V_k} \to \Omega'$ such that $\beta\beta_k' : \Omega^{V_k} \to \Omega$ is the evaluation map at the identity element and $\beta_k'\phi^{V_k}$ limits on ϕ' in the following sense:

$$\lim_{k\to\infty} \mu(\{x \in X : \beta_k'\phi^{V_k}(x) = \phi'(x)\}) = 1.$$

Note that

$$\lim_{k\to\infty} \lim_{j\to\infty} \mathrm{id} \times \beta_k' \circ \overline{Z}_j^{V_k} = \lim_{k\to\infty} \mathrm{id} \times \beta_k' \circ \overline{X} \times \overline{Y}^{V_k} = \overline{X} \times \overline{Y}',$$

$$\lim_{k\to\infty} \lim_{j\to\infty} h(\mathrm{id} \times \beta_k' \circ \overline{Z}_j^{V_k}) = h(\mathbb{K}, \overline{X}) + H(\omega).$$

So, a diagonalization argument implies that, without loss of generality, we may assume $\{V_k\}_{k=1}^{\infty}$ and $\{\beta_k'\}_{k=1}^{\infty}$ are chosen so that there is an increasing sequence $\{i(j)\}_{j=1}^{\infty}$ of positive integers with

$$\lim_{j\to\infty} \mathrm{id} \times \beta_{i(j)}' \circ \overline{Z}_{i(j)}^{V_{i(j)}} = \overline{X} \times \overline{Y}',$$

$$\lim_{j\to\infty} h(\mathrm{id} \times \beta_{i(j)}' \circ \overline{Z}_{i(j)}^{V_{i(j)}}) = h(\mathbb{K}, \overline{X}) + H(\omega).$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Let $\pi : A \times \Omega' \to \Omega'$ be the projection map. Note $\beta'_{i(j)} \circ \overline{Y}^{V_{i(j)}}_{i(j)} = \pi \circ (\mathrm{id} \times \beta'_{i(j)}) \circ \overline{Z}^{V_{i(j)}}_{i(j)}$. So, $h(\pi \circ (\mathrm{id} \times \beta'_{i(j)}) \circ \overline{Z}^{V_{i(j)}}_{i(j)}) = H(\omega)$. Therefore,

$$\lim_{j\to\infty} h(\mathrm{id} \times \beta'_{i(j)} \circ \overline{Z}^{V_{i(j)}}_{i(j)}) - h(\pi \circ (\mathrm{id} \times \beta'_{i(j)}) \circ \overline{Z}^{V_{i(j)}}_{i(j)}) = h(\mathbb{K}, \overline{X}).$$

Thus, $h(\mathbb{K}, \overline{X} \times \overline{Y}'|\overline{Y}') \geq h(\mathbb{K}, \overline{X})$, as required. □

Next, we extend Theorem 4.2 to relative entropy.

THEOREM 5.5. *Let $\overline{X}$ be a process over a group G with random sofic approximations $\mathbb{K}$ and $\mathbb{L}$. We assume A is the range of $\overline{X}$ and $\beta : A \to B$ is a map (both A and B are finite). If $\mathbb{K}$ and $\mathbb{L}$ are asymptotic then $\bar{h}(\mathbb{K}, \overline{X}|\overline{X} \circ \beta) = \bar{h}(\mathbb{L}, \overline{X}|\overline{X} \circ \beta)$ and $h(\mathbb{K}, \overline{X}|\overline{X} \circ \beta) = h(\mathbb{L}, \overline{X}|\overline{X} \circ \beta)$.*

In order to prove this, we will need a few lemmas. We say that elements $\sigma, \sigma' \in \mathrm{Sym}(m)^G$ are *conjugate* if there exists an element $\tau \in \mathrm{Sym}(m)$ such that for all $g \in G$, $\sigma(g) = \tau\sigma'(g)\tau^{-1}$. We say that probability measures κ, κ' on $\mathrm{Sym}(m)^G$ are conjugate if there exists a probability measure ϑ on $\mathrm{Sym}(m)^G \times \mathrm{Sym}(m)^G$ with marginals κ, κ' such that ϑ is supported on the set of all conjugate pairs $(\sigma, \sigma') \in \mathrm{Sym}(m)^G \times \mathrm{Sym}(m)^G$.

LEMMA 5.6. *Let κ, κ' be conjugate probability measures on $\mathrm{Sym}(m)^G$. Let η be a probability measure on $\mathrm{Sym}(m)^G \times A^{[m]}$ with projection κ and let $\overline{X}$ be the approximate process constructed from η. Let $\beta : A \to B$ be a map to a finite set B. Let $W \subset G$ be finite. Then, there is an approximate process $\overline{X}'$ constructed from a measure η' on $\mathrm{Sym}(m)^G \times A^{[m]}$ such that:*

(1) *the projection of η' to $\mathrm{Sym}(m)^G$ is κ';*
(2) *$d_W(\overline{X}, \overline{X}') = 0$; and*
(3) *$h(\overline{X}'|\beta \circ \overline{X}') \geq h(\overline{X}|\beta \circ \overline{X})$.*

Moreover, if $\overline{X} = (T, X, \mu, \phi)$ is a process over G then for any $\epsilon > 0$,

$$\eta'(\{(\sigma, \psi) : d_W((\sigma, \psi), \phi) \leq \epsilon\}) = \eta(\{(\sigma, \psi) : d_W((\sigma, \psi), \phi) \leq \epsilon\}). \tag{2}$$

Proof. Case 1. Suppose κ' is supported on a singleton $\{\sigma\}$. Then, κ-almost every (a.e.) σ' is conjugate to σ, so κ is supported on a finite set, which we denote by $\{\sigma_1, \ldots, \sigma_n\}$. For each i, there is an element $\tau_i \in \mathrm{Sym}(m)$ such that $\sigma = \tau_i\sigma_i\tau_i^{-1}$. Let η' be the measure on $\mathrm{Sym}(m)^G \times A^{[m]}$ defined by

$$\eta'(\{(\sigma, \psi)\}) := \eta(\{(\sigma_i, \psi \circ \tau_i) : 1 \leq i \leq n\}).$$

Then, η' projects to κ' and

$$\psi(\sigma(g)p) = \psi \circ \tau_i(\sigma_i(g)\tau_i^{-1}p) \quad \text{for all } p \in [m],\ \psi \in A^{[m]},\ 1 \leq i \leq n,\ g \in G$$

implies $d_W(\overline{X}, \overline{X}') = 0$, where $\overline{X}'$ is the approximate process constructed from η'. It also implies (2).

For $\gamma \in B^{[m]}$ with $\eta'(\{(\sigma, \psi') : \beta\psi' = \gamma\}) > 0$, let η'_γ be the measure on $A^{[m]}$ defined by

$$\eta'_\gamma(\{\psi\}) = \begin{cases} \dfrac{\eta'(\{(\sigma, \psi)\})}{\eta'(\{(\sigma, \psi') : \beta\psi' = \gamma\})} & \text{if } \beta\psi = \gamma, \\ 0 & \text{otherwise.} \end{cases}$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Then

$$h(\overline{X}'|\beta \circ \overline{X}') = \frac{1}{m} \sum_{\gamma} H(\eta'_{\gamma}) \eta'(\{(\sigma, \psi') : \beta\psi' = \gamma\}).$$

If $\eta(\{(\sigma_i, \psi'\tau_i) : \beta\psi' = \gamma\}) > 0$ then let $\eta_{i,\gamma}$ be the measure on $A^{[m]}$ defined by

$$\eta_{i,\gamma}(\{\psi\}) = \begin{cases} \dfrac{\eta(\{(\sigma_i, \psi\tau_i)\})}{\eta(\{(\sigma_i, \psi'\tau_i) : \beta\psi' = \gamma\})} & \text{if } \beta\psi = \gamma, \\ 0 & \text{otherwise.} \end{cases}$$

Observe that

$$\eta'_{\gamma} = \frac{\sum_{i=1}^{n} \eta(\{(\sigma_i, \psi'\tau_i) : \beta\psi' = \gamma\}) \eta_{i,\gamma}}{\eta'(\{(\sigma, \psi') : \beta\psi' = \gamma\})}.$$

By concavity of entropy,

$$H(\eta'_{\gamma}) \geq \frac{\sum_{i=1}^{n} \eta(\{(\sigma_i, \psi'\tau_i) : \beta\psi' = \gamma\}) H(\eta_{i,\gamma})}{\eta'(\{(\sigma, \psi') : \beta\psi' = \gamma\})},$$

which implies

$$\begin{aligned} h(\overline{X}'|\beta \circ \overline{X}') &\geq \frac{1}{m} \sum_{i=1}^{n} \sum_{\gamma} \eta(\{(\sigma_i, \psi'\tau_i) : \beta\psi' = \gamma\}) H(\eta_{i,\gamma}) \\ &= \frac{1}{m} \sum_{i=1}^{n} \sum_{\gamma} \eta(\{(\sigma_i, \psi'\tau_i) : \beta\psi' = \gamma\tau_i^{-1}\}) H(\eta_{i,\gamma\tau_i^{-1}}) = h(\overline{X}|\beta \circ \overline{X}). \end{aligned}$$

This proves Case 1.

Case 2. Suppose κ is supported on the singleton $\{\sigma\}$. Then κ'-a.e. σ' is conjugate to σ, so κ' is supported on a finite set, which we will denote by $\{\sigma_1, \ldots, \sigma_n\}$. For each i, there is an element $\tau_i \in \mathrm{Sym}(m)$ such that $\sigma_i = \tau_i \sigma \tau_i^{-1}$. Let η' be the measure on $\mathrm{Sym}(m)^G \times A^{[m]}$ defined by

$$\eta'(\{(\sigma_i, \psi)\}) := \kappa'(\{\sigma_i\}) \eta(\{(\sigma, \psi \circ \tau_i) : 1 \leq i \leq n\}).$$

Then, η' projects to κ' and

$$\psi(\sigma_i(g)p) = \psi \circ \tau_i(\sigma(g)\tau_i^{-1}p) \quad \text{for all } p \in [m],\ \psi \in A^{[m]},\ 1 \leq i \leq n,\ g \in G$$

implies $d_W(\overline{X}, \overline{X}') = 0$, where $\overline{X}'$ is the approximate process constructed from η'. It also implies (2). If η'_i is the probability measure obtained from η' by restricting to $\{\sigma_i\} \times A^{[m]}$ and normalizing to have total mass 1, and if $\overline{X}'_i$ is the approximate process constructed from η'_i, then $\overline{X}$ is isomorphic to $\overline{X}'_i$ in the obvious sense. Thus, $h(\overline{X}|\beta \circ \overline{X}) = h(\overline{X}'_i|\beta \circ \overline{X}'_i)$. So

$$h(\overline{X}'|\beta \circ \overline{X}') = \sum_{i=1}^{n} h(\overline{X}'_i|\beta \circ \overline{X}'_i) \kappa'(\{\sigma'_i\}) = h(\overline{X}|\beta \circ \overline{X}).$$

This proves Case 2.

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Case 3. Suppose that there is an element $\sigma \in \mathrm{Sym}(m)^G$ such that κ-a.e. σ' is conjugate to σ. Then, κ'-a.e. σ' is also conjugate to σ. By Case 1, there exists a measure η_0 on $\mathrm{Sym}(m)^G \times A^{[m]}$ whose first marginal is the Dirac measure supported on $\{\sigma\}$ such that if $\overline{X}_0$ is the approximate process constructed from η_0 then $d_W(\overline{X}, \overline{X}_0) = 0$, equation (2) holds with η' replaced by η_0 and $h(\overline{X}_0|\beta \circ \overline{X}_0) \geq h(\overline{X}|\beta \circ \overline{X})$. By Case 2, there exists a measure η' on $\mathrm{Sym}(m)^G \times A^{[m]}$ whose first marginal is κ' such that if $\overline{X}'$ is the approximate process constructed from η' then $d_W(\overline{X}', \overline{X}_0) = 0$, equation (2) holds with η replaced by η_0 and $h(\overline{X}'|\beta \circ \overline{X}') \geq h(\overline{X}_0|\beta \circ \overline{X}_0)$. So $d_W(\overline{X}', \overline{X}) = 0$ and $h(\overline{X}'|\beta \circ \overline{X}') \geq h(\overline{X}|\beta \circ \overline{X})$, and equation (2) holds.

Case 4. Now we handle the general case. This case follows from the previous one by disintegrating κ and κ' over the set of conjugacy classes. To be precise, let $[\mathrm{Sym}(m)^G]$ be the space of conjugacy classes of $\mathrm{Sym}(m)^G$. This is the quotient of $\mathrm{Sym}(m)^G$ by the conjugacy action of $\mathrm{Sym}(m)$. Let $\pi : \mathrm{Sym}(m)^G \to [\mathrm{Sym}(m)^G]$ be the quotient map. For each conjugacy class $c \in [\mathrm{Sym}(m)^G]$, let κ_c, κ'_c be the fiber measures of κ, κ' over c, respectively. Likewise, let η_c be the fiber measure of η over c (so κ_c is the projection of η_c to $\mathrm{Sym}(m)^G$).

By Case 3, if $\overline{X}_c$ is the approximate process constructed from η_c then there is a measure η'_c on $\mathrm{Sym}(m)^G \times A^{[m]}$ such that the projection of η'_c to $\mathrm{Sym}(m)^G$ is κ'_c; if $\overline{X}'_c$ is the approximate process constructed from η'_c then $d_W(\overline{X}_c, \overline{X}'_c) = 0$, equation (2) holds for η_c and η'_c in place of η and η', and $h(\overline{X}'_c|\beta \circ \overline{X}'_c) \geq h(\overline{X}_c|\beta \circ \overline{X}_c)$.

The hypothesis that κ and κ' are conjugate implies $\pi_*\kappa = \pi_*\kappa'$. Let $\eta' = \int \eta'_c \, d\pi_*\kappa(c)$ and $\overline{X}'$ be the approximate process constructed from η'. So

$$\begin{aligned} h(\overline{X}'|\beta \circ \overline{X}') &= \int h(\overline{X}'_c|\beta \circ \overline{X}'_c) \, d\pi_*\kappa(c) \\ &\geq \int h(\overline{X}_c|\beta \circ \overline{X}_c) \, d\pi_*\kappa(c) = h(\overline{X}|\beta \circ \overline{X}). \end{aligned}$$

Because $d_W(\overline{X}_c, \overline{X}'_c) = 0$, it follows that $d_W(\overline{X}, \overline{X}') = 0$. Also, equation (2) holds. This finishes the lemma. □

LEMMA 5.7. *Let κ, κ' be probability measures on $\mathrm{Sym}(m)^G, \mathrm{Sym}(m')^G$, respectively. Let η be a probability measure on $\mathrm{Sym}(m)^G \times A^{[m]}$ with projection κ and let $\overline{X}$ be the approximate process constructed from η. Suppose κ and κ' are (W, ϵ)-close, where $W \subset G$ is a finite set containing the identity and $0 \leq \epsilon < 1$. Let $\beta : A \to B$ be a map to a finite set B. Then, there is an approximate process $\overline{X}'$ constructed from a measure η' on $\mathrm{Sym}(m')^G \times A^{[m']}$ such that:*

(1) *the projection of η' to $\mathrm{Sym}(m')^G$ is κ';*
(2) *$d_W(\overline{X}, \overline{X}') \leq 6|W|\epsilon$; and*
(3) *$h(\overline{X}'|\beta \circ \overline{X}') \geq (1-\epsilon)^2 h(\overline{X}|\beta \circ \overline{X}) - 2(1-\epsilon)^2\epsilon \log |A|$.*

Moreover, if $\overline{X} = (T, X, \mu, \phi)$ is a process over G then for any $\epsilon > 0$,

$$\begin{aligned} &\eta'(\{(\sigma, \psi) : d_W((\sigma, \psi), \phi) \leq (4|W|+1)\epsilon\}) \\ &\quad \geq (1-\epsilon)\eta(\{(\sigma, \psi) : d_W((\sigma, \psi), \phi) \leq \epsilon\}) - \epsilon. \end{aligned}$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Proof. Let ϑ be as in Definition 12. So $\vartheta(\mathcal{G}(W,\epsilon)) \geq 1-\epsilon$, where $\mathcal{G}(W,\epsilon)$ is the set of all $(\sigma,\sigma') \in \mathrm{Sym}(m)^G \times \mathrm{Sym}(m')^G$ that are (W,ϵ)-close to each other. Thus, for every $(\sigma,\sigma') \in \mathcal{G}(W,\epsilon)$, there exists a bijection $\beta_{\sigma,\sigma'} : Q_{\sigma,\sigma'} \to Q'_{\sigma,\sigma'}$ between subsets $Q_{\sigma,\sigma'} \subset [m]$, $Q'_{\sigma,\sigma'} \subset [m']$ such that:

(1) $\sigma'(w)\beta_{\sigma,\sigma'}(q) = \beta_{\sigma,\sigma'}(\sigma(w)q)$ for all $w \in W$ and $q \in Q_{\sigma,\sigma'}$ with $\sigma(w)q \in Q_{\sigma,\sigma'}$;
(2) $\sigma(w)\beta^{-1}_{\sigma,\sigma'}(q') = \beta^{-1}_{\sigma,\sigma'}(\sigma'(w)q')$ for all $w \in W$ and $q' \in Q'_{\sigma,\sigma'}$ with $\sigma'(w)q' \in Q'_{\sigma,\sigma'}$; and
(3) $|Q_{\sigma,\sigma'}| \geq (1-\epsilon)m$, $|Q'_{\sigma,\sigma'}| \geq (1-\epsilon)m'$.

After removing elements from each $Q_{\sigma,\sigma'}$ and $Q'_{\sigma,\sigma'}$ if necessary, we may assume, without loss of generality, that there exists a number $m'' \leq \min(m,m')$ so that $m'' \geq (1-\epsilon)\max(m,m')$ and $m'' = |Q_{\sigma,\sigma'}| = |Q'_{\sigma,\sigma'}|$ for every $(\sigma,\sigma') \in \mathcal{G}(W,\epsilon)$. There exist elements $\tau_{\sigma,\sigma'} \in \mathrm{Sym}(m)$ and $\tau'_{\sigma,\sigma'} \in \mathrm{Sym}(m')$ such that $\tau_{\sigma,\sigma'}Q_{\sigma,\sigma'} = [m'']$, $\tau'_{\sigma,\sigma'}Q'_{\sigma,\sigma'} = [m'']$ and $\tau'_{\sigma,\sigma'}\beta_{\sigma,\sigma'}\tau^{-1}_{\sigma,\sigma'}$ is the identity map on $[m'']$. Let ϑ' be the measure obtained from ϑ by pushing forward under the map defined by $(\sigma,\sigma') \mapsto (\tau_{\sigma,\sigma'}\sigma\tau^{-1}_{\sigma,\sigma'}, \tau'_{\sigma,\sigma'}\sigma(\tau'_{\sigma,\sigma'})^{-1})$ for $(\sigma,\sigma') \in \mathcal{G}(W,\epsilon)$, and the map is equal to the identity on the complement of $\mathcal{G}(W,\epsilon)$. The marginals of ϑ' are conjugate to κ and κ', respectively. By the previous lemma, therefore, we may assume that $\vartheta' = \vartheta$.

About notation: if μ is a measure on a space X and $\pi : X \to Y$ is a Borel map, we let μ_y denote the fiber measure of μ over $y \in Y$. It is a measure on X supported on $\pi^{-1}(y)$ and $\mu = \int \mu_y \, d\pi_*\mu(y)$. Note that μ_y depends on π but this dependence is left implicit. If μ is a probability measure then μ_y is a probability measure for $\pi_*\mu$-a.e. $y \in Y$. If $\alpha : X \to Y$ and $\beta : Y \to Z$, and μ is a measure on X, then $\mu_z = \int \mu_{\alpha(x)} \, d\mu_z(x)$ and $\alpha_*(\mu_z) = (\alpha_*\mu)_z$ for $\beta_*\alpha_*\mu$-a.e. $z \in Z$.

Case 1. Suppose $\vartheta(\mathcal{G}(W,\epsilon)) = 1$. We will prove that there exists a probability measure η' on $\mathrm{Sym}(m')^G \times A^{[m']}$ which projects to κ' such that if $\overline{X}'$ is the approximate process constructed from η' then $d_W(\overline{X}', \overline{X}) \leq 4\epsilon|W|$, $h(\overline{X}'|\beta \circ \overline{X}') \geq (1-\epsilon)h(\overline{X}|\beta \circ \overline{X}) - (1-\epsilon)\epsilon \log|A|$ and

$$\eta'(\{(\sigma,\psi) : d_W((\sigma,\psi),\phi) \leq (4|W|+1)\epsilon\}) \geq \eta(\{(\sigma,\psi) : d_W((\sigma,\psi),\phi) \leq \epsilon\}).$$

Let $\pi : \mathrm{Sym}(m)^G \to \mathrm{Sym}(m'')^G$, $\pi' : \mathrm{Sym}(m')^G \to \mathrm{Sym}(m'')^G$ be Borel maps satisfying the following.

(1) If $\sigma \in \mathrm{Sym}(m)^G$, $w \in W$, $p \in [m'']$ and $\sigma(w)p \in [m'']$ then $\pi(\sigma)(w)p = \sigma(w)p$.
(2) If $\sigma' \in \mathrm{Sym}(m')^G$, $w \in W$, $p \in [m'']$ and $\sigma'(w)p \in [m'']$ then $\pi'(\sigma')(w)p = \sigma'(w)p$.
(3) If $\sigma \in \mathrm{Sym}(m)^G$ and $\sigma'(m')^G$ satisfy $\sigma(w)p = \sigma'(w)p$ for every $w \in W$ and $p \in [m'']$ such that $\sigma(w)p \in [m'']$ and $\sigma'(w)p \in [m'']$ then $\pi(\sigma) = \pi'(\sigma')$.

Because we assume $\vartheta(\mathcal{G}(W,\epsilon)) = 1$, it follows from these assumptions that $\pi_*\kappa = \pi'_*\kappa'$. Let κ'' denote $\pi_*\kappa$. Let $R : A^{[m]} \to A^{[m'']}$ be the restriction map. Let $\eta'' = (\pi \times R)_*\eta$ be the pushforward measure and $\overline{X}''$ be the approximate process constructed from η''.

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Note that if ζ is any probability measure on $A^{[m]}$ then $H(R_*\zeta) \geq H(\zeta) - (m - m'')\log|A|$. Thus,

$$\begin{aligned} h(\overline{X}|\beta \circ \overline{X}) &= \frac{1}{m}\int H(\eta_{\sigma,\beta\psi})\, d\eta(\sigma, \psi) \\ &\leq \frac{1}{m}\int H((\mathrm{id} \times R)_*(\eta_{\sigma,\beta\psi}))\, d\eta(\sigma, \psi) + \frac{m - m''}{m}\log|A|, \end{aligned}$$

where id denotes the identity map on $\mathrm{Sym}(m)^G$. By the definition of fiber measure,

$$\int \eta_{\sigma,\beta\psi'}\, d\eta_{\sigma,\beta R\psi}(\sigma, \psi') = \eta_{\sigma,\beta R\psi}.$$

Therefore,

$$\int (\mathrm{id} \times R)_*(\eta_{\sigma,\beta\psi'})\, d\eta_{\sigma,\beta R\psi}(\sigma, \psi') = ((\mathrm{id} \times R)_*\eta)_{\sigma,\beta R\psi}.$$

So, by concavity of entropy,

$$\int H((\mathrm{id} \times R)_*(\eta_{\sigma,\beta\psi'}))\, d\eta_{\sigma,\beta R\psi}(\sigma, \psi') \leq H(((\mathrm{id} \times R)_*\eta)_{\sigma,\beta R\psi}).$$

By integrating over all $(\sigma, \psi) \in \mathrm{Sym}(m)^G \times A^{[m]}$, we obtain

$$\int H((\mathrm{id} \times R)_*(\eta_{\sigma,\beta\psi}))\, d\eta(\sigma, \psi) \leq \int H(((\mathrm{id} \times R)_*\eta)_{\sigma,\beta R\psi})\, d\eta(\sigma, \psi).$$

Note $\pi \times \mathrm{id}$ is injective on the support of $((\mathrm{id} \times R)_*\eta)_{\sigma,\beta R\psi}$. So

$$H(((\mathrm{id} \times R)_*\eta)_{\sigma,\beta R\psi}) = H((\pi \times \mathrm{id})_*(((\mathrm{id} \times R)_*\eta)_{\sigma,\beta R\psi})).$$

Note that $((\mathrm{id} \times R)_*\eta)_{\sigma,\beta R\psi} = (\mathrm{id} \times R)_*(\eta_{\sigma,\beta R\psi})$. Therefore,

$$(\pi \times \mathrm{id})_*(((\mathrm{id} \times R)_*\eta)_{\sigma,\beta R\psi}) = (\pi \times R)_*(\eta_{\sigma,\beta R\psi}).$$

Because $\int \eta_{\sigma',\beta R\psi}\, d\eta_{\pi\sigma,\beta R\psi}(\sigma', \psi') = \eta_{\pi\sigma,\beta R\psi}$, it follows that

$$\int (\pi \times R)_*\eta_{\sigma',\beta R\psi}\, d\eta_{\pi\sigma,\beta R\psi}(\sigma', \psi') = (\pi \times R)_*(\eta_{\pi\sigma,\beta R\psi}) = \eta''_{\pi\sigma,\beta R\psi}.$$

By concavity of entropy,

$$\int H((\pi \times R)_*\eta_{\sigma',\beta R\psi})\, d\eta_{\pi\sigma,\beta R\psi}(\sigma', \psi') \leq H(\eta''_{\pi\sigma,\beta R\psi}).$$

So,

$$\begin{aligned} \int H(((\mathrm{id} \times R)_*\eta)_{\sigma,\beta R\psi})\, d\eta(\sigma, \psi) &= \int H((\pi \times R)_*(\eta_{\sigma,\beta R\psi}))\, d\eta(\sigma, \psi) \\ &\leq \int H(\eta''_{\pi\sigma,\beta R\psi})\, d\eta(\sigma, \psi). \end{aligned}$$

Since $(m - m'')/m \leq \epsilon$,

$$\begin{aligned} h(\overline{X}|\beta \circ \overline{X}) &\leq \epsilon \log|A| + \frac{1}{m}\int H(\eta''_{\pi\sigma,\beta R\psi})\, d\eta(\sigma, \psi) \\ &= \epsilon \log|A| + \frac{1}{m}\int H(\eta''_{\sigma'',\beta\psi''})\, d\eta''(\sigma'', \psi'') \\ &= \epsilon \log|A| + \frac{m''}{m}h(\overline{X}''|\beta \circ \overline{X}''). \end{aligned}$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Next, let $R' : A^{[m']} \to A^{[m'']}$ be the restriction map. We claim that there exists a probability measure η' on $\mathrm{Sym}(m')^G \times A^{[m']}$ such that the first marginal of η' is κ' and $(\pi' \times R')_* \eta' = \eta''$. To see this, let $a \in A$ be an arbitrary element. For $\psi \in A^{[m'']}$, define $\psi_a \in A^{[m']}$ by $\psi_a(p) = \psi(p)$ if $p \in [m'']$ and $\psi_a(p) = a$ if $p \in [m']\backslash[m'']$. For $(\sigma, \psi) \in \mathrm{Sym}(m'')^G \times A^{[m'']}$, let $\eta'_{\sigma,\psi}$ be the probability measure on $\mathrm{Sym}(m')^G \times A^{[m']}$ whose first marginal is κ'_σ and has support contained in $\{(\sigma', \psi_a) : \pi'(\sigma') = \sigma\}$, i.e., $\eta'_{\sigma,\psi}$ is the direct product of κ'_σ and the Dirac measure concentrated on ψ_a. Then, $\eta' := \int \eta'_{\sigma,\psi}\, d\eta''(\sigma, \psi)$ satisfies the claim because $\kappa'' = (\pi')_* \kappa'$ implies the first marginal of η' is κ', as required.

Let $\alpha : A^{[m'']} \to A^{[m']}$ be the map $\alpha(\psi) = \psi_a$. Let $\rho : \mathrm{Sym}(m'')^G \times A^{[m'']} \to A^{[m'']}$ be the projection map. By construction, for κ''-a.e. $\sigma'' \in \mathrm{Sym}(m'')^G$, $\eta'_{\sigma''} = \kappa'_{\sigma''} \times \alpha_* \rho_* (\eta''_{\sigma''})$. Therefore, for κ'-a.e. σ, η'_σ is the direct product of δ_σ, the Dirac measure concentrated on σ, with $\alpha_* \rho_* (\eta''_{\pi'\sigma})$. So, for η'-a.e. (σ, ψ), $\eta'_{\sigma,\beta R'\psi} = \delta_\sigma \times \alpha_* \rho_* (\eta''_{\pi'\sigma,\beta R'\psi})$ and $\alpha(R'\psi) = \psi$, which implies $\eta'_{\sigma,\beta R'\psi} = \eta'_{\sigma,\beta\psi}$. Because $R'\alpha$ is the identity map,

$$\begin{aligned}(\pi' \times R')_*(\eta'_{\sigma,\beta\psi}) &= (\pi' \times R')_*(\delta_\sigma \times \alpha_* \rho_* (\eta''_{\pi'\sigma,\beta R'\psi})) \\ &= \delta_{\pi'\sigma} \times \rho_* \eta''_{\pi'\sigma,\beta R'\psi} = \eta''_{\pi'\sigma,\beta R'\psi}.\end{aligned}$$

Let $\overline{X}'$ be the approximate process constructed from η'. Then,

$$\begin{aligned}h(\overline{X}'|\beta \circ \overline{X}') &= \frac{1}{m'} \int H(\eta'_{\sigma,\beta\psi})\, d\eta'(\sigma, \psi) \geq \frac{1}{m'} \int H(\eta''_{\pi'\sigma,\beta R'\psi})\, d\eta'(\sigma, \psi) \\ &= \frac{1}{m'} \int H(\eta''_{\pi'\sigma,\beta R'\psi})\, d\eta''(\pi'\sigma, R'\psi) = \frac{m''}{m'} h(\overline{X}''|\beta \circ \overline{X}'') \\ &\geq (1-\epsilon) h(\overline{X}|\beta \circ \overline{X}) - (1-\epsilon)\epsilon \log |A|.\end{aligned}$$

Let $\chi : \mathrm{Sym}(m)^G \times A^{[m]} \times [m] \to A$ be the map $\chi(\sigma, \psi, p) = \psi(p)$. Define $\chi'' : \mathrm{Sym}(m'')^G \times A^{[m'']} \times [m''] \to A$ similarly. Then,

$$d_W(\overline{X}, \overline{X}'') = \tfrac{1}{2} \| \chi^W_*(\eta \times u) - (\chi'')^W_*(\eta'' \times u'') \|_1,$$

where u, u'' denote the uniform probability measures on $[m]$, $[m'']$, respectively. So, if Θ is any probability measure on

$$\mathrm{Sym}(m)^G \times A^{[m]} \times [m] \times \mathrm{Sym}(m'')^G \times A^{[m'']} \times [m'']$$

with marginals $\eta \times u$ and $\eta'' \times u''$ then $d_W(\overline{X}, \overline{X}'') \leq \Theta(S)$, where

$$S = \{(\sigma, \psi, p, \sigma'', \psi'', p'') : \chi^W(\sigma, \psi, p) \neq (\chi'')^W(\sigma'', \psi'', p'')\}.$$

Let Θ'' be the pushforward of $\eta \times u''$ under the map $(\sigma, \psi, p) \mapsto (\sigma, \psi, p, \pi\sigma, R\psi, p)$. Let u_0 be the uniform probability measure on $[m]\backslash[m'']$. Then,

$$\eta \times u = \left(\frac{m - m''}{m}\right) \eta \times u_0 + \left(\frac{m''}{m}\right) \eta \times u''.$$

So, if $\Theta_0 = \eta \times u_0 \times \eta'' \times u''$ and

$$\Theta := \left(\frac{m - m''}{m}\right) \Theta_0 + \left(\frac{m''}{m}\right) \Theta''$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

then Θ has marginals $\eta \times u$ and $\eta'' \times u''$. Moreover, $\Theta(S) \leq ((m - m'')/m) + (m''/m)\,\Theta''(S)$ and, we claim, $\Theta''(S) \leq |W|\epsilon$. This is because if U is the set of all $(\sigma, \psi, p, \pi\sigma, R\psi, p)$ such that $\sigma(g)p \in [m'']$ for all $g \in W$ then $U \cap S = \emptyset$ and $\Theta''(U) \geq 1 - |W|(m - m'')/m$. Thus,

$$\begin{aligned} d_W(\overline{X}, \overline{X}'') &\leq \Theta(S) \leq \frac{m - m''}{m} + \frac{m''}{m}\Theta''(S) \\ &\leq \frac{m - m''}{m} + \frac{m''}{m}\left(\frac{m - m''}{m}\right)|W| \leq 2\epsilon|W|. \end{aligned}$$

Similarly, $d_W(\overline{X}', \overline{X}'') \leq 2\epsilon|W|$. So, $d_W(\overline{X}, \overline{X}') \leq 4\epsilon|W|$.

By applying the same argument to the Dirac measure concentrated on an arbitrary $(\sigma, \psi) \in \mathrm{Sym}(m)^G \times A^{[m]}$, we obtain

$$d_W((\sigma, \psi), (\pi\sigma, R\psi)) \leq 2\epsilon|W|.$$

A similar statement holds with $(\sigma', \psi') \in \mathrm{Sym}(m')^G \times A^{[m']}$. So,

$$\eta'(\{(\sigma, \psi) : d_W((\sigma, \psi), \phi) \leq (4|W| + 1)\epsilon\}) \geq \eta(\{(\sigma, \psi) : d_W((\sigma, \psi), \phi) \leq \epsilon\}).$$

This finishes Case 1.

Case 2. This is the general case. Let κ_1, κ_1' be the marginals of ϑ restricted to $\mathcal{G}(W, \epsilon)$ and normalized to have total mass 1. Let $\eta_1 = \int \eta_\sigma \, d\kappa_1(\sigma)$. So, η_1 is absolutely continuous to η. Let $\overline{X}_1$ be the approximate process constructed from η_1.

Observe that

$$\begin{aligned} h(\overline{X}_1|\beta \circ \overline{X}_1) &= \frac{1}{m}\int H(\eta_{\sigma,\beta\psi})\, d\eta_1(\sigma, \psi) \\ &= \frac{1}{m}\frac{1}{\vartheta(\mathcal{G}(W, \epsilon))}\int_{\mathcal{G}(W,\epsilon)}\int H(\eta_{\sigma,\beta\psi})\, d\eta_\sigma(\psi)\, d\vartheta(\sigma, \sigma') \\ &\geq -\epsilon \log|A| + \frac{1}{m}\int\int H(\eta_{\sigma,\beta\psi})\, d\eta_\sigma(\psi)\, d\vartheta(\sigma, \sigma') \\ &= -\epsilon \log|A| + h(\overline{X}|\beta \circ \overline{X}). \end{aligned}$$

It is easy to check that $d_W(\overline{X}_1, \overline{X}) \leq \epsilon$ and

$$\eta_1(\{(\sigma, \psi) : d_W((\sigma, \psi), \phi) \leq \epsilon\}) \geq \eta(\{(\sigma, \psi) : d_W((\sigma, \psi), \phi) \leq \epsilon\}) - \epsilon.$$

By Case 1, there exists a measure η_1' on $\mathrm{Sym}(m')^G \times A^{[m']}$ such that η_1' projects to κ_1', and if $\overline{X}_1'$ is the approximate process constructed from η_1' then $d_W(\overline{X}_1', \overline{X}_1) \leq 4\epsilon|W|$, $h(\overline{X}_1'|\beta \circ \overline{X}_1') \geq (1 - \epsilon)h(\overline{X}_1|\beta \circ \overline{X}_1) - (1 - \epsilon)\epsilon \log|A|$ and

$$\eta_1'(\{(\sigma, \psi) : d_W((\sigma, \psi), \phi) \leq (4|W| + 1)\epsilon\}) \geq \eta_1(\{(\sigma, \psi) : d_W((\sigma, \psi), \phi) \leq \epsilon\}).$$

Of course, κ_1' is absolutely continuous with respect to κ'. Let η' be any probability measure on $\mathrm{Sym}(m')^G \times A^{[m']}$ such that η' projects to κ' and

$$\eta_1' = \int \eta_\sigma' \, d\kappa_1'(\sigma).$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Let $\overline{X}'$ be the approximate process constructed from η'. Because $\kappa' \geq (1-\epsilon)\kappa'_1$, it follows that $\eta' \geq (1-\epsilon)\eta'_1$. So,

$$\begin{aligned} h(\overline{X}'|\beta \circ \overline{X}') &= \frac{1}{m'} \int H(\eta'_{\sigma,\beta\psi})\, d\eta'(\sigma,\psi) \geq (1-\epsilon)\frac{1}{m'} \int H(\eta'_{\sigma,\beta\psi})\, d\eta'_1(\sigma,\psi) \\ &= (1-\epsilon)h(\overline{X}'_1|\beta \circ \overline{X}'_1) \geq (1-\epsilon)^2 h(\overline{X}_1|\beta \circ \overline{X}_1) - (1-\epsilon)^2\epsilon \log|A| \\ &\geq (1-\epsilon)^2 h(\overline{X}|\beta \circ \overline{X}) - 2(1-\epsilon)^2\epsilon \log|A|. \end{aligned}$$

Of course, we also have $d_W(\overline{X}'_1, \overline{X}') \leq \epsilon$ and because $\eta' \geq (1-\epsilon)\eta'_1$,

$$\begin{aligned} &\eta'(\{(\sigma,\psi) : d_W((\sigma,\psi),\phi) \leq (4|W|+1)\epsilon\}) \\ &\quad \geq (1-\epsilon)\eta'_1(\{(\sigma,\psi) : d_W((\sigma,\psi),\phi) \leq (4|W|+1)\epsilon\}). \end{aligned}$$

Thus,

$$d_W(\overline{X}, \overline{X}') \leq d_W(\overline{X}, \overline{X}_1) + d_W(\overline{X}_1, \overline{X}'_1) + d_W(\overline{X}'_1, \overline{X}') \leq 6\epsilon|W|$$

and

$$\begin{aligned} &\eta'(\{(\sigma,\psi) : d_W((\sigma,\psi),\phi) \leq (4|W|+1)\epsilon\}) \\ &\quad \geq (1-\epsilon)\eta'_1(\{(\sigma,\psi) : d_W((\sigma,\psi),\phi) \leq (4|W|+1)\epsilon\}) \\ &\quad \geq (1-\epsilon)\eta_1(\{(\sigma,\psi) : d_W((\sigma,\psi),\phi) \leq \epsilon\}) \\ &\quad \geq (1-\epsilon)\eta(\{(\sigma,\psi) : d_W((\sigma,\psi),\phi) \leq \epsilon\}) - \epsilon. \end{aligned}$$

□

Proof of Theorem 5.5. Let $\{\overline{X}_j\}_{j=1}^{\infty}$ be a sequence of approximate processes adapted to $\mathbb{K}'$ (where $\mathbb{K}'$ is a subsequence of $\mathbb{K}$) such that $\lim_{j\to\infty} \overline{X}_j = \overline{X}$. Let η_j be the probability measure on $\mathrm{Sym}(m_j)^G \times A^{[m_j]}$ from which $\overline{X}_j$ is constructed.

Let $\epsilon > 0$ and $W \subset G$ be finite. Let $\mathbb{L} = \{\lambda_j\}_{j=1}^{\infty}$. By the previous lemma, for all sufficiently large j, there exists an approximate process $\overline{Y}_j$ adapted to λ_j such that $h(\overline{Y}_j|\beta \circ \overline{Y}_j) \geq (1-\epsilon)^2 h(\overline{X}_j|\beta \circ \overline{X}_j) - 2\epsilon \log|A|$, $d_W(\overline{Y}_j, \overline{X}_j) \leq 6|W|\epsilon$, and if η'_j is the measure on $\mathrm{Sym}(m'_j)^G \times A^{[m'_j]}$ from which $\overline{Y}_j$ is constructed then

$$\begin{aligned} &\eta'_j(\{(\sigma,\psi) : d_W((\sigma,\psi),\phi) \leq (4|W|+1)\epsilon\}) \\ &\quad \geq (1-\epsilon)\eta_j(\{(\sigma,\psi) : d_W((\sigma,\psi),\phi) \leq \epsilon\}) - \epsilon. \end{aligned}$$

So, if $\{W_n\}_{n=1}^{\infty}$ is an increasing sequence of finite subsets of G then for every n there is a $J(n) > 0$, so that for every $j \geq J(n)$ there is an approximate process $\overline{Y}_j$ adapted to λ_j such that

$$h(\overline{Y}_j|\beta \circ \overline{Y}_j) \geq \left(1 - \frac{1}{n}\right) h(\overline{X}_j|\beta \circ \overline{X}_j) - \frac{1}{n}, \quad d_{W_n}(\overline{Y}_j, \overline{X}_j) \leq \frac{1}{n}$$

and, if $\lim_{j\to\infty} \overline{X}_j = \overline{X}$ strongly then

$$\eta'_j\left(\left\{(\sigma,\psi) : d_{W_n}((\sigma,\psi),\phi) \leq \frac{1}{n}\right\}\right) \geq 1 - \frac{1}{n}.$$

Without loss of generality, we may assume $\{J(n)\}_{n=1}^{\infty}$ is an increasing sequence. So,

$$\lim_{n\to\infty} \overline{Y}_{J(n)} = \overline{X}, \{\overline{Y}_{J(n)}\}_{n=1}^{\infty}$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

is adapted to a subsequence of $\mathbb{L}$ and $\limsup_{n\to\infty} h(\overline{Y}_n|\beta\circ\overline{Y}_n) \geq \bar{h}(\overline{X}|\beta\circ\overline{X})$. So, $\bar{h}(\mathbb{L}, \overline{X}|\beta\circ\overline{X}) \geq \bar{h}(\mathbb{K}, \overline{X}|\beta\circ\overline{X})$. By symmetry, we must in fact have $\bar{h}(\mathbb{L}, \overline{X}|\beta\circ\overline{X}) = \bar{h}(\mathbb{K}, \overline{X}|\beta\circ\overline{X})$ as claimed. Moreover, if $\lim_{j\to\infty} \overline{X}_j = \overline{X}$ strongly then $\lim_{n\to\infty} \overline{Y}_{J(n)} = \overline{X}$ strongly. So, $h(\mathbb{L}, \overline{X}|\beta\circ\overline{X}) = h(\mathbb{K}, \overline{X}|\beta\circ\overline{X})$. □

Next, we extend Proposition 4.1 to the relative case.

PROPOSITION 5.8. *Let $\mathbb{K} = \{\kappa_i\}_{i=1}^{\infty}$ be a random sofic approximation of $\mathbb{Z}$. Let $\overline{X}$ be a process over $\mathbb{Z}$ with range A and let $\beta : A \to B$ be a map. Then, $\bar{h}(\mathbb{K}, \overline{X}|\beta\circ\overline{X}) = h(\mathbb{K}, \overline{X}|\beta\circ\overline{X}) = h(\overline{X}|\beta\circ\overline{X})$.*

Proof. The inequalities

$$\bar{h}(\mathbb{K}, \overline{X}|\beta\circ\overline{X}) \geq h(\mathbb{K}, \overline{X}|\beta\circ\overline{X}) \geq h(\overline{X}|\beta\circ\overline{X})$$

follow immediately from the definitions and Proposition 4.1. So, it suffices to show $h(\overline{X}|\beta\circ\overline{X}) \geq \bar{h}(\mathbb{K}, \overline{X}|\beta\circ\overline{X})$. By Theorems 5.5 and 4.3, we may assume $\mathbb{K}$ is the non-random sofic approximation $\Sigma = \{\sigma_i\}_{i=1}^{\infty}$, where $\sigma_i : \mathbb{Z} \to \mathrm{Sym}(m_i)$ is the homomorphism with $\sigma_i(1) = (1, 2, \ldots, m_i)$.

Let N be a large positive integer. If m_i' is the integer nearest to m_i that is divisible by N, and $\sigma_i' : \mathbb{Z} \to \mathrm{Sym}(m_i')$ is the homomorphism with $\sigma_i'(1) = (1, 2, \ldots, m_i')$, then $\{\sigma_i'\}_{i=1}^{\infty}$ is asymptotic to $\{\sigma_i\}_{i=1}^{\infty}$. So, without loss of generality, we may assume N divides m_i for each i.

Let $\{\overline{X}_i\}_{i=1}^{\infty}$ be a sequence of approximate processes constructed from measures η_i on $\mathrm{Sym}(m_i)^{\mathbb{Z}} \times A^{[m_i]}$ (as in Definition 6) adapted to $\mathbb{K}'$, a subsequence of $\mathbb{K}$, such that $\lim_{i\to\infty} \overline{X}_i = \overline{X}$ and $\lim_{i\to\infty} h(\overline{X}_i) - h(\beta\circ\overline{X}_i) = \bar{h}(\mathbb{K}, \overline{X}|\beta\circ\overline{X})$. Without loss of generality, we may assume $\mathbb{K}' = \mathbb{K}$. So, $\eta_i = \delta_i \times \lambda_i$ for some measure λ_i on $A^{[m_i]}$, where δ_i is the probability measure concentrated on $\{\sigma_i\}$.

Fix i for now. For $a, b \in [m_i]$, let $a + b \in [m_i]$ denote their sum modulo m_i. Also, let $[a, b]$ be the interval from a to b: $[a, b] = \{a, a + 1, a + 2, \ldots, b\}$, for example, $[m_i - 1, 1] = \{m_i - 1, m_i, 1\}$. Let $\pi^{[a,b]} : A^{[m_i]} \to A^{[a,b]}$ be the projection map and $\lambda_i^{[a,b]} = \pi_*^{[a,b]}\lambda_i$ the pushforward measure. For $N > 0$ a positive integer, let $\lambda_{i,N}'$ be the product measure

$$\lambda_{i,N}' := \lambda_i^{[1,N]} \times \lambda_i^{[N+1,2N]} \times \cdots \times \lambda_i^{[m_i-N+1,m_i]}.$$

Let $\beta^{[a,b]} : A^{[a,b]} \to B^{[a,b]}$ denote the map $(\beta^{[a,b]}\xi)(p) = \beta(\xi(p))$. In order to simplify notation, we write $\beta = \beta^{[a,b]}$ when $[a, b]$ is clear from the context. We claim that

$$\frac{H(\lambda_{i,N}') - H(\beta_*\lambda_{i,N}')}{m_i} \geq \frac{H(\lambda_i) - H(\beta_*\lambda_i)}{m_i} = h(\overline{X}_i) - h(\beta\circ\overline{X}_i). \tag{3}$$

For $\xi \in B^{[m_i]}$ and $\zeta \in B^{[a,b]}$, we let $\lambda_i(\cdot|\xi)$ and $\lambda_i(\cdot|\zeta)$ be the measure defined for $E \subset A^{[m_i]}$ by

$$\lambda_i(E|\xi) := \frac{\lambda_i(E \cap (\beta^{[1,m_i]}\pi^{[1,m_i]})^{-1}(\xi))}{\lambda_i((\beta^{[1,m_i]}\pi^{[1,m_i]})^{-1}(\xi))}, \quad \lambda_i(E|\zeta) := \frac{\lambda_i(E \cap (\beta^{[a,b]}\pi^{[a,b]})^{-1}(\zeta))}{\lambda_i((\beta^{[a,b]}\pi^{[a,b]})^{-1}(\zeta))}.$$

Other conditional measures such as $\beta_*\lambda_i(\cdot|\zeta)$ are defined similarly.

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Let $\lambda''_{i,N}$ be the measure on $A^{[m_i]}$ satisfying $\beta_*\lambda''_{i,N} = \beta_*\lambda_i$ whose fiber over $\xi \in B^{[m_i]}$ is the measure

$$\lambda''_{i,N}(\cdot|\xi) = \pi_*^{[1,N]}\lambda_i(\cdot|\xi) \times \pi_*^{[N+1,2N]}\lambda_i(\cdot|\xi) \times \cdots \times \pi_*^{[m_i-N+1,m_i]}\lambda_i(\cdot|\xi).$$

Since $\beta_*\lambda''_{i,N} = \beta_*\lambda_i$ and $H(\lambda''_{i,N}(\cdot|\xi)) \geq H(\lambda_i(\cdot|\xi))$ for any ξ, we have

$$H(\lambda''_{i,N}) - H(\beta_*\lambda''_{i,N}) \geq H(\lambda_i) - H(\beta_*\lambda_i).$$

Note

$$\begin{aligned}
&H(\lambda''_{i,N}) - H(\beta_*\lambda''_{i,N}) \\
&= \int \sum_{j=0}^{m_i/N-1} H(\pi_*^{[jN+1,jN+N]}\lambda_i(\cdot|\xi))\, d\beta_*\lambda_i(\xi) \\
&= \sum_{j=0}^{m_i/N-1} \iint H(\pi_*^{[jN+1,jN+N]}\lambda_i(\cdot|\xi))\, d\beta_*\lambda_i(\xi|\zeta)\, d\beta_*\lambda_i^{[jN+1,jN+N]}(\zeta) \\
&\leq \sum_{j=0}^{m_i/N-1} \int H(\pi_*^{[jN+1,jN+N]}\lambda_i(\cdot|\zeta))\, d\beta_*\lambda_i^{[jN+1,jN+N]}(\zeta) \\
&= H(\lambda'_{i,N}) - H(\beta_*\lambda'_{i,N}).
\end{aligned}$$

The inequality above holds by concavity of entropy. This proves the inequality in (3). The equality in (3) holds by definition.

For each $p \in \{0, \ldots, m_i - 1\}$, let $\lambda'_{i,p,N}$ be the product measure

$$\lambda'_{i,p,N} := \lambda_i^{[p+1,p+N]} \times \lambda_i^{[p+N+1,p+2N]} \times \cdots \times \lambda_i^{[p+m_i-N+1,p+m_i]}.$$

An argument similar to the one proving (3) shows

$$H(\lambda'_{i,p,N}) - H(\beta_*\lambda'_{i,p,N}) \geq H(\lambda_i) - H(\beta_*\lambda_i). \tag{4}$$

For $a, b \in [m_i]$, define $\tau^{[a,b]} : A^{[1,b-a+1]} \to A^{[a,b]}$ by $\tau^{[a,b]}(\xi)(p) = \xi(p + a - 1)$. Let $\tilde{\lambda}_i^{[a,b]}$ be the measure defined for sets $E \subset A^{[1,b-a+1]}$ by

$$\tilde{\lambda}_i^{[a,b]}(E) := \lambda_i^{[a,b]}(\tau^{[a,b]}E).$$

For $\zeta \in B^{[1,b-a+1]}$, let $\tilde{\lambda}_i^{[a,b]}(\cdot|\zeta)$ be the measure defined for sets $E \subset A^{[1,b-a+1]}$ by

$$\tilde{\lambda}_i^{[a,b]}(E|\zeta) := \lambda_i^{[a,b]}(\tau^{[a,b]}E|\tau^{[a,b]}\zeta),$$

where we have abused notation by letting $\tau^{[a,b]}$ denote the analogous map from $B^{[1,b-a+1]} \to B^{[a,b]}$.

Because $H(\tilde{\lambda}_i^{[p+1,p+N+1]}(\cdot|\zeta)) = H(\lambda_i^{[p+1,p+N+1]}(\cdot|\tau^{[p+1,p+N+1]}\zeta)$,

$$\begin{aligned}
&H(\lambda'_{i,p,N}) - H(\beta_*\lambda'_{i,p,N}) \\
&= \sum_{j=0}^{m_i/N-1} \int H(\tilde{\lambda}_i^{[jN+p+1,(j+1)N+p]}(\cdot|\zeta))\, d\beta_*\tilde{\lambda}_i^{[jN+p+1,(j+1)N+p]}(\zeta).
\end{aligned}$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

By (4),

$$\frac{H(\lambda_i) - H(\beta_*\lambda_i)}{m_i} \le \frac{1}{m_i^2}\sum_{p=0}^{m_i-1} H(\lambda'_{i,p,N}) - H(\beta_*\lambda'_{i,p,N})$$
$$= \frac{1}{Nm_i}\sum_{p=0}^{m_i-1}\int H(\tilde{\lambda}_i^{[p+1,p+N]}(\cdot|\zeta))\, d\beta_*\tilde{\lambda}_i^{[p+1,p+N]}(\zeta).$$

By concavity of entropy, if, for $\zeta \in B^{[1,N]}$, $\omega_{i,\zeta}$ is the measure on $A^{[1,N]}$ defined by

$$\omega_{i,\zeta} = \frac{\sum_{p=0}^{m_i-1} \beta_*\tilde{\lambda}_i^{[p+1,p+N]}(\{\zeta\}) \cdot \tilde{\lambda}_i^{[p+1,p+N]}(\cdot|\zeta)}{\sum_{p=0}^{m_i-1} \beta_*\tilde{\lambda}_i^{[p+1,p+N]}(\{\zeta\})}$$

then

$$\sum_{p=0}^{m_i-1}\int H(\tilde{\lambda}_i^{[p+1,p+N]}(\cdot|\zeta))\, d\beta_*\tilde{\lambda}_i^{[p+1,p+N]}(\zeta) \le \sum_{p=0}^{m_i-1}\int H(\omega_{i,\zeta})\, d\beta_*\tilde{\lambda}_i^{[p+1,p+N]}(\zeta).$$

Therefore,

$$\frac{H(\lambda_i) - H(\beta_*\lambda_i)}{m_i} \le \frac{1}{Nm_i}\sum_{p=0}^{m_i-1}\int H(\omega_{i,\zeta})\, d\beta_*\tilde{\lambda}_i^{[p+1,p+N]}(\zeta).$$

So, if $\omega_{\beta,i}$ is the measure on $B^{[1,N]}$ defined by

$$\omega_{\beta,i} := \frac{1}{m_i}\sum_{p=0}^{m_i-1} \beta_*\tilde{\lambda}_i^{[p+1,p+N]}$$

then

$$\frac{H(\lambda_i) - H(\beta_*\lambda_i)}{m_i} \le \frac{1}{N}\int H(\omega_{i,\zeta})\, d\omega_{\beta,i}(\zeta).$$

Let $\overline{X} = (T, X, \mu, \phi)$. By construction, $\omega_{i,\zeta}$ converges (as $i \to \infty$) to the measure μ_ζ defined for $E \subset A^{[1,N]}$ by

$$\mu_\zeta(E) := \frac{\mu(\{x \in X : \phi^N(x) \in E, (\beta\circ\phi)^N(x) = \zeta\})}{\mu(\{x \in X : (\beta\circ\phi)^N(x) = \zeta\})}.$$

Also, $\omega_{\beta,i}$ converges to $(\beta\circ\phi)^N_*\mu$. Therefore,

$$\bar{h}(\mathbb{K}, \overline{X}|\beta\circ\overline{X}) = \limsup_{i\to\infty} \frac{H(\lambda_i) - H(\beta_*\lambda_i)}{m_i} \le \int \frac{H(\mu_\zeta)}{N}\, d(\beta\circ\phi)^N_*\mu(\zeta)$$
$$= \frac{H(\phi^N_*\mu) - H((\beta\circ\phi)^N_*\mu)}{N}.$$

The right-hand side converges to $h(\overline{X}|\beta\circ\overline{X})$ as N tends to infinity. This proves the proposition. □

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

6. *Orbit-equivalence and entropy*

In this section, we prove Theorem 1.2 by generalizing a theorem of Rudolph and Weiss, which is explained next.

Definition 19. Let G, Γ be countable discrete groups and let $(X, \mathcal{B}, \mu)$ be a standard probability space. Let $G \curvearrowright^T (X, \mathcal{B}, \mu)$ and $\Gamma \curvearrowright^S (X, \mathcal{B}, \mu)$ be two probability measure-preserving actions with the same orbits. We assume that both actions are essentially free. Let $\rho : \Gamma \times X \to G$ be the cocycle

$$\rho(\gamma, x) := g \Leftrightarrow T_g(x) = S_\gamma(x).$$

If $\mathcal{A} \subset \mathcal{B}$ is a sub-σ-algebra such that $\rho(\gamma, \cdot)$ is $\mathcal{A}$-measurable for all $\gamma \in \Gamma$ then the *orbit change from T to S* is said to be $\mathcal{A}$-measurable. The smallest such σ-algebra is called the *orbit change σ-algebra*.

The next theorem is proven in [**RW00**].

THEOREM 6.1. *Suppose T is an essentially free ergodic action of a countable discrete amenable group G and $\mathcal{A}$ is a T-invariant sub-σ-algebra. Suppose also that S is an essentially free action of Γ with the same orbits as T (this implies Γ is amenable and S is ergodic). Suppose the orbit change from T to S is $\mathcal{A}$-measurable. Then, for any finite observable $\phi : X \to A$, we conclude*

$$h(T, \phi|\mathcal{A}) = h(S, \phi|\mathcal{A}).$$

The rest of the paper is devoted to proving a related result.

PROPOSITION 6.2. *Let G be an amenable group, $\overline{X} = (T, X, \mu, \phi)$ an essentially free G-process with finite range A and $S : (X, \mu) \to (X, \mu)$ be an essentially free measure-preserving Borel automorphism with the same orbits as T (i.e., for μ-a.e. $x \in X$, $\{T_g x : g \in G\} = \{S^n x : n \in \mathbb{Z}\}$).*

Let $\beta : A \to B$ be a map and suppose the orbit change from T to S is measurable with respect to both the T-invariant sub-σ-algebra generated by $\psi := \beta \circ \phi$ and the S-invariant sub-σ-algebra generated by ψ. Then, for any random sofic approximation $\mathbb{K}$ to G,

$$\bar{h}(\mathbb{K}, \overline{X}|\beta \circ \overline{X}) = h(\mathbb{K}, \overline{X}|\beta \circ \overline{X}) = h(S, \phi|\beta \circ \phi).$$

Given the proposition above, we prove the following theorem.

THEOREM 6.3. *Let G be a countably infinite amenable group with random sofic approximation $\mathbb{K}$. Let $\overline{X} = (T, X, \mu, \phi)$ be a G-process. Then,*

$$h(\mathbb{K}, \overline{X}) = \bar{h}(\mathbb{K}, \overline{X}) = h(\overline{X}).$$

Of course, this implies Theorem 1.2.

Proof. We will prove the statement for lower-sofic entropy only; the upper-sofic entropy case is similar. Let $\overline{Y} = (S, Y, \nu, \psi')$ be a Bernoulli process over G with base (B', ω), where B' is a finite set and ω is not supported on a singleton. This process is weakly mixing and $\psi' : Y \to B'$ is generating. By [**CFW81, Dy59, Dy63**], any ergodic essentially free probability measure-preserving action of $\mathbb{Z}$ is orbit-equivalent to (S, Y, ν). So, there exists

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

a weakly mixing automorphism $U : (Y, \nu) \to (Y, \nu)$ with the same orbits as S and a Borel map $\psi'' : Y \to B''$ to a finite set B'', which generates in the sense that the smallest U-invariant σ-algebra on which ψ'' is measurable is the Borel σ-algebra of (Y, ν) (up to measure-zero sets).

Let $\psi : Y \to B = B' \times B''$ be the map $\psi(y) = (\psi'(y), \psi''(y))$. Note that the Borel σ-algebra of (Y, ν) is the smallest σ-algebra generated by ψ and the G-action (modulo measure 0 sets) which is the smallest σ-algebra generated by ψ and U (modulo measure 0 sets).

Define $V : X \times Y \to X \times Y$ by

$$V(x, y) = (T_g x, S_g y) \Leftrightarrow Uy = S_g y.$$

Note that the orbit change from $T \times S$ to V is measurable with respect to both the G-invariant sub-σ-algebra generated by ψ and the V-invariant sub-σ-algebra generated by ψ. Proposition 6.2 and Lemma 5.4 imply

$$h_\mu(\mathbb{K}, \phi) = h_{\mu\times\nu}(\mathbb{K}, \phi \times \psi|\psi) = h_{\mu\times\nu}(V, \phi \times \psi|\psi). \tag{5}$$

Let $\mu = \int \lambda \, d\zeta(\lambda)$ be the ergodic decomposition of μ. Because $\overline{Y}$ is weakly mixing,

$$\mu \times \nu = \int \lambda \times \nu \, d\zeta(\lambda)$$

is the ergodic decomposition of $\mu \times \nu$. It is well known that the classical entropy of a process equals the integral of the entropies of its ergodic components. So, Theorem 6.1 implies

$$\begin{aligned} h_{\mu\times\nu}(V, \phi \times \psi|\psi) &= \int h_{\lambda\times\nu}(V, \phi \times \psi|\psi) \, d\zeta(\lambda) \\ &= \int h_{\lambda\times\nu}(T \times S, \phi \times \psi|\psi) \, d\zeta(\lambda) \\ &= h_{\mu\times\nu}(T \times S, \phi \times \psi|\psi) = h_\mu(T, \phi). \end{aligned}$$

So, (5) implies $h_\mu(\mathbb{K}, \phi) = h_\mu(T, \phi)$, as required. □

6.1. *Lifting factors.* Proposition 6.2 is proven by 'lifting' the orbit-equivalence to sofic approximations. But first, we 'lift' factors that do not necessarily come from composing with a map $\beta : A \to B$.

Definition 20. Let $\{\overline{X}_i\}_{i=1}^\infty$ be a sequence of approximate processes over G constructed from a sequence $\{\eta_i\}_{i=1}^\infty$ of probability measures on $\mathrm{Sym}(m_i)^G \times A^{[m_i]}$. If $W \subset G$ is finite, $\sigma \in \mathrm{Sym}(m_i)^G$ and $\xi \in A^{[m_i]}$ then let

$$\xi[\sigma, W] \in (A^W)^{[m_i]}, \quad \xi[\sigma, W](p)(w) := \xi(\sigma(w)p).$$

Let η_i^W be the measure on $\mathrm{Sym}(m_i)^G \times (A^W)^{[m_i]}$ obtained by pushing forward η_i under the map

$$(\sigma, \xi) \in \mathrm{Sym}(m_i)^G \times A^{[m_i]} \mapsto (\sigma, \xi[\sigma, W]).$$

Let $\{\overline{X}_i^W\}_{i=1}^\infty$ denote the sequence of approximate processes constructed from η_i^W.

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

The next lemma is immediate.

LEMMA 6.4. *If, in the definition above,* $\lim_{i\to\infty} \overline{X}_i = \overline{X} = (T, X, \mu, \phi)$ *then* $\lim_{i\to\infty} \overline{X}_i^W = (T, X, \mu, \phi^W)$.

Assume that $\overline{X} = (T, X, \mu, \phi)$ is a process. Let $\psi : X \to B$ be a measurable map into a finite or countable set B. For each finite $W \subset G$, let $\psi_W : A^W \to B$ be a measurable function satisfying

$$\psi_W(\xi) = b \Rightarrow \mu(\{x \in X : \psi(x) = b \text{ and } \phi^W(x) = \xi\}) \geq \mu(\{x \in X : \psi(x) = c \text{ and } \phi^W(x) = \xi\}) \quad \text{for all } c \in B.$$

Warning: do not confuse $\psi_W : A^W \to B$ with $\psi^W : X \to B^W$.

Definition 21. Let Λ be a function on the set of finite subsets of G. We write $\lim_{W\to G} \Lambda(W) = L$ if for every increasing sequence $\{W_j\}_{j=1}^\infty \subset G$ with $\bigcup_j W_j = G$, $\lim_{j\to\infty} \Lambda(W_j) = L$.

LEMMA 6.5. *Let* $\overline{X} = (T, X, \mu, \phi)$ *be a* G*-process such that* $\phi : X \to A$ *is generating. Let* $\{\overline{X}_i\}_{i=1}^\infty$ *be a sequence of approximate processes constructed from a sequence* $\{\eta_i\}_{i=1}^\infty$ *of probability measures on* $\mathrm{Sym}(m_i)^G \times A^{[m_i]}$. *Suppose that* $\lim_{i\to\infty} \overline{X}_i = \overline{X}$. *Then*

$$\lim_{i\to\infty} \psi_W \circ \overline{X}_i^W = (T, X, \mu, \psi_W \circ \phi^W),$$

$$\lim_{W\to G} (T, X, \mu, \psi_W \circ \phi^W) = (T, X, \mu, \psi).$$

Proof. The first limit follows from the previous lemma and Lemma 5.1. The second limit is a consequence of the fact that ϕ is generating. □

6.2. *Lifting orbit-equivalences.* The concepts of the previous subsection are used to 'lift' orbit-equivalences as follows. Let $\overline{X} = (T, X, \mu, \phi)$ be a process over G with range A. Suppose Γ is a (possibly different) group and $\overline{Y} = (S, X, \mu, \phi)$ is a process over Γ with the same orbits as G (up to μ-measure zero). Suppose there is a map $\beta : A \to B$ such that the orbit change from T to S is measurable with respect to the G-invariant sub-σ-algebra generated by $\psi := \beta \circ \phi$, and that both $\overline{X}$ and $\overline{Y}$ are essentially free (for example, for a.e. $x \in X$ $T_g x \neq T_h x$ if $g \neq h$). Define the cocycle $\rho : \Gamma \times X \to G$ by

$$\rho(\gamma, x) = g \Leftrightarrow S_\gamma x = T_g x.$$

LEMMA 6.6. *For each finite set* $W \subset G$, *there exists a map* $\rho_W : \Gamma \times B^W \to G$ *such that for every* $\gamma \in \Gamma$ *and* $g \in G$,

$$\lim_{W\to G} \mu(\{x \in X : \rho(\gamma, x) = g\} \,\Delta\, \{x \in X : \rho_W(\gamma, \psi^W(x)) = g\}) = 0.$$

Proof. For each $\gamma \in \Gamma$ and $g \in G$, let

$$X(\gamma, g) = \{x \in X : \rho(\gamma, x) = g\}.$$

Choose orderings $G = \{g_1, g_2, \ldots\}$, $\Gamma = \{\gamma_1, \gamma_2, \ldots\}$ of G and Γ. For each n, let P_n be the smallest partition of X containing $X(\gamma_i, g_j)$ for all $1 \leq i, j \leq n$. Because the orbit change is measurable with respect to the smallest G-invariant σ-algebra Σ on which ψ is measurable, the partitions P_n are contained in Σ (up to measure-zero sets). Therefore, for each n, there are a finite set $W_n \subset G$ and sets $X_n(\gamma_i, g_j)$ for $1 \leq i, j \leq n$ such that:

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

(1) $X_n(\gamma_i, g_j)$ is contained in the smallest σ-algebra on which ψ^{W_n} is measurable;
(2) $X_n(\gamma_i, g_j) \cap X_n(\gamma_i, g_k) = \emptyset$ if $j \neq k$ (and $1 \leq i, j, k \leq n$); and
(3) $\mu(X(\gamma_i, g_j) \,\Delta\, X_n(\gamma_i, g_j)) < 2^{-n}$ for all $1 \leq i, j \leq n$.

Let $\rho_{W_n} : \Gamma \times B^{W_n} \to G$ be a map such that $\rho_{W_n}(\gamma_i, \psi^{W_n}(x)) = g_j$ if $1 \leq i, j \leq n$ and $x \in X_n(\gamma_i, g_j)$.

For any finite $W \subset G$, if there is an $i \geq 1$ such that $W_i \subset W$ then choose a maximal such i and let $\rho_W := \rho_{W_i}$. Otherwise, define ρ_W arbitrarily. The lemma follows immediately. □

LEMMA 6.7. (Asymptotic cocycle identity) *For any* $\gamma_1, \gamma_2 \in \Gamma$,

$$\lim_{W \to G} \psi_*^W \mu(\{\xi \in B^W : \rho_W(\gamma_1\gamma_2, \xi) = \rho_V(\gamma_1, \xi')\rho_W(\gamma_2, \xi)\}) = 1,$$

where $g_2 = \rho_W(\gamma_2, \xi)$, $V = W \cap Wg_2^{-1}$, $\xi' = \psi^V(T_{g_2}x)$ *and* $x \in X$ *is any element with* $\psi^W(x) = \xi$.

Proof. The claim of the lemma is equivalent to

$$\lim_{W \to G} \mu(\{x \in X : \rho_W(\gamma_1\gamma_2, \psi^W(x)) = \rho_V(\gamma_1, \psi^V(T_{g_2}x))\rho_W(\gamma_2, \psi^W(x))\}) = 1,$$

where V, g_2 are as defined above. For $\gamma \in \Gamma$, $g \in G$, define $X(\gamma, g) := \{x \in X : \rho(\gamma, x) = g\}$ as in the previous lemma. Also, for $W \subset G$ finite, define

$$X_W(\gamma, g) := \{x \in X : \rho_W(\gamma, \psi^W(x)) = g\}.$$

The cocycle identity $\rho(\gamma_1\gamma_2, x) = \rho(\gamma_1, S_{\gamma_2}x)\rho(\gamma_2, x)$ is equivalent to the statement

$$X(\gamma_2, g_2) \cap T_{g_2}^{-1}X(\gamma_1, g_1) \subset X(\gamma_1\gamma_2, g_1g_2)$$

(mod 0) for every $g_1, g_2 \in G$. Let

$$Y(\gamma_1, \gamma_2, g_1, g_2) := X(\gamma_1\gamma_2, g_1g_2) \cap X(\gamma_2, g_2) \cap T_{g_2}^{-1}X(\gamma_1, g_1),$$
$$Y_W(\gamma_1, \gamma_2, g_1, g_2) := X_W(\gamma_1\gamma_2, g_1g_2) \cap X_W(\gamma_2, g_2) \cap T_{g_2}^{-1}X_V(\gamma_1, g_1).$$

Then, $\{Y(\gamma_1, \gamma_2, g_1, g_2) : g_1, g_2 \in G\}$ partitions X by the cocycle equation. Observe that

$$Y_W(\gamma_1, \gamma_2, g_1, g_2) \subset \{x \in X : \rho_W(\gamma_1\gamma_2, \psi^W(x)) = \rho_V(\gamma_1, \psi^V(T_{g_2}x))\rho_W(\gamma_2, \psi^W(x))\}.$$

So, it suffices to prove that

$$\lim_{W \to G} \mu\left(\bigcup_{g_1, g_2 \in G} Y_W(\gamma_1, \gamma_2, g_1, g_2)\right) = 1.$$

If $V = W \cap Wg_2^{-1}$ then as $W \to G$, $V \to G$. So, the previous lemma implies

$$0 = \lim_{W \to G} \mu(X_W(\gamma_1\gamma_2, g_1g_2) \,\Delta\, X(\gamma_1\gamma_2, g_1g_2)),$$
$$0 = \lim_{W \to G} \mu(X_W(\gamma_2, g_2) \,\Delta\, X(\gamma_2, g_2)),$$
$$0 = \lim_{W \to G} \mu(T_{g_2}^{-1}X_V(\gamma_1, g_1) \,\Delta\, T_{g_2}^{-1}X(\gamma_1, g_1)),$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

which implies

$$0 = \lim_{W \to G} \mu(Y(\gamma_1, \gamma_2, g_1, g_2) \,\Delta\, Y_W(\gamma_1, \gamma_2, g_1, g_2))$$

for any fixed $g_1, g_2 \in G$. Since $\{Y(\gamma_1, \gamma_2, g_1, g_2) : g_1, g_2 \in G\}$ partitions X, for every $\epsilon > 0$, there is a finite set $S \subset G$ such that

$$\mu\Bigg(\bigcup_{g_1, g_2 \in S} Y(\gamma_1, \gamma_2, g_1, g_2) \Bigg) \geq 1 - \epsilon.$$

Because

$$0 = \lim_{W \to G} \mu\Bigg(\bigcup_{g_1, g_2 \in S} Y(\gamma_1, \gamma_2, g_1, g_2) \,\Delta\, Y_W(\gamma_1, \gamma_2, g_1, g_2) \Bigg)$$

it follows that

$$\lim_{W \to G} \mu\Bigg(\bigcup_{g_1, g_2 \in G} Y_W(\gamma_1, \gamma_2, g_1, g_2) \Bigg) \geq 1 - \epsilon.$$

Since $\epsilon > 0$ is arbitrary, the lemma follows. □

In general, if Z is a topological space, and $\{z_{W,i}\}_{W \subset G, i \in \mathbb{N}}$ is a collection of elements of Z, then we write $\lim_{(W,i) \to (G,\infty)} z_{W,i} = z$ if, for every limit point z_W of $\{z_{W,i}\}_{i=1}^{\infty}$, $\lim_{W \to G} z_W = z$.

Let $\{\overline{X}_{\beta,i}\}_{i=1}^{\infty}$ be a sequence of approximate processes constructed from measures $\{\eta_i\}_{i=1}^{\infty}$ (adapted to a random sofic approximation $\mathbb{K}$ to G) so that $\lim_{i \to \infty} \overline{X}_{\beta,i} = \beta \circ \overline{X}$. For each $(\sigma, \xi) \in \mathrm{Sym}(m_i)^G \times B^{[m_i]}$ and $\gamma \in \Gamma$, define $S'_{W,\sigma,\xi}(\gamma) : [m_i] \to [m_i]$ by

$$S'_{W,\sigma,\xi}(\gamma)(p) := \sigma(g)p \Leftrightarrow \rho_W(\gamma, \xi[\sigma, W](p)) = g.$$

LEMMA 6.8. *For any $\gamma_1, \gamma_2 \in \Gamma$,*

$$\lim_{(W,i) \to (G,\infty)} \eta_i \times u_{m_i}(\{(\sigma, \xi, p) \in \mathrm{Sym}(m_i)^G \times B^{[m_i]} \times [m_i] : S'_{W,\sigma,\xi}(\gamma_1) S'_{W,\sigma,\xi}(\gamma_2) p = S'_{W,\sigma,\xi}(\gamma_1\gamma_2) p\}) = 1.$$

Also, for any $g_2' \in G$,

$$\lim_{(W,i) \to (G,\infty)} \eta_i \times u_{m_i}(\{(\sigma, \xi, p) \in \mathrm{Sym}(m_i)^G \times B^{[m_i]} \times [m_i] : \rho_W(\gamma_2, \xi[\sigma, W](p)) = g_2'\}) = \mu(\{x \in X : \rho(\gamma_2, x) = g_2'\}).$$

Proof. For $g_1, g_2 \in G$, let

$$Y(g_1, g_2) := \{x \in X : \rho(\gamma_2, x) = g_2, \rho(\gamma_1, T_{g_2} x) = g_1\}.$$

Note that a.e. $x \in Y(g_1, g_2)$ satisfies $\rho(\gamma_1\gamma_2, x) = g_1 g_2$ by the cocycle equation. For $i \geq 0$ and $W \subset G$ finite, let $Y_i^W(g_1, g_2)$ be the set of all $(\sigma, \xi, p) \in \mathrm{Sym}(m_i)^G \times B^{[m_i]} \times [m_i]$ such that

$$\begin{aligned} g_1 g_2 &= \rho_W(\gamma_1\gamma_2, \xi[\sigma, W](p)), \\ g_2 &= \rho_W(\gamma_2, \xi[\sigma, W](p)), \\ g_1 &= \rho_W(\gamma_1, \xi[\sigma, W](\sigma(g_2)p)), \\ \sigma(g_1 g_2)p &= \sigma(g_1)\sigma(g_2)p. \end{aligned}$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Note that if $(\sigma, \xi, p) \in Y_i^W(g_1, g_2)$ then

$$S'_{W,\sigma,\xi}(\gamma_1)S'_{W,\sigma,\xi}(\gamma_2)p = S'_{W,\sigma,\xi}(\gamma_1\gamma_2)p.$$

Therefore, it suffices to show that

$$\lim_{(W,i)\to(G,\infty)} \eta_i \times u_{m_i}\left(\bigcup\{Y_i^W(g_1, g_2) : g_1, g_2 \in G\}\right) = 1. \tag{6}$$

We claim that for every $g_1, g_2 \in G$,

$$\lim_{(W,i)\to(G,\infty)} \eta_i \times u_{m_i}(Y_i^W(g_1, g_2)) = \mu(Y(g_1, g_2)) \quad \text{for all } g_1, g_2 \in G. \tag{7}$$

To see this, for any finite $W \subset G$ and $g_1, g_2 \in G$, let $Z(W; g_1, g_2)$ be the set of all $x \in X$ such that

$$\begin{aligned} g_1g_2 &= \rho_W(\gamma_1\gamma_2, \psi^W(x)) \\ g_2 &= \rho_W(\gamma_2, \psi^W(x)) \\ g_1 &= \rho_W(\gamma_1, \psi^W(T_{g_2}x)). \end{aligned}$$

Then

$$\lim_{i\to\infty} \eta_i \times u_{m_i}(Y_i^W(g_1, g_2)) = \mu(Z(W; g_1, g_2)).$$

This follows from Lemmas 6.4 and 5.1. Lemma 6.6 implies

$$\lim_{W\to G} \mu(\{x \in X : g_1g_2 = \rho_W(\gamma_1\gamma_2, \psi^W(x))\} \,\Delta\, \{x \in X : g_1g_2 = \rho(\gamma_1\gamma_2, x)\}) = 0,$$

$$\lim_{W\to G} \mu(\{x \in X : g_2 = \rho_W(\gamma_2, \psi^W(x))\} \,\Delta\, \{x \in X : g_2 = \rho(\gamma_2, x)\}) = 0,$$

$$\lim_{W\to G} \mu(\{x \in X : g_1 = \rho_W(\gamma_1, \psi^W(T_{g_2}x))\} \,\Delta\, \{x \in X : g_1 = \rho(\gamma_1, T_{g_2}x)\}) = 0.$$

Therefore,

$$\lim_{W\to G} \mu(Z(W; g_1, g_2)) = \mu(Y(g_1, g_2)),$$

which implies the claim. Because

$$\mu\left(\bigcup\{Y(g_1, g_2) : g_1, g_2 \in G\}\right) = 1,$$

for any $\epsilon > 0$ there exists a finite set $G' \subset G$ such that

$$\mu\left(\bigcup\{Y(g_1, g_2) : g_1, g_2 \in G'\}\right) \geq 1 - \epsilon.$$

By the claim,

$$\lim_{(W,i)\to(G,\infty)} \eta_i \times u_{m_i}\left(\bigcup\{Y_i^W(g_1, g_2) : g_1, g_2 \in G'\}\right) \geq 1 - \epsilon.$$

Since this is true for every ϵ,

$$\lim_{(W,i)\to(G,\infty)} \eta_i \times u_{m_i}\left(\bigcup\{Y_i^W(g_1, g_2) : g_1, g_2 \in G\}\right) = 1,$$

which implies the first statement.

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

To see the second statement, observe

$$\{(\sigma, \xi, p) \in \mathrm{Sym}(m_i)^G \times B^{[m_i]} \times [m_i] : \rho_W(\gamma_2, \xi[\sigma, W](p)) = g_2'\}$$
$$\cap \bigcup \{Y_i^W(g_1, g_2) : g_1, g_2 \in G\} = \bigcup \{Y_i^W(g_1, g_2') : g_1 \in G\}.$$

Because $Y_i^W(g_1, g_2')$ is disjoint from $Y_i^W(g_1', g_2')$ if $g_1 \neq g_1'$, equations (6) and (7) imply

$$\liminf_{(W,i)\to(G,\infty)} \eta_i \times u_{m_i}(\{(\sigma, \xi, p) \in \mathrm{Sym}(m_i)^G$$
$$\times B^{[m_i]} \times [m_i] : \rho_W(\gamma_2, \xi[\sigma, W](p)) = g_2'\})$$
$$= \mu\left(\bigcup \{Y(g_1, g_2') : g_1 \in G\}\right) = \mu(\{x \in X : \rho(\gamma_2, x) = g_2'\}),$$

as required. □

LEMMA 6.9. *There are maps $S_{W,\sigma,\xi}(\gamma) \in \mathrm{Sym}(m_i)$ (for every W, σ, ξ, γ) such that for every $\gamma \in \Gamma$,*

$$\lim_{W\to G} \liminf_{i\to\infty} \eta_i \times u_{m_i}(\{(\sigma, \xi, p) \in \mathrm{Sym}(m_i)^G$$
$$\times B^{[m_i]} \times [m_i] : S'_{W,\sigma,\xi}(\gamma)(p) = S_{W,\sigma,\xi}(\gamma)(p)\}) = 1.$$

Proof. It follows from the previous lemma that for any $\gamma \in \Gamma$,

$$\lim_{W\to G} \liminf_{i\to\infty} \eta_i \times u_{m_i}(\{(\sigma, \xi, p) \in \mathrm{Sym}(m_i)^G$$
$$\times B^{[m_i]} \times [m_i] : S'_{W,\sigma,\xi}(\gamma) S'_{W,\sigma,\xi}(\gamma^{-1})(p) = p\}) = 1.$$

Therefore, $S'_{W,\sigma,\xi}$ is asymptotically surjective, in the sense that

$$\lim_{W\to G} \liminf_{i\to\infty} \eta_i \times u_{m_i}(\{(\sigma, \xi, p) \in \mathrm{Sym}(m_i)^G$$
$$\times B^{[m_i]} \times [m_i] : p \in \text{ image } S'_{W,\sigma,\xi}(\gamma)\}) = 1.$$

This implies the lemma. □

Let $\theta_{W,i}$ be the measure on $\mathrm{Sym}(m_i)^\Gamma \times B^{[m_i]}$ obtained by pushing η_i forward under the map

$$(\sigma, \xi) \in \mathrm{Sym}(m_i)^G \times B^{[m_i]} \mapsto (S_{W,\sigma,\xi}, \xi).$$

Let $\overline{Y}_{\beta,W,i}$ be the approximate process constructed from $\theta_{W,i}$ and $\kappa^\rho_{W,i}$ be the projection of $\theta_{W,i}$ to $\mathrm{Sym}(m_i)^\Gamma$.

LEMMA 6.10. *$\{\kappa^\rho_{W,i}\}_{i\in\mathbb{N}, W\subset G}$ is a random sofic approximation to Γ in the following sense. For every $\gamma_1, \gamma_2 \in \Gamma$,*

$$\lim_{(W,i)\to(G,\infty)} \kappa^\rho_{W,i} \times u_{m_i}(\{(\sigma, p) \in \mathrm{Sym}(m_i)^\Gamma \times [m_i] : \sigma(\gamma_1)\sigma(\gamma_2)p = \sigma(\gamma_1\gamma_2)p\}) = 1$$

and for every $\gamma_1 \neq \gamma_2 \in \Gamma$,

$$\lim_{(W,i)\to(G,\infty)} \kappa^\rho_{W,i} \times u_{m_i}(\{(\sigma, p) \in \mathrm{Sym}(m_i)^\Gamma \times [m_i] : \sigma(\gamma_1)p \neq \sigma(\gamma_2)p\}) = 1.$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Proof. For the first assertion, note that by definition of $\kappa^{\rho}_{W,i}$ and the previous lemma

$$\begin{aligned}
&\lim_{(W,i)\to(G,\infty)} \kappa^{\rho}_{W,i} \times u_{m_i}(\{(\sigma,\, p) \in \mathrm{Sym}(m_i)^{\Gamma} \times [m_i] : \sigma(\gamma_1)\sigma(\gamma_2)p = \sigma(\gamma_1\gamma_2)p\}) \\
&\quad = \lim_{(W,i)\to(G,\infty)} \eta_i \times u_{m_i}(\{(\sigma, \xi,\, p) \in \mathrm{Sym}(m_i)^{G} \\
&\qquad \times B^{[m_i]} \times [m_i] : S_{W,\sigma,\xi}(\gamma_1)S_{W,\sigma,\xi}(\gamma_2)p = S_{W,\sigma,\xi}(\gamma_1\gamma_2)p\}) \\
&\quad = \lim_{(W,i)\to(G,\infty)} \eta_i \times u_{m_i}(\{(\sigma, \xi,\, p) \in \mathrm{Sym}(m_i)^{G} \\
&\qquad \times B^{[m_i]} \times [m_i] : S'_{W,\sigma,\xi}(\gamma_1)S'_{W,\sigma,\xi}(\gamma_2)p = S'_{W,\sigma,\xi}(\gamma_1\gamma_2)p\}) \\
&\quad = 1.
\end{aligned}$$

The last equality follows from Lemma 6.8.

Because the first statement is true, to prove the second statement it suffices to show that for any $\gamma \neq e$,

$$\lim_{(W,i)\to(G,\infty)} \kappa^{\rho}_{W,i} \times u_{m_i}(\{(\sigma,\, p) \in \mathrm{Sym}(m_i)^{\Gamma} \times [m_i] : p \neq \sigma(\gamma)p\}) = 1.$$

Equivalently,

$$\lim_{(W,i)\to(G,\infty)} \eta_i \times u_{m_i}(\{(\sigma, \xi,\, p) \in \mathrm{Sym}(m_i)^{G} \times B^{[m_i]} \times [m_i] : p \neq S_{W,\sigma,\xi}(\gamma)p\}) = 1.$$

By Lemma 6.9 and the definition of $S'_{W,\sigma,\xi}$, it suffices to prove

$$\begin{aligned}
&\lim_{(W,i)\to(G,\infty)} \eta_i \times u_{m_i}(\{(\sigma, \xi,\, p) \in \mathrm{Sym}(m_i)^{G} \\
&\qquad \times B^{[m_i]} \times [m_i] : p \neq \sigma(\rho_W(\gamma, \xi[\sigma, W](p)))p\}) = 1.
\end{aligned}$$

If $g \in G\backslash\{e\}$ then, because $\mathbb{K}$ is a random sofic approximation to G,

$$\lim_{i\to\infty} \eta_i \times u_{m_i}(\{(\sigma, \xi,\, p) \in \mathrm{Sym}(m_i)^{G} \times B^{[m_i]} \times [m_i] : p \neq \sigma(g)p\}) = 1.$$

Also, if γ is not the identity element then, by Lemma 6.8 and because $\Gamma \curvearrowright (X, \mu)$ is essentially free, for every $\epsilon > 0$, there is a finite set $V \subset G\backslash\{e\}$ such that

$$\begin{aligned}
&\liminf_{(W,i)\to(G,\infty)} \eta_i \times u_{m_i}(\{(\sigma, \xi,\, p) \in \mathrm{Sym}(m_i)^{G} \\
&\qquad \times B^{[m_i]} \times [m_i] : \rho_W(\gamma, \xi[\sigma, W](p)) \in V\}) \geq 1 - \epsilon.
\end{aligned}$$

Because $\epsilon > 0$ is arbitrary, this inequality and the previous equation imply the lemma. □

LEMMA 6.11. *The following hold.*

(1) $\lim_{(W,i)\to(G,\infty)} \overline{Y}_{\beta,W,i} = \beta \circ \overline{Y}$.

(2) *If* $\lim_{i\to\infty} \overline{X}_{\beta,i} = \beta \circ \overline{X}$ *strongly then* $\lim_{(W,i)\to(G,\infty)} \overline{Y}_{\beta,W,i} = \beta \circ \overline{Y}$ *strongly in the sense that for all finite* $V \subset \Gamma$ *and* $\epsilon > 0$,

$$\lim_{(W,i)\to(G,\infty)} \theta_{W,i}(\{(\sigma, \xi) \in \mathrm{Sym}(m_i)^{\Gamma} \times B^{[m_i]} : d_V((\sigma, \xi), \beta \circ \phi) < \epsilon\}) = 1.$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Proof. The first statement is equivalent to stating that the V-local statistics of $\overline{Y}_{\beta,W,i}$ converges to the V-local statistics of $\beta \circ \overline{Y}$ for every finite $V \subset \Gamma$.

For $W \subset G$ finite, define

$$L_{V,W,i} : \mathrm{Sym}(m_i)^G \times B^{[m_i]} \times [m_i] \to B^V$$

by

$$L_{V,W,i}(\sigma, \xi, p)(\gamma) := \xi(S_{W,\sigma,\xi}(\gamma)p) \quad \text{for all } \gamma \in V.$$

By definition, the V-local statistics of $\overline{Y}_{\beta,W,i}$ is $(L_{V,W,i})_*\eta_i \times u_{m_i}$.

Let $\omega_{W,i}$ be the W-local statistics of $\overline{X}_{\beta,i}$. To be precise, if

$$\chi_{W,i} : \mathrm{Sym}(m_i)^G \times B^{[m_i]} \times [m_i] \to B^W$$

is defined by

$$\chi_{W,i}(\sigma, \xi, p)(g) := \xi(\sigma(g)p) \quad \text{for all } g \in W$$

then $\omega_{W,i} = (\chi_{W,i})_*\eta_i \times u_{m_i}$.

Fix $b_0 \in B$. Define $\Phi_{W,V} : B^W \to B^V$ by

$$\Phi_{W,V}(\xi)(\gamma) := \xi(\rho_W(\gamma, \xi))$$

if $\gamma \in V$ and $\rho_W(\gamma, \xi) \in W$. Set $\Phi_{W,V}(\xi)(\gamma) := b_0$ otherwise. We claim that $(\Phi_{W,V})_*\omega_{W,i}$ is asymptotic to $(L_{V,W,i})_*\eta_i \times u_{m_i}$ in the sense that

$$\lim_{(W,i)\to(G,\infty)} \|(\Phi_{W,V})_*\omega_{W,i} - (L_{V,W,i})_*\eta_i \times u_{m_i}\|_1 = 0. \tag{8}$$

According to Lemma 6.9 and the definition of $S'_{W,\sigma,\xi}$, for all $\gamma \in V$,

$$\lim_{(W,i)\to(G,\infty)} \eta_i \times u_{m_i}(\{(\sigma, \xi, p) : S_{W,\sigma,\xi}(\gamma)p = \sigma(\rho_W(\gamma, \xi[\sigma, W](p)))p\}) = 1.$$

Because $\lim_{i\to\infty} \overline{X}_{\beta,i} = \beta \circ \overline{X}$, $\lim_{i\to\infty} \omega_{W,i} = \psi_*^W\mu$. So, by Lemmas 6.4 and 6.6, for all $\gamma \in V$,

$$\lim_{(W,i)\to(G,\infty)} \eta_i \times u_{m_i}(\{(\sigma, \xi, p) : \rho_W(\gamma, \xi[\sigma, W](p)) \in W\}) = 1.$$

Therefore,

$$\lim_{(W,i)\to(G,\infty)} \eta_i \times u_{m_i}(\{(\sigma, \xi, p) : L_{V,W,i}(\sigma, \xi, p) = \Phi_{W,V}\chi_{W,i}(\sigma, \xi, p)\}) = 1.$$

This implies (8).

By Lemmas 6.4 and 5.1,

$$\lim_{i\to\infty} \Phi_{W,V} \circ \overline{X}^W_{\beta,i} = (T, X, \mu, \Phi_{W,V}\psi^W).$$

Because $\Phi_{W,V}\psi_*^W\mu$ converges to $\psi_*^V\mu$ as $W \to G$,

$$\lim_{(W,i)\to(G,\infty)} \Phi_{W,V} \circ \overline{X}^W_{\beta,i} = (T, X, \mu, \psi^V).$$

By (8), this means that the V-local statistics of $\overline{Y}_{\beta,W,i}$ converge to $(\psi^V)_*\mu$, which is the V-local statistics of $\beta \circ \overline{Y}$. This proves the first statement of the lemma. The second statement is similar. □

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

In the previous lemma, we used a sequence $\{\overline{X}_{\beta,i}\}_{i=1}^{\infty}$ of approximate processes over G converging to $\beta \circ \overline{X}$ to construct a sequence $\{\overline{Y}_{\beta,W,i}\}_{i\in\mathbb{N},W\subset G}$ that converges to $\beta \circ \overline{Y}$ (there is a slight abuse of notation here since W varies over all *finite* subsets of G instead of all subsets; we will continue this abuse below). In the next lemma, a sequence $\{\overline{X}_i\}_{i=1}^{\infty}$ of approximate processes over G such that $\beta \circ \overline{X}_i = \overline{X}_{\beta,i}$ is used to construct a new sequence $\{\overline{Y}_{W,i}\}_{i\in\mathbb{N},W\subset G}$ of approximate processes over Γ satisfying various properties.

LEMMA 6.12. *Given a sequence $\{\overline{X}_i\}_{i=1}^{\infty}$ of approximate processes over G such that $\beta \circ \overline{X}_i = \overline{X}_{\beta,i}$ there exists a collection $\{\overline{Y}_{W,i}\}_{i\in\mathbb{N},W\subset G}$ of approximate processes over Γ such that:*

(1) *$\beta \circ \overline{Y}_{W,i}$ is equivalent to $\overline{Y}_{\beta,W,i}$;*
(2) *if $\lim_{i\to\infty} \overline{X}_i = \overline{X}$ then $\lim_{(W,i)\to(G,\infty)} \overline{Y}_{W,i} = \overline{Y}$;*
(3) *if $\lim_{i\to\infty} \overline{X}_i = \overline{X}$ strongly then $\lim_{(W,i)\to(G,\infty)} \overline{Y}_{W,i} = \overline{Y}$ strongly; and*
(4) *$h(\overline{Y}_{W,i}) - h(\beta \circ \overline{Y}_{W,i}) \geq h(\overline{X}_i) - h(\beta \circ \overline{X}_i)$ for all W, i.*

Item (3) means that for all finite $V \subset \Gamma$ and $\epsilon > 0$,

$$\lim_{(W,i)\to(G,\infty)} \tilde{\theta}_{W,i}(\{(\sigma,\xi) \in \mathrm{Sym}(m_i)^{\Gamma} \times A^{[m_i]} : d_V((\sigma,\xi),\phi) < \epsilon\}) = 1,$$

where $\overline{Y}_{W,i}$ is constructed from measures $\tilde{\theta}_{W,i}$ on $\mathrm{Sym}(m_i)^{\Gamma} \times A^{[m_i]}$.

Proof. Let $\tilde{\eta}_i$ be the probability measure on $\mathrm{Sym}(m_i)^G \times A^{[m_i]}$ such that $\overline{X}_i$ is constructed from $\tilde{\eta}_i$. Let $\tilde{\theta}_{W,i}$ be obtained by pushing $\tilde{\eta}_i$ forward under the map $J = J_{W,i}$ defined by

$$J(\sigma,\xi) = (S_{W,\sigma,\beta\circ\xi}, \xi).$$

Let $\overline{Y}_{W,i}$ be the approximate process over Γ constructed from $\tilde{\theta}_{W,i}$.

Since $\beta \circ \overline{X}_i = \overline{X}_{\beta,i}$, it follows that $\beta \circ \overline{Y}_{W,i}$ is equivalent to $\overline{Y}_{\beta,W,i}$. The proofs of statements (2) and (3) are similar to the proofs of (1) and (2) of Lemma 6.11.

To prove the last statement, fix i and W. Let $\tilde{\beta} : \mathrm{Sym}(m_i)^G \times A^{[m_i]} \to \mathrm{Sym}(m_i)^G \times B^{[m_i]}$ be the map $\tilde{\beta}(\sigma,\xi) = (\sigma, \beta \circ \xi)$. Let $\tilde{\eta}_{i,\tilde{\beta}(\sigma,\xi)}$ be the fiber measure of $\tilde{\eta}_i$ over $\tilde{\beta}(\sigma,\xi)$. Thus,

$$\tilde{\eta}_i = \int \tilde{\eta}_{i,\tilde{\beta}(\sigma,\xi)}\, d\tilde{\beta}_*\tilde{\eta}_i(\tilde{\beta}(\sigma,\xi)),$$

$$h(\overline{X}_i) - h(\beta \circ \overline{X}_i) = \frac{1}{m_i} \int H(\tilde{\eta}_{i,\tilde{\beta}(\sigma,\xi)})\, d\tilde{\beta}_*\tilde{\eta}_i(\tilde{\beta}(\sigma,\xi)).$$

By abuse of notation, we also let $\tilde{\beta}$ denote the same map with G replaced by Γ. We define the fiber measure $\tilde{\theta}_{W,i,\tilde{\beta}J(\sigma,\xi)}$ of $\tilde{\theta}_{W,i}$ over $\tilde{\beta}(J(\sigma,\xi))$ similarly. Thus,

$$\begin{aligned} h(\overline{Y}_{W,i}) - h(\beta \circ \overline{Y}_{W,i}) &= \frac{1}{m_i} \int H(\tilde{\theta}_{W,i,\tilde{\beta}J(\sigma,\xi)})\, d\tilde{\theta}_{W,i}(J(\sigma,\xi)) \\ &= \frac{1}{m_i} \int H(\tilde{\theta}_{W,i,\tilde{\beta}J(\sigma,\xi)})\, d\tilde{\eta}_i(\sigma,\xi). \end{aligned}$$

Let $\alpha : \mathrm{Sym}(m_i)^G \times A^{[m_i]} \to A^{[m_i]}$ denote the projection map. By abuse of notation, we also let α denote the projection map from $\mathrm{Sym}(m_i)^{\Gamma} \times A^{[m_i]}$ to $A^{[m_i]}$. Because $\alpha J = \alpha$ and $J_*(\tilde{\eta}_{i,\tilde{\beta}J(\sigma,\xi)}) = \tilde{\theta}_{W,i,\tilde{\beta}J(\sigma,\xi)}$,

$$\alpha_*(\tilde{\eta}_{i,\tilde{\beta}J(\sigma,\xi)}) = \alpha_*(\tilde{\theta}_{W,i,\tilde{\beta}J(\sigma,\xi)}) \quad \text{for all } (\sigma,\xi) \in \mathrm{Sym}(m_i)^G \times A^{[m_i]}.$$

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Since α is injective on the support of $\tilde{\theta}_{W,i,\tilde{\beta}J(\sigma,\xi)}$, it follows that

$$H(\tilde{\theta}_{W,i,\tilde{\beta}J(\sigma,\xi)}) = H(\alpha_*\tilde{\theta}_{W,i,\tilde{\beta}J(\sigma,\xi)}) = H(\alpha_*\tilde{\eta}_{i,\tilde{\beta}J(\sigma,\xi)}).$$

We claim that

$$\tilde{\eta}_{i,\tilde{\beta}J(\sigma,\xi)} = \int \tilde{\eta}_{i,\sigma',\beta\xi'}\, d\tilde{\eta}_{i,\tilde{\beta}J(\sigma,\xi)}(\sigma',\xi').$$

To see this, suppose X, Y, Z are any three Borel spaces, μ is a probability measure on X and $\pi_1 : X \to Y$, $\pi_2 : Y \to Z$ are Borel maps. Then, for $(\pi_2\pi_1)_*\mu$-a.e. $z \in Z$, $\mu_z = \int \mu_{\pi_1(x)}\, d\mu_z(x)$. This follows from the fact that if Σ_Y, Σ_Z are the σ-algebras on X obtained from pulling back the σ-algebras on Y, Z, respectively, and f is a bounded function on X, then $\mathbb{E}[f|\Sigma_Z] = \mathbb{E}[\mathbb{E}[f|\Sigma_Y]|\Sigma_Z]$. To see how this applies to the equation above, set $X = \mathrm{Sym}(m_i)^G \times A^{[m_i]}$, $Y = \mathrm{Sym}(m_i)^G \times B^{[m_i]}$ and $Z = \mathrm{Sym}(m_i)^\Gamma \times B^{[m_i]}$. Let $\pi_1 = \tilde{\beta}$ and choose π_2 so that $\pi_2\pi_1 = \tilde{\beta}J$.

Therefore,

$$\alpha_*\tilde{\eta}_{i,\tilde{\beta}J(\sigma,\xi)} = \int \alpha_*\tilde{\eta}_{i,\sigma',\beta\xi'}\, d\tilde{\eta}_{i,\tilde{\beta}J(\sigma,\xi)}(\sigma',\xi').$$

By concavity of entropy,

$$H(\alpha_*\tilde{\eta}_{i,\tilde{\beta}J(\sigma,\xi)}) \geq \int H(\alpha_*\tilde{\eta}_{i,\sigma',\beta\xi'})\, d\tilde{\eta}_{i,\tilde{\beta}J(\sigma,\xi)}(\sigma',\xi').$$

Because α is injective on the support of $\tilde{\eta}_{i,\sigma',\beta\xi'}$, $H(\alpha_*\tilde{\eta}_{i,\sigma',\beta\xi'}) = H(\tilde{\eta}_{i,\sigma',\beta\xi'})$. Therefore,

$$H(\tilde{\theta}_{W,i,\tilde{\beta}J(\sigma,\xi)}) \geq \int H(\tilde{\eta}_{i,\sigma',\beta\xi'})\, d\tilde{\eta}_{i,\tilde{\beta}J(\sigma,\xi)}(\sigma',\xi').$$

Thus,

$$\begin{aligned} h(\overline{Y}_{W,i}) - h(\beta \circ \overline{Y}_{W,i}) &\geq \frac{1}{m_i} \iint H(\tilde{\eta}_{i,\sigma',\beta\xi'})\, d\tilde{\eta}_{i,\tilde{\beta}J(\sigma,\xi)}(\sigma',\xi')\, d\tilde{\eta}_i(\sigma,\xi) \\ &= \frac{1}{m_i} \int H(\tilde{\eta}_{i,\sigma,\beta\xi})\, d\tilde{\eta}_i(\sigma,\xi) \\ &= h(\overline{X}_i) - h(\beta \circ \overline{X}_i). \end{aligned}$$

□

We can now prove Proposition 6.2 whose statement is the following.

PROPOSITION 6.2. *Let G be an amenable group, $\overline{X} = (T, X, \mu, \phi)$ an essentially free G-process with finite range A and $S : (X, \mu) \to (X, \mu)$ be an essentially free measure-preserving Borel automorphism with the same orbits as T (i.e., for μ-a.e. $x \in X$, $\{T_g x : g \in G\} = \{S^n x : n \in \mathbb{Z}\}$).*

Let $\beta : A \to B$ be a map and suppose the orbit change from T to S is measurable with respect to both the T-invariant sub-σ-algebra generated by $\psi := \beta \circ \phi$ and the S-invariant sub-σ-algebra generated by ψ. Then, for any random sofic approximation $\mathbb{K}$ to G,

$$\bar{h}(\mathbb{K}, \overline{X}|\beta \circ \overline{X}) = h(\mathbb{K}, \overline{X}|\beta \circ \overline{X}) = h(S, \phi|\beta \circ \phi).$$

Proof. Let $\Gamma = \mathbb{Z}$. Let $\{\overline{X}_i\}_{i=1}^\infty$ be a sequence of approximate processes adapted to $\mathbb{K}'$, a subsequence of $\mathbb{K}$, so that:

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

(1) $\lim_{i\to\infty} \overline{X}_i = \overline{X}$; and
(2) $\lim_{i\to\infty} h(\overline{X}_i) - h(\beta \circ \overline{X}_i) = \bar{h}(\mathbb{K}, \overline{X}|\beta \circ \overline{X})$.

Let $\overline{X}_{\beta,i} = \beta \circ \overline{X}_i$. Let $\{\overline{Y}_{\beta,W,i}\}_{i\in\mathbb{N},W\subset G}$ be constructed as in the paragraph before Lemma 6.10. Let $\{\overline{Y}_{W,i}\}_{i\in\mathbb{N},W\subset G}$ be the collection of approximate processes given by Lemma 6.12.

A diagonalization argument implies that there exist increasing sequences $\{W_i\}_{i=1}^{\infty}$ and $\{k_i\}_{i=1}^{\infty}$ so that if $\overline{Y}_i := \overline{Y}_{W_i,k_i}$ then $\lim_{i\to\infty} \overline{Y}_i = \overline{Y} = (S, X, \mu, \phi)$. Moreover, we may assume (using Lemma 6.10) that if $\mathbb{K}^{\rho} := \{\kappa^{\rho}_{W_i,k_i}\}_{i=1}^{\infty}$ then $\mathbb{K}^{\rho}$ is a random sofic approximation to $\mathbb{Z}$. So, Lemma 6.12 implies

$$\begin{aligned}\bar{h}(\mathbb{K}, \overline{X}|\beta \circ \overline{X}) &= \lim_{i\to\infty} h(\overline{X}_i) - h(\beta \circ \overline{X}_i) \le \limsup_{i\to\infty} h(\overline{Y}_i) - h(\beta \circ \overline{Y}_i)\\ &\le \bar{h}(\mathbb{K}^{\rho}, \overline{Y}|\beta \circ \overline{Y}) = h(\overline{Y}|\beta \circ \overline{Y}).\end{aligned}$$

The last equality follows from Proposition 5.8. Since $\bar{h}(\mathbb{K}, \overline{X}|\beta \circ \overline{X}) \ge h(\mathbb{K}, \overline{X}|\beta \circ \overline{X})$, it now suffices to prove $h(\mathbb{K}, \overline{X}|\beta \circ \overline{X}) \ge h(\overline{Y}|\beta \circ \overline{Y})$.

Let integers m_i be given so that $\mathbb{K}' = \{\kappa'_i\}_{i=1}^{\infty}$, where each κ'_i is a probability measure on $\mathrm{Sym}(m_i)^G$. Let $\mathbb{L} = \{\lambda_i\}_{i=1}^{\infty}$ be a random sofic approximation to $\mathbb{Z}$, where each λ_i is a probability measure on $\mathrm{Sym}(m_i)^{\mathbb{Z}}$. By Proposition 5.8, there exists a sequence $\{\overline{Y}'_i\}_{i\in\mathbb{N}}$ of approximate processes over $\mathbb{Z}$ adapted to $\mathbb{L}'$, a subsequence of $\mathbb{L}$ such that:

(1) $\lim_{i\to\infty} \overline{Y}'_i = \overline{Y}$ strongly; and
(2) $\lim_{i\to\infty} h(\overline{Y}'_i) - h(\beta \circ \overline{Y}'_i) = h(\mathbb{L}, \overline{Y}|\beta \circ \overline{Y}) = h(\overline{Y}|\beta \circ \overline{Y})$.

Let $\{\overline{X}_{\beta,V,i}\}_{i\in\mathbb{N},V\subset\mathbb{Z}}$ be constructed as in the paragraph before Lemma 6.10 with the roles of X and Y swapped. This construction is possible because the orbit change is, by hypothesis, measurable with respect to the smallest σ-algebra generated by ψ and the $\mathbb{Z}$-action. Let $\{\overline{X}_{V,i}\}_{i\in\mathbb{N},V\subset\mathbb{Z}}$ be the collection of approximate processes given by Lemma 6.12. According to that lemma:

(1) $\lim_{(V,i)\to(\mathbb{Z},\infty)} \overline{X}'_{V,i} = \overline{X}$ strongly; and
(2) $h(\overline{X}'_{V,i}) - h(\beta \circ \overline{X}'_{V,i}) \ge h(\overline{Y}'_i) - h(\beta \circ \overline{Y}'_i)$ for all V, i.

So,

$$h(\overline{Y}|\beta \circ \overline{Y}) = \lim_{i\to\infty} h(\overline{Y}'_i) - h(\beta \circ \overline{Y}'_i) \le \liminf_{(V,i)\to(\mathbb{Z},\infty)} h(\overline{X}'_{V,i}) - h(\beta \circ \overline{X}'_{V,i}).$$

A diagonalization argument and Lemma 6.10 imply that there exist increasing sequences $\{V_i\}_{i=1}^{\infty}$ and $\{k_i\}_{i=1}^{\infty}$ so that if $\overline{X}'_i := \overline{X}'_{V_i,k_i}$ then $\{\overline{X}'_i\}_{i=1}^{\infty}$ is adapted to a random sofic approximation $\mathbb{K}''$ of G. So, the inequalities above imply

$$h(\overline{Y}|\beta \circ \overline{Y}) \le \limsup_{i\to\infty} h(\overline{X}'_i) - h(\beta \circ \overline{X}'_i) \le h(\mathbb{K}'', \overline{X}|\beta \circ \overline{X}).$$

Observe that $\mathbb{K}'' = \{\kappa''_i\}_{i=1}^{\infty}$, where each κ''_i is a probability measure on $\mathrm{Sym}(n_i)^G$ and $\{n_i\}_{i=1}^{\infty}$ is a subsequence of $\{m_i\}_{i=1}^{\infty}$. It follows from Theorem 4.3 that $\mathbb{K}''$ is asymptotic to a subsequence of $\mathbb{K}'$. It follows from Theorem 5.5 that

$$h(\mathbb{K}'', \overline{X}|\beta \circ \overline{X}) \le h(\mathbb{K}', \overline{X}|\beta \circ \overline{X}) \le h(\mathbb{K}, \overline{X}|\beta \circ \overline{X}).$$

So, $h(\mathbb{K}, \overline{X}|\beta \circ \overline{X}) \ge h(\overline{Y}|\beta \circ \overline{Y})$ as required. □

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press

Acknowledgements. I would like to thank Gabor Elek for providing a rough outline of a proof of the main theorem based on quasi-tiling machinery. The proof presented here does not use his outline but it helped get me started. I would also like to thank David Kerr for encouragement and especially Hanfeng Li for finding many errors in previous versions. The author was supported in part by NSF grant DMS-0968762 and NSF CAREER Award DMS-0954606.

References

[Bo10a] L. Bowen. Measure conjugacy invariants for actions of countable sofic groups. *J. Amer. Math. Soc.* **23** (2010), 217–245.

[Bo10b] L. Bowen. Entropy for expansive algebraic actions of residually finite groups. *Ergod. Th. & Dynam. Sys.* to appear, arXiv:0909.4770.

[CFW81] A. Connes, J. Feldman and B. Weiss. An amenable equivalence relation is generated by a single transformation. *Ergod. Th. & Dynam. Sys.* **1**(4) (1982), 431–450.

[Dy59] H. A. Dye. On groups of measure preserving transformation. I. *Amer. J. Math.* **81** (1959), 119–159.

[Dy63] H. A. Dye. On groups of measure preserving transformations. II. *Amer. J. Math.* **85** (1963), 551–576.

[ES10] G. Elek and E. Szabó. Sofic representations of amenable groups. *Preprint*, arXiv:1010.3424.

[Gl03] E. Glasner. *Ergodic Theory via Joinings (Mathematical Surveys and Monographs, 101)*. American Mathematical Society, Providence, RI, 2003.

[Gro99] M. Gromov. Endomorphisms of symbolic algebraic varieties. *J. Eur. Math. Soc.* **1**(2) (1999), 109–197.

[KL1] D. Kerr and H. Li. Entropy and the variational principle for actions of sofic groups. *Preprint*, arXiv:1005.0399.

[KL2] D. Kerr and H. Li. Soficity, amenability and dynamical entropy. *Preprint*, arXiv:1008.1429.

[Ma40] A. I. Mal'cev. On faithful representations of infinite groups of matrices. *Mat. Sb.* **8** (1940), 405–422, *Amer. Math. Soc. Transl.* (2) **45** (1965), 1–18.

[Ol85] J. Moulin Ollagnier. *Ergodic Theory and Statistical Mechanics (Lecture Notes in Mathematics, 1115)*. Springer, Berlin, 1985.

[OW80] D. 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. Ornstein and B. Weiss. Entropy and isomorphism theorems for actions of amenable groups. *J. Anal. Math.* **48** (1987), 1–141.

[Pe08] V. Pestov. Hyperlinear and sofic groups: a brief guide. *Bull. Symbolic Logic* **14**(4) (2008), 449–480.

[Ro88] A. Rosenthal. Finite uniform generators for ergodic, finite entropy, free actions of amenable groups. *Probab. Theory Related Fields* **77**(2) (1988), 147–166.

[RW00] D. J. Rudolph and B. Weiss. Entropy and mixing for amenable group actions. *Ann. of Math.* (2) **151**(3) (2000), 1119–1150.

[We00] B. Weiss. Sofic groups and dynamical systems. *Ergodic Theory and Harmonic Analysis, Mumbai, 1999*. *Sankhyā Ser. A* **62**(3) (2000), 350–359.

[WZ92] T. Ward and Q. Zhang. The Abramov–Rohlin entropy addition formula for amenable group actions. *Monatsh. Math.* **114**(3–4) (1992), 317–329.

https://doi.org/10.1017/S0143385711000253 Published online by Cambridge University Press