KRIEGER'S FINITE GENERATOR THEOREM FOR ACTIONS OF COUNTABLE GROUPS I

BRANDON SEWARD

ABSTRACT. For an ergodic p.m.p. action $G \curvearrowright (X, \mu)$ of a countable group G, we define the Rokhlin entropy $h_G^{\mathrm{Rok}}(X, \mu)$ to be the infimum of the Shannon entropies of countable generating partitions. It is known that for free ergodic actions of amenable groups this notion coincides with classical Kolmogorov–Sinai entropy. It is thus natural to view Rokhlin entropy as a close analogue to classical entropy. Under this analogy we prove that Krieger's finite generator theorem holds for all countably infinite groups. Specifically, if $h_G^{\mathrm{Rok}}(X, \mu) < \log(k)$ then there exists a generating partition consisting of k sets.

1. Introduction

Let (X, μ) be a standard probability space, meaning X is a standard Borel space with Borel σ-algebra $\mathcal{B}(X)$ and μ is a Borel probability measure. Let G be a countable group, and let $G \curvearrowright (X, \mu)$ be a probability-measure-preserving (p.m.p.) action. For a collection $\xi \subseteq \mathcal{B}(X)$, we let $\sigma\text{-alg}_G(\xi)$ denote the smallest G-invariant σ-algebra containing ξ. A countable Borel partition α is *generating* if $\sigma\text{-alg}_G(\alpha) = \mathcal{B}(X)$ (equality modulo μ-null sets). The *Shannon entropy* of a countable Borel partition α is

$$\mathrm{H}(\alpha) = \sum_{A \in \alpha} -\mu(A) \cdot \log(\mu(A)).$$

A *probability vector* is a finite or countable ordered tuple $\bar{p} = (p_i)$ of positive real numbers which sum to 1 (a more general definition will appear in Section 2). We write $|\bar{p}|$ for the length of $\bar{p}$ and $\mathrm{H}(\bar{p}) = \sum -p_i \cdot \log(p_i)$ for the Shannon entropy of $\bar{p}$.

Countable Borel partitions, and generating partitions in particular, have long played an important role in classical entropy theory. Generating partitions greatly simplify the definition and computation of Kolmogorov–Sinai entropy, and the proofs of two of the most well known results in entropy theory, Sinai's factor theorem and Ornstein's isomorphism theorem, relied upon deep, intricate constructions in which partitions played a starring role. Furthermore, generating partitions are more than merely a tool in entropy theory, but in fact are intimately connected with the notion of entropy itself. This fact is demonstrated by the following fundamental theorems of Rokhlin and Krieger.

1

Theorem (Rokhlin's generator theorem [31], 1967). *If $\mathbb{Z} \curvearrowright (X, \mu)$ is a free ergodic p.m.p. action then its Kolmogorov–Sinai entropy $h_{\mathbb{Z}}^{\mathrm{KS}}(X, \mu)$ satisfies*

$$h_{\mathbb{Z}}^{\mathrm{KS}}(X, \mu) = \inf \Big\{ \mathrm{H}(\alpha) \, : \, \alpha \text{ is a countable generating partition} \Big\}.$$

Theorem (Krieger's finite generator theorem [25], 1970). *If $\mathbb{Z} \curvearrowright (X, \mu)$ is a free ergodic p.m.p. action and $h_{\mathbb{Z}}^{\mathrm{KS}}(X, \mu) < \log(k)$ then there exists a generating partition α consisting of k sets.*

Both of the above theorems were later superseded by the following result of Denker.

Theorem (Denker [11], 1974). *If $\mathbb{Z} \curvearrowright (X, \mu)$ is a free ergodic p.m.p. action and $\bar{p}$ is a finite probability vector with $h_{\mathbb{Z}}^{\mathrm{KS}}(X, \mu) < \mathrm{H}(\bar{p})$, then for every $\epsilon > 0$ there is a generating partition $\alpha = \{A_0, \ldots, A_{|\bar{p}|-1}\}$ with $|\mu(A_i) - p_i| < \epsilon$ for every $0 \le i < |\bar{p}|$.*

Grillenberger and Krengel [17] obtained a further strengthening of these results which roughly says that, under the assumptions of Denker's theorem, one can control the joint distribution of α and finitely many of its translates. In particular, they showed that under the assumptions of Denker's theorem there is a generating partition α with $\mu(A_i) = p_i$ for every $0 \le i < |\bar{p}|$.

Over the years, Krieger's theorem acquired much fame and underwent various generalizations. In 1972, Katznelson and Weiss [18] outlined a proof of Krieger's theorem for free ergodic actions of $\mathbb{Z}^d$. Roughly a decade later, Šujan [41] stated Krieger's theorem for amenable groups but only outlined the proof. The first proof for amenable groups to appear in the literature was obtained in 1988 by Rosenthal [32] who proved Krieger's theorem under the more restrictive assumption that $h_G^{\mathrm{KS}}(X, \mu) < \log(k-2) < \log(k)$. This was not improved until 2002 when Danilenko and Park [10] proved Krieger's theorem for amenable groups under the assumption $h_G^{\mathrm{KS}}(X, \mu) < \log(k-1) < \log(k)$. It is none-the-less a folklore unpublished result that Krieger's theorem holds for amenable groups, i.e. if $G \curvearrowright (X, \mu)$ is a free ergodic p.m.p. action of an amenable group and $h_G^{\mathrm{KS}}(X, \mu) < \log(k)$ then there is a generating partition consisting of k sets. Our much more general investigations here yield this as a consequence. We believe that this is the first explicit proof of this fact.

Krieger's theorem was also generalized to a relative setting. The relative version of Krieger's theorem for $\mathbb{Z}$ actions was first proven by Kifer and Weiss [24] in 2002. It states that if $\mathbb{Z} \curvearrowright (X, \mu)$ is a free ergodic p.m.p. action, $\mathcal{F}$ is a $\mathbb{Z}$-invariant sub-σ-algebra, and the relative entropy satisfies $h_{\mathbb{Z}}^{\mathrm{KS}}(X, \mu | \mathcal{F}) < \log(k)$, then there is a Borel partition α consisting of k sets such that $\sigma\text{-alg}_{\mathbb{Z}}(\alpha) \vee \mathcal{F} = \mathcal{B}(X)$. This result was later extended by Danilenko and Park [10] to free ergodic actions of amenable groups under the assumption that $\mathcal{F}$ induces a class-bijective factor.

Rokhlin's theorem was generalized to actions of abelian groups by Conze [8] in 1972 and was just recently extended to amenable groups by Seward and Tucker-Drob [39]. Specifically, if $G \curvearrowright (X, \mu)$ is a free ergodic p.m.p. action of an amenable group then the entropy $h_G^{\mathrm{KS}}(X, \mu)$ is equal to the infimum of $\mathrm{H}(\alpha)$ over all countable generating partitions α. Denker's theorem on the other hand has not been extended beyond actions of $\mathbb{Z}$.

In this paper we consider arbitrary countable groups, but we are particularly interested in the case of non-amenable groups. This is due to the recent breathtaking development of an entropy theory for actions of certain non-amenable groups.

Specifically, groundbreaking work of Bowen in 2008 [4], followed with improvements by Kerr and Li [21], has created the notion of *sofic entropy* for p.m.p. actions of sofic groups. We remind the reader that the class of sofic groups contains the countable amenable groups, and it is an open question if every countable group is sofic. Sofic entropy in fact extends Kolmogorov–Sinai entropy as the two notions coincide for actions of amenable groups [5, 22]. The sofic entropy of a Bernoulli shift (L^G, λ^G) is equal to the Shannon entropy of its base $H(\lambda)$ [4, 23]. Consequently, for sofic groups containing an infinite amenable subgroup, sofic entropy classifies Bernoulli shifts up to isomorphism [26, 27, 29, 40]. Bowen has also made significant progress on classifying Bernoulli shifts over general countable groups [6], but a full classification does not yet exist.

This paper is motivated by some of the challenges facing sofic entropy theory. For instance, is Sinai's factor theorem true: do free ergodic actions of positive sofic entropy factor onto Bernoulli shifts? Is the Ornstein isomorphism theorem true: are Bernoulli shifts over sofic groups classified up to isomorphism by their sofic entropy? For both questions, the classical proofs involving elaborate constructions of partitions cannot be carried out. This is because vital properties of actions of amenable groups such as the Rokhlin lemma, the Shannon–McMillan–Breiman theorem, and the monotone decreasing property of entropy under factor maps are simply false for non-amenable groups. Additionally, the formula for sofic entropy involves counting auxiliary objects which are external to the original action, making it unclear if the language of sofic entropy is sufficiently rich to uncover the delicate constructions of partitions which are needed for answering these questions. We also do not know if sofic entropy satisfies either the Rokhlin generator theorem or the Krieger finite generator theorem. These are pertinent questions since sofic entropy is easier to define, compute, and understand when there exists a finite generating partition. Finally, a significant challenge to sofic entropy is that it is restricted to the realm of actions of sofic groups (in fact, to the realm of "sofic actions"). If countable, non-sofic groups exist, how will we understand and classify their Bernoulli shifts?

Motivated by these issues, we introduce a new notion of entropy which is defined for all p.m.p. actions of all countable groups. For a p.m.p. action $G \curvearrowright (X, \mu)$ we define the *Rokhlin entropy* to be

$$h_G^{\mathrm{Rok}}(X, \mu) = \inf \left\{ H(\alpha | \mathscr{I}) : \alpha \text{ is a countable partition and } \sigma\text{-alg}_G(\alpha) \vee \mathscr{I} = \mathcal{B}(X) \right\},$$

where $\mathscr{I}$ is the σ-algebra of G-invariant Borel sets. We study this invariant in this three-part series, but in the present paper we only consider ergodic actions, in which case the Rokhlin entropy simplifies to

$$h_G^{\mathrm{Rok}}(X, \mu) = \inf \left\{ H(\alpha) : \alpha \text{ is a countable Borel generating partition} \right\}.$$

We name this invariant in honor of Rokhlin's generator theorem. For free actions of amenable groups, Rokhlin entropy coincides with Kolmogorov–Sinai entropy [3, 39]. Thus Rokhlin entropy is a simple and natural extension of Kolmogorov–Sinai entropy.

More generally, if $\mathcal{F} \subseteq \mathcal{B}(X)$ is a G-invariant σ-algebra, then we define the *Rokhlin entropy of $G \curvearrowright (X, \mu)$ relative to $\mathcal{F}$*, denoted $h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F})$, to be

$$\inf \left\{ H(\alpha | \mathcal{F} \vee \mathscr{I}) : \alpha \text{ is a countable Borel partition and } \sigma\text{-alg}_G(\alpha) \vee \mathcal{F} \vee \mathscr{I} = \mathcal{B}(X) \right\}.$$

Again, $\mathscr{I}$ will always be trivial in this paper because we will only consider ergodic actions. We refer the reader to Section 2 for the definition of the conditional Shannon entropy $\mathrm{H}(\alpha|\mathcal{F})$, but we remark that when $\mathcal{F} = \{X, \varnothing\}$ we have $\mathrm{H}(\alpha|\mathcal{F}) = \mathrm{H}(\alpha)$. In particular $h_G^{\mathrm{Rok}}(X, \mu|\{X, \varnothing\}) = h_G^{\mathrm{Rok}}(X, \mu)$.

We show in Proposition 9.1 that for free ergodic actions of amenable groups relative Rokhlin entropy coincides with relative Kolmogorov–Sinai entropy (this is extended to non-ergodic actions in Part III [3]). We also observe in Proposition 3.4 that $h_G^{\mathrm{Rok}}(X, \mu|\mathcal{F})$ is invariant under orbit equivalences for which the orbit-change cocycle is $\mathcal{F}$-measurable, generalizing a similar property of Kolmogorov–Sinai entropy discovered by Rudolph and Weiss [33].

Our main theorem is the following finite generator theorem which applies to all ergodic actions of countably infinite groups. (We in fact prove a stronger result; see Theorem 2.3).

Theorem 1.1. *Let G be a countably infinite group acting ergodically, but not necessarily freely, by measure-preserving bijections on a non-atomic standard probability space (X, μ). Let $\mathcal{F}$ be a G-invariant sub-σ-algebra. If $\bar{p} = (p_i)$ is any finite or countable probability vector with $h_G^{\mathrm{Rok}}(X, \mu|\mathcal{F}) < \mathrm{H}(\bar{p})$, then there is a Borel partition $\alpha = \{A_i : 0 \le i < |\bar{p}|\}$ with $\mu(A_i) = p_i$ for every $0 \le i < |\bar{p}|$ and with $\sigma\text{-alg}_G(\alpha) \vee \mathcal{F} = \mathcal{B}(X)$.*

This theorem greatly supersedes previous work of the author in [34] which, under the assumption $h_G^{\mathrm{Rok}}(X, \mu) < \infty$, constructed a finite generating partition without any control over its cardinality or distribution. The major difficulty which the present work overcomes is that all prior arguments for controlling the cardinality and distribution of generating partitions rely critically upon the classical Rokhlin lemma and Shannon–McMillan–Breiman theorem, and these tools do not exist for actions of general countable groups.

We remark that in order for a partition α to exist as described in Theorem 1.1, it is necessary that $h_G^{\mathrm{Rok}}(X, \mu|\mathcal{F}) \le \mathrm{H}(\bar{p})$. So the above theorem is optimal since in general there are actions where the infimum $h_G^{\mathrm{Rok}}(X, \mu)$ is not achieved, such as free ergodic actions which are not isomorphic to any Bernoulli shift [36].

If $h_G^{\mathrm{Rok}}(X, \mu) < \log(k)$ then using $\bar{p} = (p_0, \ldots, p_{k-1})$ where each $p_i = 1/k$ we obtain the following generalization of the (relative) Krieger finite generator theorem:

Corollary 1.2. *Let G be a countably infinite group acting ergodically, but not necessarily freely, by measure-preserving bijections on a non-atomic standard probability space (X, μ), and let $\mathcal{F}$ be a G-invariant sub-σ-algebra. If $h_G^{\mathrm{Rok}}(X, \mu|\mathcal{F}) < \log(k)$, then there is a partition α with $|\alpha| = k$ and $\sigma\text{-alg}_G(\alpha) \vee \mathcal{F} = \mathcal{B}(X)$.*

We mention that Corollary 1.2 is the first version of Krieger's finite generator theorem for non-free actions. Furthermore, we believe that Corollary 1.2 (together with the Rokhlin generator theorem for amenable groups [39]) is the first explicit proof of Krieger's finite generator theorem for free ergodic actions of countable amenable groups. In fact, we obtain the following strong form of Denker's theorem for amenable groups:

Corollary 1.3. *Let G be a countably infinite amenable group and let $G \curvearrowright (X, \mu)$ be a free ergodic p.m.p. action. If $\bar{p} = (p_i)$ is any finite or countable probability vector with $h_G^{\mathrm{KS}}(X, \mu) < \mathrm{H}(\bar{p})$ then there exists a generating partition $\alpha = \{A_i : 0 \le i < |\bar{p}|\}$ with $\mu(A_i) = p_i$ for every $0 \le i < |\bar{p}|$.*

Returning to our points of motivation, we point out that Theorem 1.1 implies that sofic entropy will satisfy the Krieger finite generator theorem provided it satisfies the Rokhlin generator theorem (i.e. provided sofic entropy coincides with Rokhlin entropy). A simple consequence of the definitions is that sofic entropy is bounded above by Rokhlin entropy [3, 4], and its an important open problem to determine if they are equal (assuming a free action with sofic entropy not minus infinity).

Since sofic entropy is a lower bound to Rokhlin entropy, we know that Bernoulli shifts (L^G, λ^G) over sofic groups G have Rokhlin entropy $h_G^{\mathrm{Rok}}(L^G, \lambda^G) = \mathrm{H}(\lambda)$. Since the definition of Rokhlin entropy does not mention soficity, this suggests the equality $h_G^{\mathrm{Rok}}(L^G, \lambda^G) = \mathrm{H}(\lambda)$ may hold for all countably infinite groups. If so, Rokhlin entropy could be capable of classifying Bernoulli shifts over all countably infinite groups up to isomorphism (in particular, over non-sofic groups if they exist). We further investigate this question in Part II [36].

In the long-term, we hope that Rokhlin entropy will not only be useful in its own right, but that its study will develop in parallel with sofic entropy theory and that the two theories will be mutually beneficial to one another. In particular, we hope Rokhlin entropy will enrich the language and framework available for studying entropy-type problems and that it will progress our techniques for constructing partitions, possibly leading the way to generalizations of the many deep results of Ornstein.

A hidden significance of Theorem 1.1 is that it opens the door to developing a theory of Rokhlin entropy. It should be pointed out that the definition of Rokhlin entropy is both quite natural and immediately suggested by Rokhlin's generator theorem, and the idea of its definition had certainly occurred to researchers beforehand. However, the abstract nature of the definition, an infimum over *all* generating partitions, seems to prevent any viable means of study. Our main theorem changes this situation. It reveals, as a consequence, a sub-additive property of Rokhlin entropy. To properly state this property in its strongest form requires additional definitions and is postponed to Part II [36], but we mention here a simple corollary to give an indication.

Corollary 1.4. *Let G be a countably infinite group acting ergodically, but not necessarily freely, by measure-preserving bijections on a non-atomic standard probability space (X, μ). If $G \curvearrowright (Y, \nu)$ is a factor of $G \curvearrowright (X, \mu)$ and $\mathcal{F}$ is the sub-σ-algebra of X associated to Y then*

$$h_G^{\mathrm{Rok}}(X, \mu) \leq h_G^{\mathrm{Rok}}(Y, \nu) + h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F}).$$

For example, if α and β are partitions with $\sigma\text{-alg}_G(\alpha \vee \beta) = \mathcal{B}(X)$, then the above corollary implies $h_G^{\mathrm{Rok}}(X, \mu) \leq \mathrm{H}(\alpha) + \mathrm{H}(\beta | \sigma\text{-alg}_G(\alpha))$. The inequality in Corollary 1.4 can be strict, such as when $h_G^{\mathrm{Rok}}(X, \mu) < h_G^{\mathrm{Rok}}(Y, \nu)$. A strict inequality is common for actions of non-amenable groups [35].

The sub-additive property turns out to be tremendously useful, and it is absolutely critical to our study of Rokhlin entropy in Parts II and III.

Update. Between this article's first appearance on the arXiv and it reaching its final form, a few developments have occurred.

Rokhlin entropy theory was ultimately successful in providing a framework for generalizing the Sinai factor theorem to all countably infinite groups. Specifically, if $G \curvearrowright (X, \mu)$ is a free ergodic action of any countably infinite group G, then $G \curvearrowright$

(X, μ) factors onto the Bernoulli shift $G \curvearrowright (L^G, \lambda^G)$ whenever $h_G^{\mathrm{Rok}}(X, \mu) \geq \mathrm{H}(\lambda)$ [38]. In particular, actions which do not admit any finite generating partitions must factor onto all Bernoulli shifts. Also, since sofic entropy is bounded above by Rokhlin entropy, this implies that sofic entropy satisfies the Sinai factor theorem as well.

In [37] two fairly computable expressions were found which provide upper bounds to Rokhlin entropy. In particular, it was shown that if $G \curvearrowright (X, \mu)$ is a free p.m.p. action and α is a generating partition with $\mathrm{H}(\alpha) < \infty$, then the Rokhlin entropy satisfies

$$h_G^{\mathrm{Rok}}(X, \mu) \leq \inf_{T \subseteq G} \frac{1}{|T|} \cdot \mathrm{H}(\bigvee_{t \in T} t \cdot \alpha),$$

where the infimum is over all finite $T \subseteq G$. When G is amenable, the right-hand side coincides with Kolmogorov–Sinai entropy (it is common to use a sequence of Følner sets T, but this isn't necessary [29]). Thus, in some sense this upper bound should be no more difficult to compute in practice than Kolmogorov–Sinai entropy. We point out that this upper bound makes our finite generator theorem, Theorem 1.1, easier to apply.

In addition to Parts II and III [36, 3] and the papers [37, 38] mentioned above, additional study of Rokhlin entropy has been undertaken in [2, 7, 13]. The sub-additive property of Rokhlin entropy continues to serve as the foundation for all of these new results.

Outline. The proof of Theorem 1.1 is essentially self-contained as it relies on technical constructions carried out by hand. The proof makes significant use of the pseudo-group of the induced orbit-equivalence relation. We review basic properties of the pseudo-group in Section 3. In Section 4 we review and strengthen a construction of the author used in [34] for replacing countably infinite partitions with finite ones. Sections 3 and 4 thus reprove the main theorem of [34] which states that finite Rokhlin entropy implies the existence of a finite generating partition.

The real difficulty of the present work is constructing a generating partition while controlling its cardinality and distribution. The classical Rokhlin lemma and Shannon–McMillan–Breiman theorem were critical to this task in all prior proofs of Krieger's theorem. The important advantage we obtain by working with the pseudo-group is that we are able to develop a replacement to the Rokhlin lemma and the Shannon–McMillan–Breiman theorem which is suitable to our needs. We present this replacement in Section 5.

A new, significant difficulty is created from our use of the pseudo-group. Specifically, the notion of a "generating" partition cannot be expressed in the language of the pseudo-group. Ultimately, we must build a single, efficient partition which both codes information for certain transformations in the pseudo-group, and simultaneously codes information for the action of G. An obstacle in this simultaneous-coding problem is that there is no geometric relationship between our pseudo-group transformations and the G-action. This is the most challenging part of the proof. The coding machinery needed for this task is presented in Section 7. Then in Section 8 we prove the main theorem. Finally, in Section 9 we show that relative Rokhlin entropy and relative Kolmogorov–Sinai entropy coincide for free actions of amenable groups.

2. Preliminaries

Every probability space (X, μ) which we consider will be assumed to be standard. In particular, X will be a standard Borel space. For $\xi \subseteq \mathcal{B}(X)$, we let $\sigma\text{-alg}(\xi)$ denote the smallest sub-σ-algebra containing ξ (not to be confused with the notation $\sigma\text{-alg}_G(\xi)$ from the introduction). At times, we will consider the space of all Borel probability measures on X. Recall that the space of Borel probability measures on X has a natural standard Borel structure which is generated by the maps $\lambda \mapsto \lambda(A)$ for $A \subseteq X$ Borel [19, Theorem 17.24].

An action $G \curvearrowright (Y, \nu)$ is a *factor* of $G \curvearrowright (X, \mu)$ if there exists a measure-preserving G-equivariant map $\pi : (X, \mu) \to (Y, \nu)$. Every factor $\pi : (X, \mu) \to (Y, \nu)$ is uniquely associated (mod μ-null sets) to a G-invariant sub-σ-algebra $\mathcal{F}$ of X, and conversely every G-invariant sub-σ-algebra $\mathcal{F}$ of (X, μ) is uniquely associated (up to isomorphism) to a factor $\pi : (X, \mu) \to (Y, \nu)$ [16, Theorem 2.15]. If $\pi : (X, \mu) \to (Y, \nu)$ is a factor map, then there is an essentially unique Borel map associating each $y \in Y$ to a Borel probability measure μ_y on X such that $\mu = \int \mu_y \, d\nu(y)$ and $\mu_y(\pi^{-1}(y)) = 1$. We call this the *disintegration* of μ over ν. Note that for any Borel set $A \subseteq X$, the map $y \mapsto \mu_y(A)$ is Borel.

Let $G \curvearrowright (X, \mu)$ be a p.m.p. action, and let $\mathcal{F}$ be a G-invariant sub-σ-algebra. Let $\pi : (X, \mu) \to (Y, \nu)$ be the associated factor, and let $\mu = \int \mu_y \, d\nu(y)$ be the disintegration of μ over ν. For a countable Borel partition α of X, the *conditional Shannon entropy* of α relative to $\mathcal{F}$ is

$$\mathrm{H}(\alpha|\mathcal{F}) = \int_Y \sum_{A \in \alpha} -\mu_y(A) \cdot \log(\mu_y(A)) \, d\nu(y) = \int_Y \mathrm{H}_{\mu_y}(\alpha) \, d\nu(y).$$

If $\mathcal{F} = \{X, \varnothing\}$ is the trivial σ-algebra then $\mathrm{H}(\alpha|\mathcal{F}) = \mathrm{H}(\alpha)$. For a countable partition β of X we set $\mathrm{H}(\alpha|\beta) = \mathrm{H}(\alpha|\sigma\text{-alg}(\beta))$. We write $\alpha \geq \beta$ if α is finer than β. We will need the following standard properties of Shannon entropy (proofs can be found in [12]):

Lemma 2.1. *Let (X, μ) be a standard probability space, let α and β be countable Borel partitions of X, and let $\mathcal{F}$ and Σ be sub-σ-algebras. Then*

   (i) $\mathrm{H}(\alpha|\mathcal{F}) \leq \log|\alpha|$;

   (ii) *if $\alpha \geq \beta$ then $\mathrm{H}(\alpha|\mathcal{F}) \geq \mathrm{H}(\beta|\mathcal{F})$;*

   (iii) *if $\Sigma \subseteq \mathcal{F}$ then $\mathrm{H}(\alpha|\Sigma) \geq \mathrm{H}(\alpha|\mathcal{F})$;*

   (iv) $\mathrm{H}(\alpha \vee \beta) = \mathrm{H}(\beta) + \mathrm{H}(\alpha|\beta) \geq \mathrm{H}(\beta)$;

   (v) $\mathrm{H}(\alpha \vee \beta|\mathcal{F}) = \mathrm{H}(\beta|\mathcal{F}) + \mathrm{H}(\alpha|\sigma\text{-alg}(\beta) \vee \mathcal{F}) \leq \mathrm{H}(\beta|\mathcal{F}) + \mathrm{H}(\alpha)$;

   (vi) $\mathrm{H}(\alpha|\mathcal{F}) = \sup_\xi \mathrm{H}(\xi|\mathcal{F})$, *where the supremum is over all finite partitions ξ coarser than α;*

   (vii) *if $\mathrm{H}(\alpha) < \infty$ then $\mathrm{H}(\alpha|\mathcal{F}) = \inf_\xi \mathrm{H}(\alpha|\xi)$, where the infimum is over all finite partitions $\xi \subseteq \mathcal{F}$.*

Throughout this paper, whenever working with a probability space (X, μ) we will generally ignore sets of measure zero. In particular, we write $A = B$ for $A, B \subseteq X$ if their symmetric difference is null. We similarly write $\mathcal{F} = \Sigma$ for sub-σ-algebras $\mathcal{F}$, Σ if they agree up to null sets. Also, by a partition of X we will mean a countable collection of pairwise-disjoint Borel sets whose union is conull. In particular, we allow partitions to contain the empty set. Similarly, we will use the term *probability vector* more freely than described in the introduction. A probability vector $\bar{p} = (p_i)$ will be any finite or countable ordered tuple of non-negative real numbers which sum to 1 (so some terms p_i may be 0). We say that another probability vector $\bar{q}$ is *coarser* than $\bar{p}$ if there is a partition $\mathcal{Q} = \{Q_j \ : \ 0 \leq j < |\bar{q}|\}$ of the integers $\{0 \leq i < |\bar{p}|\}$ such that for every $0 \leq j < |\bar{q}|$

$$q_j = \sum_{i \in Q_j} p_i.$$

A *pre-partition* of X is a countable collection of pairwise-disjoint subsets of X. We say that another pre-partition β extends α, written $\beta \sqsupseteq \alpha$, if there is an injection $\iota : \alpha \to \beta$ with $A \subseteq \iota(A)$ for every $A \in \alpha$. Equivalently, $\beta \sqsupseteq \alpha$ if and only if $\cup \alpha \subseteq \cup \beta$ and the restriction of β to $\cup \alpha$ coincides with α.

For a Borel pre-partition α, we define the *reduced σ-algebra* $\sigma\text{-alg}_G^{\mathrm{red}}(\alpha)$ to be the collection of Borel sets $R \subseteq X$ such that there is a conull $X' \subseteq X$ satisfying:

> for every $r \in R \cap X'$ and $x \in X' \setminus R$ there is $g \in G$ with $g \cdot r, g \cdot x \in \cup \alpha$
> and with $g \cdot r$ and $g \cdot x$ lying in distinct classes of α.

It is a basic exercise to verify that $\sigma\text{-alg}_G^{\mathrm{red}}(\alpha)$ is indeed a σ-algebra. We note the following basic property.

Lemma 2.2. *Let $G \curvearrowright (X, \mu)$ be a p.m.p. action, and let α be a pre-partition. If β is a pre-partition and $\beta \sqsupseteq \alpha$ then $\sigma\text{-alg}_G^{red}(\beta) \supseteq \sigma\text{-alg}_G^{red}(\alpha)$. In particular, if β is a partition and $\beta \sqsupseteq \alpha$ then $\sigma\text{-alg}_G(\beta) \supseteq \sigma\text{-alg}_G^{red}(\alpha)$.*

Proof. Fix $R \in \sigma\text{-alg}_G^{\mathrm{red}}(\alpha)$. By definition of $\sigma\text{-alg}_G^{\mathrm{red}}(\alpha)$, there is a conull $X' \subseteq X$ such that for all $r \in R \cap X'$ and $x \in X' \setminus R$ there is $g \in G$ with $g \cdot r, g \cdot x \in \cup \alpha$ and such that α separates $g \cdot r$ and $g \cdot x$. Since the restriction of β to $\cup \alpha$ is equal to α, we also have that $g \cdot r, g \cdot x \in \cup \beta$ and β separates $g \cdot r$ and $g \cdot x$. We conclude that $R \in \sigma\text{-alg}_G^{\mathrm{red}}(\beta)$. $\square$

The definition of reduced σ-algebra may seem a bit odd at first, but comes about naturally from our work here and will significantly simplify some of the proofs in Part II and Part III [36, 3]. A key property of this definition is that if β is any partition extending α then one automatically has $\sigma\text{-alg}_G(\beta) \supseteq \sigma\text{-alg}_G^{\mathrm{red}}(\alpha)$. Another important property is that if $G \curvearrowright (Y, \nu)$ is a factor of (X, μ) via $\phi : (X, \mu) \to (Y, \nu)$, then for any pre-partition α of Y we have $\phi^{-1}(\sigma\text{-alg}_G^{\mathrm{red}}(\alpha)) \subseteq \sigma\text{-alg}_G^{\mathrm{red}}(\phi^{-1}(\alpha))$. These properties can be quite useful for specialized constructions. For example, one could imagine constructing two pre-partitions α^1 and α^2 which achieve different goals. If $\cup \alpha^1$ is disjoint from $\cup \alpha^2$, then one can choose a common extension partition α and automatically have $\sigma\text{-alg}_G(\alpha) \supseteq \sigma\text{-alg}_G^{\mathrm{red}}(\alpha^1) \vee \sigma\text{-alg}_G^{\mathrm{red}}(\alpha^2)$. This type of construction will be performed in Part II.

Below is the statement of the main theorem of this paper in its strongest form. It is a strengthening of Theorem 1.1 mentioned in the introduction. This theorem is new even in the case $G = \mathbb{Z}$.

Theorem 2.3. *Let G be a countably infinite group acting ergodically, but not necessarily freely, by measure-preserving bijections on a non-atomic standard probability space (X, μ). Let $\mathcal{F}$ be a G-invariant sub-σ-algebra of X. If $0 < r \leq 1$ and $\bar{p} = (p_i)$ is any finite or countable probability vector with $h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F}) < r \cdot \mathrm{H}(\bar{p})$, then there is a Borel pre-partition $\alpha = \{A_i \, : \, 0 \leq i < |\bar{p}|\}$ with $\mu(\cup \alpha) = r$, $\mu(A_i) = r \cdot p_i$ for every $0 \leq i < |\bar{p}|$, and $\sigma\text{-alg}_G^{red}(\alpha) \vee \mathcal{F} = \mathcal{B}(X)$.*

3. The pseudo-group of an ergodic action

For a p.m.p. action $G \curvearrowright (X, \mu)$ we let E_G^X denote the induced orbit equivalence relation:
$$E_G^X = \{(x, y) \, : \, \exists g \in G, \;\; g \cdot x = y\}.$$
The *pseudo-group* of E_G^X, denoted $[[E_G^X]]$, is the set of all Borel bijections $\theta : \mathrm{dom}(\theta) \to \mathrm{rng}(\theta)$ where $\mathrm{dom}(\theta), \mathrm{rng}(\theta) \subseteq X$ are Borel and $\theta(x) \in G \cdot x$ for every $x \in \mathrm{dom}(\theta)$. The *full group* of E_G^X, denoted $[E_G^X]$, is the set of all $\theta \in [[E_G^X]]$ with $\mathrm{dom}(\theta) = \mathrm{rng}(\theta) = X$ (i.e. conull in X).

For every $\theta \in [[E_G^X]]$ there is a Borel partition $\{Z_g^\theta \, : \, g \in G\}$ of $\mathrm{dom}(\theta)$ such that $\theta(x) = g \cdot x$ for every $x \in Z_g^\theta$. Thus, an important fact which we will use repeatedly is that every $\theta \in [[E_G^X]]$ is measure-preserving. We mention that the sets Z_g^θ are in general not uniquely determined from θ since the action of G might not be free. It will be necessary to keep record of such decompositions $\{Z_g^\theta\}$ for $\theta \in [[E_G^X]]$. The precise notion we need is the following.

Definition 3.1. Let $G \curvearrowright (X, \mu)$ be a p.m.p. action, let $\theta \in [[E_G^X]]$, and let $\mathcal{F}$ be a G-invariant sub-σ-algebra. We say that θ is $\mathcal{F}$-*expressible* if $\mathrm{dom}(\theta), \mathrm{rng}(\theta) \in \mathcal{F}$ and there is a $\mathcal{F}$-measurable partition $\{Z_g^\theta \, : \, g \in G\}$ of $\mathrm{dom}(\theta)$ such that $\theta(x) = g \cdot x$ for every $x \in Z_g^\theta$ and all $g \in G$.

We observe two simple facts on the notion of expressibility.

Lemma 3.2. *Let $G \curvearrowright (X, \mu)$ be a p.m.p. action and let $\mathcal{F}$ be a G-invariant sub-σ-algebra. If $\theta \in [[E_G^X]]$ is $\mathcal{F}$-expressible and $A \subseteq X$, then $\theta(A) = \theta(A \cap \mathrm{dom}(\theta))$ is $\sigma\text{-alg}_G(\{A\}) \vee \mathcal{F}$-measurable. In particular, if $A \in \mathcal{F}$ then $\theta(A) \in \mathcal{F}$.*

Proof. Fix a $\mathcal{F}$-measurable partition $\{Z_g^\theta \, : \, g \in G\}$ of $\mathrm{dom}(\theta)$ such that $\theta(x) = g \cdot x$ for all $x \in Z_g^\theta$. Then
$$\theta(A) = \bigcup_{g \in G} g \cdot (A \cap Z_g^\theta) \in \sigma\text{-alg}_G(\{A\}) \vee \mathcal{F}. \qquad \square$$

Lemma 3.3. *Let $G \curvearrowright (X, \mu)$ be a p.m.p. action and let $\mathcal{F}$ be a G-invariant sub-σ-algebra. If $\theta, \phi \in [[E_G^X]]$ are $\mathcal{F}$-expressible then so are θ^{-1} and $\theta \circ \phi$.*

Proof. Fix $\mathcal{F}$-measurable partitions $\{Z_g^\theta \, : \, g \in G\}$ and $\{Z_g^\phi \, : \, g \in G\}$ of $\mathrm{dom}(\theta)$ and $\mathrm{dom}(\phi)$, respectively, satisfying $\theta(x) = g \cdot x$ for all $x \in Z_g^\theta$ and $\phi(x) = g \cdot x$ for all $x \in Z_g^\phi$. Define for $g \in G$
$$Z_g^{\theta^{-1}} = g^{-1} \cdot Z_{g^{-1}}^\theta.$$
Then each $Z_g^{\theta^{-1}}$ is $\mathcal{F}$-measurable since $\mathcal{F}$ is G-invariant. It is easily checked that $\{Z_g^{\theta^{-1}} \, : \, g \in G\}$ partitions $\mathrm{rng}(\theta)$ and satisfies $\theta^{-1}(x) = g \cdot x$ for all $x \in Z_g^{\theta^{-1}}$. Thus θ^{-1} is $\mathcal{F}$-expressible.

Observe that by the previous lemma, $\phi^{-1}(Z_g^\theta) \in \mathcal{F}$ for every $g \in G$ since ϕ^{-1} is $\mathcal{F}$-expressible. Notice that the sets $Z_g^\phi \cap \phi^{-1}(Z_h^\theta)$ partition $\mathrm{dom}(\theta \circ \phi)$. Define for $g \in G$

$$Z_g^{\theta \circ \phi} = \bigcup_{h \in G} \left(Z_{h^{-1}g}^\phi \cap \phi^{-1}(Z_h^\theta) \right).$$

These sets are $\mathcal{F}$-measurable and pairwise-disjoint and we have $\theta \circ \phi(x) = g \cdot x$ for all $x \in Z_g^{\theta \circ \phi}$. $\qquad\square$

With the aid of Lemma 3.2, we observe a basic property of relative Rokhlin entropy. The proposition below resembles a theorem of Rudolph and Weiss from classical entropy theory [33]. Note that if G and Γ act on (X, μ) with the same orbits then $E_G^X = E_\Gamma^X$ and $[[E_G^X]] = [[E_\Gamma^X]]$. In this situation, we say that $\theta \in [[E_G^X]]$ is $(G, \mathcal{F})$-expressible if it is $\mathcal{F}$-expressible with respect to the G-action $G \curvearrowright (X, \mu)$.

Proposition 3.4. *Let G and Γ be countable groups, and let $G \curvearrowright (X, \mu)$ and $\Gamma \curvearrowright (X, \mu)$ be p.m.p. ergodic actions having the same orbits. Suppose that $\mathcal{F}$ is a G and Γ invariant sub-σ-algebra such that the transformation associated to each $g \in G$ is $(\Gamma, \mathcal{F})$-expressible and similarly the transformation associated to each $\gamma \in \Gamma$ is $(G, \mathcal{F})$-expressible. Then*

$$h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F}) = h_\Gamma^{\mathrm{Rok}}(X, \mu | \mathcal{F}).$$

Proof. It suffices to show that for every countable partition α, $\sigma\text{-alg}_G(\alpha) \vee \mathcal{F} = \sigma\text{-alg}_\Gamma(\alpha) \vee \mathcal{F}$. Indeed, since the transformation associated to each $g \in G$ is $(\Gamma, \mathcal{F})$-expressible, it follows from Lemma 3.2 that the σ-algebra $\sigma\text{-alg}_\Gamma(\alpha) \vee \mathcal{F}$ is G-invariant and contains α. Therefore $\sigma\text{-alg}_G(\alpha) \vee \mathcal{F} \subseteq \sigma\text{-alg}_\Gamma(\alpha) \vee \mathcal{F}$. With the same argument we obtain the reverse containment. $\qquad\square$

The lemma below and the corollaries which follow it provide us with all elements of the pseudo-group $[[E_G^X]]$ which will be needed in forthcoming sections.

Lemma 3.5. *Let $G \curvearrowright (X, \mu)$ be an ergodic p.m.p. action. Let $A, B \subseteq X$ be Borel sets with $0 < \mu(A) \leq \mu(B)$. Then there exists a $\sigma\text{-alg}_G(\{A, B\})$-expressible function $\theta \in [[E_G^X]]$ with $\mathrm{dom}(\theta) = A$ and $\mathrm{rng}(\theta) \subseteq B$.*

Proof. Let $g_0, g_1, \ldots$ be an enumeration of G. Set $Z_{g_0}^\theta = A \cap g_0^{-1} \cdot B$ and inductively define

$$Z_{g_n}^\theta = \left(A \setminus \left(\bigcup_{i=0}^{n-1} Z_{g_i}^\theta \right) \right) \bigcap g_n^{-1} \cdot \left(B \setminus \left(\bigcup_{i=0}^{n-1} g_i \cdot Z_{g_i}^\theta \right) \right).$$

Define $\theta : \bigcup_{n \in \mathbb{N}} Z_{g_n}^\theta \to B$ by setting $\theta(x) = g_n \cdot x$ for $x \in Z_{g_n}^\theta$. Clearly θ is $\sigma\text{-alg}_G(\{A, B\})$-expressible.

Set $C = A \setminus \mathrm{dom}(\theta)$. Towards a contradiction, suppose that $\mu(C) > 0$. Then we have

$$\mu(\mathrm{rng}(\theta)) = \mu(\mathrm{dom}(\theta)) < \mu(A) \leq \mu(B).$$

So $\mu(B \setminus \mathrm{rng}(\theta)) > 0$ and by ergodicity there is $n \in \mathbb{N}$ with

$$\mu\left(C \cap g_n^{-1} \cdot (B \setminus \mathrm{rng}(\theta)) \right) > 0.$$

However, this implies that $\mu(C \cap Z_{g_n}^\theta) > 0$, a contradiction. We conclude that, up to a null set, $\mathrm{dom}(\theta) = A$. $\qquad\square$

Corollary 3.6. *Let $G \curvearrowright (X, \mu)$ be a p.m.p. ergodic action. If $C \subseteq B \subseteq X$ and $\mu(C) = \frac{1}{n} \cdot \mu(B)$ with $n \in \mathbb{N}$, then there is a σ-$\mathrm{alg}_G(\{C, B\})$-measurable partition ξ of B into n pieces with each piece having measure $\frac{1}{n} \cdot \mu(B)$ and with $C \in \xi$.*

Proof. Set $C_1 = C$. Once σ-$\mathrm{alg}_G(\{C, B\})$-measurable subsets $C_1, \ldots, C_{k-1}$ of B, each of measure $\frac{1}{n} \cdot \mu(B)$, have been defined, we apply Lemma 3.5 to get a σ-$\mathrm{alg}_G(\{C, B\})$-expressible function $\theta \in [[E_G^X]]$ with $\mathrm{dom}(\theta) = C$ and

$$\mathrm{rng}(\theta) \subseteq B \setminus (C_1 \cup \cdots \cup C_{k-1}).$$

We set $C_k = \theta(C)$. We note that $\mu(C_k) = \frac{1}{n} \cdot \mu(B)$ and $C_k \in \sigma$-$\mathrm{alg}_G(\{C, B\})$ by Lemma 3.2. Finally, set $\xi = \{C_1, \ldots, C_n\}$. $\square$

In the corollary below we write $\mathrm{id}_A \in [[E_G^X]]$ for the identity function on A for $A \subseteq X$.

Corollary 3.7. *Let $G \curvearrowright (X, \mu)$ be an ergodic p.m.p. action. If $\xi = \{C_1, \ldots, C_n\}$ is a collection of pairwise disjoint Borel sets of equal measure, then there is a σ-$\mathrm{alg}_G(\xi)$-expressible function $\theta \in [[E_G^X]]$ which cyclically permutes the members of ξ, meaning that $\mathrm{dom}(\theta) = \mathrm{rng}(\theta) = \cup \xi$, $\theta(C_k) = C_{k+1}$ for $1 \le k < n$, $\theta(C_n) = C_1$, and $\theta^n = \mathrm{id}_{\cup \xi}$.*

Proof. By Lemma 3.5, for each $2 \le k \le n$ there is a σ-$\mathrm{alg}_G(\xi)$-expressible function $\phi_k \in [[E_G^X]]$ with $\mathrm{dom}(\phi_k) = C_1$ and $\mathrm{rng}(\phi_k) = C_k$. We define $\theta : \cup \xi \to \cup \xi$ by

$$\theta(x) = \begin{cases} \phi_2(x) & \text{if } x \in C_1 \\ \phi_{k+1} \circ \phi_k^{-1}(x) & \text{if } x \in C_k \text{ and } 1 < k < n \\ \phi_n^{-1}(x) & \text{if } x \in C_n. \end{cases}$$

Then θ cyclically permutes the members of ξ and has order n. Finally, each restriction $\theta \upharpoonright C_k$ is σ-$\mathrm{alg}_G(\xi)$-expressible by Lemma 3.3 and thus θ is σ-$\mathrm{alg}_G(\xi)$-expressible. $\square$

4. Countably infinite partitions

In this section, we show how to replace countably infinite partitions by finite ones. This will allow us to carry-out counting arguments in proving the main theorem. Our work in this section retraces and improves upon methods used by the author in [34]. We improve upon [34] in two ways. First, we work in a relative setting where a G-invariant sub-σ-algebra $\mathcal{F}$ is given, and second we show that one can control the Shannon entropy of the newly constructed partition.

For a finite set S we let $S^{<\omega}$ denote the set of all finite words with letters in S (the ω in the superscript denotes the first infinite ordinal). For $z \in S^{<\omega}$ we let $|z|$ denote the length of the word z. The lemma below is a relativized version of a similar result due to Krieger [25].

Lemma 4.1. *Let (X, μ) be a probability space, let $\mathcal{F}$ be a sub-σ-algebra, let (Y, ν) be the associated factor of (X, μ), and let $\mu = \int_Y \mu_y \, d\nu(y)$ be the disintegration of μ over ν. If ξ is a countable Borel partition of X with $\mathrm{H}(\xi|\mathcal{F}) < \infty$, then there is a Borel function $L : Y \times \xi \to \{0, 1, 2\}^{<\omega}$ which has finite average length*

$$\int_Y \sum_{C \in \xi} |L(y, C)| \cdot \mu_y(C) \, d\nu(y) < \infty$$

and such that ν-almost-every restriction $L(y, \cdot) : \xi \to \{0, 1, 2\}^{<\omega}$ is essentially injective in the sense that $L(y, C) = L(y, C')$ and $C \neq C'$ implies $\mu_y(C) \cdot \mu_y(C') = 0$.

Proof. If ξ is finite then we can simply fix an injection $L : \xi \to \{0, 1, 2\}^k$ for some $k \in \mathbb{N}$. So suppose that ξ is infinite. Say $\xi = \{C_1, C_2, \ldots\}$. For $y \in Y$ let $\sigma(y) : \mathbb{N} \to \mathbb{N}$ be the unique bijection satisfying for all $n \in \mathbb{N}$: either $\mu_y(C_{\sigma(y)(n+1)}) < \mu_y(C_{\sigma(y)(n)})$ or else $\mu_y(C_{\sigma(y)(n+1)}) = \mu_y(C_{\sigma(y)(n)})$ and $\sigma(y)(n+1) > \sigma(y)(n)$. Since each map $y \mapsto \mu_y(C_k)$ is Borel (see §2), we see that $\sigma : Y \to \mathbb{N}^{\mathbb{N}}$ is Borel.

For each n let $t(n) \in \{0, 1, 2\}^{<\omega}$ be the ternary expansion of n. Note that $|t(n)| \leq \log_3(n) + 1$. For $y \in Y$ define

$$L(y, C_{\sigma(y)(n)}) = t(n) \text{ for } n \in \mathbb{N} \quad \text{and} \quad L(y, C_k) = t(1) \text{ for } k \in \mathbb{N} \setminus \sigma(y)(\mathbb{N}).$$

If $|t(n)| = |L(y, C_{\sigma(y)(n)})| > -\log \mu_y(C_{\sigma(y)(n)})$ then for all $k \leq n$

$$\mu_y(C_{\sigma(y)(k)}) \geq \mu_y(C_{\sigma(y)(n)}) > e^{-|t(n)|} \geq \frac{1}{e} \cdot e^{-\log_3(n)} = \frac{1}{e} \cdot n^{-\log_3(e)}.$$

Thus

$$\frac{1}{e} \cdot n^{1-\log_3(e)} = n \cdot \frac{1}{e} \cdot n^{-\log_3(e)} < \sum_{k=1}^{n} \mu_y(C_{\sigma(y)(k)}) \leq 1,$$

and hence $n \leq \exp(1/(1 - \log_3(e)))$. Letting m be the least integer greater than $\exp(1/(1 - \log_3(e)))$, we have that $|L(y, C_{\sigma(y)(n)})| \leq -\log \mu_y(C_{\sigma(y)(n)})$ for all $y \in Y$ and all $n > m$. Therefore, recalling that $\mu_y(C_k) = 0$ for all $k \in \mathbb{N} \setminus \sigma(y)(\mathbb{N})$, we have

$$\sum_{n \in \mathbb{N}} |L(y, C_n)| \cdot \mu_y(C_n) = \sum_{n \in \mathbb{N}} |L(y, C_{\sigma(y)(n)})| \cdot \mu_y(C_{\sigma(y)(n)})$$

$$\leq m \cdot |t(m)| + \sum_{n > m} |L(y, C_{\sigma(y)(n)})| \cdot \mu_y(C_{\sigma(y)(n)})$$

$$\leq m \cdot |t(m)| + \sum_{n \in \mathbb{N}} -\mu_y(C_n) \log \mu_y(C_n)$$

$$= m \cdot |t(m)| + \mathrm{H}_{\mu_y}(\xi).$$

Integrating both sides over Y and using $\int_Y \mathrm{H}_{\mu_y}(\xi) \, d\nu(y) = \mathrm{H}(\xi | \mathcal{F}) < \infty$ completes the proof. $\square$

Proposition 4.2. *Let $G \curvearrowright (X, \mu)$ be an ergodic p.m.p. action, let $\mathcal{F}$ be a G-invariant sub-σ-algebra, and let ξ be a countable Borel partition with $\mathrm{H}(\xi | \mathcal{F}) < \infty$. Then for every $\epsilon > 0$ there is a finite Borel partition α with $\sigma\text{-alg}_G(\alpha) \vee \mathcal{F} = \sigma\text{-alg}_G(\xi) \vee \mathcal{F}$ and $\mathrm{H}(\alpha | \mathcal{F}) < \mathrm{H}(\xi | \mathcal{F}) + \epsilon$.*

Proof. Let $\pi : (X, \mu) \to (Y, \nu)$ be the factor map associated to $\mathcal{F}$, and let $\mu = \int \mu_y \, d\nu(y)$ be the disintegration of μ over ν. Apply Lemma 4.1 to obtain a Borel function $L : Y \times \xi \to \{0, 1, 2\}^{<\omega}$ such that ν-almost-every restriction $L(y, \cdot) : \xi \to \{0, 1, 2\}^{<\omega}$ is essentially injective and

$$\int_Y \sum_{C \in \xi} |L(y, C)| \cdot \mu_y(C) \, d\nu(y) < \infty.$$

We define $\ell : X \to \{0, 1, 2\}^{<\omega}$ by

$$\ell(x) = L(\pi(x), C)$$

for $x \in C \in \xi$. Observe that ℓ is $\sigma\text{-alg}(\xi) \vee \mathcal{F}$-measurable and

$$\int_X |\ell(x)| \, d\mu(x) = \int_Y \int_X |\ell(x)| \, d\mu_y(x) \, d\nu(y) = \int_Y \sum_{C \in \xi} |L(y, C)| \cdot \mu_y(C) \, d\nu(y) < \infty.$$

For $n \in \mathbb{N}$ let $\mathcal{P}_n = \{P_n, X \setminus P_n\}$ where

$$P_n = \{x \in X : |\ell(x)| \geq n\}.$$

Then the P_n's are decreasing and have empty intersection. Refine $\mathcal{P}_n$ to $\beta_n = \{X \setminus P_n, B_n^0, B_n^1, B_n^2\}$ where for $i \in \{0, 1, 2\}$

$$B_n^i = \{x \in P_n : \ell(x)(n) = i\}.$$

For $n \in \mathbb{N}$ define

$$\gamma_n = \bigvee_{k \leq n} \beta_k.$$

Since each restriction $L(y, \cdot) : \xi \to \{0, 1, 2\}^{<\omega}$ is essentially injective we have that

(4.1) $$\xi \subseteq \mathcal{F} \vee \bigvee_{n \in \mathbb{N}} \sigma\text{-alg}(\gamma_n).$$

Fix $0 < \delta < \min(1/4, \epsilon/2)$ with

$$-\delta \cdot \log(\delta) - (1 - \delta) \cdot \log(1 - \delta) + \delta \cdot \log(7) < \epsilon.$$

Since

$$\sum_{n \in \mathbb{N}} \mu(P_n) = \int_X |\ell(x)| \, d\mu(x) < \infty$$

we may fix $N \in \mathbb{N}$ so that $\sum_{n=N}^{\infty} \mu(P_n) < \delta$. Observe that in particular $\mu(P_N) < \delta$ and thus

$$\mu(P_N) + \sum_{n=N}^{\infty} \mu(P_n) < 2\delta < 1/2.$$

For $n \geq N$ we seek to build $\sigma\text{-alg}_G(\mathcal{P}_n \vee \gamma_{n-1})$-expressible functions $\theta_n \in [[E_G^X]]$ with $\text{dom}(\theta_n) = P_n$ and

$$\text{rng}(\theta_n) \subseteq X \setminus \left(P_N \cup \bigcup_{k=N}^{n-1} \theta_k(P_k) \right).$$

We build the θ_n's by induction on $n \geq N$. To begin we note that $\mu(P_N) < \mu(X \setminus P_N)$ and we apply Lemma 3.5 to obtain θ_N. Now assume that $\theta_N, \ldots, \theta_{n-1}$ have been defined and posses the properties stated above. Then since γ_{n-1} refines $\mathcal{P}_k \vee \gamma_{k-1}$ for every $k < n$, we obtain from Lemma 3.2

$$P_N \cup \bigcup_{k=N}^{n-1} \theta_k(P_k) \in \sigma\text{-alg}_G(\gamma_{n-1}).$$

Also, by our choice of N we have that

$$\mu(P_n) \leq \mu(P_N) < \frac{1}{2} < 1 - 2\delta < 1 - \mu(P_N) - \sum_{k=N}^{n-1} \mu(P_k)$$

$$= \mu \left(X \setminus \left(P_N \cup \bigcup_{k=N}^{n-1} \theta_k(P_k) \right) \right).$$

Therefore we may apply Lemma 3.5 to obtain θ_n. This defines the functions θ_n, $n \geq N$.

Define the partition $\beta = \{X \setminus P, B^0, B^1, B^2\}$ of X by

$$P = \bigcup_{n \geq N} \theta_n(P_n);$$

$$B^i = \bigcup_{n \geq N} \theta_n(B_n^i).$$

Note that the above expressions do indeed define a partition of X since the images of the θ_n's are pairwise disjoint. Also define $\mathcal{Q} = \{Q, X \setminus Q\}$ where

$$Q = \bigcup_{n \geq N} \theta_n(P_{n+1}).$$

Note that Q is contained in P and so β might not refine $\mathcal{Q}$. Set $\alpha = \gamma_N \vee \beta \vee \mathcal{Q}$. Then α is finite. Using Lemma 2.1 and the facts that $X \setminus P \in \beta \vee \mathcal{Q}$, $\mu(P) < \delta$, and $H_{\mu_y}(\gamma_N) \leq H_{\mu_y}(\xi)$ for ν-almost-every $y \in Y$ (since ξ μ_y-almost-everywhere refines γ_N), we obtain

$$\begin{aligned}
H(\alpha|\mathcal{F}) &\leq H(\gamma_N|\mathcal{F}) + H(\beta \vee \mathcal{Q}) \\
&= H(\gamma_N|\mathcal{F}) + H(\{P, X \setminus P\}) + H(\beta \vee \mathcal{Q}|\{P, X \setminus P\}) \\
&\leq H(\gamma_N|\mathcal{F}) - \mu(P) \cdot \log \mu(P) - \mu(X \setminus P) \log \mu(X \setminus P) + \mu(P) \cdot \log(7) \\
&< H(\gamma_N|\mathcal{F}) + \epsilon \\
&= \int_Y H_{\mu_y}(\gamma_N) \, d\nu(y) + \epsilon \\
&\leq \int_Y H_{\mu_y}(\xi) \, d\nu(y) + \epsilon \\
&= H(\xi|\mathcal{F}) + \epsilon.
\end{aligned}$$

Thus it only remains to check that $\sigma\text{-alg}_G(\alpha) \vee \mathcal{F} = \sigma\text{-alg}_G(\xi) \vee \mathcal{F}$.

First notice that the function ℓ and all of the partitions γ_n and $\mathcal{P}_n$ are $\sigma\text{-alg}_G(\xi) \vee \mathcal{F}$-measurable and therefore each θ_k is $\sigma\text{-alg}_G(\xi) \vee \mathcal{F}$-expressible. It follows from Lemma 3.2 that β, $\mathcal{Q}$, and α are $\sigma\text{-alg}_G(\xi) \vee \mathcal{F}$-measurable. Thus $\sigma\text{-alg}_G(\alpha) \vee \mathcal{F} \subseteq \sigma\text{-alg}_G(\xi) \vee \mathcal{F}$. Now we consider the reverse inclusion. By induction and by (4.1) it suffices to assume that $\gamma_k \subseteq \sigma\text{-alg}_G(\alpha)$ and prove that $\gamma_{k+1} \subseteq \sigma\text{-alg}_G(\alpha)$ as well. This is immediate when $k \leq N$. So assume that $k \geq N$ and that $\gamma_k \subseteq \sigma\text{-alg}_G(\alpha)$. Since θ_k is expressible with respect to $\sigma\text{-alg}_G(\gamma_k) \subseteq \sigma\text{-alg}_G(\alpha)$, we have that

$$P_{k+1} = \theta_k^{-1}(Q) \in \sigma\text{-alg}_G(\alpha)$$

by Lemmas 3.2 and 3.3. Therefore $\mathcal{P}_{k+1} \subseteq \sigma\text{-alg}_G(\alpha)$. Now since θ_{k+1} is expressible with respect to $\sigma\text{-alg}_G(\mathcal{P}_{k+1} \vee \gamma_k) \subseteq \sigma\text{-alg}_G(\alpha)$ we have that for $i \in \{0, 1, 2\}$

$$B_{k+1}^i = \theta_{k+1}^{-1}(B^i) \in \sigma\text{-alg}_G(\alpha)$$

by Lemmas 3.2 and 3.3. Thus $\beta_{k+1} \subseteq \sigma\text{-alg}_G(\alpha)$ and we conclude that $\gamma_{k+1} \subseteq \sigma\text{-alg}_G(\alpha)$. This completes the proof. $\square$

5. Finite subequivalence relations

The methods of the previous section produce finite generating partitions but do not provide any control over the cardinality or distribution of the partition constructed. Overcoming this difficulty is the main focus of this paper and requires entirely new techniques. We develop these techniques in this section and in Section

7. The goal of this section is to construct finite subequivalence relations which will ultimately be used to replace the traditional role of the Rokhlin lemma and the Shannon–McMillan–Breiman theorem.

For an equivalence relation E on X and $x \in X$, we write $[x]_E$ for the E-class of x. Recall that a set $T \subseteq X$ is a *transversal* for E if $|T \cap [x]_E| = 1$ for almost-every $x \in X$. We will work with equivalence relations which are generated by an element of the pseudo-group in the following sense.

Definition 5.1. Let $G \curvearrowright (X, \mu)$ be a p.m.p. action, let $B \subseteq X$ be a Borel set of positive measure, and let E be an equivalence relation on B with $E \subseteq E_G^X \cap B \times B$. We say that E is *generated by* $\theta \in [[E_G^X]]$ if $\mathrm{dom}(\theta) = \mathrm{rng}(\theta) = B$ and $[x]_E = \{\theta^i(x) : i \in \mathbb{Z}\}$ for almost-all $x \in B$. In this case, we write $E = E_\theta$.

Lemma 5.2. *Let $G \curvearrowright (X, \mu)$ be an ergodic p.m.p. action, let $B \subseteq X$ have positive measure, let α be a finite partition of X, and let $\epsilon > 0$. Then there is an equivalence relation E on B with $E \subseteq E_G^X \cap B \times B$ and $n \in \mathbb{N}$ so that for μ-almost-every $x \in B$, the E-class of x has cardinality n and*

$$\forall A \in \alpha \qquad \frac{\mu(A \cap B)}{\mu(B)} - \epsilon < \frac{|A \cap [x]_E|}{|[x]_E|} < \frac{\mu(A \cap B)}{\mu(B)} + \epsilon.$$

Moreover, E admits a σ-$\mathrm{alg}_G(\alpha \cup \{B\})$-measurable transversal and is generated by a σ-$\mathrm{alg}_G(\alpha \cup \{B\})$-expressible function $\theta : B \to B$ in $[[E_G^X]]$ which satisfies $\theta^n = \mathrm{id}_B$.

Proof. Let $\pi : (X, \mu) \to (Y, \nu)$ be the factor map associated to the G-invariant sub-σ-algebra generated by $\alpha \cup \{B\}$. Enumerate α as $\alpha = \{A_1, A_2, \ldots, A_p\}$. Set $B' = \pi(B)$ and $\alpha' = \{A_i' : 1 \leq i \leq p\}$ where $A_i' = \pi(A_i)$. Note that α' is a partition of (Y, ν) and that $\nu(A_i' \cap B') = \mu(A_i \cap B)$.

First, let's suppose that (Y, ν) is non-atomic. Pick $n \in \mathbb{N}$ satisfying $p/n < \epsilon$, and for $1 \leq i \leq p$ set

$$r_i = \lfloor n\nu(A_i' \cap B')/\nu(B') \rfloor.$$

Since (Y, ν) is non-atomic, we can find a partition ξ' of B' into n pieces each of measure $\frac{1}{n} \cdot \nu(B')$ such that for every i, $A_i' \cap B'$ contains at least r_i many classes of ξ'. Then at most p many classes of ξ' are not contained in any $A_i' \cap B'$. Set $\xi = \pi^{-1}(\xi')$. Then the classes of ξ lie in σ-$\mathrm{alg}_G(\alpha \cup \{B\})$. Apply Corollary 3.7 to get a σ-$\mathrm{alg}_G(\alpha \cup \{B\})$-expressible function $\theta \in [[E_G^X]]$ which cyclically permutes the classes of ξ. Set $E = E_\theta$. Then for μ-almost-every $x \in B$, the E-class of x has cardinality n and

$$\forall i \qquad -\epsilon \leq -\frac{1}{n} < \frac{|A_i \cap [x]_E|}{|[x]_E|} - \frac{\mu(A_i \cap B)}{\mu(B)} \leq \frac{p}{n} < \epsilon.$$

In the case that (Y, ν) has an atom, we deduce by ergodicity that, modulo a null set, Y is finite. Say $|Y| = m$ and each point in Y has measure $\frac{1}{m}$. Set $n = |B'|$. Clearly there are integers $k_i \in \mathbb{N}$, with $\sum_{i=1}^p k_i = n$ and

$$\frac{\mu(A_i \cap B)}{\mu(B)} = \frac{\nu(A_i' \cap B')}{\nu(B')} = \frac{k_i/m}{n/m} = \frac{k_i}{n}.$$

Let ξ' be the partition of B' into points, and pull back ξ' to a partition ξ of B. Now apply Corollary 3.7 and follow the argument from the non-atomic case. $\square$

Corollary 5.3. *Let $G \curvearrowright (X, \mu)$ be an ergodic p.m.p. action, let $B \subseteq X$ have positive measure, let $\epsilon > 0$, and let $F = \{f : B \to \mathbb{R}\}$ be a finite collection of*

finite valued Borel functions. Then there is an equivalence relation E on B with $E \subseteq E_G^X \cap B \times B$ and $n \in \mathbb{N}$ so that for μ-almost-every $x \in B$, the E-class of x has cardinality n and

$$\forall f \in F \qquad \frac{1}{\mu(B)} \cdot \int_B f \, d\mu - \epsilon < \frac{1}{|[x]_E|} \cdot \sum_{y \in [x]_E} f(y) < \frac{1}{\mu(B)} \cdot \int_B f \, d\mu + \epsilon.$$

Moreover, if each $f \in F$ is $\mathcal{F}$-measurable then E admits a σ-alg$_G(\mathcal{F} \cup \{B\})$-measurable transversal and is generated by a σ-alg$_G(\mathcal{F} \cup \{B\})$-expressible function $\theta : B \to B$ in $[[E_G^X]]$ which satisfies $\theta^n = \mathrm{id}_B$.

Proof. Define a partition α of B so that $x, y \in B$ lie in the same piece of α if and only if $f(x) = f(y)$ for all $f \in F$. Then α is a finite partition. Now the desired equivalence relation E is obtained from Lemma 5.2. $\square$

The conclusions of the previous lemma and corollary are not too surprising since you are allowed to "see" the sets which you wish to mix, i.e. you are allowed to use σ-alg$_G(\alpha \cup \{B\})$. The following proposition however is unexpected. It roughly says that you can achieve the same conclusion even if you are restricted to only seeing a very small sub-σ-algebra. We will use the proposition below in the same fashion one typically uses the Rokhlin lemma and the Shannon–McMillan–Breiman theorem, although technically the proposition below bears more similarity with the Rokhlin lemma and the ergodic theorem.

Let us say a few words on the Rokhlin lemma to highlight the similarity. For a free p.m.p. action $\mathbb{Z} \curvearrowright (X, \mu)$, $n \in \mathbb{N}$, and $\epsilon > 0$, the Rokhlin lemma provides a Borel set $S \subseteq X$ such that the sets $i \cdot S$, $0 \leq i \leq n - 1$, are pairwise disjoint and union to a set having measure at least $1 - \epsilon$. The set S naturally produces a subequivalence relation E defined as follows. For $x \in X$ set $x_S = (-i) \cdot x$ where $(-i) \cdot x \in S$ and $(-j) \cdot x \notin S$ for all $0 \leq j < i$. We set $x \, E \, y$ if and only if $x_S = y_S$. Clearly every E class has cardinality at least n, and a large measure of E-classes have cardinality precisely n. A key fact which is frequently used in classical results such as Krieger's theorem is that the equivalence relation E is easily described. Specifically, S is small since $\mu(S) \leq 1/n$, and so E can be defined by using the small sub-σ-algebra σ-alg$_{\mathbb{Z}}(\{S\})$.

Proposition 5.4. *Let $G \curvearrowright (X, \mu)$ be an ergodic p.m.p. action with (X, μ) non-atomic, let α be a finite collection of Borel subsets of X, let $\epsilon > 0$, and let $N \in \mathbb{N}$. Then there are $n \geq N$, Borel sets $S_1, S_2 \subseteq X$ with $\mu(S_1) + \mu(S_2) < \epsilon$, and a σ-alg$_G(\{S_1, S_2\})$-expressible $\theta \in [E_G^X]$ such that E_θ admits a σ-alg$_G(\{S_1, S_2\})$-measurable transversal, and for almost-every $x \in X$ we have $|[x]_{E_\theta}| = n$ and*

$$\forall A \in \alpha \qquad \mu(A) - \epsilon < \frac{|A \cap [x]_{E_\theta}|}{|[x]_{E_\theta}|} < \mu(A) + \epsilon.$$

Proof. Pick $m > \max(4/\epsilon, \ N)$ with $m \in \mathbb{N}$ and

$$|\alpha| \cdot \log_2(m + 1) < \frac{\epsilon}{4} \cdot m.$$

Let $S_1 \subseteq X$ be any Borel set with $\mu(S_1) = \frac{1}{m} < \frac{\epsilon}{4}$. Apply Corollaries 3.6 and 3.7 to obtain a σ-alg$_G(\{S_1\})$-expressible function $h \in [E_G^X]$ such that $\mathrm{dom}(h) = \mathrm{rng}(h) = X$, $h^m = \mathrm{id}_X$, and such that $\{h^i(S_1) : 0 \leq i < m\}$ is a partition of X. The induced Borel equivalence relation E_h is finite, in fact almost-every E_h-class

has cardinality m, and it has S_1 as a transversal. We imagine the classes of E_h as extending horizontally to the right, and we visualize S_1 as a vertical column.

We consider the distribution of $\alpha \upharpoonright [s]_{E_h}$ for each $s \in S_1$. For $A \in \alpha$ define $d_A : S_1 \to \mathbb{R}$ by

$$d_A(s) = \frac{|A \cap [s]_{E_h}|}{|[s]_{E_h}|} = \frac{1}{m} \cdot \left| A \cap [s]_{E_h} \right|.$$

Note that for each $A \in \alpha$

$$\int_{S_1} d_A \ d\mu = \frac{1}{m} \cdot \mu(A) = \mu(S_1) \cdot \mu(A).$$

By Corollary 5.3 there is $k \in \mathbb{N}$ and an equivalence relation $E_v \subseteq E_G^X \cap S_1 \times S_1$ on S_1 such that for almost every $s \in S_1$, the E_v-class of s has cardinality k and

$$\forall A \in \alpha \qquad \mu(A) - \epsilon < \frac{1}{|[s]_{E_v}|} \cdot \sum_{s' \in [s]_{E_v}} d_A(s') < \mu(A) + \epsilon.$$

Moreover, if we let $\mathcal{F}$ denote the G-invariant sub-σ-algebra generated by the functions d_A, $A \in \alpha$, then E_v admits a $\sigma\text{-alg}_G(\mathcal{F} \cup \{S_1\})$-measurable transversal T and is generated by a $\sigma\text{-alg}_G(\mathcal{F} \cup \{S_1\})$-expressible function $v \in [[E_G^X]]$ which satisfies $\mathrm{dom}(v) = \mathrm{rng}(v) = S_1$ and $v^k = \mathrm{id}_{S_1}$.

Let $E = E_v \vee E_h$ be the equivalence relation generated by E_v and E_h. Then $T \subseteq S_1$ is a transversal for E, and for every $s \in T$

$$\left| [s]_E \right| = \sum_{s' \in [s]_{E_v}} \left| [s']_{E_h} \right| = k \cdot m.$$

Setting $n = k \cdot m \geq N$, we have that almost every E-class has cardinality n. Also, for every $A \in \alpha$ and $s \in T$ we have

$$\frac{|A \cap [s]_E|}{|[s]_E|} = \frac{1}{k \cdot m} \cdot \sum_{s' \in [s]_{E_v}} \left| A \cap [s']_{E_h} \right| = \frac{1}{|[s]_{E_v}|} \cdot \sum_{s' \in [s]_{E_v}} d_A(s').$$

It follows that for μ-almost-every $x \in X$

$$\forall A \in \alpha \qquad \mu(A) - \epsilon < \frac{|A \cap [x]_E|}{|[x]_E|} < \mu(A) + \epsilon.$$

Now consider the partition $\xi = \{T_{i,j} : 0 \leq i < k, \ 0 \leq j < m\}$ of X where

$$T_{i,j} = h^j \circ v^i(T).$$

Note that $T_{i,j} \in \sigma\text{-alg}_G(\mathcal{F} \cup \{S_1\})$ by Lemmas 3.2 and 3.3. We will define a function $\theta \in [E_G^X]$ which generates E by defining θ on each piece of ξ. We define

$$\theta \upharpoonright T_{i,j} = \begin{cases} h \upharpoonright T_{i,j} & \text{if } j + 1 < m \\ v \circ h \upharpoonright T_{i,j} & \text{if } j + 1 = m. \end{cases}$$

In regard to the second case above, one should observe that $h(T_{i,m-1}) = T_{i,0}$ since $h^m = \mathrm{id}_X$. Since v satisfies $v^k = \mathrm{id}_{S_1}$ and $n = k \cdot m$, we see that θ satisfies $\theta^n = \mathrm{id}_X$. We also have $E = E_\theta$. Finally, θ is $\sigma\text{-alg}_G(\mathcal{F} \cup \{S_1\})$-expressible since each restriction $\theta \upharpoonright T_{i,j}$ is $\sigma\text{-alg}_G(\mathcal{F} \cup \{S_1\})$-expressible by Lemma 3.3.

To complete the proof, we must find a Borel set $S_2 \subseteq X$ with $\mu(S_2) < \frac{3}{4} \cdot \epsilon <$ $\epsilon - \mu(S_1)$ such that $\mathcal{F} \subseteq \sigma\text{-alg}_G(\{S_1, S_2\})$. Notice that $|\text{rng}(d_A)| \leq m+1$ for every $A \in \alpha$ and therefore the product map

$$d_\alpha = \prod_{A \in \alpha} d_A : S_1 \to \left\{ 0, \frac{1}{m}, \frac{2}{m}, \ldots, 1 \right\}^\alpha$$

has an image of cardinality at most $(m+1)^{|\alpha|}$. Set $\ell = \lceil (\epsilon/4) \cdot m \rceil$ (i.e. the least integer greater than or equal to $(\epsilon/4) \cdot m$). Since $(\epsilon/4) \cdot m > 1$ we have that $\ell < (\epsilon/2) \cdot m$. By our choice of m we have

$$(m+1)^{|\alpha|} < 2^{(\epsilon/4) \cdot m} \leq 2^\ell.$$

Therefore there is an injection

$$r : \{0, 1/m, \ldots, 1\}^\alpha \to \{0,1\}^\ell.$$

Now we will define S_2 so that, for every $s \in S_1$, the integers $\{1 \leq i \leq \ell : h^i(s) \in S_2\}$ will encode the value $r \circ d_\alpha(s)$. Specifically, we define

$$S_2 = \{h^i(s) : 1 \leq i \leq \ell, \ s \in S_1, \ r(d_\alpha(s))(i) = 1\}.$$

We have that $S_2 \subseteq \bigcup_{1 \leq i \leq \ell} h^i(S_1)$ and therefore

$$\mu(S_2) \leq \ell \cdot \mu(S_1) < \left(\frac{\epsilon}{2} \cdot m \right) \cdot \frac{1}{m} = \frac{\epsilon}{2}$$

as required. Finally, we check that $\mathcal{F} \subseteq \sigma\text{-alg}_G(\{S_1, S_2\})$. Fix $p \in \{0, 1/m, \ldots, 1\}^\alpha$. Set

$$I_p^0 = \{1 \leq i \leq \ell : r(p)(i) = 0\} \quad \text{and} \quad I_p^1 = \{1 \leq i \leq \ell : r(p)(i) = 1\}.$$

Then for $s \in S_1$ we have

$$d_\alpha(s) = p \Longleftrightarrow r(d_\alpha(s)) = r(p)$$

$$\Longleftrightarrow (\forall i \in I_p^0) \ h^i(s) \notin S_2 \quad \text{and} \quad (\forall i \in I_p^1) \ h^i(s) \in S_2$$

$$\Longleftrightarrow s \in S_1 \cap \left(\bigcap_{i \in I_p^0} h^{-i}(X \setminus S_2) \right) \cap \left(\bigcap_{i \in I_p^1} h^{-i}(S_2) \right).$$

So $d_\alpha^{-1}(p) \in \sigma\text{-alg}_G(\{S_1, S_2\})$ by Lemmas 3.2 and 3.3. Thus $\mathcal{F} \subseteq \sigma\text{-alg}_G(\{S_1, S_2\})$. $\qquad\square$

6. Distributions on finite sets

In this section we present a few counting lemmas from information theory which we will need. These facts are well known and were used in classical proofs of Krieger's finite generator theorem. At the end of this section we will briefly sketch why replacing the Rokhlin lemma and the Shannon–McMillan–Breiman theorem in the classical proof of Krieger's theorem with Proposition 5.4 does not (yet) result in a proof of our main theorem. This will illustrate what new techniques are required and will motivate the technical constructions in the next section.

For a finite probability vector $\bar{p}$, $n \in \mathbb{N}$, and $\epsilon \geq 0$, we let $L_{\bar{p}, \epsilon}^n$ be the set of functions $\ell : \{0, \ldots, n-1\} \to \{0, \ldots, |\bar{p}| - 1\}$ which approximate the distribution of $\bar{p}$ in the sense that

$$\forall 0 \leq t < |\bar{p}| \qquad \left| \frac{|\ell^{-1}(t)|}{n} - p_t \right| \leq \epsilon.$$

Similarly, if (X, μ) is a probability space and ξ is a finite partition of X, then we let $L_{\xi,\epsilon}^n$ be the set of functions $\ell : \{0, 1, \ldots, n-1\} \to \xi$ such that

$$\forall C \in \xi \qquad \left| \frac{|\ell^{-1}(C)|}{n} - \mu(C) \right| \leq \epsilon.$$

If ξ is a finite partition of (X, μ) and $\theta : (X, \mu) \to (X, \mu)$ is a measure-preserving bijection with every θ-orbit having cardinality n, then we associate to each $x \in X$ its (ξ, θ)-*name* $\mathcal{N}_\xi^\theta(x) \in L_{\xi,\infty}^n$ defined by setting $\mathcal{N}_\xi^\theta(x)(i) = C$ if $\theta^i(x) \in C \in \xi$.

If ξ and ζ are finite partitions of (X, μ) and ξ is finer than ζ, then we define the *coarsening map* $\pi_\zeta : \xi \to \zeta$ to be the unique map satisfying $C \subseteq \pi_\zeta(C)$ for all $C \in \xi$. By applying π_ζ coordinate-wise, we obtain a map $\pi_\zeta : L_{\xi,\infty}^n \to L_{\zeta,\infty}^n$.

Lemma 6.1. *Let (X, μ) be a probability space and let ξ and ζ be finite partitions of X. Suppose that ξ refines ζ and let $\pi_\zeta : \xi \to \zeta$ be the coarsening map. Then for every $\kappa > 0$ there is $\epsilon_0 > 0$ so that for all $0 < \epsilon < \epsilon_0$, all sufficiently large n, and every $z \in L_{\zeta,\epsilon}^n$*

$$\exp\left(n \cdot \mathrm{H}(\xi|\zeta) - n \cdot \kappa \right) \leq \left| \left\{ c \in L_{\xi,\epsilon}^n \ : \ \pi_\zeta(c) = z \right\} \right| \leq \exp\left(n \cdot \mathrm{H}(\xi|\zeta) + n \cdot \kappa \right).$$

Proof. This is a well known fact from information theory which can be quickly deduced from Stirling's formula. See [9, Lemma 2.13]. $\qquad \square$

By taking ζ to be the trivial partition in the previous lemma, we obtain the following.

Corollary 6.2. *Let $\bar{p}$ be a finite probability vector. Then for every $\kappa > 0$ there is $\epsilon_0 > 0$ so that for all $0 < \epsilon < \epsilon_0$ and all sufficiently large n*

$$\exp\left(n \cdot \mathrm{H}(\bar{p}) - n \cdot \kappa \right) \leq \left| L_{\bar{p},\epsilon}^n \right| \leq \exp\left(n \cdot \mathrm{H}(\bar{p}) + n \cdot \kappa \right).$$

For $x \in \mathbb{R}$ we write $\lfloor x \rfloor$ and $\lceil x \rceil$ for the greatest integer less than or equal to x and the least integer greater than or equal to x, respectively.

Corollary 6.3. *Fix $0 < \delta < 1$. Then for every $\kappa > 0$ and for all sufficiently large n we have*

$$\binom{n}{\lfloor \delta \cdot n \rfloor} \leq \exp\left(n \cdot \mathrm{H}(\delta, 1 - \delta) + n \cdot \kappa \right).$$

Proof. Set $\bar{p} = (1 - \delta, \delta)$. By definition $\binom{n}{\lfloor \delta \cdot n \rfloor}$ is the number of subsets of $\{0, \ldots, n-1\}$ having cardinality $\lfloor \delta \cdot n \rfloor$. Such subsets naturally correspond, via their characteristic functions, to elements of $L_{\bar{p},\epsilon}^n$ when $n > 1/\epsilon$. Thus when $n > 1/\epsilon$ we have $\binom{n}{\lfloor \delta \cdot n \rfloor} \leq |L_{\bar{p},\epsilon}^n|$. Now apply Corollary 6.2. $\qquad \square$

The normalized Hamming metric d_{Ham} on the set $L_{\bar{p},\infty}^n$ is defined by

$$d_{\mathrm{Ham}}(\ell, \ell') = \frac{1}{n} \cdot |\{i \ : \ \ell(i) \neq \ell'(i)\}|.$$

Corollary 6.4. *Let $\bar{p}$ be a finite probability vector. Then for every $\kappa > 0$ there are $\delta, \epsilon_0 > 0$ so that for all $0 < \epsilon < \epsilon_0$ and all sufficiently large n there exists $K \subseteq L_{\bar{p},\epsilon}^n$ satisfying $d_{\mathrm{Ham}}(k, k') > 2\delta$ for all $k \neq k' \in K$ and $|K| \geq \exp(n \cdot \mathrm{H}(\bar{p}) - n \cdot \kappa)$.*

Proof. Fix $\delta, \epsilon_0 > 0$ so that

$$\mathrm{H}(2\delta, 1 - 2\delta) + 2\delta \cdot \log |\bar{p}| < \kappa/3$$

and $|L_{\bar{p},\epsilon}^n| \geq \exp(n \cdot \mathrm{H}(\bar{p}) - n \cdot \kappa/3)$ for all $0 < \epsilon < \epsilon_0$ and for sufficiently large n. For $V \subseteq L_{\bar{p},\infty}^n$ let

$$B(V; \rho) = \{\ell \in L_{\bar{p},\infty}^n \; : \; \exists v \in V \;\; d_{\mathrm{Ham}}(\ell, v) \leq \rho\}$$

be the ball about V of radius ρ. Basic combinatorics implies that for $0 < \rho < 1$

$$\left| B(V; \rho) \right| \leq |V| \cdot \binom{n}{\lfloor \rho \cdot n \rfloor} \cdot |\bar{p}|^{\rho \cdot n}.$$

Fix $0 < \epsilon < \epsilon_0$. For each n let $K_n \subseteq L_{\bar{p},\epsilon}^n$ be maximal with the property that $d(k, k') > 2\delta$ for all $k \neq k' \in K_n$. Then by maximality of K_n we have $L_{\bar{p},\epsilon}^n \subseteq B(K_n; 2\delta)$ and thus

$$|L_{\bar{p},\epsilon}^n| \leq |B(K_n; 2\delta)| \leq |K_n| \cdot \binom{n}{\lfloor 2\delta \cdot n \rfloor} \cdot |\bar{p}|^{2\delta \cdot n}.$$

Solving for $|K_n|$ and letting n be sufficiently large gives

$$\begin{aligned}
|K_n| &\geq |L_{\bar{p},\epsilon}^n| \cdot \binom{n}{\lfloor 2\delta \cdot n \rfloor}^{-1} \cdot |\bar{p}|^{-2\delta \cdot n} \\
&\geq \exp(n \cdot \mathrm{H}(\bar{p}) - n \cdot \kappa/3 - n \cdot \mathrm{H}(2\delta, 1 - 2\delta) - n \cdot \kappa/3 - n \cdot 2\delta \cdot \log |\bar{p}|) \\
&\geq \exp(n \cdot \mathrm{H}(\bar{p}) - n\kappa). \qquad \qquad \qquad \qquad \qquad \qquad \qquad \square
\end{aligned}$$

We present one more technical lemma we will need.

Lemma 6.5. *Let $\bar{p}$ be a finite probability vector, let $\epsilon, \delta > 0$, and let $n \in \mathbb{N}$. If $\ell \in L_{\bar{p},\epsilon}^n$ then there is $J \subseteq \{0, \ldots, n-1\}$ such that $|J| \leq (\epsilon n + 1) \cdot |\bar{p}| + \delta n$ and*

$$\forall 0 \leq t < |\bar{p}| \qquad \frac{1}{n} \cdot \left| \{i : \ell(i) = t\} \setminus J \right| \leq (1 - \delta)p_t.$$

Proof. For each t we have $|\{i : \ell(i) = t\}| \leq np_t + n\epsilon$. So we may choose J so that for every t

$$|J \cap \{i : \ell(i) = t\}| = \max(0, \; \lceil |\{i : \ell(i) = t\}| - (1 - \delta)np_t \rceil) \leq \lceil \epsilon n + \delta np_t \rceil.$$

Then J will have the desired property and

$$|J| \leq \sum_{t=0}^{|\bar{p}|-1} \lceil \epsilon n + \delta np_t \rceil \leq \epsilon n \cdot |\bar{p}| + \delta n + |\bar{p}| = (\epsilon n + 1)|\bar{p}| + \delta n. \qquad \square$$

Before closing this section we briefly clarify to the reader what new methods are required in order to prove the main theorem. Let us consider the simplest setting where partitions and probability vectors are finite, $\mathcal{F} = \{X, \varnothing\}$ is trivial, and $r = 1$. The argument we present below is simply a sketch intended to give some intuition and motivation.

Let $G \curvearrowright (X, \mu)$ be a p.m.p. ergodic action with μ non-atomic, let ξ be a finite generating partition, and let $\bar{p}$ be a finite probability vector with $\mathrm{H}(\xi) < \mathrm{H}(\bar{p})$. We would like to construct a generating partition $\alpha = \{A_i : 0 \leq i < |\bar{p}|\}$ with $\mu(A_i) = p_i$ for all i. Pick $n_0 \in \mathbb{N}$ and $\epsilon > 0$. By Proposition 5.4 there are $n \geq n_0$, Borel sets $S_1, S_2 \subseteq X$ with $\mu(S_1) + \mu(S_2) < \epsilon$, and a $\sigma\text{-alg}_G(\{S_1, S_2\})$-expressible

$\theta \in [E_G^X]$ such that E_θ admits a σ-alg$_G(\{S_1, S_2\})$-measurable transversal Y and such that for μ-almost-every $x \in X$, the E_θ class of x has cardinality n, and

$$\forall C \in \xi \qquad \mu(C) - \epsilon < \frac{|C \cap [x]_{E_\theta}|}{|[x]_{E_\theta}|} < \mu(C) + \epsilon.$$

So we have $\mathcal{N}_\xi^\theta(y) \in L_{\xi,\epsilon}^n$ for almost-every $y \in Y$.

Since $|L_{\xi,\epsilon}^n| \approx \exp(n \cdot \mathrm{H}(\xi)) < \exp(n \cdot \mathrm{H}(\bar{p})) \approx |L_{\bar{p},\epsilon}^n|$, by beginning with a sufficiently small ϵ and sufficiently large n_0, we can conclude from Corollary 6.2 that

$$|L_{\xi,\epsilon}^n| < |L_{\bar{p},\epsilon}^n|.$$

So there is an injection $f : L_{\xi,\epsilon}^n \to L_{\bar{p},\epsilon}^n$. For $y \in Y$ set $c_y = \mathcal{N}_\xi^\theta(y)$ and $a_y = f(c_y)$. Since $\{\theta^i(Y) : 0 \le i < n\}$ is a partition of X, there is a unique partition $\alpha = \{A_i : 0 \le i < |\bar{p}|\}$ of X satisfying $\mathcal{N}_\alpha^\theta(y) = a_y$ for all $y \in Y$. Specifically, $x \in A_t$ if and only if $a_y(i) = t$ where $y \in Y$ and $0 \le i < n$ satisfy $\theta^i(y) = x$. The condition $\mathcal{N}_\alpha^\theta(y) = a_y \in L_{\bar{p},\epsilon}^n$ implies that $|\mu(A_i) - p_i| \le \epsilon$ for all i. Since f is an injection, $c_y = f^{-1}(a_y)$ is determined from a_y. It can be deduced from this fact that

(6.1) $\xi \subseteq \sigma\text{-alg}_{\langle\theta\rangle}(\alpha \vee \{Y, X \setminus Y\}).$

Now it becomes clear what is missing in order to prove our main theorem. We do want $\mu(A_i) = p_i$ instead of $|\mu(A_i) - p_i| \le \epsilon$, but this is only a minor problem. The major problem is that instead of (6.1) we want $\xi \subseteq \sigma\text{-alg}_G(\alpha)$, so that α is a generating partition. For this, it would suffice to have both (6.1) hold and $S_1, S_2 \in \sigma\text{-alg}_G(\alpha)$. In this case, Y would be measurable and θ would be expressible with respect to $\sigma\text{-alg}_G(\{S_1, S_2\}) \subseteq \sigma\text{-alg}_G(\alpha)$ and therefore

$$\xi \subseteq \sigma\text{-alg}_{\langle\theta\rangle}(\alpha \vee \{Y, X \setminus Y\}) \subseteq \sigma\text{-alg}_G(\alpha).$$

So we want to have both (6.1) hold and $S_1, S_2 \in \sigma\text{-alg}_G(\alpha)$ simultaneously. This first requirement on α uses θ-translates and the second uses G-translates. The difficulty is that we must build an α which simultaneously encodes messages under the θ-action and encodes messages under the G-action, and these two actions are almost completely unrelated. We solve this problem in the next section.

7. Coding small sets

The goal of this section is to construct, for any two sets S_1, S_2 having small measure, a pre-partition $\beta = \{B_0, B_1\}$ with the property that $\mu(\cup\beta)$ is small and $S_1, S_2 \in \sigma\text{-alg}_G^{\mathrm{red}}(\beta)$. This task is vital to the proof of the main theorem, as it will connect the G-action and the transformation θ used in Proposition 5.4.

We first focus our attention on building a pre-partition β and a set R, $0 < \mu(R) < 1$, with $R \in \sigma\text{-alg}_G^{\mathrm{red}}(\beta)$. The pre-partition β will consist of two (disjoint) sets B_0, B_1. On an intuitive level, it is likely helpful to imagine points in B_0 as "labeled with 0," points in B_1 as "labeled with 1", and points in $X \setminus (B_0 \cup B_1)$ as "unlabeled." The condition $R \in \sigma\text{-alg}_G^{\mathrm{red}}(\beta)$ then roughly means that no matter how the unlabeled points are later labeled, for every $x \in X$ it is possible to determine from the labeling of its orbit whether $x \in R$. This condition is similar to the notions of locally recognizable functions, membership tests, and recognizable sets appearing in [14, 15, 39]. It is the similarity with recognizable sets which led us to use the letter R.

A naive but suggestive idea for building R, B_0, B_1 is to fix a finite window $W \subseteq G$ and a finite set F containing W^2, and label all points in $W \cdot R$ with 1 (i.e. set
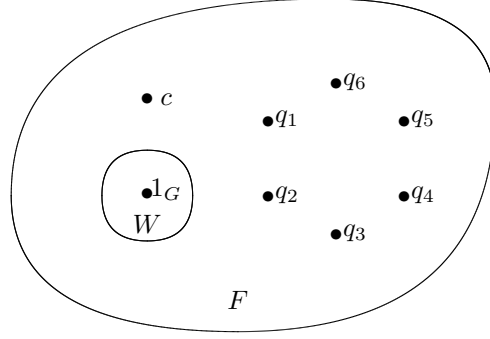
FIGURE 1. The structure in G which will be needed in constructing R and $\beta = \{B_0, B_1\}$.

$B_1 = W \cdot R$), label all points in $(F \setminus W) \cdot R$ with 0, and label more points 0 as needed so that for every $x \notin R$ there is a point in $W \cdot x$ labeled 0. It seems plausible then that $x \in R$ if and only if $W \cdot x$ is labeled identically 1. If so, $R \in \sigma\text{-alg}_G^{\mathrm{red}}(\beta)$ as desired. This naive approach is the right idea but does not quite work. For example, this may fail if W has too much symmetry, such as if W is a finite subgroup, for one might not be able to distinguish R from $W \cdot R$. This problem can be easily fixed by wisely choosing a "checkpoint" $c \in F \setminus W$ and labeling $c \cdot R$ with 1 in order to break any potential symmetries. In the case of free actions nothing more is required, but for non-free actions additional problems emerge which are a bit tedious to handle. There are a few interacting problems in the case of a non-free action, but in brief the primary problem is that we may have $W \cdot x = \{x\}$ for some $x = c \cdot r$ with $r \in R$. We will overcome this problem by introducing two new points $q_1, q_2 \in F \setminus W$. In fact, we will construct a set $Q = \{q_1, \ldots, q_6\} \subseteq F \setminus W$ of "query points" whose labels will hold important information. However $q_3, \ldots, q_6$ will not be needed until the second half of this section. See Figure 1 for an illustration.

The following lemma is well known.

Lemma 7.1. *Let $G \curvearrowright (X, \mu)$ be a p.m.p. action. If $Y \subseteq X$ is Borel and $F \subseteq G$ is finite, then there exists a Borel set $D \subseteq Y$ such that $Y \subseteq F^{-1}F \cdot D$ and $F \cdot d \cap F \cdot d' = \varnothing$ for all $d \neq d' \in D$. In particular, if $\mu(Y) > 0$ then $\mu(D) > 0$.*

Proof. Define the Borel graph $\Gamma \subseteq Y \times Y$ by $(y, y') \in \Gamma$ if and only if $y \neq y'$ and $F \cdot y \cap F \cdot y' \neq \varnothing$. Since every vertex has finite degree in Γ, a result of Kechris–Solecki–Todorcevic [20, Prop 4.2 and Prop. 4.5] states that there is a maximal (with respect to containment) Borel set D which is Γ-independent (i.e. no two elements of D are adjacent). Since D is Γ-independent we have $F \cdot d \cap F \cdot d' = \varnothing$ for all $d \neq d'$, and since D is maximal we have $Y \subseteq F^{-1}F \cdot D$. $\square$

In the outline above we mentioned labeling points 0 as needed so that for most $x \in X$ there is a point in $W \cdot x$ labeled 0. The lemma below will help us determine where to place these 0's.

Lemma 7.2. *Let $G \curvearrowright (X, \mu)$ be a p.m.p. ergodic action with (X, μ) non-atomic, and let $\delta > 0$. Then there exists a finite symmetric set $W \subseteq G$ with $1_G \in W$ and a Borel set $D \subseteq X$ such that $0 < \mu(D) < \delta$ and $W \cdot x \cap D \neq \varnothing$ for all $x \in X$.*

Proof. Since (X, μ) is non-atomic, we can fix a Borel set $D_1 \subseteq X$ with $0 < \mu(D_1) < \delta/2$. As $G \curvearrowright (X, \mu)$ is ergodic, we have $\mu(G \cdot D_1) = 1$. So there must be a sufficiently large finite symmetric set $W \subseteq G$ with $1_G \in W$ satisfying $\mu(W \cdot D_1) > 1 - \delta/2$. Now set $D = D_1 \cup (X \setminus W \cdot D_1)$. Then $\mu(D) < \delta$ and $W \cdot D = X$. $\square$

We will need the following fact from group theory.

Lemma 7.3 (B.H. Neumann, [30]). *Let G be a group, and let H_i, $1 \le i \le n$, be subgroups of G. Suppose there are group elements $g_i \in G$ so that*

$$G = \bigcup_{i=1}^{n} g_i \cdot H_i.$$

Then there is i such that $|G : H_i| < \infty$.

The corollary below will allow us to construct the checkpoint c.

Corollary 7.4. *Let $G \curvearrowright (X, \mu)$ be a p.m.p. ergodic action with (X, μ) non-atomic. Let $R \subseteq X$ have positive measure and let $W, T \subseteq G$ be finite. Then there are a Borel set $R' \subseteq R$ with $\mu(R') > 0$ and $c \in G$ such that $cW \cdot R' \cap T \cdot R' = \varnothing$.*

Proof. Our assumptions imply that almost-every orbit is infinite. So for μ-almost-every $r \in R$ the stability group $\mathrm{Stab}(r) = \{g \in G : g \cdot r = r\}$ has infinite index in G and thus by Lemma 7.3

$$T \cdot \mathrm{Stab}(r) \cdot W^{-1} = \bigcup_{t \in T} \bigcup_{w \in W} tw^{-1} \cdot (w\mathrm{Stab}(r)w^{-1}) \ne G.$$

As G is countable, there is $c \in G$ and a non-null Borel set $R_0 \subseteq R$ with

$$c \notin T \cdot \mathrm{Stab}(r) \cdot W^{-1}$$

for all $r \in R_0$. It follows that $cW \cdot r \cap T \cdot r = \varnothing$ for all $r \in R_0$. Now apply Lemma 7.1 to get positive measure Borel set $R' \subseteq R_0$ with $(cW \cup T) \cdot r \cap (cW \cup T) \cdot r' = \varnothing$ for all $r \ne r' \in R'$. $\square$

The next lemma will later be applied six times in order to build $Q = \{q_1, \ldots, q_6\}$.

Lemma 7.5. *Let $G \curvearrowright (X, \mu)$ be a p.m.p. ergodic action with (X, μ) non-atomic. Let $R, Y \subseteq X$ be positive measure Borel sets and let $T \subseteq G$ be finite. Then there are $q \in G$ and a Borel set $R' \subseteq R$ of positive measure such that $q \cdot R' \subseteq Y$ and $q \cdot R' \cap T \cdot R' = \varnothing$.*

Proof. Let $R_0 \subseteq R$ be a Borel set with $\mu(R_0) > 0$ and $\mu(Y \setminus T \cdot R_0) > 0$. By ergodicity, there is $q \in G$ such that $R_0 \cap q^{-1} \cdot (Y \setminus T \cdot R_0)$ has positive measure. Set

$$R' = R_0 \cap q^{-1} \cdot (Y \setminus T \cdot R_0).$$ $\square$

Now we are ready to build $W, F \subseteq G$ and $c, q_1, \ldots, q_6 \in G$ as pictured in Figure 1. We mention that when we will later apply this lemma, the set Y will be defined as $Y = \{y \in X : |W \cdot y| > 1\}$. We also mention that the equalities in clauses (v) and (vi) below are mostly due to the fact that $1_G \in W$, while in clause (vii) it would be preferable that the intersection be empty but due to non-trivial stabilizers the stated containment is the best one can hope for. Following this lemma we will immediately build a pre-partition β with $R \in \sigma\text{-alg}_G^{\mathrm{red}}(\beta)$.

Lemma 7.6. *Let $G \curvearrowright (X, \mu)$ be a p.m.p. ergodic action with (X, μ) non-atomic. Let $Y \subseteq X$ be a Borel set of positive measure, let $W \subseteq G$ be finite and symmetric with $1_G \in W$, and let $m \in \mathbb{N}$. Then there exist $n \in \mathbb{N}$, $F \cup Q \cup \{c\} \subseteq G$, and a Borel set $R \subseteq X$ with $Q = \{q_1, \ldots, q_6\}$, $\mu(R) = \frac{1}{n}$, $n > m \cdot |F|$, and satisfying the following:*

   (i) *$Q \cdot R \subseteq Y$;*
   (ii) *$|(\{c\} \cup Q) \cdot r \setminus W \cdot r| = 7$ for all $r \in R$;*
  (iii) *$(W \cup \{c\} \cup Q)^2 \subseteq F$;*
  (iv) *$F \cdot r \cap F \cdot r' = \varnothing$ for all $r \neq r' \in R$;*
   (v) *$Wq \cdot R \cap (W \cup \{c\} \cup Q) \cdot R = q \cdot R$ for every $q \in Q$;*
  (vi) *$cW \cdot R \cap (W \cup \{c\} \cup Q) \cdot R = c \cdot R$;*
 (vii) *$Qc \cdot R \cap (W \cup \{c\} \cup Q) \cdot R \subseteq c \cdot R$;*
(viii) *for all $r \in R$, either $q_1 c \cdot r \neq c \cdot r$ or $q_2 c \cdot r = c \cdot r$.*

Proof. Set $R_0 = X$. By induction on $1 \leq i \leq 6$ we choose $q_i \in G$ and a Borel set $R_i \subseteq R_{i-1}$ such that $\mu(R_i) > 0$, $q_i \cdot R_i \subseteq Y$, and

$$q_i \cdot R_i \cap W(W \cup \{q_j \,:\, j < i\}) \cdot R_i = \varnothing.$$

Both the base case and the inductive steps are taken care of by Lemma 7.5. Set $Q = \{q_1, q_2, \ldots, q_6\}$. Then $q_i \cdot R_6 \subseteq q_i \cdot R_i \subseteq Y$ and $|Q \cdot r \setminus W \cdot r| = 6$ for all $r \in R_6$. Now apply Corollary 7.4 to obtain $c \in G$ and a Borel set $R_c \subseteq R_6$ with $\mu(R_c) > 0$ and

$$cW \cdot R_c \cap (\{1_G\} \cup Q^{-1})(W \cup Q \cup WQ) \cdot R_c = \varnothing.$$

Set $F = (W \cup \{c\} \cup Q)^2$ so that (iii) is satisfied.

If there is $q \in Q$ with $qc \cdot r = c \cdot r$ for all $r \in R_c$, then set $R' = R_c$ and re-index the elements of Q so that $q_2 = q$. Otherwise, we may re-index Q and find a Borel set $R' \subseteq R_c$ of positive measure with $q_1 c \cdot r \neq c \cdot r$ for all $r \in R'$. Now apply Lemma 7.1 to obtain a positive measure Borel set $R \subseteq R'$ with $F \cdot r \cap F \cdot r' = \varnothing$ for all $r \neq r' \in R$. By shrinking R if necessary, we may suppose that $\mu(R) = \frac{1}{n}$ for some $n > m \cdot |F|$. Then (iv) is immediately satisfied, (viii) is satisfied since $R \subseteq R'$, and (i) is satisfied since $R \subseteq R_6$. Clause (ii) also holds since $c \cdot r \in cW \cdot r$ is disjoint from $(W \cup Q) \cdot r$ for every $r \in R$.

Recall that $W = W^{-1}$ and $1_G \in W$. Fix $1 \leq i \leq 6$. By the definition of q_i we have $Wq_i \cdot R \cap W \cdot R = \varnothing$, and if $j \neq i$ then $Wq_i \cdot R \cap q_j \cdot R = \varnothing$. Also, the definition of c implies that $Wq_i \cdot R \cap c \cdot R = \varnothing$. Therefore

$$Wq_i \cdot R \cap (W \cup \{c\} \cup Q) \cdot R \subseteq q_i \cdot R.$$

This establishes (v) since $1_G \in W$. By definition of c we have $Qc \cdot R \cap (W \cup Q) \cdot R = \varnothing$. So (vii) follows. Similarly, $cW \cdot R \cap (W \cup Q) \cdot R = \varnothing$ which gives one inclusion in (vi), and the reverse inclusion follows from $1_G \in W$. $\qquad\square$

Now we construct a pre-partition β with $R \in \sigma\text{-alg}_G^{\mathrm{red}}(\beta)$.

Lemma 7.7. *Let $G \curvearrowright (X, \mu)$ be a p.m.p. ergodic action with (X, μ) non-atomic. Let $W \subseteq G$ be finite and symmetric with $1_G \in W$, and let $D \subseteq X$ be a Borel set with $W \cdot x \cap D \neq \varnothing$ for all $x \in X$. Assume that the set $Y = \{x \in X \,:\, |W \cdot x| \geq 2\}$ has positive measure, and let $F \cup Q \cup \{c\} \subseteq G$ and $R \subseteq X$ be as in Lemma 7.6. If*

$\beta = \{B_0, B_1\}$ *is a pre-partition satisfying*

(7.1) $B_1 \supseteq (W \cup \{c, q_1\}) \cdot R, \quad and$

(7.2) $B_0 \supseteq (D \setminus F \cdot R) \bigcup \left(F \cdot R \setminus (W \cup \{c\} \cup Q) \cdot R \right) \bigcup q_2 \cdot R$

then $R \in \sigma\text{-alg}_G^{red}(\beta)$.

Proof. We will use the roman numerals (i) through (viii) to refer to clauses of Lemma 7.6. If $r \in R$ then it is immediate from the definitions that $(W \cup \{c, q_1\}) \cdot r \subseteq B_1$ and $q_2 \cdot r \in B_0$. So it suffices to show that if $x \notin R$ then either $(W \cup \{c, q_1\}) \cdot x \cap B_0 \neq \varnothing$ or $q_2 \cdot x \in B_1$.

Fix $x \notin R$. Let $w \in W$ be such that $w \cdot x \in D$. If $w \cdot x \in B_0$ then we are done. So suppose that $w \cdot x \notin B_0$. Since $w \cdot x \in D \setminus B_0$, by (7.2) we must have $w \cdot x \in F \cdot R$. Now $w \cdot x \in F \cdot R \setminus B_0$ so again by (7.2) we obtain $w \cdot x \in (W \cup \{c\} \cup Q) \cdot R$. If $x \in B_0$ then we are done since $1_G \in W$. So suppose that $x \notin B_0$. Since W is symmetric, $x \in W \cdot w \cdot x$ and hence $x \in F \cdot R$ by (iii). Again, $x \in F \cdot R \setminus B_0$ so (7.2) implies $x \in (W \cup \{c\} \cup Q) \cdot R$.

The previous paragraph shows that if $W \cdot x$ is disjoint with B_0 then we must have $x \in (W \cup \{c\} \cup Q) \cdot R$. We will divide the remainder of the argument into three cases: $x \in W \cdot R \setminus R$, $x \in c \cdot R$, and $x \in Q \cdot R$. Along the way we will illuminate why c, q_1, q_2 are needed in this construction.

Suppose that $x \in W \cdot R \setminus R$. Here is the problem in which W together with stabilizers may possess too much symmetry: even though $x \notin R$, it may be that $W \cdot x \subseteq W \cdot R \subseteq B_1$. This problem is resolved by using the checkpoint c. Since $x \notin R$, it follows from (vi) that $c \cdot x \notin (W \cup \{c\} \cup Q) \cdot R$. By (iii) we have $c \cdot x \in F \cdot R$, so by (7.2) we find that $c \cdot x \in B_0$. Thus in the case $x \in W \cdot R \setminus R$ we are done.

Now suppose that $x \in c \cdot R$. The problem in this case is that we may have $W \cdot x = \{x\}$, and based on the conditions imposed on c this situation is unavoidable to the best knowledge of the author. We handle this problem by using the points q_1 and q_2. Fix $r \in R$ with $x = c \cdot r$. By (viii) we have that either $q_1 c \cdot r \neq c \cdot r$ or $q_2 c \cdot r = c \cdot r$. In the latter case, (7.1) gives

$$q_2 \cdot x = q_2 c \cdot r = c \cdot r \in B_1$$

and we are done. So assume that $q_1 c \cdot r \neq c \cdot r$. Then $q_1 c \cdot r \notin c \cdot R$ by (iii) and (iv) and so by (vii)

$$q_1 c \cdot r \notin (W \cup \{c\} \cup Q) \cdot R.$$

As $q_1 c \in F$ by (iii), we find that

$$q_1 \cdot x = q_1 c \cdot r \in F \cdot R \setminus (W \cup \{c\} \cup Q) \cdot R \subseteq B_0$$

which finishes this case.

Finally, suppose that $x \in Q \cdot R$. Thankfully Q does not create any new problems and the argument can terminate here. Fix $r \in R$ and $q \in Q$ with $x = q \cdot r$. By (i) $q \cdot r \in Y$ and hence there is $w \in W$ with $wq \cdot r \neq q \cdot r$. It follows $wq \cdot r \notin q \cdot R$ by (iii) and (iv) and so by (v)

$$wq \cdot r \notin (W \cup \{c\} \cup Q) \cdot R.$$

Therefore

$$w \cdot x = wq \cdot r \in F \cdot R \setminus (W \cup \{c\} \cup Q) \cdot R \subseteq B_0. \qquad \square$$

We are now ready for the main result of this section.

Proposition 7.8. *Let $G \curvearrowright (X, \mu)$ be a p.m.p. ergodic action with (X, μ) non-atomic and let $0 < \delta < 1$. Then there are $\epsilon > 0$ and a Borel set $M \subseteq X$ with $\mu(M) = \delta$ with the following property: for any $S_1, S_2 \subseteq X$ satisfying $\mu(S_1) + \mu(S_2) < \epsilon$ there is a two-piece partition $\beta = \{B_0, B_1\}$ of M with $S_1, S_2 \in \sigma\text{-alg}_G^{red}(\beta)$.*

Proof. By Lemma 7.2, there is a finite symmetric set $W \subseteq G$ with $1_G \in W$ and a Borel set $D \subseteq X$ with $\mu(D) < \delta/2$ such that $W \cdot x \cap D \neq \varnothing$ for all $x \in X$. Note that if $|W \cdot x| = 1$ then $x \in D$. Thus the set $Y = \{x \in X : |W \cdot x| \geq 2\}$ has positive measure. Apply Lemma 7.6 to obtain $F \cup \{c\} \cup Q \subseteq G$ with $Q = \{q_1, \ldots, q_6\}$ and $R \subseteq X$ with $\mu(R) = \frac{1}{n}$, where $n > 2|F|/\delta$. Fix $k \in \mathbb{N}$ with

$$\log_2(2nk) < k - 1$$

and let Z_1 and Z_2 be disjoint Borel subsets of R with $\mu(Z_1) = \mu(Z_2) = \frac{1}{2nk}$. Set $Z = Z_1 \cup Z_2$ and note that $\mu(Z) = \frac{1}{nk} = \frac{1}{k} \cdot \mu(R)$. Fix $\epsilon > 0$ with $\epsilon < \frac{1}{6nk}$. Let $M \subseteq X$ be any Borel set with $D \cup F \cdot R \subseteq M$ and $\mu(M) = \delta$.

Apply Corollaries 3.6 and 3.7 to obtain a $\sigma\text{-alg}_G(\{Z, R\})$-expressible function $\rho \in [[E_G^X]]$ such that $\text{dom}(\rho) = \text{rng}(\rho) = R$, $\rho^k = \text{id}_R$, and such that $\{\rho^i(Z) : 0 \leq i < k\}$ is a partition of R. For each $j = 1, 2$, again apply these corollaries to obtain a $\sigma\text{-alg}_G(\{Z_j\})$-expressible function $\psi_j \in [E_G^X]$ such that $\text{dom}(\psi_j) = \text{rng}(\psi_j) = X$, $\psi_j^{2nk} = \text{id}_X$, and such that $\{\psi_j^i(Z_j) : 0 \leq i < 2nk\}$ is a partition of X. We mention that there are no assumed relationships between ψ_1, ψ_2, and ρ.

Let $S_1, S_2 \subseteq X$ be Borel sets with $\mu(S_1) + \mu(S_2) < \epsilon$. Our intention will be to encode how the sets $\psi_1^i(Z_1)$ meet S_1 and similarly how the sets $\psi_2^i(Z_2)$ meet S_2. For $1 \leq m \leq 2nk$ and $j = 1, 2$, let Z_j^m be the set of $z \in Z_j$ such that

$$|\{0 \leq i < 2nk : \psi_j^i(z) \in S_j\}| \geq m.$$

Then $Z_j^1 \supseteq Z_j^2 \supseteq \cdots \supseteq Z_j^{2nk}$ and

$$\sum_{m=1}^{2nk} \mu(Z_1^m \cup Z_2^m) = \mu(S_1) + \mu(S_2) < \epsilon.$$

Setting $Z_j^* = Z_j \setminus Z_j^1$, we have

$$\mu(Z_j^*) = \mu(Z_j) - \mu(Z_j^1) > \frac{1}{2nk} - \epsilon > \frac{1}{3nk} > 2\epsilon.$$

In particular,

$$(7.3) \qquad \mu(Z_1^* \cup Z_2^*) - \sum_{m=1}^{2nk} \mu(Z_1^m \cup Z_2^m) > 4\epsilon - \epsilon = 3\epsilon.$$

Set $Z^m = Z_1^m \cup Z_2^m$ and $Z^* = Z_1^* \cup Z_2^*$.

For each $1 \leq m \leq 2nk$ we wish to build a function $\theta_m \in [[E_G^X]]$ which is expressible with respect to $\sigma\text{-alg}_G(\{Z^*, Z^1, \ldots, Z^m\})$ and satisfies $\text{dom}(\theta_m) = Z^m$ and

$$\text{rng}(\theta_m) \subseteq Z^* \setminus \bigcup_{k=1}^{m-1} \theta_k(Z^k).$$

We construct these functions inductively. When $m = 1$, we have $\mu(Z^1) < \epsilon < \mu(Z^*)$ and thus θ_1 is obtained immediately from Lemma 3.5. Now suppose that θ_1 through

θ_{m-1} have been defined. Then

$$Z^* \setminus \bigcup_{k=1}^{m-1} \theta_k(Z^k)$$

lies in σ-$\mathrm{alg}_G(\{Z^*, Z^1, \ldots, Z^{m-1}\})$ by Lemma 3.2. By (7.3) we have

$$\mu(Z^m) < \epsilon < \mu(Z^*) - \sum_{k=1}^{m-1} \mu(Z^k) = \mu\left(Z^* \setminus \bigcup_{k=1}^{m-1} \theta_k(Z^k)\right).$$

Therefore we may apply Lemma 3.5 to obtain θ_m. This completes the construction.

Define $f : \bigcup_{m=1}^{2nk} \mathrm{rng}(\theta_m) \to \{0, 1, \ldots, 2nk - 1\}$ by setting $f(\theta_m(z)) = \ell$ for $z \in Z_j^m$ if and only if $\psi_j^\ell(z) \in S_j$, and

$$|\{0 \leq i \leq \ell \,:\, \psi_j^i(z) \in S_j\}| = m.$$

For $i, t \in \mathbb{N}$ we let $\mathbb{B}_i(t) \in \{0, 1\}$ denote the ith digit in the binary expansion of t (so $\mathbb{B}_i(t) = 0$ for all $i > \log_2(t) + 1$). Now define a Borel set $B_1 \subseteq X$ by the rule

$$x \in B_1 \Longleftrightarrow \begin{cases} x \in W \cdot R & \text{or} \\ x \in c \cdot R & \text{or} \\ x \in q_1 \cdot R & \text{or} \\ x \in q_3 \cdot Z & \text{or} \\ x \in q_4 \cdot Z_1 & \text{or} \\ x \in q_5 \cdot Z^1 & \text{or} \\ x \in q_6 \cdot \theta_m(Z^{m+1}) & \text{for some } 1 \leq m < 2nk, \text{ or} \\ x = q_6 \cdot \rho^i(z) & \text{where } 1 \leq i < k, \ z \in \mathrm{dom}(f), \\ & \qquad \text{and } \mathbb{B}_i(f(z)) = 1. \end{cases}$$

It is important to note that $B_1 \subseteq (W \cup \{c\} \cup Q) \cdot R$. In particular, $B_1 \subseteq F \cdot R$ by Lemma 7.6.(iii). We also define the Borel set

$$B_0 = M \setminus B_1 \supseteq (D \setminus F \cdot R) \cup \Big(F \cdot R \setminus (W \cup \{c\} \cup Q) \cdot R\Big) \cup \Big((W \cup \{c\} \cup Q) \cdot R \setminus B_1\Big).$$

Note that clauses (iii) and (iv) of Lemma 7.6 imply that for every $r \neq r' \in R$

$$(W \cup \{c\} \cup Q) \cdot r \cap (W \cup \{c\} \cup Q) \cdot r' = \varnothing.$$

Thus from clause (ii) of Lemma 7.6 we obtain the following one-way implications

$$x \in B_0 \Longleftarrow \begin{cases} x \in q_2 \cdot R & \text{or} \\ x \in q_3 \cdot (R \setminus Z) & \text{or} \\ x \in q_4 \cdot (R \setminus Z_1) & \text{or} \\ x \in q_5 \cdot (R \setminus Z^1) & \text{or} \\ x \in q_6 \cdot \bigcap_{m=1}^{2nk-1}(Z \setminus \theta_m(Z^{m+1})) & \text{or} \\ x = q_6 \cdot \rho^i(z) & \text{where } 1 \leq i < k, \ z \in \mathrm{dom}(f), \\ & \qquad \text{and } \mathbb{B}_i(f(z)) = 0. \end{cases}$$

In particular, $q_2 \cdot R \subseteq B_0$. Therefore $\beta = \{B_0, B_1\}$ satisfies the assumptions of Lemma 7.7.

We will now check that $S_1, S_2 \in \sigma$-$\mathrm{alg}_G^{\mathrm{red}}(\beta)$. By Lemma 7.7 we have $R \in \sigma$-$\mathrm{alg}_G^{\mathrm{red}}(\beta)$. By G-invariance of σ-$\mathrm{alg}_G^{\mathrm{red}}(\beta)$, we have $q_i \cdot R \in \sigma$-$\mathrm{alg}_G^{\mathrm{red}}(\beta)$ for $1 \leq i \leq$

6. Since $q_i \cdot R \subseteq B_0 \cup B_1$, it immediately follows from the definition of $\sigma\text{-alg}_G^{\mathrm{red}}(\beta)$ that $B_0 \cap q_i \cdot R$ and $B_1 \cap q_i \cdot R$ lie in $\sigma\text{-alg}_G^{\mathrm{red}}(\beta)$. Defining the partition

$$\gamma = \Big\{ R, X \setminus (R \cup Q \cdot R) \Big\} \cup \Big\{ B_0 \cap q_i \cdot R : 1 \leq i \leq 6 \Big\} \cup \Big\{ B_1 \cap q_i \cdot R : 1 \leq i \leq 6 \Big\},$$

we have $\sigma\text{-alg}_G(\gamma) \subseteq \sigma\text{-alg}_G^{\mathrm{red}}(\beta)$. It suffices to show that $S_1, S_2 \in \sigma\text{-alg}_G(\gamma)$.

We have $x \in Z$ if and only if $q_3 \cdot x \in B_1 \cap q_3 \cdot R \in \gamma$. Thus $Z \in \sigma\text{-alg}_G(\gamma)$. Similarly, $x \in Z_1$ if and only if $q_4 \cdot x \in B_1 \cap q_4 \cdot R$. As $R \in \gamma$, we conclude that $R, Z, Z_1, Z_2 = Z \setminus Z_1 \in \sigma\text{-alg}_G(\gamma)$. It follows that ρ, ψ_1, and ψ_2 are $\sigma\text{-alg}_G(\gamma)$-expressible.

We prove by induction on $1 \leq m \leq 2nk$ that $Z^m, Z_1^m, Z_2^m \in \sigma\text{-alg}_G(\gamma)$ and that θ_m is $\sigma\text{-alg}_G(\gamma)$-expressible. Since $x \in Z^1$ if and only if $q_5 \cdot x \in B_1 \cap q_5 \cdot R$, we have $Z^1 \in \sigma\text{-alg}_G(\gamma)$. Also $Z_1^1 = Z^1 \cap Z_1$ and $Z_2^1 = Z^1 \cap Z_2$ are in $\sigma\text{-alg}_G(\gamma)$. So $Z^* = Z \setminus Z^1$, $Z_1^* = Z_1 \setminus Z_1^1$, and $Z_2^* = Z_2 \setminus Z_2^1$ are in $\sigma\text{-alg}_G(\gamma)$ as well. It follows that θ_1 is $\sigma\text{-alg}_G(\gamma)$-expressible. Now inductively suppose that $Z^i \in \sigma\text{-alg}_G(\gamma)$ and that θ_i is $\sigma\text{-alg}_G(\gamma)$-expressible for all $1 \leq i \leq m$. Then $z \in Z^{m+1}$ if and only if $z \in Z^m$ and $q_6 \cdot \theta_m(z) \in B_1 \cap q_6 \cdot R$. In other words,

$$Z^{m+1} = \theta_m^{-1}\Big(q_6^{-1} \cdot (B_1 \cap q_6 \cdot R) \Big).$$

Thus $Z^{m+1} \in \sigma\text{-alg}_G(\gamma)$ by Lemmas 3.2 and 3.3. Similarly, $Z_1^{m+1} = Z^{m+1} \cap Z_1$ and $Z_2^{m+1} = Z^{m+1} \cap Z_2$ are in $\sigma\text{-alg}_G(\gamma)$. Finally, θ_{m+1} is expressible with respect to $\sigma\text{-alg}_G(\{Z^*, Z^1, \ldots, Z^{m+1}\}) \subseteq \sigma\text{-alg}_G(\gamma)$. This completes the inductive argument.

Now to complete the proof we show that $S_1, S_2 \in \sigma\text{-alg}_G(\gamma)$. We first argue that f is $\sigma\text{-alg}_G(\gamma)$-measurable. It follows from the previous paragraph and Lemma 3.2 that $\mathrm{dom}(f) \in \sigma\text{-alg}_G(\gamma)$. Observe that the numbers $\ell \in \mathrm{rng}(f)$ are distinguished by their first $(k-1)$-binary digits $\mathbb{B}_i(\ell)$, $1 \leq i < k$, since by construction $\log_2(2nk) < k - 1$. So for $0 \leq \ell < 2nk$, if we set $I_0 = \{1 \leq i < k : \mathbb{B}_i(\ell) = 0\}$ and $I_1 = \{1 \leq i < k\} \setminus I_0$ then we have

$$f^{-1}(\ell) = \mathrm{dom}(f) \cap \bigcap_{i \in I_0} \rho^{-i}\Big(q_6^{-1} \cdot (B_0 \cap q_6 \cdot R) \Big) \cap \bigcap_{i \in I_1} \rho^{-i}\Big(q_6^{-1} \cdot (B_1 \cap q_6 \cdot R) \Big).$$

Thus $f^{-1}(\ell) \in \sigma\text{-alg}_G(\gamma)$ by Lemmas 3.2 and 3.3. Now suppose that $x \in S_j$. Then there is $z \in Z_j$ and $0 \leq \ell < 2nk$ with $x = \psi_j^\ell(z)$. It follows that $z \in Z_j^m$ where

$$m = |\{0 \leq i \leq \ell : \psi_j^i(z) \in S_j\}|.$$

Furthermore, $\ell = f(\theta_m(z))$. Conversely, if there is $1 \leq m \leq 2nk$, $z \in Z_j^m$, and $0 \leq \ell < 2nk$ with $x = \psi_j^\ell(z)$ and $f(\theta_m(z)) = \ell$, then $x \in S_j$. Therefore

$$S_j = \bigcup_{\ell=0}^{2nk-1} \bigcup_{m=1}^{2nk} \psi_j^\ell\Big(Z_j \cap \theta_m^{-1}(f^{-1}(\ell)) \Big) \in \sigma\text{-alg}_G(\gamma) \subseteq \sigma\text{-alg}_G^{\mathrm{red}}(\beta). \qquad \square$$

8. Krieger's finite generator theorem

We now present the main theorem.

Theorem 8.1. *Let $G \curvearrowright (X, \mu)$ be a p.m.p. ergodic action with (X, μ) non-atomic, and let $\mathcal{F}$ be a G-invariant sub-σ-algebra. If ξ is a countable Borel partition of X, $0 < r \leq 1$, and $\bar{p}$ is a probability vector with $\mathrm{H}(\xi | \mathcal{F}) < r \cdot \mathrm{H}(\bar{p})$, then there is a Borel pre-partition $\alpha = \{A_i : 0 \leq i < |\bar{p}|\}$ with $\mu(A_i) = rp_i$ for every i and $\sigma\text{-alg}_G(\xi) \vee \mathcal{F} \subseteq \sigma\text{-alg}_G^{\mathrm{red}}(\alpha) \vee \mathcal{F}$.*

Proof. Apply Proposition 4.2 to obtain a finite Borel partition ξ' with $\sigma\text{-alg}_G(\xi') \vee \mathcal{F} = \sigma\text{-alg}_G(\xi) \vee \mathcal{F}$ and $\mathrm{H}(\xi'|\mathcal{F}) < r \cdot \mathrm{H}(\bar{p})$. Since ξ' is finite, by Lemma 2.1 we have that $\mathrm{H}(\xi'|\mathcal{F})$ is equal to the infimum of $\mathrm{H}(\xi'|\zeta)$ over finite $\mathcal{F}$-measurable partitions ζ of X. So fix a finite $\mathcal{F}$-measurable partition ζ with $\mathrm{H}(\xi'|\zeta) < r \cdot \mathrm{H}(\bar{p})$. Since $\mathrm{H}(\xi' \vee \zeta|\zeta) = \mathrm{H}(\xi'|\zeta)$ and $\sigma\text{-alg}_G(\xi' \vee \zeta) \vee \mathcal{F} = \sigma\text{-alg}_G(\xi') \vee \mathcal{F}$, we may replace ξ' with $\xi' \vee \zeta$ if necessary and assume that ξ' refines ζ. Let $\pi_\zeta : \xi' \to \zeta$ be the coarsening map. Finally, by Lemma 2.1 we may let $\bar{q}$ be a finite probability vector which coarsens $\bar{p}$ and satisfies $\mathrm{H}(\xi'|\zeta) < r \cdot \mathrm{H}(\bar{q}) \leq r \cdot \mathrm{H}(\bar{p})$. Since $\mathrm{H}(\bar{q}) > 0$, by permuting the coordinates of $\bar{q}$ if necessary we may assume that $0 < q_0 \leq q_1$.

Fix $\kappa > 0$ with $\mathrm{H}(\xi'|\zeta) + \kappa < r\mathrm{H}(\bar{q}) - r\kappa$. Apply Corollary 6.4 to $\bar{q}, \kappa$ to obtain $\delta_0, \epsilon_0 > 0$ with the property that for all $0 < \epsilon < \epsilon_0$ and all sufficiently large n

$$(8.1) \quad \exists K \subseteq L^n_{\bar{q},\epsilon} \quad |K| \geq \exp(n \cdot \mathrm{H}(\bar{q}) - n \cdot \kappa) \quad \text{and} \quad \forall k \neq k' \in K \; d_{\mathrm{Ham}}(k, k') > 2\delta_0.$$

By Lemma 6.1 we may shrink ϵ_0 so that for all $0 < \epsilon < \epsilon_0$, all sufficiently large n, and all $z \in L^n_{\zeta,\epsilon}$

$$(8.2) \qquad\qquad |\{c \in L^n_{\xi',\epsilon} : \pi_\zeta(c) = z\}| \leq \exp(n \cdot \mathrm{H}(\xi'|\zeta) + n \cdot \kappa).$$

Set $\delta = r \cdot q_0 \cdot \delta_0/6$. Let M with $\mu(M) = \delta$ and $\epsilon > 0$ be given by Proposition 7.8. By shrinking ϵ if necessary, we may assume that $\epsilon < \epsilon_0$ and $\epsilon|\bar{q}| < \delta_0/6$. Let $n_0 \in \mathbb{N}$ be such that for all $n \geq \lfloor rn_0 \rfloor$: statement (8.1) holds, for all $z \in L^n_{\zeta,\epsilon}$ inequality (8.2) holds,

$$(\epsilon n + 1) \cdot |\bar{q}| < (\delta_0/6)n, \quad \text{and}$$

$$(8.3) \qquad\qquad n \cdot \mathrm{H}(\xi'|\zeta) + n\kappa < \lfloor rn \rfloor \cdot \mathrm{H}(\bar{q}) - \lfloor rn \rfloor \cdot \kappa.$$

By Proposition 5.4 there are $n \geq n_0$, Borel sets $S_1, S_2 \subseteq X$ with $\mu(S_1) + \mu(S_2) < \epsilon$, and a $\sigma\text{-alg}_G(\{S_1, S_2\})$-expressible $\theta \in [E^X_G]$ such that E_θ admits a $\sigma\text{-alg}_G(\{S_1, S_2\})$-measurable transversal Y and such that for μ-almost-every $x \in X$, the E_θ class of x has cardinality n,

$$\forall C \in \xi' \cup \zeta \qquad \mu(C) - \epsilon < \frac{|C \cap [x]_{E_\theta}|}{|[x]_{E_\theta}|} < \mu(C) + \epsilon,$$

$$\text{and} \qquad \frac{|M \cap [x]_{E_\theta}|}{|[x]_{E_\theta}|} < \mu(M) + \delta = 2\delta.$$

So we have $\mathcal{N}^\theta_{\xi'}(y) \in L^n_{\xi',\epsilon}$ and $\mathcal{N}^\theta_\zeta(y) \in L^n_{\zeta,\epsilon}$ for almost-every $y \in Y$.

We set $m = \lfloor r \cdot n \rfloor$ and encourage the reader to pay attention to the distinction between m and n. Let $K \subseteq L^m_{\bar{q},\epsilon}$ be as given by (8.1) so that $d_{\mathrm{Ham}}(k, k') > 2\delta_0$ for all $k \neq k' \in K$. By (8.1), (8.2), and (8.3) we have that for every $z \in L^n_{\zeta,\epsilon}$

$$\left|\{c \in L^n_{\xi',\epsilon} : \pi_\zeta(c) = z\}\right| \leq \exp\left(n \cdot \mathrm{H}(\xi'|\zeta) + n \cdot \kappa\right) < \exp\left(m \cdot \mathrm{H}(\bar{q}) - m \cdot \kappa\right) \leq |K|.$$

Thus for every $z \in L^n_{\zeta,\epsilon}$ we may fix an injection $f_z : \{c \in L^n_{\xi',\epsilon} : \pi_\zeta(c) = z\} \to K \subseteq L^m_{\bar{q},\epsilon}$.

For $y \in Y$ set $z_y = \mathcal{N}^\theta_\zeta(y)$, $c_y = \mathcal{N}^\theta_{\xi'}(y)$, and $\tilde{a}_y = f_{z_y}(c_y) \in K \subseteq L^m_{\bar{q},\epsilon}$. Also define $M_y = \{0 \leq i < n : \theta^i(y) \in M\}$. Then

$$|M_y| < 2\delta \cdot n = 2(q_0 \cdot \delta_0/6)rn < (\delta_0/3)(m+1) \leq (2\delta_0/3)m$$

for μ-almost-every $y \in Y$. Since $\tilde{a}_y \in L^m_{\bar{q},\epsilon}$, Lemma 6.5 provides a set $J_y \subseteq \{0, 1, \ldots, m-1\}$ with $|J_y| < (\epsilon m + 1) \cdot |\bar{q}| + (\delta_0/6)m < (\delta_0/3)m$ such that for

all $0 \le t < |\bar{q}|$

$$(8.4) \qquad \frac{1}{m} \cdot \left| \tilde{a}_y^{-1}(t) \setminus (M_y \cup J_y) \right| \le \frac{1}{m} \cdot \left| \tilde{a}_y^{-1}(t) \setminus J_y \right| < (1 - \delta_0/6)q_t.$$

Since there are only finitely many choices for J_y, it is easy to arrange the map $y \mapsto J_y$ to be Borel. We then let J be the Borel set $J = \{\theta^j(y) : y \in Y, \ j \in J_y\}$.

Define the pre-partition $\alpha^0 = \{A_t^0 : 0 \le t < |\bar{q}|\}$ by setting

$$A_t^0 = \{\theta^i(y) : y \in Y, \ 0 \le i < m, \ i \notin M_y \cup J_y, \ \text{and} \ \tilde{a}_y(i) = t\}.$$

Observe that $\mu(Y) = 1/n$ since Y is a transversal for E_θ. By (8.4) we have that for every $0 \le t < |\bar{q}|$

$$\mu(A_t^0) = \int_Y |\tilde{a}_y^{-1}(t) \setminus (M_y \cup J_y)| \ d\mu(y)$$

$$< \frac{m}{n} \cdot (1 - \delta_0/6)q_t \le r(1 - \delta_0/6) \cdot q_t = \left(r - \frac{\delta}{q_0} \right) q_t.$$

Thus $\mu(A_t^0) < r \cdot q_t$ and, since $q_0 \le q_1$, we have $\mu(A_0^0) < r \cdot q_0 - \delta = r \cdot q_0 - \mu(M)$ and similarly $\mu(A_1^0) \le r \cdot q_1 - \mu(M)$.

Apply Proposition 7.8 to get a partition $\beta = \{B_0, B_1\}$ of M with $S_1, S_2 \in \sigma\text{-alg}_G^{\mathrm{red}}(\beta)$. Since M is disjoint from $\cup \alpha^0$ and μ is non-atomic, there is a pre-partition $\alpha' = \{A_t' : 0 \le t < |\bar{q}|\}$ with $\mu(A_t') = r \cdot q_t$ for every $0 \le t < |\bar{q}|$, $A_0^0 \cup B_0 \subseteq A_0'$, $A_1^0 \cup B_1 \subseteq A_1'$, and $A_t^0 \subseteq A_t'$ for $2 \le t < |\bar{q}|$. The pre-partition α' extends both α^0 and β. In particular, $S_1, S_2 \in \sigma\text{-alg}_G^{\mathrm{red}}(\beta) \subseteq \sigma\text{-alg}_G^{\mathrm{red}}(\alpha')$ by Lemma 2.2. We have that θ is expressible and Y is measurable with respect to $\sigma\text{-alg}_G(\{S_1, S_2\}) \subseteq \sigma\text{-alg}_G^{\mathrm{red}}(\alpha')$. By Lemma 3.3 it follows that θ^i is $\sigma\text{-alg}_G^{\mathrm{red}}(\alpha')$-expressible for all $i \in \mathbb{Z}$. We will show that $\xi' \subseteq \sigma\text{-alg}_G^{\mathrm{red}}(\alpha') \vee \mathcal{F}$.

We claim that the map $y \in Y \mapsto \tilde{a}_y$ is $\sigma\text{-alg}_G^{\mathrm{red}}(\alpha')$-measurable. We check this via the definition of reduced σ-algebras. Fix $y \in Y$ and $x \in X$ with either $x \notin Y$ or $\tilde{a}_x \ne \tilde{a}_y$. If $x \notin Y$ then we are done since $Y \in \sigma\text{-alg}_G^{\mathrm{red}}(\alpha')$ (since then there is $g \in G$ with $g \cdot x, g \cdot y \in \cup \alpha'$ and with α' separating $g \cdot x$ from $g \cdot y$). So suppose that $x \in Y$ and $\tilde{a}_x \ne \tilde{a}_y$. Then $d_{\mathrm{Ham}}(\tilde{a}_y, \tilde{a}_x) > 2\delta_0$ since $\tilde{a}_x, \tilde{a}_y \in K$. Set $I = \{0 \le i < m : \tilde{a}_y(i) \ne \tilde{a}_x(i)\}$ and note $|I| > 2\delta_0 \cdot m$. Since

$$\left| M_y \cup J_y \cup M_x \cup J_x \right| < 2 \cdot (2\delta_0/3)m + 2 \cdot (\delta_0/3)m = 2\delta_0 m,$$

we may fix $i \in I \setminus (M_y \cup J_y \cup M_x \cup J_x)$. Since θ^i is $\sigma\text{-alg}_G^{\mathrm{red}}(\alpha')$-expressible, there is a $\sigma\text{-alg}_G^{\mathrm{red}}(\alpha')$-measurable partition $\{Z_g : g \in G\}$ of X such that $\theta^i(z) = g \cdot z$ for all $g \in G$ and $z \in Z_g$. If y and x are separated by the partition $\{Z_g : g \in G\}$ then, since this partition is $\sigma\text{-alg}_G^{\mathrm{red}}(\alpha')$-measurable, there must be $h \in G$ with both $h \cdot y$ and $h \cdot x$ lying in $\cup \alpha'$ and separated by α'. We are done in this case. So assume there is $g \in G$ with $y, x \in Z_g$. Then $g \cdot y = \theta^i(y)$ lies in $A_t^0 \subseteq A_t'$ where $t = \tilde{a}_y(i)$ and similarly $g \cdot x = \theta^i(x)$ lies in $A_s^0 \subseteq A_s'$ where $s = \tilde{a}_x(i)$. As $t \ne s$ we have that $g \cdot y$ and $g \cdot x$ lie in $\cup \alpha'$ and are separated by α'. This proves the claim.

We observe that the map $y \in Y \mapsto z_y$ is $\sigma\text{-alg}_G^{\mathrm{red}}(\alpha') \vee \mathcal{F}$-measurable since $\zeta \subseteq \mathcal{F}$ and the value of z_y is entirely determined by the location of y in the partition $\bigvee_{i=0}^{n-1} \theta^{-i}(\zeta) \upharpoonright Y$ of Y. This partition is $\sigma\text{-alg}_G^{\mathrm{red}}(\alpha') \vee \mathcal{F}$-measurable by Lemmas 3.2 and 3.3. So the map $y \in Y \mapsto (z_y, \tilde{a}_y)$ is $\sigma\text{-alg}_G^{\mathrm{red}}(\alpha') \vee \mathcal{F}$-measurable. Since $c_y = f_{z_y}^{-1}(\tilde{a}_y)$, it follows that the map $y \in Y \mapsto c_y$ is $\sigma\text{-alg}_G^{\mathrm{red}}(\alpha') \vee \mathcal{F}$-measurable

as well. For $C \in \xi'$ we have

$$C = \{\theta^i(y) \,:\, y \in Y,\ 0 \le i < n,\ \text{and}\ c_y(i) = C\} = \bigcup_{i=0}^{n-1} \theta^i\Big(\{y \in Y \,:\, c_y(i) = C\}\Big).$$

Therefore $\xi' \subseteq \sigma\text{-alg}_G^{\mathrm{red}}(\alpha') \vee \mathcal{F}$ by Lemmas 3.2 and 3.3. We conclude that

$$\sigma\text{-alg}_G(\xi) \vee \mathcal{F} = \sigma\text{-alg}_G(\xi') \vee \mathcal{F} \subseteq \sigma\text{-alg}_G^{\mathrm{red}}(\alpha') \vee \mathcal{F}.$$

Finally, since (X, μ) is non-atomic, $\mu(A'_t) = r \cdot q_t$, and $\bar{q}$ is a coarsening of $\bar{p}$, there is a refinement α of α' with $\mu(A_t) = r \cdot p_t$ for all $0 \le t < |\bar{p}|$. Clearly we still have $\sigma\text{-alg}_G(\xi) \vee \mathcal{F} \subseteq \sigma\text{-alg}_G^{\mathrm{red}}(\alpha) \vee \mathcal{F}$. $\qquad\square$

Note that Theorem 2.3 follows from the above theorem by choosing a partition ξ with $\mathrm{H}(\xi|\mathcal{F})$ close to $h_G^{\mathrm{Rok}}(X, \mu|\mathcal{F})$ and with $\sigma\text{-alg}_G(\xi) \vee \mathcal{F} = \mathcal{B}(X)$.

Corollary 8.2. *Let $G \curvearrowright (X, \mu)$ be a p.m.p. ergodic action with (X, μ) non-atomic, and let $\mathcal{F}$ be a G-invariant sub-σ-algebra. If $G \curvearrowright (Y, \nu)$ is a factor of $G \curvearrowright (X, \mu)$ and Σ is the sub-σ-algebra of X associated to Y then*

$$h_G^{\mathrm{Rok}}(X, \mu|\mathcal{F}) \le h_G^{\mathrm{Rok}}(Y, \nu) + h_G^{\mathrm{Rok}}(X, \mu|\mathcal{F} \vee \Sigma).$$

Proof. This is immediate if either $h_G^{\mathrm{Rok}}(Y, \nu)$ or $h_G^{\mathrm{Rok}}(X, \mu|\mathcal{F} \vee \Sigma)$ is infinite, so suppose that both are finite. Fix $\epsilon > 0$ and fix a generating partition β' for $G \curvearrowright (Y, \nu)$ with $\mathrm{H}(\beta') < h_G^{\mathrm{Rok}}(Y, \nu) + \epsilon/2$. Pull back β' to a partition β of X. Then $\mathrm{H}(\beta) = \mathrm{H}(\beta')$ and $\sigma\text{-alg}_G(\beta) = \Sigma$. By definition of $h_G^{\mathrm{Rok}}(X, \mu|\mathcal{F} \vee \Sigma)$, there is a partition γ' of X with

$$\mathrm{H}(\gamma'|\mathcal{F} \vee \Sigma) < h_G^{\mathrm{Rok}}(X, \mu|\mathcal{F} \vee \Sigma) + \epsilon/2$$

and $\sigma\text{-alg}_G(\gamma') \vee \mathcal{F} \vee \Sigma = \mathcal{B}(X)$. Apply Theorem 8.1 to get a partition γ of X with

$$\mathrm{H}(\gamma) < h_G^{\mathrm{Rok}}(X, \mu|\mathcal{F} \vee \Sigma) + \epsilon/2$$

and $\sigma\text{-alg}_G(\gamma) \vee \mathcal{F} \vee \Sigma = \mathcal{B}(X)$. Then

$$\mathcal{B}(X) = \sigma\text{-alg}_G(\gamma) \vee \mathcal{F} \vee \Sigma = \sigma\text{-alg}_G(\gamma \vee \beta) \vee \mathcal{F},$$

and hence

$$h_G^{\mathrm{Rok}}(X, \mu|\mathcal{F}) \le \mathrm{H}(\beta \vee \gamma|\mathcal{F}) \le \mathrm{H}(\beta) + \mathrm{H}(\gamma) < h_G^{\mathrm{Rok}}(Y, \nu) + h_G^{\mathrm{Rok}}(X, \mu|\mathcal{F} \vee \Sigma) + \epsilon. \ \square$$

9. Relative Rokhlin entropy and amenable groups

We verify that for free ergodic actions of amenable groups, relative Rokhlin entropy and relative Kolmogorov–Sinai entropy agree. This result was previously established in the non-relative case by the author and Tucker-Drob [39].

We first recall the definition of relative Kolmogorov–Sinai entropy. Let G be a countably infinite amenable group, and let $G \curvearrowright (X, \mu)$ be a free p.m.p. action. For a partition α and a finite set $T \subseteq G$, we write α^T for the join $\bigvee_{t \in T} t \cdot \alpha$, where $t \cdot \alpha = \{t \cdot A \,:\, A \in \alpha\}$. Given a G-invariant sub-σ-algebra $\mathcal{F}$, the relative Kolmogorov–Sinai entropy is defined as

$$h_G^{\mathrm{KS}}(X, \mu|\mathcal{F}) = \sup_\alpha \inf_{T \subseteq G} \frac{1}{|T|} \cdot \mathrm{H}(\alpha^T|\mathcal{F}),$$

where α ranges over all finite partitions and T ranges over finite subsets of G [10]. Equivalently, one can replace the infimum with a limit over a Følner sequence (T_n) [29]. Recall that a sequence $T_n \subseteq G$ of finite sets is a *Følner sequence* if

$$\lim_{n \to \infty} \frac{|\partial_K(T_n)|}{|T_n|} = 0$$

for every finite $K \subseteq G$, where $\partial_K(T) = \{t \in T : tK \not\subseteq T\}$. We also write $\mathcal{I}_K(T)$ for $T \setminus \partial_K(T)$.

Proposition 9.1. *Let G be a countably infinite amenable group, let $G \curvearrowright (X, \mu)$ be a free ergodic action, and let $\mathcal{F}$ be a G-invariant sub-σ-algebra. Then the relative Kolmogorov–Sinai entropy and relative Rokhlin entropy coincide:*

$$h_G^{\mathrm{KS}}(X, \mu | \mathcal{F}) = h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F}).$$

Proof. We first show that $h_G^{\mathrm{KS}}(X, \mu | \mathcal{F}) \leq h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F})$. If $h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F}) = \infty$ then there is nothing to show. So suppose that $h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F}) < \infty$ and fix $\epsilon > 0$. Let α be a countable partition with $\sigma\text{-alg}_G(\alpha) \vee \mathcal{F} = \mathcal{B}(X)$ and $\mathrm{H}(\alpha | \mathcal{F}) < h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F}) + \epsilon$. Let β be any finite partition of X and let (T_n) be a Følner sequence. Then by Lemma 2.1

$$0 = \mathrm{H}(\beta | \sigma\text{-alg}_G(\alpha) \vee \mathcal{F}) = \inf_{K \subseteq G} \mathrm{H}(\beta | \alpha^K \vee \mathcal{F}),$$

where K ranges over finite subsets of G. Fix $K \subseteq G$ so that $\mathrm{H}(\beta | \alpha^K \vee \mathcal{F}) < \epsilon$. Note that $\mathrm{H}(t \cdot \beta | \alpha^{tK} \vee \mathcal{F}) < \epsilon$ for all $t \in G$. Therefore

$$\lim_{n \to \infty} \frac{1}{|T_n|} \cdot \mathrm{H}(\beta^{T_n} | \mathcal{F})$$

$$\leq \lim_{n \to \infty} \frac{1}{|T_n|} \cdot \mathrm{H}(\alpha^{T_n} \vee \beta^{T_n} | \mathcal{F})$$

$$= \lim_{n \to \infty} \frac{1}{|T_n|} \cdot \mathrm{H}(\alpha^{T_n} | \mathcal{F}) + \frac{1}{|T_n|} \cdot \mathrm{H}(\beta^{T_n} | \alpha^{T_n} \vee \mathcal{F})$$

$$\leq \lim_{n \to \infty} \frac{1}{|T_n|} \cdot \sum_{t \in T_n} \left(\mathrm{H}(t \cdot \alpha | \mathcal{F}) + \mathrm{H}(t \cdot \beta | \alpha^{T_n} \vee \mathcal{F}) \right)$$

$$< \lim_{n \to \infty} h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F}) + \epsilon + \frac{|\mathcal{I}_K(T_n)|}{|T_n|} \cdot \epsilon + \frac{|\partial_K(T_n)|}{|T_n|} \cdot \mathrm{H}(\beta)$$

$$= h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F}) + 2\epsilon.$$

Now let ϵ tend to 0 and then take the supremum over all β.

Now we argue that $h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F}) \leq h_G^{\mathrm{KS}}(X, \mu | \mathcal{F})$. Again this is immediate if $h_G^{\mathrm{KS}}(X, \mu | \mathcal{F}) = \infty$, so we assume $h_G^{\mathrm{KS}}(X, \mu | \mathcal{F}) < \infty$. Since the action of G is free, a theorem of Seward and Tucker-Drob [39] states that there is a factor action $G \curvearrowright (Z, \eta)$ of (X, μ) such that the action of G on Z is free and $h_G^{\mathrm{Rok}}(Z, \eta) < \epsilon$. Let Σ be the G-invariant sub-σ-algebra of X associated to Z, and let $G \curvearrowright (Y, \nu)$ be the factor of (X, μ) associated to $\mathcal{F} \vee \Sigma$. Then G acts freely on (Y, ν) since (Y, ν) factors onto (Z, η). By the Ornstein–Weiss theorem [28], all free ergodic actions of countably infinite amenable groups are orbit equivalent. In particular, there is a free ergodic p.m.p. action $\mathbb{Z} \curvearrowright (Y, \nu)$ which has the same orbits as $G \curvearrowright (Y, \nu)$ and has 0 Kolmogorov–Sinai entropy, $h_{\mathbb{Z}}^{\mathrm{KS}}(Y, \nu) = 0$. By the Rokhlin generator theorem [31], we have $h_{\mathbb{Z}}^{\mathrm{Rok}}(Y, \nu) = 0$ as well.

Let's say $\mathbb{Z} = \langle t \rangle$. Define $c : Y \to G$ by

$$c(y) = g \Longleftrightarrow t \cdot y = g \cdot y.$$

Let $f : (X, \mu) \to (Y, \nu)$ be the factor map, and let $\mathbb{Z}$ act on (X, μ) by setting

$$t \cdot x = c(f(x)) \cdot x.$$

Then $\mathcal{F} \vee \Sigma$ and the actions of G and $\mathbb{Z}$ on (X, μ) satisfy the assumptions of Proposition 3.4. Equivalently, in the terminology of Rudolph–Weiss [33], the orbit-change cocycles between the actions of G and $\mathbb{Z}$ on X are $\mathcal{F} \vee \Sigma$-measurable. Thus $h_G^{\mathrm{KS}}(X, \mu | \mathcal{F} \vee \Sigma) = h_{\mathbb{Z}}^{\mathrm{KS}}(X, \mu | \mathcal{F} \vee \Sigma)$ by [33, Theorem 2.6]. Also, since $h_{\mathbb{Z}}^{\mathrm{Rok}}(X, \mu | \mathcal{F} \vee \Sigma) \leq h_{\mathbb{Z}}^{\mathrm{Rok}}(X, \mu)$ and $h_{\mathbb{Z}}^{\mathrm{Rok}}(Y, \nu) = 0$, it follows from Corollary 8.2 that

$$(9.1) \qquad h_{\mathbb{Z}}^{\mathrm{Rok}}(X, \mu | \mathcal{F} \vee \Sigma) = h_{\mathbb{Z}}^{\mathrm{Rok}}(X, \mu).$$

We have

$$
\begin{array}{llll}
h_G^{\mathrm{KS}}(X, \mu | \mathcal{F} \vee \Sigma) & = & h_{\mathbb{Z}}^{\mathrm{KS}}(X, \mu | \mathcal{F} \vee \Sigma) & \text{by the Rudolph–Weiss theorem [33]} \\
& = & h_{\mathbb{Z}}^{\mathrm{KS}}(X, \mu) - h_{\mathbb{Z}}^{\mathrm{KS}}(Y, \nu) & \text{by the Abramov–Rokhlin theorem [1]} \\
& = & h_{\mathbb{Z}}^{\mathrm{KS}}(X, \mu) & \text{since } h_{\mathbb{Z}}^{\mathrm{KS}}(Y, \nu) = 0 \\
& = & h_{\mathbb{Z}}^{\mathrm{Rok}}(X, \mu) & \text{by the Rokhlin generator theorem [31]} \\
& = & h_{\mathbb{Z}}^{\mathrm{Rok}}(X, \mu | \mathcal{F} \vee \Sigma) & \text{by Equation 9.1} \\
& = & h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F} \vee \Sigma) & \text{by Proposition 3.4}
\end{array}
$$

So $h_G^{\mathrm{KS}}(X, \mu | \mathcal{F} \vee \Sigma) = h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F} \vee \Sigma)$. Also, it is immediate from the definitions that $h_G^{\mathrm{KS}}(X, \mu | \mathcal{F} \vee \Sigma) \leq h_G^{\mathrm{KS}}(X, \mu | \mathcal{F})$. Finally, by Corollary 8.2 we have

$$h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F}) \leq h_G^{\mathrm{Rok}}(Z, \eta) + h_G^{\mathrm{Rok}}(X, \mu | \mathcal{F} \vee \Sigma) < \epsilon + h_G^{\mathrm{KS}}(X, \mu | \mathcal{F} \vee \Sigma) \leq \epsilon + h_G^{\mathrm{KS}}(X, \mu | \mathcal{F}).$$

Now let ϵ tend to 0. $\qquad \square$

References

[1] L. M. Abramov and V. A. Rohlin, *Entropy of a skew product of mappings with invariant measure*, Vestnik Leningrad. Univ. 17 (1962), no. 7, 5–13.

[2] A. Alpeev, *On Pinsker factors for Rokhlin entropy*, Journal of Mathematical Sciences 209 (2015), no. 6, 826–829.

[3] A. Alpeev and B. Seward, *Krieger's finite generator theorem for actions of countable groups III*, preprint. https://arxiv.org/abs/1705.09707.

[4] L. Bowen, *Measure conjugacy invariants for actions of countable sofic groups*, Journal of the American Mathematical Society 23 (2010), 217–245.

[5] L. Bowen, *Sofic entropy and amenable groups*, Ergod. Th. & Dynam. Sys. 32 (2012), no. 2, 427–466.

[6] L. Bowen, *Every countably infinite group is almost Ornstein*, in Dynamical Systems and Group Actions, Contemp. Math., 567, Amer. Math. Soc., Providence, RI, 2012, 67–78.

[7] L. Bowen, *Zero entropy is generic*, Entropy 18 (2016), no. 6.

[8] J. P. Conze, *Entropie d'un groupe abélien de transformations*, Z. Wahrscheinlichkeitstheorie verw. Geb. 15 (1972), 11–30.

[9] I. Csiszár and J. Körner, Information Theory: Coding Theorems for Discrete Memoryless Systems. Cambridge University Press, New York, 2011.

[10] A. Danilenko and K. Park, *Generators and Bernoullian factors for amenable actions and cocycles on their orbits*, Ergod. Th. & Dynam. Sys. 22 (2002), 1715–1745.

[11] M. Denker, *Finite generators for ergodic, measure-preserving transformations*, Prob. Th. Rel. Fields 29 (1974), no. 1, 45–55.

[12] T. Downarowicz, Entropy in Dynamical Systems. Cambridge University Press, New York, 2011.

[13] D. Gaboriau and B. Seward, *Cost, ℓ^2-Betti numbers, and the sofic entropy of some algebraic actions*, preprint. http://arxiv.org/abs/1509.02482.

[14] S. Gao, S. Jackson, and B. Seward, *A coloring property for countable groups*, Mathematical Proceedings of the Cambridge Philosophical Society 147 (2009), no. 3, 579–592.

[15] S. Gao, S. Jackson, and B. Seward, *Group colorings and Bernoulli subflows*, Memoirs of the American Mathematical Society 241 (2016), no. 1141, 1–241.

[16] E. Glasner, *Ergodic theory via joinings*. Mathematical Surveys and Monographs, 101. American Mathematical Society, Providence, RI, 2003. xii+384 pp.

[17] C. Grillenberger and U. Krengel, *On marginal distributions and isomorphisms of stationary processes*, Math. Z. 149 (1976), no. 2, 131–154.

[18] Y. Katznelson and B. Weiss, *Commuting measure preserving transformations*, Israel J. Math. 12 (1972), 161–173.

[19] A. Kechris, Classical Descriptive Set Theory. Springer-Verlag, New York, 1995.

[20] A. Kechris, S. Solecki, and S. Todorcevic, *Borel chromatic numbers*, Adv. in Math. 141 (1999), 1–44.

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

[22] D. Kerr and H. Li, *Soficity, amenability, and dynamical entropy*, American Journal of Mathematics 135 (2013), 721–761.

[23] D. Kerr and H. Li, *Bernoulli actions and infinite entropy*, Groups Geom. Dyn. 5 (2011), 663–672.

[24] Y. Kifer and B. Weiss, *Generating partitions for random transformations*, Ergod. Th. & Dynam. Sys. 22 (2002), 1813–1830.

[25] W. Krieger, *On entropy and generators of measure-preserving transformations*, Trans. Amer. Math. Soc. 149 (1970), 453–464.

[26] D. Ornstein, *Bernoulli shifts with the same entropy are isomorphic*, Advances in Math. 4 (1970), 337–348.

[27] D. Ornstein, *Two Bernoulli shifts with infinite entropy are isomorphic*, Advances in Math. 5 (1970), 339–348.

[28] D. Ornstein and B. Weiss, *Ergodic theory of amenable group actions. I : The Rohlin lemma*, Bull. Amer. Math. Soc. 2 (1980), no. 1, 161–164.

[29] D. Ornstein and B. Weiss, *Entropy and isomorphism theorems for actions of amenable groups*, Journal d'Analyse Mathématique 48 (1987), 1–141.

[30] B.H. Neumann, *Groups covered by permutable subsets*, J. London Math. Soc. (1954), no. 2, 236–248.

[31] V. A. Rokhlin, *Lectures on the entropy theory of transformations with invariant measure*, Uspehi Mat. Nauk 22 (1967), no. 5, 3–56.

[32] A. Rosenthal, *Finite uniform generators for ergodic, finite entropy, free actions of amenable groups*, Prob. Th. Rel. Fields 77 (1988), 147–166.

[33] D. J. Rudolph and B. Weiss, *Entropy and mixing for amenable group actions*, Annals of Mathematics (151) 2000, no. 2, 1119–1150.

[34] B. Seward, *Ergodic actions of countable groups and finite generating partitions*, Groups, Geometry, and Dynamics 9 (2015), no. 3, 793–810.

[35] B. Seward, *Every action of a non-amenable group is the factor of a small action*, Journal of Modern Dynamics 8 (2014), no. 2, 251–270.

[36] B. Seward, *Krieger's finite generator theorem for actions of countable groups II*, preprint. http://arxiv.org/abs/1501.03367.

[37] B. Seward, *Weak containment and Rokhlin entropy*, preprint. http://arxiv.org/abs/1602.06680.

[38] B. Seward, *Positive entropy actions of countable groups factor onto Bernoulli shifts*, preprint. https://arxiv.org/abs/1804.05269.

[39] B. Seward and R. D. Tucker-Drob, *Borel structurability on the 2-shift of a countable group*, Annals of Pure and Applied Logic 167 (2016), no. 1, 1–21.

[40] A. M. Stepin, *Bernoulli shifts on groups*, Dokl. Akad. Nauk SSSR 223 (1975), no. 2, 300–302.

[41] Š. Šujan, *Generators for amenable group actions*, Mh. Math. 95 (1983), no. 1, 67–79.

Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10003, U.S.A.

E-mail address: b.m.seward@gmail.com