ISRAEL JOURNAL OF MATHEMATICS 78 (1992), 145–176

VERY WEAK BERNOULLI FOR AMENABLE GROUPS

BY

SCOT ADAMS

*Department of Mathematics, University of Chicago
Chicago, Illinois 60637, U.S.A.*

ABSTRACT

We strengthen Følner Independence (defined in an earlier paper) to a criterion (which we call Very Weak Bernoulli) and prove that this new condition is equivalent to Finitely Determined.

0. Introduction

Donald Ornstein and Benjamin Weiss [O-W], in collaboration with Jacob Feldman and Dan Rudolph have generalized much of classical ergodic theory to actions of amenable groups. However, there is one important part of the theory which has not been extended: the Very Weak Bernoulli (VWB) criterion. In the original theory, this condition was useful in verifying that many processes that arise are, in fact, Bernoulli. For examples, see [O, pp. 115-125] and [Sh, pp. 104-105]. In addition, the thesis of J. Steif ([St1] and [St2]) contains applications of the VWB criterion to interacting particle systems.

For **Z**-actions, the VWB condition was developed in by Ornstein in [O]. In [O-W-3], it was shown by Ornstein and Weiss that VWB is equivalent to Finitely Determined.

As it turns out, VWB is very hard to define in a general amenable group, because the definition for **Z**-actions relies so heavily on a concept of the "past"; in the case of the integers, the "past" refers to the negative integers. Even for $\mathbf{Z}^d$, a notion of the past has been developed (see [K-W]). For a general amenable group, however, it is not at all clear what the correct definition is.

In [A], we avoided much of this difficulty by considering a stronger condition which we named "Følner Independence". Even for **Z**-actions, however, Følner

Received October 28, 1990 and in revised form October 2, 1991

Independence implies but is not implied by Finitely Determined (see [O-W-2]). It is, however, suggestive of another criterion. It is the point of this work that this other definition (Definition 3.1), which we call Very Weak Bernoulli, is equivalent to Finitely Determined.

In [A-S], an application of VWB for amenable groups will be given. We will show that an attractive interacting particle system (with the usual lattice $\mathbf{Z}^d$ replaced by an arbitrary discrete amenable group) is Bernoulli.

In §1, we prove two technical lemmas needed to estimate how close quasi-tilings are to being true tilings. Then in §2, we define closeness in entropy and finite distribution for finite processes, then prove (Theorem 2.8) that large but finite portions of a Finitely Determined process inherit a property similar to Finitely Determined. Finally, in §3, we give an analogue (Definition 3.1) of Very Weak Bernoulli for amenable groups. In Theorem 3.7, we prove, using the main result (Theorem 2.8) of §2, that this condition is equivalent to Bernoulli. Much of the proof of Theorem 3.7 is modelled on [O-W-3].

I would like to thank Don Ornstein for useful insights, especially the idea of using extremality (see (1C7) of the proof of Theorem 2.8). Jeff Steif also helped out by finding a simplification to an earlier definition of VWB. I would also like to thank the referee for many useful suggestions and for catching a serious error in the original version of this paper.

1. Technical estimates

We collect here some technical results.

Recall from [A, §1] that if I is any set, then an **I-process** is a family of random variables, indexed by I, together with a grand coupling. If each variable takes values in $\{+1,-1\}$, then we call the process an **I-spin system.** If G is a countable, discrete group then a G-process is **stationary** if the grand coupling is invariant under right translation by elements of G. (Here, G acts on $\{+1,-1\}^G$ by $(\sigma g)(g')=\sigma(g'g^{-1})$.)

We adopt some of the notation of [A, §3]. If $K\subseteq G$ is finite and $\delta>0$, then a finite subset $F\subseteq G$ is **(K,δ)-invariant** if, for all $k\in K$, $|kF\Delta F|/|F|<\delta$. (Note that this does *not* agree with [O-W, p. 24, l.+5].) Recall that a group G is **amenable** iff for all (K,δ), there exists a (K,δ)-invariant subset of G [O-W, p. 14, l.–5]. A property $P(F)$ of finite subsets $F\subseteq G$ is said to **hold for sufficiently invariant sets** if there exists a (K,δ) such that $P(F)$ holds whenever F is

(K,δ)-invariant. If $\mathcal{F}$ is a collection of finite subsets of G, then we say that $\mathcal{F}$ is a **Følner family** if, for every (K,δ), there exists a (K,δ)-invariant set in $\mathcal{F}$.

Throughout this paper, G is a countable, discrete, amenable group with identity element e. (See [A, §3] for the definition of **amenable**.) Recall from [A, §3] that if $K, F \subseteq G$ are finite subsets, and if $e \in F = F^{-1}$, then the F**-boundary of** K is the set B of $a \in G$ such that $Fa \cap K \neq \emptyset \neq Fa \cap (G \backslash K)$; recall also that the F-**interior of** K is $K \backslash B$. Recall [O-W, p. 21, l.–3] that a collection $\bar{A}_1, \dots, \bar{A}_m \subseteq G$ is said to be δ**-disjoint** if there exist subsets $\tilde{A}_1 \subseteq \bar{A}_1, \dots, \tilde{A}_m \subseteq \bar{A}_m$ such that $p \neq q \Rightarrow \tilde{A}_p \cap \tilde{A}_q = \emptyset$ and such that $|\bar{A}_p \backslash \tilde{A}_p| < \delta |\bar{A}_p|$, for $p = 1, \dots, m$. A collection of subsets $A_1, \dots, A_k$ is a δ**-quasi-tiling system** if $e \in A_1 \subseteq \cdots \subseteq A_k$ and if these sets δ-quasi-tile G, in the sense of [O-W, p. 24]. A δ-disjoint collection $\bar{A}_1, \dots, \bar{A}_m$ of right translates of $A_1, \dots, A_k$ is a δ**-quasi-tiling of** a finite subset $K \subseteq G$ if $|K \backslash (\bar{A}_1 \cup \cdots \cup \bar{A}_m)| < \delta |K|$ and if $K \cap \bar{A}_p \neq \emptyset$, for $p = 1, \dots, m$. Complications arise since the sets $\bar{A}_1, \dots, \bar{A}_m$ are not pairwise-disjoint. The elements of K lying in two of the $\bar{A}_p$s are to be thought of as "bad" points. We now estimate the number of such points:

LEMMA 1.1: *Assume $0 < \delta < 1/2$. Let $A_1, \dots, A_k \subseteq G$ be a δ-quasi-tiling system. Let K be a finite subset of G and let $\mathrm{BDRY}(K)$ denote the union (over $i = 1, \dots, k$) of the $(A_i A_i^{-1})$-boundaries of K. Assume:*

(A) *$|\mathrm{BDRY}(K)| < \delta |K|$; and*

(B) *$\bar{A}_1, \dots, \bar{A}_m$ is a δ-quasi-tiling of K by right translates of $A_1, \dots, A_k$.*

For $p = 1, \dots, m$, let $\bar{B}_p := \bar{A}_p \cap \left(\bigcup_{p' \neq p} \bar{A}_{p'} \right)$. (These represent the "bad" points in $\bar{A}_p$.) Let J be the set of integers $1 \leq p \leq m$ such that $|\bar{B}_p| > \sqrt{\delta} |\bar{A}_p|$. Then

(C) $\left| |K| - \sum_{p=1}^{m} |\bar{A}_p| \right| < 4\delta |K|$;

(D) $\sum_{p=1}^{m} |\bar{B}_p| < 10\delta |K|$; *and*

(E) $\sum_{p \in J} |\bar{A}_p| < 10\sqrt{\delta} |K|$.

Proof: Let $M := \max\{|A_1|, \dots, |A_k|\}$. By (B), there exists a pairwise-disjoint collection of sets $\tilde{A}_1, \dots, \tilde{A}_m$ such that, for $p = 1, \dots, m$, $\tilde{A}_p \subseteq \bar{A}_p$ and $|\bar{A}_p \backslash \tilde{A}_p| < \delta |\bar{A}_p|$. For $p = 1, \dots, m$, let $A_p^* := \bar{A}_p \backslash \tilde{A}_p$. Let

$$\bar{A} := \bigcup_{p=1}^{m} \bar{A}_p, \quad \tilde{A} := \bigcup_{p=1}^{m} \tilde{A}_p, \quad A^* := \bigcup_{p=1}^{m} A_p^*, \quad \bar{B} := \bigcup_{p=1}^{m} \bar{B}_p.$$

Let

$$\mathrm{SUMA} := \sum_{p=1}^{m} |\bar{A}_p|, \qquad \mathrm{SUMB} := \sum_{p=1}^{m} |\bar{B}_p|.$$

Because $\tilde{A}_1, \dots, \tilde{A}_m$ are pairwise-disjoint, $|\tilde{A}| = \sum_{p=1}^{m} |\tilde{A}_p|$. Further, for $p = 1, \dots, m$, $(1-\delta)|\bar{A}_p| < |\tilde{A}_p|$ and $\delta|\bar{A}_p| > |A_p^*|$. Summing these over p, we obtain

$$(1-\delta)(\mathrm{SUMA}) < \sum_{p=1}^{m} |\tilde{A}_p| = |\tilde{A}| \qquad \text{and} \qquad \delta(\mathrm{SUMA}) > \sum_{p=1}^{m} |A_p^*| \geq |A^*|.$$

Subtracting the two inequalities displayed above, we have $(1-2\delta)\mathrm{SUMA} < |\tilde{A}| - |A^*|$. Now $|\tilde{A}| - |A^*| \leq |\tilde{A} \backslash A^*|$, so

(F) $(1-2\delta)(\mathrm{SUMA}) < |\tilde{A} \backslash A^*|$.

By (B), $\bar{A}$ covers all but a δ-fraction of K, i.e.,

(G) $(1-\delta)|K| \leq |\bar{A}|$.

Multiplying this by $1-2\delta$, we get

$$(1-3\delta)|K| < (1-3\delta+2\delta^2)|K| \leq (1-2\delta)|\bar{A}| \leq (1-2\delta)(\mathrm{SUMA}).$$

So, by (F), we have

(H) $(1-3\delta)|K| < |\tilde{A} \backslash A^*|$.

Fix any i, $1 \leq i \leq k$. By definition of a δ-quasi-tiling system, $e \in A_i$. Thus, any right translate of A_i which intersects both K and $G \backslash K$ must be contained in the $(A_i A_i^{-1})$-boundary of K. By definition of a δ-quasi-tiling, every $\bar{A}_p$ intersects K. Consequently, every $\bar{A}_p$ either lies in K or lies in some $(A_i A_i^{-1})$-boundary of K. That is, $\bar{A} \subseteq K \cup \mathrm{BDRY}(K)$. Since $\tilde{A}_1, \dots, \tilde{A}_m$ are pairwise disjoint, and since $\tilde{A} \subseteq \bar{A} \subseteq K \cup \mathrm{BDRY}(K)$, we have $\sum_{p=1}^{m} |\tilde{A}_p| = |\tilde{A}| \leq |K| + |\mathrm{BDRY}(K)|$.

Then, by (A),

(I) $\sum_{p=1}^{m} |\tilde{A}_p| \leq (1+\delta)|K|$.

For $p = 1, \dots, m$, $|\bar{A}_p| - |\tilde{A}_p| = |\bar{A}_p \backslash \tilde{A}_p| \leq \delta|\bar{A}_p|$, so $(1-\delta)|\bar{A}_p| \leq |\tilde{A}_p|$. Summing this over p, we obtain $(1-\delta)(\mathrm{SUMA}) \leq \sum_{p=1}^{m} |\tilde{A}_p|$. So, by (I),

(J) $\mathrm{SUMA} \leq \dfrac{1+\delta}{1-\delta}|K|$.

Proof of (C): By (G), by the inequality $|\bar{A}| \leq \text{SUMA}$ and by (J),

$$(1-\delta)|K| \leq \text{SUMA} \leq \frac{1+\delta}{1-\delta}|K|.$$

Subtracting $|K|$ from this inequality,

$$-\delta|K| \leq \text{SUMA} - |K| \leq \frac{(1+\delta)-(1-\delta)}{(1-\delta)}|K| = \frac{2\delta}{1-\delta}|K|.$$

By assumption, $0 < \delta < 1/2$, so $-4\delta < -\delta$ and $1-\delta > 1/2$, so

$$-4\delta|K| < \text{SUMA} - |K| < \frac{2\delta}{1-(1/2)}|K| = 4\delta|K|,$$

which is equivalent to (C).

Proof of (D): If an element of $\tilde{A}$ lies in two of the $\bar{A}_p$s, then it must lie in some A^*_p, for otherwise it would lie in two $\tilde{A}_p$s, which is impossible by pairwise-disjointness. So any element of $\tilde{A}\backslash A^*$ lies in exactly one $\bar{A}_p$. Therefore, $\tilde{A}\backslash A^* \subseteq \bigcup_{p=1}^{m} \bar{A}_p\backslash\bar{B}_p$, and so $|\tilde{A}\backslash A^*| \leq \sum_{p=1}^{m}[|\bar{A}_p| - |\bar{B}_p|] = \text{SUMA} - \text{SUMB}$. Then, by (J) and (H),

$$\text{SUMB} \leq \text{SUMA} - |\tilde{A}\backslash A^*| \leq |K|\left[\frac{1+\delta}{1-\delta} - (1-3\delta)\right] = |K|\frac{5\delta-3\delta^2}{1-\delta}.$$

Since $0 < \delta < 1/2$, we then have $\text{SUMB} \leq 2|K|(5\delta - 3\delta^2) < 2|K|(5\delta) = 10\delta|K|$, which is (D).

Proof of (E): Notice that, with J defined as in the statement of Lemma 1.1,

$$\text{SUMB} \geq \sum_{p\in J}|\bar{B}_p| \geq \sum_{p\in J}\sqrt{\delta}|\bar{A}_p| = \sqrt{\delta}\sum_{p\in J}|\bar{A}_p|.$$

So, if (E) were to fail, then $\text{SUMB} \geq 10\delta|K|$, which would contradict (D). ∎

LEMMA 1.2: *For every $\zeta > 0$ and every positive integer k, there exists $\tau > 0$ such that, for all $N = 1, 2, \ldots$,*

$$\sum_{1\leq m\leq \tau N}\binom{N}{m}k^m < 2^{\zeta N}.$$

Proof: This is an easy consequence of Stirling's formula, as both Russell Lyons and Doug Jungreis have pointed out to me. There is also an interesting probabilistic proof of this result which we omit. ∎

We will need a way of coding up quasi-tilings.

Definition 1.3: If $A_1, \ldots, A_k$ is a finite collection of subsets of G, and if $\sigma \in \{0, \ldots, k\}^G$ then the **set of right translates defined by σ via $A_1, \ldots, A_k$** is, by definition, $\{A_i g \mid g \in G, i = \sigma(g) > 0\}$. We denote by $\mathcal{QT}_\delta(K)$ the set of elements $\sigma \in \{0, \ldots, k\}^G$ such that $\sigma(g) = 0$ for $g \notin K$ and such that σ corresponds to a δ-quasi-tiling of K.

If no A_i is a right translate of any other, then we say that $A_1, \ldots, A_k$ is **reduced**, in which case there is a one-to-one correspondence between such σ and such collections of right translates. If $\sigma \in \{0, \ldots, k\}^G$ and $g \in G$, then we define $\sigma g \in \{0, \ldots, k\}^G$ by $\sigma g(g') = \sigma(g' g^{-1})$. If $\sigma \in \{0, \ldots, k\}^G$ and $F \subseteq G$, then we define $\sigma_F \in \{0, \ldots, k\}^G$ by

$$\sigma_F(g) = \begin{cases} \sigma(g), & \text{if } g \in F \\ 0, & \text{if } g \notin F. \end{cases}$$

As in [A, §1], the notation $\mathcal{C}(Z)$ means the set of those configurations of the process Z which occur with positive probability.

THEOREM 1.4: *Let $0 < \delta < 1/2$ and let $A_1, \ldots, A_k$ be a reduced $(\delta^2/9)$-quasi-tiling system for G. For $F \subseteq G$ finite, let* $\mathrm{BDRY}(F)$ *denote the union over $i = 1, \ldots, k$ of the $(A_i A_i^{-1})$-boundaries of F. Let $F \subseteq G$ be a finite subset satisfying $e \in F = F^{-1}$ and*

(A) $|\mathrm{BDRY}(F)| < \delta |F|/3$.

Let $K \subseteq G$ be sufficiently invariant that

(B) *the F-boundary of K has fewer than $5\delta^2 |K|/9$ elements.*

Let $\tilde{k}$ be a random variable taking values in K, with each value having probability $1/|K|$. Let $\sigma \in \mathcal{QT}_{\delta^2/9}(K)$. Then

(C) $\Pr[(\sigma \tilde{k}^{-1})_F \in \mathcal{QT}_\delta(F)] > 1 - \delta$; *and*

(D) *if $M \leq \min\{|A_1|, \ldots, |A_k|\}$, then* $|\mathcal{C}((\sigma \tilde{k}^{-1})_F)| < \displaystyle\sum_{1 \leq m \leq 2|F|/M} \binom{|F|}{m} k^m.$

The point of conclusion (C) is that if we take a very good quasi-tiling of the very large set K, translate it by the inverse of a random element of K and intersect with the fairly large set F, then, with high probability, we will end up with a fairly good quasi-tiling of F. This produces a "random quasi-tiling" of F.

PROOF OF THEOREM 1.4: *Proof of (C).* Let $L := \{g \in G \mid Fg \subseteq K\}$ denote the F-interior of K. Let $\bar{A}_1, \ldots, \bar{A}_m$ correspond to σ. Let $\bar{A} := \bigcup_{p=1}^{m} \bar{A}_p$. Let

$$T := \{k_1 \in K \mid (\sigma k_1^{-1})_F \in \mathcal{QT}_\delta(F)\}.$$

We must show that

(E) $|K\backslash T| < \delta|K|$.

Let

$$S := \{(k_1, g) \in K \times G \,|\, g \in \bar{A} \cap Fk_1\}.$$

For $k_1 \in K$, $g \in G$, let

$$S^{k_1} := \{g \in G \,|\, (k_1, g) \in S\}, \quad S_g := \{k_1 \in K \,|\, (k_1, g) \in S\}.$$

Note that,

(E') for all $k_1 \in K$, $S^{k_1} = \bar{A} \cap Fk_1$.

By definition of F-interior, for all $g \in L$, $Fg \subseteq K$. Recall that $F = F^{-1}$; thus, for all $g \in \bar{A} \cap L$,

$$S_g = \{k_1 \in K \,|\, g \in Fk_1\} = F^{-1}g \cap K = Fg \cap K = Fg,$$

so $|S_g| = |F|$. Summing this over $g \in \bar{A} \cap L$,

(F) $|S| \geq |\bar{A} \cap L||F|$.

Now $K\backslash L$ is contained in the F-boundary of K, so, by (B),

$$|K\backslash L| < (5\delta^2/9)|K|.$$

Also, by definition of a $(\delta^2/9)$-quasi-tiling, $|K\backslash\bar{A}| < (\delta^2/9)|K|$, so

$$|K| - |\bar{A} \cap L| = |K\backslash(\bar{A} \cap L)| = |(K\backslash\bar{A}) \cup (K\backslash L)| < \left[\frac{\delta^2}{9} + \frac{5\delta^2}{9}\right]|K| = \frac{2\delta^2}{3}|K|.$$

So $|\bar{A} \cap L| > [1 - (2\delta^2/3)]|K|$. We therefore conclude from (F) that

(G) $|S| > [1 - (2\delta^2/3)]|F||K|$.

By (E'), for all $k_1 \in K$, $S^{k_1} = \bar{A} \cap Fk_1 \subseteq Fk_1$, so $|Fk_1\backslash S^{k_1}| = |Fk_1| - |S^{k_1}| = |F| - |S^{k_1}|$. Summing this over $k_1 \in K$ and using (G), we have

$$\sum_{k_1 \in K} |Fk_1\backslash S^{k_1}| = |F||K| - |S| < \frac{2\delta^2}{3}|F||K|.$$

Therefore, if

$$R := \left\{k_1 \in K \,\middle|\, |Fk_1\backslash S^{k_1}| \geq \frac{2\delta}{3}|F|\right\},$$

then $|R| < \delta|K|$. To prove (E), we will show that $K\backslash T \subseteq R$, i.e., that $K \subseteq T \cup R$, i.e., that $K\backslash R \subseteq T$.

So fix $k_1 \in K\backslash R$. To show that $k_1 \in T$, i.e., that $(\sigma k_1^{-1})_F \in \mathcal{QT}_\delta(F)$, it suffices, by right translation, to show $\sigma_{Fk_1} \in \mathcal{QT}_\delta(Fk_1)$. Reorder $\bar{A}_1, \ldots, \bar{A}_m$ so that the collection $\bar{A}_1, \ldots, \bar{A}_t$ corresponds to σ_{Fk_1}, while $\bar{A}_{t+1}, \ldots, \bar{A}_m$ corresponds to $\sigma_{G\backslash Fk_1}$. Since the larger collection $\bar{A}_1, \ldots, \bar{A}_m$ is δ-disjoint, it is immediate that the smaller one $\bar{A}_1, \ldots, \bar{A}_t$ is as well. Thus, it remains to show that the sets $\bar{A}_1, \ldots, \bar{A}_t$ cover all but a δ-fraction of Fk_1, i.e., that

(H) $|Fk_1\backslash(\bar{A}_1 \cup \cdots \cup \bar{A}_t)| < \delta|Fk_1|$.

Now $S^{k_1} = \bar{A} \cap Fk_1$, so $Fk_1\backslash\bar{A} = Fk_1\backslash(\bar{A} \cap Fk_1) = Fk_1\backslash S^{k_1}$. Since $k_1 \notin R$, this implies that

(I) $|Fk_1\backslash\bar{A}| < 2\delta|F|/3$.

Fix any integer p in the range $t+1 \leq p \leq m$. Since $\bar{A}_{t+1}, \ldots, \bar{A}_m$ corresponds to $\sigma_{G\backslash Fk_1}$, we find that $\bar{A}_p \cap (G\backslash Fk_1) \neq \emptyset$. So, by definition of BDRY(Fk_1), either $\bar{A}_p \cap Fk_1 = \emptyset$ or $\bar{A}_p \subseteq$ BDRY(Fk_1). In either case, $\bar{A}_p \cap Fk_1 \subseteq$ BDRY(Fk_1). This holds for $p = t+1, \ldots, m$, so $Fk_1 \cap (\bar{A}_{t+1} \cup \cdots \cup \bar{A}_m) \subseteq$ BDRY(Fk_1).

Now let $Q := Fk_1\backslash(\bar{A}_1 \cup \cdots \cup \bar{A}_t)$. Then

$$\begin{aligned} Q &= [Q\backslash(\bar{A}_{t+1} \cup \cdots \cup \bar{A}_m)] \cup [Q \cap (\bar{A}_{t+1} \cup \cdots \cup \bar{A}_m)] \\ &\subseteq [Fk_1\backslash(\bar{A}_1 \cup \cdots \cup \bar{A}_m)] \cup [Fk_1 \cap (\bar{A}_{t+1} \cup \cdots \cup \bar{A}_m)] \\ &\subseteq (Fk_1\backslash\bar{A}) \cup [\mathrm{BDRY}(Fk_1)]. \end{aligned}$$

By translation invariance $|\mathrm{BDRY}(Fk_1)| = |\mathrm{BDRY}(F)|$, so, by (I) and (A),

$$|Q| \leq |Fk_1\backslash\bar{A}| + |\mathrm{BDRY}(F)| < \frac{2\delta|F|}{3} + \frac{\delta|F|}{3} = \delta|F|.$$

By definition of the set Q, this gives (H), finishing the proof of (C).

Proof of (D): Fix $k_1 \in K$. Let SUPP $:= \{g \in F \,|\, \sigma(gk_1) > 0\}$ denote the support of $(\sigma k_1^{-1})_F$. The sets corresponding to $(\sigma k_1^{-1})_F$ form a $(\delta^2/9)$-disjoint collection of subsets; further, they are all contained in $U := F \cup \mathrm{BDRY}(F)$ and they all have cardinality $\geq \min\{|A_1|, \ldots, |A_k|\} \geq M$. By definition of $(\delta^2/9)$-disjointness, there is then a disjoint collection of subsets of U, indexed by SUPP, each of size $\geq M[1-(\delta^2/9)]$. Thus $|\mathrm{SUPP}|M[1-(\delta^2/9)] \leq |U|$. By (A), $|U| < |F|(1+\delta)$, so

$$|\mathrm{SUPP}| < \frac{|F|}{M}\frac{1+\delta}{1-(\delta^2/9)}.$$

Since $\delta < 1/2$, we conclude that $|\mathrm{SUPP}| < (|F|/M)(3/2)(35/36)^{-1} < 2|F|/M$.

Thus each configuration of $(\sigma\tilde{k}^{-1})_F$ has fewer than $2|F|/M$ non-zero values. These values can range through the set $\{1,\dots,k\}$, so

$$\left|\mathcal{C}((\sigma\tilde{k}^{-1})_F)\right| < \sum_{1\leq m\leq 2|F|/M} \binom{|F|}{m} k^m. \qquad \blacksquare$$

2. Finitely Determined and finite processes

Again, G denotes a discrete, countable, amenable group with identity element e.

As in [A, Definition 3.6], two stationary G-spin systems X and X' are (δ, F)-**close in entropy and finite distribution** if

$$|H(X) - H(X')| < \delta \text{ and } d(X_F, X'_F) < \delta.$$

As in [A, Definition 3.7], a stationary G-spin system X is **Finitely Determined** if, for all ϵ, there exists (δ, F) such that: any stationary G-spin system X' which is (δ, F)-close to X in entropy and finite distribution satisfies $\bar{d}(X_K, X'_K) < \epsilon$, for all sufficiently invariant finite subsets $K \subseteq G$.

In this section, we describe a way of comparing two finite processes in entropy and distribution. We use this to give a criterion for a G-process to be Finitely Determined in terms of the finite subprocesses of the process obtained by restricting to Følner sets (see Theorem 2.8).

We define a right action of G on $\{+1,-1\}^G$ by $(\sigma g)(g') = \sigma(g'g^{-1})$. In the following definition, if $F, K \subseteq G$ are finite and if $\sigma \in \{+1,-1\}^K$, then we define $\sigma_F^+ \in \{+1,-1\}^F$ by

$$\sigma_F^+(g) = \begin{cases} \sigma(g), & \text{if } g \in K \cap F \\ +1, & \text{if } g \in F \backslash K. \end{cases}$$

Note that if $g \in G$, if $F, K \subseteq G$ are finite and if $\sigma \in \{+1,-1\}^K$, then $\sigma g \in \{+1,-1\}^{Kg}$, so $(\sigma g)_F^+ \in \{+1,-1\}^F$ is defined by

$$(\sigma g)_F^+(g') = \begin{cases} \sigma(g'g^{-1}), & \text{if } g' \in (Kg) \cap F \\ +1, & \text{if } g' \in F \backslash (Kg). \end{cases}$$

Definition 2.1: Let $F, K \subseteq G$ be finite. Let $Y := \{Y_k\}_{k\in K}$ be a K-spin system. Let $\tilde{k}$ be a K-valued random variable taking each value of K with probability $1/|K|$. Couple Y and $\tilde{k}$ independently. Then we define $\operatorname*{Avg}_F(Y) := (Y\tilde{k}^{-1})_F^+$.

Recall from [A, end of §1 and Definition 2.4] the definitions of the entropy $H(X)$ of a finite process X and of the d metric for finite processes.

Definition 2.2: Let $F, K \subseteq G$ be finite. Let X and Y be K-spin systems. We say that X and Y are **(δ, F)-close in entropy and finite distribution** if

(1) $|H(X) - H(Y)| < \delta|K|$; and

(2) $d\Big(\operatorname*{Avg}_F(X), \operatorname*{Avg}_F(Y)\Big) < \delta$.

There is the possibility of confusion here in case G is a finite group and $K = G$, since $H(X)$ then has two possible interpretations, one as the entropy of a group-process (see [A, Lemma 3.5]), the other as the entropy of a set-process (see [A, §1]). In (1), above, $H(X)$ and $H(Y)$ are to be interpreted as entropies of *set*-processes. By constrast, in (1) of [A, Definition 3.6], $H(X)$ and $H(X')$ are to be interpreted as entropies of *group*-processes. Furthermore, if G is finite and $K = G$, then, for any $F \subseteq G$, we have $\operatorname*{Avg}_F(X) = X_F$. So there is no inconsistency between [A, Definition 3.6] and Definition 2.2 above.

Recall that if Z is any process, then $\mathcal{C}(Z)$ denotes the set of configurations of Z which occur with positive probability.

LEMMA 2.3: *If K is any set, if Z a K-process and if $\mathcal{C}(Z)$ is a finite set, then $H(Z) \leq \log_2 |\mathcal{C}(Z)|$.*

Proof: This is a basic property of entropy. ∎

Recall the d and $\bar{d}$ metrics of [A, Definition 2.4].

LEMMA 2.4: *Let I be any set. Let X be an I-process and let $\tau > 0$. Let $C \subseteq \mathcal{C}(X)$ satisfy $\Pr[X \notin C] < \tau$. Let $Y := X|(X \in C)$ denote the conditional process of X conditioned on $X \in C$. Then $d(X, Y) < \tau$.*

Proof: For every $\sigma \in C$, we have $\Pr[X = \sigma] = (1 - \tau)\Pr[Y = \sigma] \leq \Pr[Y = \sigma]$. Therefore there exists a coupling m of X and Y under which: for every $\sigma \in C$, $X = \sigma \Rightarrow Y = \sigma$. Under this coupling, $X \neq Y \Rightarrow X \notin C$, so $\Pr^m[X \neq Y] < \tau$, as desired. ∎

LEMMA 2.5: *Let $\delta > 0$, let K be a finite set and let X, Y and Z be K-processes. Let c be a coupling of X and Y. For all $\sigma \in \mathcal{C}(Y)$, let X_σ denote the conditional process $X|(Y = \sigma)$ with respect to c. If, for all $\sigma \in \mathcal{C}(Y)$, $\bar{d}(X_\sigma, Z) < \delta$, then $\bar{d}(X, Z) < \delta$.*

Proof: Any family $\{m_\sigma\}_{\sigma \in \mathcal{C}(Y)}$ of couplings of X_σ with Z induces a coupling m of X with Z. Let $\mathcal{P}_\sigma := \Pr[Y = \sigma]$, for all $\sigma \in \mathcal{C}(Y)$. Then

$$\bar{d}(X, Z) \leq \bar{d}^m(X, Z) = \sum_{\sigma \in \mathcal{C}(Y)} \mathcal{P}_\sigma \bar{d}^{m_\sigma}(X_\sigma, Z).$$

The result now follows by taking the infimum over families $\{m_\sigma\}$. ∎

Definition 2.6: If K is a finite set and X is a K-process, then we say that X is (ϵ,δ)-**extremal** if: for all K-processes Z satisfying $\log_2 |\mathcal{C}(Z)| < \delta|K|$, for all couplings of X with Z, we have

$$\sum_{\sigma\in\mathcal{C}(Z)} \Pr[Z=\sigma]\bar{d}(X|Z=\sigma, X) < \epsilon.$$

The inequality in Definition 2.6 is meant to indicate that X and Z are almost independent in the process sense. In fact, if X and Z were distinct, then this inequality would be equivalent to saying that X and Z were ϵ-process independent (see [A, Definition 2.8] and [A, Lemma 2.6]).

THEOREM 2.7: *If X is a Finitely Determined stationary G-spin system, then, for all $\epsilon > 0$, there exists $\delta > 0$ such that: for any sufficiently invariant $K \subseteq G$, the process X_K is (ϵ,δ)-extremal.*

Proof: This is a rephrasing of [O-W, Proposition 8, p. 117]. ∎

The following is the main result of this section.

THEOREM 2.8: *Let X be a Finitely Determined stationary G-spin system. For every $\epsilon > 0$, there exists a $\delta > 0$ and a finite subset $F_0 \subseteq G$ such that: if K is sufficiently invariant and if Y is a K-process which is (δ, F_0)-close to X_K in entropy and finite distribution (Definition 2.2), then $\bar{d}(X_K, Y) < \epsilon$.*

Proof: The proof of this theorem occupies the remainder of this section. We divide it into five steps.

STEP 1: *The construction of Y.* We may assume $\epsilon < 1/2$. Then

(1A) $\epsilon^8 + \epsilon^4 < \epsilon < 1$.

As X is Finitely Determined ([A, Definition 3.7]), we may choose $\eta > 0$ and $F_0 \subseteq G$ finite such that:

(1B) if Z is a G-process which is (η, F_0)-close to X in entropy and finite distribution (see [A, Definition 3.6]), then $\bar{d}(X_K, Z_K) < \epsilon^8$, for all sufficiently invariant $K \subseteq G$.

Replacing F_0 by $F_0 \cup \{e\} \cup F_0^{-1}$, we may assume that $e \in F_0 = F_0^{-1}$. Using (1A), we may choose δ satisfying:

(1C1) $0 < \delta < 1/100$;

(1C2) $10\sqrt{\delta} + \sqrt{(\delta + \epsilon^4 + \epsilon^8)/(1-\delta)^2} < 1$;

(1C3) $\delta + \sqrt{\delta} + \sqrt{(\delta + \epsilon^4 + \epsilon^8)/(1-\delta)^2} < \epsilon$;

(1C4) $5\delta H(X) + 19\delta < \eta$; and

(1C5) $13\delta + 20\sqrt{\delta} + 2\sqrt{\delta}(1 + |F_0 F_0^{-1}|) < \eta$.

It follows from (1C2) that

(1C6) $\delta + \epsilon^4 + \epsilon^8 < 1$.

By Theorem 2.7, we may choose δ so small that

(1C7) for all sufficiently invariant $K \subseteq G$, X_K is (ϵ^{16}, δ)-extremal.

Let $\mathcal{F}$ denote the collection of all subsets $A \subseteq G$ such that:

(1D1) there exists an A-process Y^A such that X_A and Y^A are (δ, F_0)-close in entropy and finite distribution (Definition 2.2); and

(1D2) $\bar{d}(X_A, Y^A) \geq \epsilon$.

To prove Theorem 2.8, we need to show that $\mathcal{F}$ is *not* a Følner family (defined at the beginning of [A, §3]). We assume **for a contradiction** that $\mathcal{F}$ is a Følner family.

A Følner family $\mathcal{G}$ will be said to be **translation-invariant** if, for all $A \in \mathcal{G}$, $g \in G$, we have $Ag \in \mathcal{G}$. By the proof of [O-W, Theorem 6, p. 24], we may choose a positive integer k such that:

(1E) any translation-invariant Følner family contains a $(\delta^2/9)$-quasi-tiling system with k sets.

By Lemma 1.2, choose $M > 0$ such that:

(1F) $$\sum_{1 \leq m \leq 2N/M} \binom{N}{m} k^m < 2^{\delta N}, \text{ for all } N = 1, 2, \ldots.$$

Let $\mathcal{G}$ denote the set of $A \in \mathcal{F}$ such that

(1G1) $(1-\delta)|A| > M$;

(1G2) the $(F_0 F_0^{-1})$-boundary of A has cardinality $< \sqrt{\delta}|A|$; and

(1G3) $\Big| H(X_A) - |A| H(X) \Big| < \delta |A|$.

Since $\mathcal{F}$ is a Følner family, $\mathcal{G}$ is again a Følner family. All of the conditions (1D1), (1D2), (1G1), (1G2), (1G3) are translation-invariant, so, by (1E), we may choose a finite subsets $A_1, \ldots, A_k \subseteq G$ such that:

(1H1) for $i \in \{1, \ldots, k\}$, the set $A = A_i$ satisfies (1D1), (1D2), (1G1), (1G2) and (1G3); and

(1H2) $A_1, \ldots, A_k$ is a $(\delta^2/9)$-quasi-tiling system for G.

By translation-invariance, we conclude from (1H1) that

(1H3) if A is a right translate of some A_i, then A satisfies (1D1), (1D2), (1G1), (1G2) and (1G3).

By definition of a quasi-tiling system, $e \in A_1 \subseteq A_2 \subseteq \cdots \subseteq A_k$, so

(1H4) $e \in A_1 \cap \cdots \cap A_k$.

Since $A_1 \subseteq \cdots \subseteq A_k$ the system $A_1, \ldots, A_k$ is reduced. Hence the $(A_1, \ldots, A_k)$-correspondence is well-defined. (See the text following Definition 1.3.)

For $i = 1, \ldots, k$, let $A = A_i \in \mathcal{G} \subseteq \mathcal{F}$, choose Y^A as in (1D1) and set $Y^i := Y^A$. Then, for all $i = 1, \ldots, k$, we have:

(1I) X_{A_i} and Y^i are (δ, F_0)-close in entropy and distribution.

Let $K_1, K_2, \ldots$ be any **Følner sequence in** G, i.e., a sequence such that

(1J) if $\eta_1 > 0$, and $F_1 \subseteq G$ is finite, then K_s is (η_1, F_1)-invariant, for all but a finite number of s.

(It follows immediately from the definition of amenability that Følner sequences exist.) For $s = 1, 2, \ldots$, let $\tilde{k}_s$ be a random variable taking values in K_s, with each element of K_s having probability $1/|K_s|$. Choose a $(\delta^2/9)$-quasi-tiling of K_s by right translates of $A_1, \ldots, A_k$ and let this quasi-tiling be represented by $\sigma_s \in \{0, \ldots, k\}^G$ in the $(A_1, \ldots, A_k)$-correspondence.

Fix a positive integer s. Let $\bar{A}_1, \ldots, \bar{A}_m$ correspond to σ_s. Let K_0 be the set of $k \in K$ such that k lies in a unique $\bar{A}_p$. If $\tau_p \in \{+1, -1\}^{\bar{A}_p}$, for $p = 1, \ldots, m$ then define $\left(\mathop{*}_{p=1}^{m} \tau_p\right) \in \{+1, -1\}^G$ by:

$$\left(\mathop{*}_{p=1}^{m} \tau_p\right)(k) = \begin{cases} \tau_p(k), & \text{if } k \in K_0 \cap \bar{A}_p \\ +1, & \text{if } k \in G \backslash K_0. \end{cases}$$

Now recall that G acts on $\{+1, -1\}^G$ via $(\sigma g)(g') = \sigma(g' g^{-1})$. For each $p = 1, \ldots, m$, choose an integer i, $1 \le i \le k$, and a $g \in G$ such that $\bar{A}_p = A_i g$; then define $\bar{Y}^p := Y^i g$. Let Y_s denote the G-process $\left(\mathop{*}_{p=1}^{m} \bar{Y}^p\right)$.

For each $s = 1, 2, \ldots$, we couple Y_s and $\tilde{k}_s$ independently. Recall from [A] the notation: if $F \subseteq F' \subseteq G$ and if $Z = \{Z_g\}_{g \in F'}$ is any F'-process, then $Z_F := \{Z_f\}_{f \in F}$ denotes the restriction of Z to F. For each finite $F \subseteq G$, we obtain a sequence of F-processes $(Y_s \tilde{k}_s^{-1})_F$. Each of these processes is coupled with $(\sigma_s \tilde{k}_s^{-1})_F$ in an obvious way; call this coupling m_s. Using a Cantor diagonalization argument, we may pass to a subsequence of $K_1, K_2, \ldots$ and assume that: for each finite $F \subseteq G$, as $s \to \infty$, the processes $(\sigma_s \tilde{k}_s^{-1})_F$ and $(Y_s \tilde{k}_s^{-1})_F$ both converge. Call the limits of these two sequences Q^F and Y^F, respectively. The couplings m_s (after passing to another subsequence) tend toward a coupling m^F of Q^F and Y^F. It is routine to check the consistency conditions of the Kolmogorov existence theorem, which shows that there exist G-processes Q and Y

such that $Q_F = Q^F$ and $Y_F = Y^F$, for all finite $F \subseteq G$. This existence theorem also gives a coupling m of Q with Y. It is straightforward to check that Q and Y are stationary (under right translation by elements of G).

Intuitively, Q is a random quasi-tiling and Y is a process which is obtained by running Q and then independently running the translated Y^i-processes corresponding to each tile of the output of Q. Since we have a quasi-tiling and not a perfect tiling, there may be some elements of G covered by no tiles or covered by more than one of the tiles. We assign the output of the Y-process to be $+1$ at such ambiguous points.

Steps 3, 4 and 5 verify the following three statements:

(1M1) if K is sufficiently invariant and $e \in K = K^{-1}$, then $|H(Y_K) - H(X_K)| < \eta|K|$;

(1M2) $d(X_{F_0}, Y_{F_0}) < \eta$; and

(1M3) the finite sets $K \subseteq G$ for which $e \in K = K^{-1}$ and $\bar{d}(X_K, Y_K) \geq \epsilon^8$ form a Følner family.

These together will **contradict** (1B), and will prove the theorem.

STEP 2: *Preliminaries.* As in Theorem 1.4, for any finite subset $F \subseteq G$, let $\mathrm{BDRY}(F)$ denote the union of the $(A_i A_i^{-1})$-boundaries of F, over $i = 1, \ldots, k$ and let $\mathcal{QT}_\delta(F)$ denote the set of elements of $\{0, \ldots, k\}^G$ which correspond to δ-quasi-tilings of F under the $(A_1, \ldots, A_k)$-correspondence (Definition 1.3).

By conclusion (C) of Theorem 1.4: for all sufficiently invariant $F \subseteq G$, if $e \in F = F^{-1}$, then

$$\Pr[(\sigma_s \tilde{k}_s^{-1})_F \in \mathcal{QT}_\delta(F)] > 1 - \delta,$$

for all sufficiently large integers s. Consequently, this is true in the limit $Q^F = Q_F$, so we conclude:

(2K1) if $e \in F = F^{-1}$ and $F \subseteq G$ is sufficiently invariant, then $\Pr[Q_F \in \mathcal{QT}_\delta(F)] > 1 - \delta$.

Further, $\sigma_s \tilde{k}_s | F$ almost surely corresponds to a collection of sets which are δ-**disjoint** (i.e., such that each set can be shrunk by a δ-fraction with the resulting sets pairwise-disjoint, see the start of §1). Further, if a right translate of an A_i meets F, then, using (1H4), we see that it is contained in $F \cup \mathrm{BDRY}(F)$. So:

(2K2) for all finite $F \subseteq G$, Q_F defines a δ-disjoint collection of subsets of $F \cup \mathrm{BDRY}(F)$, almost surely.

We **claim** that

(2L) if $K \subseteq G$ is sufficiently invariant, then $\log_2 |\mathcal{C}(Q_K)| < \delta|K|$.

In fact, by [A, Lemma 3.2], choose K sufficiently invariant that $|\mathrm{BDRY}(K)| < |K|$. If $\bar{A}_1, \dots, \bar{A}_m$ is a δ-disjoint collection of right translates of $A_1, \dots, A_k$ and if $\bar{A}_p \cap K \neq \emptyset$, for $p = 1, \dots, m$, then each $\bar{A}_p$ may be shrunk by a δ-fraction with the resulting sets forming a disjoint collection, each one a subset of $K \cup \mathrm{BDRY}(K)$. Using (1G1) and the fact that $\delta < 1/100$ (see (1C1)), we then have

$$mM \leq |\bar{A}_1|(1-\delta) + \cdots + |\bar{A}_m|(1-\delta) < |K \cup \mathrm{BDRY}(K)| < |K| + |K| = 2|K|.$$

Then, by (2K2), each configuration of Q_K has fewer than $2|K|/M$ non-zero values. These values can range through the set $\{1, \dots, k\}$, so

$$|\mathcal{C}(Q_K)| < \sum_{1 \leq m \leq 2|K|/M} \binom{|K|}{m} k^m.$$

By (1F), $\log_2 |\mathcal{C}(Q_K)| < \delta|K|$, proving the claim.

STEP 3: *The entropy calculation.* In this step, we verify (1M1).

It is easily seen that the collection of finite subsets $K \subseteq G$ satisfying $e \in K = K^{-1}$ is a Følner collection. So, by (2K1), choose $e \in K = K^{-1}$ sufficiently invariant that

(3A1) $\Pr[Q_K \in \mathcal{QT}_\delta(K)] > 1 - \delta$.

Recall that $\mathrm{BDRY}(K)$ denotes the union of the $(A_i A_i^{-1})$-boundaries of K, over $i = 1, \dots, k$. By [A, Lemma 3.2], choose $K \subseteq G$ sufficiently invariant that

(3A2) $|\mathrm{BDRY}(K)| < \delta|K|$.

By conclusion (D) of Lemma 1.1,

(3A3) in any δ-quasi-tiling of K, the cardinality of the collection of points lying in two or more tiles is $< 10\delta|K|$.

Recall that the number M was chosen so that (1F) holds. We now assume that K is sufficiently invariant that

(3A5) $M|\mathrm{BDRY}(K)| < |K|$.

By (2L) and Lemma 2.3, we may choose K sufficiently invariant that

(3A6) $H(Q_K) < \delta|K|$.

Finally, by the defnition of entropy ([A, Lemma 3.5]), assume K is sufficiently invariant that

(3A7) $\left|H(X_K) - |K|H(X)\right| < \delta|K|$.

Let $L := K \backslash \mathrm{BDRY}(K)$. Since $K \backslash L \subseteq \mathrm{BDRY}(K)$, it follows from (3A2) that

(3A8) $0 < |K\backslash L| < \delta|K|$.

To finish Step 3, we must now show that

(3B) $|H(Y_K) - H(X_K)| < \eta|K|$.

Fix $\sigma \in \mathcal{C}(Q_K)$. Let Y'_K be the conditional process $Y_K|(Q_K = \sigma)$ and let Y'_L be the conditional process $Y_L|(Q_K = \sigma)$. Then

$$H(Y'_K) - H(Y'_L) = H(Y'_{K\backslash L}|Y'_L) \leq H(Y'_{K\backslash L}) \leq |K\backslash L|.$$

For $\sigma \in \mathcal{C}(Q_K)$, let $H_\sigma := H(Y_K|Q_K = \sigma)$. Thus, by (3A8), for all $\sigma \in \mathcal{C}(Q_K)$,

(3C) $\left|H_\sigma - H(Y_L|Q_K = \sigma)\right| < \delta|K|$.

Now fix any $\sigma \in QT_\delta(K)$. Let $\bar{A}_1, \ldots, \bar{A}_m$ be the sets corresponding to σ. Recall that G acts on $\{+1,-1\}^G$ via $(\sigma g)(g') = \sigma(g'g^{-1})$. For $p = 1, \ldots, m$, write $\bar{A}_p = A_i g$ and define $\bar{Y}^p := Y^i g$. Let $H_p := H(\bar{Y}^p)$. The conditional process $Y_L|(Q_K = \sigma)$ is a factor of the joint process obtained by independently coupling $\bar{Y}^1, \ldots, \bar{Y}^m$: one runs $\bar{Y}^1, \ldots, \bar{Y}^m$ independently, then sets to $+1$ the values at elements of G which lie in more than one of the tiles $\bar{A}_1, \ldots, \bar{A}_m$, and then ignores the values in BDRY(K). Thus, by (3A2) and (3A3),

(3D) $\left|H_1 + \cdots + H_m - H(Y_L|Q_K = \sigma)\right| < 10\delta|K| + \delta|K| + \delta|K|$.

Let $H := H(X)$. For $p = 1, \ldots, m$, by (1H3), (1D1) and (1G3), $|H(\bar{Y}^p) - H(X_{\bar{A}_p})| < \delta|\bar{A}_p|$ and $\left|H(X_{\bar{A}_p}) - |\bar{A}_p|H\right| < \delta|\bar{A}_p|$, so

(3E) $\left|H_p - |\bar{A}_p|H\right| < 2\delta|\bar{A}_p|$.

Let SUMA $:= |\bar{A}_1| + \cdots + |\bar{A}_m|$. Since $\delta < 1/100$ (see (1C1)), conclusion (C) of Lemma 1.1 implies that SUMA $\leq |K| + 4\delta|K| < 2|K|$, so adding the inequalities of (3E) for $p = 1, \ldots, m$ gives

(3F) $\left|H_1 + \cdots + H_m - H[\text{SUMA}]\right| < 4\delta|K|$.

Finally, by conclusion (C) of Lemma 1.1, $\left|\text{SUMA} - |K|\right| < 4\delta|K|$, so

(3G) $\left|H[\text{SUMA}] - H|K|\right| < 4\delta|K|H$.

Combining (3C), (3D), (3F) and (3G) we see that if $\sigma \in QT_\delta(K)$, then

(3H) $\left|H_\sigma - H|K|\right| < 16\delta|K| + 5\delta|K|H$.

Let $QT := QT_\delta(K)$. For each $\sigma \in \mathcal{C}(Q_K)$, let $\mathcal{P}_\sigma := \Pr[Q_K = \sigma]$. By (3A1),

(3I) $0 \leq \sum_{\sigma \notin QT} \mathcal{P}_\sigma \leq \delta$.

By [A, Lemma 2.22], $H_\sigma \leq |K|$, so (3I) implies

(3J) $0 \leq \sum_{\sigma \notin QT} \mathcal{P}_\sigma \cdot H_\sigma < \delta|K|$.

By (3H) and the fact that $0 \le \sum_{\sigma \in QT} \mathcal{P}_\sigma \le 1$, we get

(3K) $\sum_{\sigma\in QT} \mathcal{P}_\sigma \cdot \Big| H_\sigma - |K|H \Big| \le 16\delta|K| + 4\delta|K|H.$

The basic properties of entropy imply that $H(Y_K|Q_K) = \sum_\sigma \mathcal{P}_\sigma \cdot H_\sigma$. Thus, using (3I), (3J) and (3K), we conclude

$$\begin{aligned}
\Big|H(Y_K|Q_K) - |K|H\Big| &\le \left|\left[\sum_{\sigma\in QT} \mathcal{P}_\sigma \cdot H_\sigma\right] - |K|\cdot H\right| + \left|\sum_{\sigma\notin QT} \mathcal{P}_\sigma \cdot H_\sigma\right| \\
&< \left|\left[\sum_{\sigma\in QT} \mathcal{P}_\sigma \cdot H_\sigma\right] - \left[\sum_{\sigma\in QT} \mathcal{P}_\sigma\right]\cdot|K|\cdot H + \left[1-\sum_{\sigma\in QT} \mathcal{P}_\sigma\right]\cdot|K|\cdot H\right| + \delta|K| \\
&\le \left[\sum_{\sigma\in QT} \mathcal{P}_\sigma \cdot \Big|H_\sigma - |K|H\Big|\right] + \left[\sum_{\sigma\notin QT} \mathcal{P}_\sigma\right]|K|H + \delta|K| \\
&< (16\delta|K| + 4\delta|K|H) + \delta|K|H + \delta|K|.
\end{aligned}$$

Thus

(3L) $\Big|H(Y_K|Q_K) - |K|H\Big| < 17\delta|K| + 5\delta|K|H.$

The basic properties of conditional entropy imply that $H(Y_K) = H(Q_K) + H(Y_K|Q_K)$. By (3A6), (3L), (3A7) and (1C4), we conclude

$$\begin{aligned}
|H(Y_K) - H(X_K)| &= |H(Q_K) + [H(Y_K|Q_K) - H(X_K)]| \\
&\le H(Q_K) + \Big|H(Y_K|Q_K) - |K|H\Big| + \Big||K|H - H(X_K)\Big| \\
&< \delta|K| + \big(17\delta|K| + 5\delta|K|H\big) + \delta|K| \\
&= (5\delta H + 19\delta)|K| < \eta|K|,
\end{aligned}$$

establishing (3B), as desired. ∎

STEP 4: *The distribution calculation.* In this step, we will verify (1M2).

By (2K1), choose K satisfying $e \in K = K^{-1}$ sufficiently invariant that

(4A1) $\Pr[Q_K \in \mathcal{QT}_\delta(K)] > 1 - \delta.$

Recall that if $K \subseteq G$ is finite, then $\mathrm{BDRY}(K)$ denotes the union over $i = 1, \dots, k$ of the $(A_i A_i^{-1})$-boundaries of K. By [A, Lemma 3.2], choose K sufficiently invariant that

(4A2) $|\mathrm{BDRY}(K)| < \delta|K|.$

Let $\tilde{k}$ be a K-valued random variable, with each value of K having probability $1/|K|$. Couple $\tilde{k}$ independently with the joint process (Q, Y). (Note that there is, by construction in Step 1, a coupling m of Q and Y.)

As before, G acts on the right of $\{+1,-1\}^G$ via $(\sigma g)(g') = \sigma(g'g^{-1})$. Recall that if $Z = \{Z_g\}_{g\in G}$ is a G-process and if $S \subseteq G$, then Z_S denotes the restriction $\{Z_s\}_{s\in S}$ of Z to S. Since Y is stationary, we conclude that Y is isomorphic to $Y\tilde{k}^{-1}$, so Y_{F_0} is isomorphic to $(Y\tilde{k}^{-1})_{F_0}$. Thus, to conclude this step, it suffices to show that

$$d\big(X_{F_0}, (Y\tilde{k}^{-1})_{F_0}\big) < \eta.$$

Let $(\tilde{k}_1, Y_1, Q_1)$ be the conditional joint process $(\tilde{k}, Y, Q)$ conditioned on

(4B) $Q_K \in \mathcal{QT}_\delta(K)$.

By (4A1), the event of (4B) fails with probability $< \delta$, so, by Lemma 2.4, it suffices to show

$$d\big(X_{F_0}, (Y_1\tilde{k}_1^{-1})_{F_0}\big) < \eta - \delta.$$

Fix any $\sigma_0 \in \mathcal{QT}_\delta(K)$. Let $\bar{A}_1, \ldots, \bar{A}_m$ denote the sets in this δ-quasi-tiling. By (1H3), for $p = 1, \ldots, m$,

(4C) the set $A = \bar{A}_p$ satisfies (1D1), (1D2), (1G1), (1G2) and (1G3).

Let $(\tilde{k}_2, Y_2, Q_2)$ denote $(\tilde{k}_1, Y_1, Q_1)$ conditioned on

(4D) $(Q_1)_K = \sigma_0$.

By Lemma 2.5, it suffices to show that

$$d\big(X_{F_0}, (Y_2\tilde{k}_2^{-1})_{F_0}\big) < \eta - \delta.$$

Then $\tilde{k}$, $\tilde{k}_1$ and $\tilde{k}_2$ are all isomorphic, so $\tilde{k}_2$ takes each value of K with probability $1/|K|$. Further Y_2 is isomorphic to the conditional process $Y|(Q_K = \sigma)$. Finally, $\tilde{k}_i$ and Y_i are independent, for $i = 1, 2$.

Let K_0 denote the set of elements of K which are contained in exactly one $\bar{A}_p$. Let $(\tilde{k}_3, Y_3, Q_3)$ denote $(\tilde{k}_2, Y_2, Q_2)$ conditioned on

(4E) $\tilde{k}_2 \in K_0$.

Note that $\tilde{k}_3$ takes each value of K_0 with probability $1/|K_0|$. The inequality

$$\left| K \setminus \bigcup_{p=1}^{m} \bar{A}_p \right| < \delta |K|$$

in conjunction with conclusion (D) of Lemma 1.1 allows us to conclude that $|K\setminus K_0| < 10\delta|K| + \delta|K|$, so

(4E') $|K\backslash K_0| < 11\delta|K|$.

Thus (4E) fails with probability $< 11\delta$. Thus, by Lemma 2.4, it suffices to show that

$$d(X_{F_0}, (Y_3\tilde{k}_3^{-1})_{F_0}) < \eta - \delta - 11\delta = \eta - 12\delta.$$

Let J denote the set of p, $1 \leq p \leq m$ such that

$$\left|\bar{A}_p \cap \left(\bigcup_{p' \neq p} \bar{A}_{p'}\right)\right| \leq \sqrt{\delta}|\bar{A}_p|.$$

Let $\tilde{p}$ be the $\{1, \ldots, m\}$-valued random variable defined by $\tilde{k}_3 \in \bar{A}_{\tilde{p}}$. Note that, for all $p = 1, \ldots, m$,

$$\Pr[\tilde{p}] = \frac{|K_0 \cap \bar{A}_p|}{|K_0|}.$$

Next, let $(\tilde{k}_4, Y_4, Q_4)$ denote $(\tilde{k}_3, Y_3, Q_3)$ conditioned on

(4F) $\tilde{p} \in J$.

Note that $\tilde{k}_4$ takes each value of $K_0 \cap (\bigcup_{p \in J} \bar{A}_p)$ with equal probability. By (4E'), we have $|K|(1 - 11\delta) < |K_0|$. So, by conclusion (E) of Lemma 1.1, we conclude that (4F) fails with probability

$$< (10\sqrt{\delta}|K|)/|K_0| \leq 10\sqrt{\delta}/(1 - 11\delta) < 20\sqrt{\delta},$$

as $\delta < 1/100$ (see (1C1)). Thus, by Lemma 2.4, it suffices to show that

$$d(X_{F_0}, (Y_4\tilde{k}_4^{-1})_{F_0}) < \eta - 12\delta - 20\sqrt{\delta}.$$

Fix some number $p_0 \in J$. Let $(\tilde{k}_5, Y_5, Q_5)$ denote $(\tilde{k}_4, Y_4, Q_4)$ conditioned on

(4G) $\tilde{p} = p_0$.

Note that $\tilde{k}_5$ takes each value of $K_0 \cap \bar{A}_{p_0}$ with equal probability. By Lemma 2.5, it suffices to show that

$$d(X_{F_0}, (Y_5\tilde{k}_5^{-1})_{F_0}) < \eta - 12\delta - 20\sqrt{\delta}.$$

Let $\bar{B} := \bar{A}_{p_0} \cap \left(\bigcup_{p' \neq p_0} \bar{A}_{p'}\right)$. Since $p_0 \in J$,

(4H) $|\bar{B}| \leq \sqrt{\delta}|\bar{A}_{p_0}|$.

By definition of K_0, $K_0 \cap \bar{A}_{p_0} = \bar{A}_{p_0}\backslash\bar{B}$. Then $\tilde{k}_5$ is a random variable taking values in $\bar{A}_{p_0}\backslash\bar{B}$ (and taking each value with probability equal to all the others).

Let $\bar{C}$ denote the $(F_0F_0^{-1})$-interior of $\bar{A}_{p_0}\backslash\bar{B}$. Let $(\tilde{k}_6, Y_6, Q_6)$ denote $(\tilde{k}_5, Y_5, Q_5)$ conditioned on

(4I) $\tilde{k}_5 \in \bar{C}$.

Note that $\tilde{k}_6$ takes each value of $\bar{C}$ with probability $1/|\bar{C}|$. By (1H3) and (1G2), the $F_0F_0^{-1}$-boundary $\partial_{F_0F_0^{-1}}(\bar{A}_{p_0})$ of $\bar{A}_{p_0}$ satisfies

$$|\partial_{F_0F_0^{-1}}(\bar{A}_{p_0})| < \sqrt{\delta}|\bar{A}_{p_0}|.$$

Now

$$\bar{A}_{p_0}\backslash\bar{C} \subseteq \partial_{F_0F_0^{-1}}(\bar{A}_{p_0}) \cup \left(\bigcup_{b\in\bar{B}} F_0F_0^{-1}b\right).$$

Thus, by (4H),

$$|\bar{A}_{p_0}| - |\bar{C}| \leq \sqrt{\delta}|\bar{A}_{p_0}| + \sqrt{\delta}|\bar{A}_{p_0}||F_0F_0^{-1}|.$$

So

(4J) $\dfrac{|\bar{C}|}{|\bar{A}_{p_0}\backslash\bar{B}|} \geq \dfrac{|\bar{C}|}{|\bar{A}_{p_0}|} \geq 1 - \sqrt{\delta}(1 + |F_0F_0^{-1}|).$

So the event (4I) fails with probability $\leq \sqrt{\delta}(1 + |F_0F_0^{-1}|)$. Thus, by Lemma 2.4, it suffices to show that

$$d(X_{F_0}, (Y_6\tilde{k}_6^{-1})_{F_0}) < \eta - 12\delta - 20\sqrt{\delta} - \sqrt{\delta}(1 + |F_0F_0^{-1}|).$$

Write $\bar{A}_{p_0} = A_{i_0}g_0$, for some $1 \leq i_0 \leq k$ and some $g_0 \in G$. Let $A_0 := A_{i_0}$. Let $Y^0 := Y^{A_0}$. We know from (4C) and (1D1), by Definition 2.2, since X is stationary, that

$$d\left(X_{F_0}, \operatorname*{Avg}_{F_0}(Y^0)\right) < \delta.$$

Let $\tilde{a}_0$ be a random variable taking each value of A_0 with probability $1/|A_0|$. Couple Y^0 and $\tilde{a}_0$ independently. Then, by Definition 2.1, we know that

$$d(X_{F_0}, (Y^0\tilde{a}_0^{-1})^+_{F_0}) < \delta.$$

Thus it suffices to show that

$$d((Y_6\tilde{k}_6^{-1})_{F_0}, (Y^0\tilde{a}_0^{-1})^+_{F_0}) < \eta - 13\delta - 20\sqrt{\delta} - \sqrt{\delta}(1 + |F_0F_0^{-1}|).$$

Let $(Y^1, \tilde{a}_1)$ denote $(Y^0, \tilde{a}_0)$ conditioned on

(4K) $\tilde{a}_0 \in \bar{C}g_0^{-1}$.

By (4J), the probability that (4K) fails is $< \sqrt{\delta}(1+|F_0F_0^{-1}|)$. Thus, by Lemma 2.4, it suffices to show that

$$d\big((Y_6\tilde{k}_6^{-1})_{F_0}, (Y^0\tilde{a}_1^{-1})_{F_0}\big) < \eta - 13\delta - 20\sqrt{\delta} - 2\sqrt{\delta}(1+|F_0F_0^{-1}|).$$

(We have used that $(Y^0\tilde{a}_1^{-1})^+_{F_0} = (Y^0\tilde{a}_1^{-1})_{F_0}$, which follows from the fact that $F_0\tilde{a}_1 \subseteq A_0$, almost surely.)

Now Y_2, Y_3, Y_4, Y_5 and Y_6 are all isomorphic to the conditional process $Y|(Q_K = \sigma_0)$. Further, $\tilde{k}_i$ and Y_i are independent, for $i = 1, \ldots, 6$.

Similarly, Y^0 and Y^1 are isomorphic. Similarly, $\tilde{a}_j$ and Y^j are independent, for $j = 0, 1$.

Finally, $\tilde{a}_1$ is isomorphic to $\tilde{k}_6 g_0^{-1}$, since the former is equidistributed over $\bar{C}g_0^{-1}$ and the latter is equidistributed over $\bar{C}$.

It follows from the definitions of Y, Q and $\bar{C}$ that $(Y_6)_{F_0F_0^{-1}c}$ is isomorphic to $(Y^0 g_0)_{F_0F_0^{-1}c}$, for all $c \in \bar{C}$. Since $e \in F_0$ and since Y^0 and Y^1 are isomorphic, it follows that $(Y_6)_{F_0c}$ is isomorphic to $(Y^1 g_0)_{F_0c}$, for all $c \in \bar{C}$. Right translating by c^{-1}, we conclude that $(Y_6c^{-1})_{F_0}$ is isomorphic to $\big(Y^1(cg_0^{-1})^{-1}\big)_{F_0}$, for all $c \in \bar{C}$. Averaging these isomorphisms over $c \in \bar{C}$, we find that $(Y_6\tilde{k}_6^{-1})_{F_0}$ is isomorphic to $(Y^1\tilde{a}_1^{-1})_{F_0}$.

That is, we know:

$$d\big((Y_6\tilde{k}_6^{-1})_{F_0}, (Y^1\tilde{a}_1^{-1})_{F_0}\big) = 0.$$

Therefore, it suffices to show

$$0 < \eta - 13\delta - 20\sqrt{\delta} - 2\sqrt{\delta}(1+|F_0F_0^{-1}|).$$

This is true, by (1C5).

STEP 5: *The $\bar{d}$ calculation.* In this step, we verify (1M3). In fact, we will show that $\bar{d}(X_K, Y_K) \geq \epsilon^8$, for all sufficiently invariant $K \subseteq G$ satisfying $e \in K = K^{-1}$. Since the collection of finite $K \subseteq G$ satisfying $e \in K = K^{-1}$ is a Følner family, (1M3) will then follow.

By (1C7), we may choose $e \in K = K^{-1}$ sufficiently invariant that

(5A1) X_K is (ϵ^{16}, δ)-extremal.

By (2L), we may choose K sufficiently invariant that

(5A2) $\log_2 |\mathcal{C}(Q_K)| < \delta|K|$.

Recall the definition of $\mathcal{QT}_\delta(K)$ given following Definition 1.3. Using (2K1), we also require that K be sufficiently invariant that

(5A3) $\Pr[Q_K \in \mathcal{QT}_\delta(K)] > 1 - \delta$.

Recall that, if $K \subseteq G$ is finite, then BDRY(K) denotes the union over $i = 1, \dots, k$ of the $(A_i A_i^{-1})$-boundaries of K. By [A, Lemma 3.2], choose K so invariant that

(5A4) $|\mathrm{BDRY}(K)| < \delta |K|$.

It now suffices to show that $\bar{d}(X_K, Y_K) \geq \epsilon^8$. So assume **for a contradiction** that there exists a coupling m of X_K and Y_K such that

(5A5) $\bar{d}^m(X_K, Y_K) < \epsilon^8$.

Recall from (5A1), (5A5) and Definition 2.6 that

$$\sum_{\sigma \in \mathcal{C}(Q_K)} \Pr[Q_K = \sigma] \bar{d}(X_K | Q_K = \sigma, X_K) < \epsilon^{16}.$$

Let S_1 denote the set of $\sigma \in \mathcal{C}(Q_K)$ such that $\bar{d}(X_K\,,\, X_K | Q_K = \sigma) < \epsilon^8$. Then

(5B1) $\Pr[Q_K \in S_1] > 1 - \epsilon^8$.

Consider next the set S_2 of $\sigma \in \mathcal{C}(Q_K)$ such that

$$\bar{d}^m(X_K | Q_K = \sigma\,,\, Y_K | Q_K = \sigma) < \epsilon^4.$$

If $\Pr[Q_K \notin S_2] \geq \epsilon^4$ then (5A5) would fail. Therefore

(5B2) $\Pr[Q_K \in S_2] > 1 - \epsilon^4$.

Then, by (5A3), (5B1), (5B2) and (1C6), we may choose $\sigma_0 \in \mathcal{C}(Q_K)$ such that:

(5C1) $\sigma_0 \in \mathcal{QT}_\delta(K)$; and

(5C2) $\sigma_0 \in S_1 \cap S_2$.

From (5C2) and the definition of S_1 and S_2 we obtain

(5C3) $\bar{d}(X_K\,,\, Y_K | Q_K = \sigma_0) < \epsilon^4 + \epsilon^8$.

Let $\bar{A}_1, \dots, \bar{A}_m$ denote the sets in the quasi-tiling corresponding to σ_0. By (5C1), we may choose pairwise-disjoint $\tilde{A}_1, \dots, \tilde{A}_m$ with $\tilde{A}_p \subseteq \bar{A}_p$ and with $|\bar{A}_p / \tilde{A}_p| < \delta |\bar{A}_p|$, for $p = 1, \dots, m$. Let $\bar{A} := \bigcup_{p=1}^{m} \bar{A}_p$, $\tilde{A} := \bigcup_{p=1}^{m} \tilde{A}_p$. Then $\tilde{A} \backslash K \subseteq$ BDRY(K), so, by (5A4), we have $|\tilde{A} \backslash K| \leq \delta |K|$. Thus, by (5C3),

$$|\tilde{A}|\, \bar{d}(X_{\tilde{A}}\,,\, Y_{\tilde{A}} | Q_K = \sigma_0) < \delta |K| + (\epsilon^4 + \epsilon^8)|K|.$$

Now, by definition of a δ-quasi-tiling and by summing the inequalities $|\tilde{A}_p| > (1 - \delta)|\bar{A}_p|$ over $p = 1, \dots, m$, we obtain

$$|\tilde{A}| > (1 - \delta)|\bar{A}| > (1 - \delta)^2 |K|.$$

Let $\lambda := (\delta + \epsilon^4 + \epsilon^8)/(1-\delta)^2$. From the last two displayed inequalities, we conclude that

(5D) $\bar{d}(X_{\tilde{A}}, Y_{\tilde{A}} | Q_K = \sigma_0) < \lambda$.

Let J_1 denote the set of integers $1 \le p \le m$ such that $\bar{d}(X_{\tilde{A}_p}, Y_{\tilde{A}_p} | Q_K = \sigma_0) \ge \sqrt{\lambda}$. By (5D) above,

(5E) $\sum_{p \in J_1} |\bar{A}_p| < \sqrt{\lambda}|K|$.

Let J_2 denote the set of integers $1 \le p \le m$ such that

$$\left| \bar{A}_p \cap \left(\bigcup_{p' \neq p} \bar{A}_{p'} \right) \right| > \sqrt{\delta}|\bar{A}_p|.$$

By conclusion (E) of Lemma 1.1,

(5F) $\sum_{p \in J_2} |\bar{A}_p| < 10\sqrt{\delta}|K|$.

By (5E), (5F) and (1C2), we may fix an integer $1 \le p \le m$ such that $p \notin J_1 \cup J_2$. Then [A, Lemma 2.3] allows us to conclude from $p \notin J_1$ and from $|\tilde{A}_p|/|\bar{A}_p| > (1-\delta)|\bar{A}_p|$ that

(5G) $\bar{d}(X_{\bar{A}_p}, Y_{\bar{A}_p} | Q_K = \sigma_0) < \sqrt{\lambda} + \delta$.

Write $\bar{A}_p = A_i c$, for some integer $1 \le i \le k$ and some $c \in G$. Then $X_{A_i} c$ is isomorphic to $X_{\bar{A}_p}$. Further, as $p \notin J_2$, we see that $\bar{d}(Y^i c, Y_{\bar{A}_p} | Q_K = \sigma_0) < \sqrt{\delta}$. Thus, by (5G), $\bar{d}(X_{A_i} c, Y^i c) < \sqrt{\lambda} + \delta + \sqrt{\delta}$. By (1C3), $\bar{d}(X_{A_i} c, Y^i c) < \epsilon$, so, by translation invariance, $\bar{d}(X_{A_i}, Y^i) < \epsilon$. But $Y^i = Y^{A_i}$, so this contradicts (1H1) and (1D2), concluding Step 5.

This completes the proof of Theorem 2.8. ∎

3. Very Weak Bernoulli and Finitely Determined

Definition 3.1: We say that a stationary G-spin system X is **Very Weak Bernoulli** if, for all $\epsilon > 0$, for some disjoint ϵ-quasi-tiling system $A_1, \ldots, A_k$, for every sufficiently invariant finite set $K \subseteq G$, there exists an ordered disjoint ϵ-quasi-tiling $\bar{A}_1, \ldots, \bar{A}_m$ of K such that:

For $p = 2, \ldots, m$, let $\bar{P}_p := \bar{A}_1 \cup \cdots \cup \bar{A}_{p-1}$. Then, for all $p = 2, \ldots, m$, the process $X_{\bar{A}_p}$ is ϵ-process independent of $X_{\bar{P}_p}$ (see [A, Definition 2.8]).

An ordered disjoint ϵ-quasi-tiling with this property will be said to be **ϵ-almost independent under X**.

To obtain the definition of VWB for $\mathbf{Z}$-actions from Definition 3.1, we choose a large integer n, define $k := 1$, $A_1 := \{1 \ldots, n\}$. Now suppose, for example, that

$K = \{1, \ldots, mn\}$. We then use the ordered tiling of K by the sets

$$\{1, \ldots, n\}, \qquad \{n+1, \ldots, 2n\}, \qquad \ldots \qquad , \{(m-1)n+1, \ldots, mn\}.$$

For a moment, say that X is strong VWB if it satisfies the above definition, but with "for some disjoint ϵ-quasi-tiling system" replaced by "for any disjoint ϵ-quasi-tiling system consisting of sufficiently invariant sets". In Theorem 3.7 we will actually show that VWB implies Finitely Determined and that Finitely Determined implies strong VWB. Since strong VWB clearly implies VWB, we see that the two are, in fact, equivalent.

While we will not pursue this here, B. Weiss has pointed out to me that this formulation of Very Weak Bernoulli can be rephrased in terms of "almost block independence", a concept which is fundamental to the proof of [A, Theorem 4.2].

Fix a finite subset $A \subseteq G$. Let $\tilde{a}$ be an A-valued random variable which takes on each value of A with probability $1/|A|$. Let $\sigma \in \{+1, -1\}^A$. Define $\underset{F}{\mathrm{Avg}}(\sigma)$ to be the F-spin system $(\sigma\tilde{a}^{-1})_F^+$ (see Definition 2.1).

If $K \subseteq G$ is finite, if X is a K-process and if $P(\sigma)$ is a property of configurations $\sigma \in \mathcal{C}(X)$, then we will say that $P(\sigma)$ **holds for ϵ-a.e.** $\sigma \in \mathcal{C}(X)$ if

$$\Pr[P(X) \text{ holds}] > 1 - \epsilon.$$

A stationary G-process X is **ergodic** if: for any G-invariant Borel subset $E \subseteq \{+1, -1\}^G$, we have either $\Pr[X \in E] = 0$ or $\Pr[X \in E] = 1$.

THEOREM 3.2: *Suppose X is an ergodic, stationary G-spin system, $\epsilon > 0$ and $F \subseteq G$ is finite. Then, for all sufficiently invariant $K \subseteq G$, for ϵ-a.e. $\sigma \in \mathcal{C}(X_K)$, $d\big(X_F, \underset{F}{\mathrm{Avg}}(\sigma)\big) < \epsilon$.*

Proof: This is a consequence of the Mean Ergodic Theorem for amenable groups. ([Ol, von Neumann's theorem 3.2.4, p. 43]) ∎

THEOREM 3.3: *Suppose X is an ergodic, stationary G-spin system and $\lambda > 0$. Then, for all sufficiently invariant $L \subseteq G$, for λ-a.e. $\sigma \in \mathcal{C}(X_L)$,*

$$(1+\lambda)^{-1} H(X)|L| < -\log_2 \Pr[X_L = \sigma] < (1+\lambda) H(X)|L|.$$

Proof: This is the Shannon-McMillan Theorem for discrete amenable groups, cf. [Ol, Theorem 4.4.2, p. 64]. ∎

We use the terminology introduced in [A, Definition 3.3]. Let X be a stationary G-process. Let $A_1, \ldots, A_k$ be a disjoint ϵ-quasi-tiling system. Let $\bar{A}_1, \ldots, \bar{A}_m$ be a disjoint ϵ-quasi-tiling of some set K by right translates of $A_1, \ldots, A_k$. For $p = 2, \ldots, m$, let

$$\bar{P}_p := \bar{A}_1 \cup \cdots \cup \bar{A}_{p-1}.$$

A configuration $\sigma \in \mathcal{C}(X_{\bar{P}_p})$ is called a **p-past**, and its **weight** is defined to be $\mathcal{P}_\sigma |\bar{A}_p|/|K|$, where, as usual, $\mathcal{P}_\sigma := \Pr[X_{\bar{P}_p} = \sigma]$. The **remainder weight** is defined to be

$$R := \frac{|K \backslash (\bar{A}_1 \cup \cdots \cup \bar{A}_m)|}{|K|}.$$

A **past** is a p-past for some p.

Let P be a property of pasts. The **naïve weight of** P is the sum of the weights of the pasts for which P holds. The **weight of** P is the sum of $|\bar{A}_1|/|K|$ and the naïve weight of P. (We are making the convention here that the set of 1-pasts is empty and has weight $|\bar{A}_1|/|K|$ and that any property holds for all 1-pasts.) We say that P **holds for δ-a.e. past** if the weight of P exceeds $1-\delta$. Similarly, if P is a property of p-pasts, then we say that P **holds for δ-a.e. p-past** if the weight of P exceeds $(1-\delta)W_p$, where W_p denotes the total weight of the p-pasts. If σ is a p-past, then we will frequently denote $\bar{A}_p$ by $\bar{A}_\sigma$. We use $X^\sigma_{\bar{A}_\sigma}$ to denote the conditional process $X_{\bar{A}_p} | X_{\bar{P}_p} = \sigma$, i.e., to denote the process $X_{\bar{A}_p}$ conditioned on $X_{\bar{P}_p} = \sigma$.

Definition 3.4: Let $A_1, \ldots, A_k$ be a disjoint δ-quasi-tiling system for G. Assume $K \subseteq G$ is disjoint δ-quasi-tilable ([A, Definition 3.3]) by right translates $\bar{A}_1, \ldots, \bar{A}_m$ of $A_1, \ldots, A_k$. Let X be a K-process. Then we say that the ordered disjoint quasi-tiling $\bar{A}_1, \ldots, \bar{A}_m$ is **(δ, F_0)-VWB** if: for δ-a.e. past σ, $X^\sigma_{\bar{A}_\sigma}$ and $X_{\bar{A}_\sigma}$ are (δ, F_0)-close in entropy and finite distribution. ∎

It is interesting to note that the next result does not require X to be Very Weak Bernoulli. Thus, all ergodic processes have a certain amount of Very Weak Bernoulliness to them.

THEOREM 3.5: *Let X be an ergodic, stationary G-process. Let $\delta > 0$ and let $F_0 \subseteq G$ be finite. Then any Følner family $\mathcal{F}$ contains a disjoint δ-quasi-tiling system $A_1, \ldots, A_k$ such that, for every sufficiently invariant $K \subseteq G$, there exists an ordered (δ, F_0)-VWB disjoint quasi-tiling of K via right translates of $A_1, \ldots, A_k$.*

Proof: Replacing δ with $\min\{\delta, 1/2\}$, we may assume that $\delta < 1$. Let $H := H(X)$ denote the entropy of X.

Choose $\lambda > 0$ such that

(A1) $\lambda + (8\lambda/\delta) < \delta/16$.

Since $\delta < 1$, it follows that

(A2) $3\lambda < \delta/2$.

By the Shannon-McMillan Theorem (Theorem 3.3) and by the definition of entropy [A, Lemma 3.5], we may choose $\eta_1 > 0$ and $F_1 \subseteq G$ finite satisfying: if $L \subseteq G$ is (F_1, η_1)-invariant (defined at the start of [A, §3]), then

(B1) $\Big|H(X_L) - H|L|\Big| < (\delta^3/256)|L|$; and

(B2) for λ^2-a.e. $\sigma \in \mathcal{C}(X_L)$: $H|L|(1+\lambda)^{-1} < -\log_2 \Pr[X_L = \sigma] < H|L|(1+\lambda)$.

For $A \subseteq G$ finite, let $\mathcal{R}(A)$ denote the collection of all $\sigma \in \{+1,-1\}^A$ such that $d\big(\operatorname*{Avg}_{F_0}(\sigma), X_{F_0}\big) < \delta/2$. By the Mean Ergodic Theorem (Theorem 3.2), for all sufficiently invariant $A \subseteq G$:

$$\Pr[X_A \in \mathcal{R}(A)] > 1 - (\delta^2/16).$$

Let $\mathcal{F}_1$ denote the Følner family of all $A \in \mathcal{F}$ such that

(C1) $\Pr[X_A \in \mathcal{R}(A)] > 1 - (\delta^2/16)$;

(C2) A is (F_1, η_1)-invariant; and

(C3) $1 < \lambda|A|$.

We may choose a disjoint δ-quasi-tiling system $A_1, \ldots, A_k \in \mathcal{F}_1$ [A, Lemma 3.4]. If two of the A_is are right translates of one another, then we may eliminate one of them; we may therefore assume that $A_1, \ldots, A_k$ is reduced. Then the $(A_1, \ldots, A_k)$-correspondence (see Definition 1.3) is well-defined.

Let K be sufficiently invariant that there exists a disjoint $(\delta/2)$-quasi-tiling of K by right translates $\bar{A}_1, \ldots, \bar{A}_m$ of $A_1, \ldots, A_k$. It suffices to show, for δ-a.e. past σ, that

(D) $d\big(\operatorname*{Avg}_{F_0}(X^{\sigma}_{\bar{A}_\sigma}), X_{F_0}\big) < \delta$; and

(E) $|H(X^{\sigma}_{\bar{A}_\sigma}) - H(X_{\bar{A}_\sigma})| < \delta|\bar{A}_\sigma|$.

If $R := |K \backslash (\bar{A}_1 \cup \cdots \cup \bar{A}_m)|/|K|$ denotes the remainder weight, then, by definition of a disjoint $(\delta/2)$-quasi-tiling ([A, Definition 3.3]), we have $R < \delta/2$.

Define $H_1 := H(X_{\bar{A}_1})/|\bar{A}_1|$. For $p = 2, \ldots, m$, define $\bar{P}_p := \bar{A}_1 \cup \cdots \cup \bar{A}_{p-1}$ and $H_p := H(X_{\bar{A}_p}|X_{\bar{P}_p})/|\bar{A}_p|$. Let S_0 be the set of all $p \in \{1, \ldots, m\}$ such that $H_p > H - (\delta^2/32)$. Let $\bar{W} := \bar{A}_1 \cup \cdots \cup \bar{A}_m$.

By (C2), for all $p = 1, \ldots, m$, $\bar{A}_p$ is (F_1, η_1)-invariant, so, by (B1), we have

$$H_p \leq \frac{H(X_{\bar{A}_p})}{|\bar{A}_p|} < H + (\delta^3/256).$$

As $\bar{W}$ is (F_1, η_1)-invariant, we conclude from (B1) that

$$H(X_{\bar{W}}) > [H - (\delta^3/256)]|\bar{W}|.$$

By standard properties of entropy,

$$H(X_{\bar{W}}) = \sum_{p=1}^{m} H_p|\bar{A}_p| = \left[\sum_{p\in S_0} H_p|\bar{A}_p|\right] + \left[\sum_{p\notin S_0} H_p|\bar{A}_p|\right].$$

Combining these last three observations, we have

$$[H - (\delta^3/256)]|\bar{W}| \leq \left[\sum_{p\in S_0} [H + (\delta^3/256)]|\bar{A}_p|\right] + \left[\sum_{p\notin S_0} [H - (\delta^2/32)]|\bar{A}_p|\right].$$

Subtracting $H|\bar{W}| = \left[\sum_{p\in S_0} H|\bar{A}_p|\right] + \left[\sum_{p\notin S_0} H|\bar{A}_p|\right]$, we get

$$-(\delta^3/256)|\bar{W}| \leq (\delta^3/256)\left[\sum_{p\in S_0} |\bar{A}_p|\right] - (\delta^2/32)\left[\sum_{p\notin S_0} |\bar{A}_p|\right].$$

Using $\sum_{p\in S_0} |\bar{A}_p| \leq |\bar{W}|$, and solving for $\sum_{p\notin S_0} |\bar{A}_p|$, we find

$$\sum_{p\notin S_0} |\bar{A}_p| \leq (32/\delta^2)\big[2(\delta^3/256)|\bar{W}|\big] = (\delta/4)|\bar{W}|.$$

Since $[1-R][1-(\delta/4)][1-(\delta/4)] > [1-(\delta/2)][1-(\delta/4)]^2 > 1-\delta$, it suffices to show, for all $p \in S_0\backslash\{1\}$, that:

for $(\delta/4)$-a.e. p-past σ: both (D) and (E) hold.

So fix some $p \in S_0\backslash\{1\}$. Recall that if σ is a p-past, then $\bar{A}_\sigma := \bar{A}_p$. It suffices to show that

(F) for $(\delta/8)$-a.e. p-past σ: $d\big(\operatorname*{Avg}_{F_0}(X^\sigma_{\bar{A}_\sigma}), X_{F_0}\big) < \delta$; and

(G) for $(\delta/8)$-a.e. p-past σ: $|H(X^\sigma_{\bar{A}_\sigma}) - H(X_{\bar{A}_\sigma})| < \delta|\bar{A}_\sigma|$.

Since X is stationary, we see from (C1) that $\bar{A}_p \in \mathcal{F}$ implies

$\Pr[X_{\bar{A}_p} \in \mathcal{R}(\bar{A}_p)] > 1 - (\delta^2/16) = 1 - (\delta/8)(\delta/2).$

Thus, for $(\delta/8)$-a.e. p-past σ,

$\Pr[X^\sigma_{\bar{A}_p} \in \mathcal{R}(\bar{A}_p)] > 1 - (\delta/2).$

Then, by definition of $\mathcal{R}(\bar{A}_p)$, for $(\delta/8)$-a.e. p-past σ,

$d\big(\operatorname*{Avg}_{F_0}(X^\sigma_{\bar{A}_p}), X_{F_0}\big) < \delta/2 + (1-\delta/2)(\delta/2) < \delta,$

verifying (F).

Define $\bar{P} := \bar{P}_p$, $\bar{A} := \bar{A}_p$. By (C2), $\bar{A}$ is (F_1, η_1)-invariant, so, by (B2),

(H) for λ^2-a.e. $\tau \in \mathcal{C}(X_{\bar{A}})$: $\quad \Pr[X_{\bar{A}} = \tau] > 2^{-H|\bar{A}|(1+\lambda)}$.

Then there exists a subset $C_0 \subseteq \mathcal{C}(X_{\bar{A}})$ such that

(I1) $|C_0| < 2^{H|\bar{A}|(1+\lambda)}$; and

(I2) $\Pr[X_{\bar{A}} \in C_0] > 1 - \lambda^2$.

By (I2), we have

(J) for λ-a.e. $\sigma \in \mathcal{C}(X_{\bar{P}})$: $\quad \Pr[X^\sigma_{\bar{A}} \in C_0] > 1 - \lambda$.

By Lemma 2.3 and by (I1), for all $\sigma \in \mathcal{C}(X_{\bar{P}})$,

$$H(X^\sigma_{\bar{A}} | X^\sigma_{\bar{A}} \in C_0) \le H|\bar{A}|(1+\lambda),$$
$$H(X^\sigma_{\bar{A}} | X^\sigma_{\bar{A}} \notin C_0) \le |\bar{A}|.$$

By standard properties of entropy, these two estimates and (J) imply that: for λ-a.e. $\sigma \in \mathcal{C}(X_{\bar{P}})$,

$$H(X^\sigma_{\bar{A}}) \le [H|\bar{A}|(1+\lambda)](1-\lambda) + [|\bar{A}|]\lambda + 2.$$

By (C3), $2 < 2\lambda|\bar{A}|$; further, $(1+\lambda)(1-\lambda) = 1 - \lambda^2 < 1$, so we obtain

(K) for λ-a.e. $\sigma \in \mathcal{C}(X_{\bar{P}})$: $\quad H(X^\sigma_{\bar{A}}) < (H + 3\lambda)|\bar{A}|$.

Define

$$C_1 := \{\sigma \in \mathcal{C}(X_{\bar{P}}) \,|\, H(X^\sigma_{\bar{A}}) \ge [H + 3\lambda]|\bar{A}|\},$$
$$C_2 := \{\sigma \in \mathcal{C}(X_{\bar{P}}) \,|\, H(X^\sigma_{\bar{A}}) \le [H - (\delta/2)]|\bar{A}|\},$$
$$\lambda_1 := \Pr[X_{\bar{P}} \in C_1],$$
$$\lambda_2 := \Pr[X_{\bar{P}} \in C_2].$$

By Lemma 2.3, $H(X^\sigma_{\bar{A}}) \le |\bar{A}|$, for all $\sigma \in \mathcal{C}(X_{\bar{P}})$.

Since $p \in S_0$, we have

$$[H - (\delta^2/32)]|\bar{A}| < H(X_{\bar{A}}|X_{\bar{P}}) = \sum_{\sigma \in \mathcal{C}(X_P)} H(X^\sigma_{\bar{A}}) \Pr[X_{\bar{P}} = \sigma].$$

Breaking this sum into $\sum_{\sigma\in C_1} + \sum_{\sigma\in C_2} + \sum_{\sigma\notin C_1\cup C_2}$, we obtain

$$\begin{aligned}[H-(\delta^2/32)]|\bar{A}| &< |\bar{A}|\lambda_1 + [H-(\delta/2)]|\bar{A}|\lambda_2 + [H+3\lambda]|\bar{A}|(1-\lambda_1-\lambda_2)\\ &\le |\bar{A}|\lambda_1 + H|\bar{A}|\lambda_2 - (\delta/2)|\bar{A}|\lambda_2 + H|\bar{A}|(1-\lambda_2-0) + 3\lambda|\bar{A}|(1-0-0)\end{aligned}$$

Dividing by $|\bar{A}|$ and subtracting $H = H\lambda_2 + H(1-\lambda_2)$, we have

$$-\delta^2/32 < \lambda_1 - (\delta/2)\lambda_2 + 3\lambda.$$

Solving the estimate above for λ_2 gives $\lambda_2 < (2/\delta)[(\delta^2/32) + \lambda_1 + 3\lambda]$. By (K), we see that $\lambda_1 < \lambda$, so

$$\Pr[X_{\bar{P}} \in C_1 \cup C_2] = \lambda_1 + \lambda_2 < \lambda + (2/\delta)[(\delta^2/32) + 4\lambda] = [\lambda + (8\lambda/\delta)] + (\delta/16).$$

So, by (A1), $\Pr[X_{\bar{P}} \in C_1 \cup C_2] < \delta/8$. By definition of C_1 and C_2, we conclude, for $(\delta/8)$-a.e. $\sigma \in \mathcal{C}(X_{\bar{P}})$, that

$$[H-(\delta/2)]|\bar{A}| < H(X_{\bar{A}}^{\sigma}) < (H+3\lambda)|\bar{A}|.$$

By (A2), $3\lambda < \delta/2$, so: for $(\delta/8)$-a.e. $\sigma \in \mathcal{C}(X_{\bar{P}})$,

$$\left|H(X_{\bar{A}}^{\sigma}) - H|\bar{A}|\right| < (\delta/2)|\bar{A}|.$$

Now, by (C2), $\bar{A}$ is (F_1, η_1)-invariant, so, by (B1),

$$\left|H|\bar{A}| - H(X_{\bar{A}})\right| < (\delta^3/256)|\bar{A}|.$$

Since $\delta < 1$, we have $\delta^3/256 < \delta/2$, so the last two displayed estimates imply: for $(\delta/8)$-a.e. $\sigma \in \mathcal{C}(X_{\bar{P}})$,

$$|H(X_{\bar{A}}^{\sigma}) - H(X_{\bar{A}})| < \delta|\bar{A}|.$$

But $\bar{A} = \bar{A}_p = \bar{A}_\sigma$ and $\bar{P} = \bar{P}_p$, so this verifies (G). ∎

LEMMA 3.6: *Let X be a stationary G-spin system. Assume, for all $\epsilon > 0$, that there exists a disjoint ϵ-quasi-tiling system $A_1, \ldots, A_k$ such that: for every sufficiently invariant $K \subseteq G$, there exists an ordered disjoint ϵ-quasi-tiling of*

$\bar{A}_1, \ldots, \bar{A}_m$ of K such that for ϵ-a.e. past σ, $\bar{d}(X^{\sigma}_{\bar{A}_\sigma}, X_{\bar{A}_\sigma}) < \epsilon$. Then X is Very Weak Bernoulli.

Proof: Given $\epsilon > 0$, we try to verify the condition of Definition 3.1. Choose $0 < \delta < 1$ such that

$$\delta + \sqrt{\delta} < \epsilon.$$

Replace all "ϵ"s by "δ"s in the hypothesis of Lemma 3.6 and choose $A_1, \ldots, A_k$, K, and $\bar{A}_1, \ldots, \bar{A}_m$ as described. Then, for δ-a.e. past σ, $\bar{d}(X^{\sigma}_{\bar{A}_\sigma}, X_{\bar{A}_\sigma}) < \delta$.

For $p = 2, \ldots, m$, define $\bar{P}_p := \bar{A}_1 \cup \cdots \cup \bar{A}_{p-1}$; if σ is a p-past, let $\mathcal{P}_\sigma := \Pr[X_{\bar{A}_p} = \sigma]$. Let I denote the set of all $1 \le p \le m$ such that: for $\sqrt{\delta}$-a.e. p-past σ, $\bar{d}(X^{\sigma}_{\bar{A}_\sigma}, X_{\bar{A}_\sigma}) < \delta$. Then $\left| \bigcup_{p \in I} \bar{A}_p \right| > (1 - \sqrt{\delta})|K| > (1 - \epsilon)|K|$.

We **claim** that, for all $p \in I$, the process $X_{\bar{A}_p}$ is ϵ-process independent of $X_{\bar{P}_p}$. So fix $p \in I$. Let $\mathcal{C} := \mathcal{C}(X_{\bar{A}_p})$, let $\mathcal{C}_1$ denote those $\sigma \in \mathcal{C}$ such that $\bar{d}(X^{\sigma}_{\bar{A}_p}, X_{\bar{A}_p}) < \delta$. By definition of I, $\sum_{\sigma \in \mathcal{C}_1} \mathcal{P}_\sigma > 1 - \sqrt{\delta}$. Therefore, by [A, Lemma 2.6],

$$\begin{aligned} \bar{d}_{\bar{P}_p}(X_{\bar{P}_p} \vee X_{\bar{A}_p}, X_{\bar{P}_p} \| X_{\bar{A}_p}) &= \sum_{\sigma \in \mathcal{C}} \mathcal{P}_\sigma \bar{d}(X^{\sigma}_{\bar{A}_p}, X_{\bar{A}_p}) \\ &= \sum_{\sigma \in \mathcal{C}_1} + \sum_{\sigma \in \mathcal{C} \setminus \mathcal{C}_1} < (1)(\delta) + (\sqrt{\delta})(1) < \epsilon. \end{aligned}$$

This **proves the claim**, by definition of ϵ-process independence ([A, Definition 2.8]).

Now let $\bar{A}'_1, \ldots, \bar{A}'_{m'}$ be the ordered list of sets obtained from $\bar{A}_1, \ldots, \bar{A}_m$ by eliminating those $\bar{A}_p$ for which $p \notin I$. Then

$$|\bar{A}'_1 \cup \cdots \cup \bar{A}'_{m'}| > (1 - \epsilon)|K|.$$

By [A, Lemma 2.14], for all $p = 2, \ldots, m'$, the process $X_{\bar{A}'_p}$ is ϵ-process independent of $X_{\bar{P}'_p}$, where $\bar{P}'_p := \bar{A}'_1 \cup \cdots \cup \bar{A}'_{p-1}$. ∎

We can now state and prove the main theorem of this paper.

THEOREM 3.7: *Let X be a stationary G-process. Then X is Very Weak Bernoulli if and only if X is Finitely Determined.*

Proof: The proof that Very Weak Bernoulli implies Finitely Determined is almost the same as the proof of [A, Theorem 4.2]. The main difference is that statement (C) of the proof should be changed to say: "there exists an ordered disjoint ϵ-quasi-tiling of B_p by right translates of $A_1, \ldots, A_k$ which is ϵ-almost independent under X." (See Definition 3.1.) We eliminate statements (E) and (I). By the new version of (C), we may choose the traversal $C_1, \ldots, C_t$ of the proof in such a way that: if $C_p, \ldots, C_q$ is a block with union B, then $C_p, \ldots, C_{q-1}$ is an ordered disjoint ϵ-quasi-tiling of B which is ϵ-almost independent under X. This insures that statement (W) remains true. The rest of the proof remains unchanged.

We now turn to the proof that Finitely Determined implies Very Weak Bernoulli. Assume that X is a Finitely Determined, stationary G-process. Let $\epsilon > 0$ be given. We will verify the condition of Lemma 3.6, for this ϵ. Choose δ, F_0 as in Theorem 2.8; replacing δ by the minimum of δ and ϵ, we may assume that $\delta \leq \epsilon$. Let $\mathcal{F}$ denote the Følner family of all finite $K \subseteq G$ for which the conclusion of Theorem 2.8 holds. Note that, by [O-W, Theorem 8, p. 93], X is isomorphic to a Bernoulli process and is therefore ergodic. Choose $A_1, \ldots, A_k \in \mathcal{F}$ as in Theorem 3.5. Choose K sufficiently invariant that the conclusion of Theorem 3.5 holds. Consequently, there exists a (δ, F_0)-VWB ordered disjoint quasi-tiling $\bar{A}_1, \ldots, \bar{A}_m$ as in Definition 3.4. Note that each $\bar{A}_p$ is a right translate of some A_i, so $\bar{A}_1, \ldots, \bar{A}_m \in \mathcal{F}$. Then, for δ-a.e. past σ, $X^{\sigma}_{\bar{A}_\sigma}$ and $X_{\bar{A}_\sigma}$ are (δ, F_0)-close in entropy and finite distribution. Since $\delta \leq \epsilon$, and since δ, F_0 were chosen as in Theorem 2.8, we conclude that: for ϵ-a.e. past, $\bar{d}(X^{\sigma}_{\bar{A}_\sigma}, X_{\bar{A}_\sigma}) < \epsilon$. This is exactly what was needed to complete the hypotheses of Lemma 3.6; we conclude that X is Very Weak Bernoulli. ∎

References

[A] S. Adams, *Følner independence and the amenable Ising model,* Ergodic Theory and Dynamical Systems, to appear

[A-S] S. Adams and J. Steif, *An application of the Very Weak Bernoulli condition for amenable groups,* Pacific J. Math., to appear.

[C] J. Chatard, *Applications des propriétés de moyenne d'un groupe localement compact à la théorie ergodique,* Ann. Institut J. Poincaré **6** (1970), Sect. B, 307-326.

[K-W] Y. Katznelson and B. Weiss, *Commuting measure-preserving transformations,* Israel Journal of Mathematics **12** (1972), 161–173.

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

[O-W-2] D. Ornstein and B. Weiss, *Every transformation is bilaterally deterministic,* Israel Journal of Mathematics **21** (1975), 154–158.

[O-W-3] D. Ornstein and B. Weiss, *Finitely Determined implies Very Weak Bernoulli,* Israel Journal of Mathematics **17** (1974), 94–104.

[O] D. Ornstein, *Ergodic Theory, Randomness and Dynamical Systems,* Yale University Press, New Haven, 1974.

[O1] J. M. Ollagnier, *Ergodic Theory and Statistical Mechanics,* Springer Lecture Notes 1115, 1985.

[Sh] P. Shields, *The theory of Bernoulli shifts,* University of Chicago Press, Chicago, 1973.

[St1] J. Steif, *Space-time Bernoullicity of the lower and upper stationary processes for attractive spin systems,* Annals of Probability **19** (1991), 609–635.

[St2] J. Steif, *$\bar{d}$-convergence to equilibrium and space-time Bernoullicity for spin systems in the $M < \epsilon$ case,* Ergodic Theory and Dynamical Systems, to appear.