Lecture 18:

Recall:

Let $\alpha = \{A_1, \ldots, A_k\}$ a finite measble. partition

and $\mathcal{B}$ a sub-σ-algebra

Defn:

$$I_{\alpha|\mathcal{B}}(x) := -\sum_i \chi_{A_i}(x) \log \mu(A_i|\mathcal{B}) = -\sum_i \chi_{A_i}(x) \log E(\chi_{A_i}|\mathcal{B})$$

Defn: $H(\alpha|\mathcal{B}) := \int I_{\alpha|\mathcal{B}} d\mu$.

Prop: Let α, β be finite measble. partitions. Let $\mathcal{B}$ be the σ-algebra generated by β, i.e., the collection of all finite unions of elements of β. Then

$$H(\alpha|\mathcal{B}) = H(\alpha|\beta)$$

Continuity of Conditional entropy:

Let $\mathcal{B}_1 \subset \mathcal{B}_2 \subset \mathcal{B}_3 \ldots$ be σ-algebras

and $\mathcal{B} = \sigma(\cup_{n=1}^{\infty} \mathcal{B}_n)$.

Let α be a finite measble. partition.

Then

$$I_{\alpha|\mathcal{B}_n} \to I_{\alpha|\mathcal{B}} \text{ a.e. and in } L^1$$

and

$$H(\alpha|\mathcal{B}_n) \to H(\alpha|\mathcal{B})$$

Proof:

By Martingale Convergence Theorem

$$\mu(A_i|\mathcal{B}_n) \to \mu(A_i|\mathcal{B})\ a.e. \text{ and in } L^1$$

Thus,

$$I_{\alpha|\mathcal{B}_n} : -\sum_i \chi_{A_i} \log \mu(A_i|\mathcal{B}_n) \to \sum_i \chi_{A_i} \log \mu(A_i|\mathcal{B}) := I_{\alpha|\mathcal{B}}$$

(a.e. and in L^1)

Can show $\sup_n I_{\alpha|\mathcal{B}_n} \in L^1$ (may be done in student talk on Shannon-McMillan-Breiman Thoerem). Apply DCT to get L^1 convergence of the conditional information function and convergence of conditional entropy. □

Alternative proof only of convergence of conditional entropy:

Fact:

$$H(\alpha|\mathcal{B}) = \int -\sum_i \mu(A_i|\mathcal{B}) \log \mu(A_i|\mathcal{B}) d\mu$$

because whenever g is $\mathcal{B}$-measble, we have:

$$\int f g d\mu = \int E(f|\mathcal{B}) g d\mu$$

And integrand above is bounded by $\log k$ because $\{\mu(A_i|\mathcal{B})(x)\}_{i=1}^k$ is a probability vector.

And $-\sum_i \mu(A_i|\mathcal{B}_n) \log \mu(A_i|\mathcal{B}_n) \rightarrow -\sum_i \mu(A_i|\mathcal{B}) \log \mu(A_i|\mathcal{B})$ a.e.

Apply bounded convergence theorem. □

We can formulate another list of entropy inequalities involving conditioning on sub-σ-algebras of $\mathcal{A}$, instead of conditioning on finite measble partitions.

Defn: A measure space has a *countable basis* if there is a countable collection of measurable sets $\mathcal{C}$ s.t. for all $A \in \mathcal{A}$ and $\epsilon > 0$, there exists $C \in \mathcal{C}$ s.t. $\mu(A \Delta C) < \epsilon$.

Fact: The Borel sigma-algebra for any separable metric space has a countable basis. And so does its completion.

Proof: The σ-algebra $\mathcal{A}$ is generated by finite unions of open balls of radius $1/n$ centered at the points of a countable dense set.

Fact: $(M, \mathcal{A}, \mu)$ has a countable basis iff $L^2(M, \mathcal{A}, \mu)$ is separable.

Proof:

only if: given a countable basis $\mathcal{C}$, the collection $\sum_{i=1}^{n} a_i \chi_{A_i}$ with $a_i \in \mathbb{Q}$ and $A_i \in \mathcal{C}$ is L^2-dense.

if: given a countable L^2-dense subset $\mathcal{D}$, the collection $f^{-1}(a, b)$ with $a, b \in \mathbb{Q}$ and $f \in \mathcal{D}$ is a countable basis.

Fact: Given $\mathcal{B} \subseteq \mathcal{A}$, the map

$$L^2(M, \mathcal{A}, \mu) \to L^2(M, \mathcal{B}, \mu), \quad f \mapsto E(f|\mathcal{B})$$

is continuous.

Proof:

$$\int |E(f|\mathcal{B}) - E(g|\mathcal{B})|^2 d\mu = \int E(f-g|\mathcal{B})^2 d\mu \le \int E(|f-g|^2|\mathcal{B}) d\mu$$
$$= \int |f-g|^2 d\mu$$

Corollary: Let $\mathcal{B} \subseteq \mathcal{A}$, If $(M, \mathcal{A}, \mu)$ has a countable basis, so does $(M, \mathcal{B}, \mu)$.

Tool to develop conditional entropy inequalities, conditioning on sub-sigma-algebras:

Let $(M, \mathcal{A}, \mu)$ have a countable basis. Let $\mathcal{B}$ be sub-sigma-algebra.

Fix $\{C_1, C_2, \ldots\}$, a countable basis for $\mathcal{B}$.

Let $\mathcal{B}_n$ be the sigma-algebra (equiv.., the algebra) generated by $\{C_1, C_2, \ldots, C_n\}$.

Then $\mathcal{B}_n \uparrow \mathcal{B}$.

Let β_n denote the partition determined by $\mathcal{B}_n$.

By continuity theorem, $H(\alpha|\beta_n) = H(\alpha|\mathcal{B}_n) \to H(\alpha|\mathcal{B})$.

Since $\beta_n \preceq \beta_{n+1}$, $H(\alpha|\mathcal{B}_n) \downarrow H(\alpha|\mathcal{B})$

Properties of entropy:

1. $H(\alpha \vee \beta) = H(\alpha) + H(\beta|\alpha)$
2. $H(\alpha|\mathcal{B}) \leq H(\alpha)$ with equality iff $\alpha \perp \mathcal{B}$
3. $H(\alpha|\mathcal{B}) \geq 0$ with equality iff $\alpha \subseteq \mathcal{B}$
4. $H(\alpha \vee \beta) = H(\alpha)$ iff $\beta \preceq \alpha$
5. $H(\alpha_1 \vee \cdots \vee \alpha_n) = \sum_{i=1}^{n} H(\alpha_i|\alpha_1 \vee \cdots \vee \alpha_{i-1})$
6. $H(\alpha_1 \vee \cdots \vee \alpha_n)) \leq \sum_i H(\alpha_i)$ with equality iff α_i are independent.
7. Assume $\beta \preceq \alpha$. Then $H(\beta) \leq H(\alpha)$ and equality holds iff $\beta = \alpha$.
8. Assume $\mathcal{B} \preceq \mathcal{C}$. Then $H(\alpha|\mathcal{C}) \leq H(\alpha|\mathcal{B})$
9. $H(\alpha \vee \beta|\mathcal{C}) = H(\alpha|\mathcal{C}) + H(\beta|\sigma(\alpha \cup \mathcal{C})) \leq H(\alpha|\mathcal{C}) + H(\beta|\mathcal{C})$
10. $H(\alpha|\sigma(\mathcal{B} \cup \mathcal{C})) \leq H(\alpha|\mathcal{B})$ with equality iff $\alpha \perp_{\mathcal{B}} \mathcal{C}$
11. If $\mathcal{B} \perp_{\mathcal{C}} (\sigma(\beta \cup \mathcal{D})$, then $H(\beta|\sigma(\mathcal{D} \cup \mathcal{C} \cup \mathcal{B}) = H(\beta|\sigma(\mathcal{D} \cup \mathcal{C}))$.
12. If $\gamma \preceq \alpha$, then $H(\gamma|\mathcal{B}) \leq H(\alpha|\mathcal{B})$
13. If T is an MPT, then for all i, j, k,
    $H(T^{-i}\alpha \vee \ldots \vee T^{-j}\alpha) = H(T^{-i-k}\alpha \vee \ldots \vee T^{-j-k}\alpha)$
14. If T is an MPT, then for all i, j, k,
    $H(T^{-i}\alpha \vee \ldots \vee T^{-j}\alpha|\mathcal{B}) = H(T^{-i-k}\alpha \vee \ldots \vee T^{-j-k}\alpha|T^{-k}(\mathcal{B}))$

Proof of 2:

$H(\alpha|\mathcal{B}) \leq H(\alpha)$:

Proof: $\mathcal{B}_n \uparrow \mathcal{B}$

So, $H(\alpha|\mathcal{B}) = \lim_n H(\alpha|\mathcal{B}_n) = \lim_n H(\alpha|\beta_n) \leq H(\alpha)$.

$H(\alpha|\mathcal{B}) = H(\alpha)$ iff $\mathcal{B} \perp \alpha$.

Proof:

If: $\mathcal{B} \perp \alpha$ implies for all n, we have $\beta_n \perp \alpha$.

Thus, $H(\alpha|\beta_n) = H(\alpha)$,

Thus, $H(\alpha|\mathcal{B}) = \lim_n H(\alpha|\beta_n) = H(\alpha)$.

Only If:

$H(\alpha|\beta_n) \downarrow H(\alpha|\mathcal{B})$.

So, $H(\alpha|\mathcal{B}) \leq H(\alpha|\beta_n) \leq H(\alpha)$.

So, if $H(\alpha|\mathcal{B}) = H(\alpha)$, then for all n, $H(\alpha|\beta_n) = H(\alpha)$.

So, for all n, $\alpha \perp \beta_n$. Thus, $\alpha \perp \mathcal{B}$. QED

Proof of 3:

If:

If each $A_i \in \mathcal{B}$, then $E(\chi_{A_i}|\mathcal{B}) = \chi_{A_i}$. Thus,

$$H(\alpha|\mathcal{B}) = \int \sum_i -\chi_{A_i}(x) \log \chi_{A_i}(x) d\mu = 0$$

since χ_{A_i} takes on only 0 and 1 values.

Only if:

If $H(\alpha|\mathcal{B}) = 0$, then since the integrand is nonnegative in

$$H(\alpha|\mathcal{B}) = \int -\sum_i \mu(A_i|\mathcal{B}) \log \mu(A_i|\mathcal{B}) d\mu$$

the integrand must be zero a.e., and so each $\mu(A_i|\mathcal{B})(x)$ is 0/1 -valued. So, $\mu(A_i|\mathcal{B}) = \chi_C$ for some $C \in \mathcal{B}$.

Claim: $A_i = C \in \mathcal{B}$ (and so $\alpha \subset \mathcal{B}$)

Proof:

$$\mu(A_i \cap C) = \int_C \chi_{A_i} = \int_C \mu(A_i|\mathcal{B}) = \int_C \chi_C = \mu(C)$$

and

$$\mu(A_i \cap C^c) = \int_{C^c} \chi_{A_i} = \int_{C^c} \mu(A_i|\mathcal{B}) = \int_{C^c} \chi_C = 0$$

So, $A_i = C \bmod 0$. $\square$

Lecture 19:

Let $\mathcal{T} := (\bar{i} \mapsto T^{\bar{i}})$ be a MP $\mathbb{Z}_+^d$ action. Let α be a finite measble. partition. For a finite $G \subset \mathbb{Z}_+^d$, let

$$\alpha^G := \vee_{g \in G} T^{-g} \alpha$$

Let $D_n = [0, n-1]^d \cap \mathbb{Z}^d$.

Defn:

$$h(\mathcal{T}, \alpha) := \lim_{n \to \infty} (1/n^d) H(\alpha^{D_n})$$

Note: this defines entropy of $\mathbb{Z}^d$ actions as well.

Theorem: The limit exists. In fact, for $\overline{n} = (n_1, \ldots, n_d) \in \mathbb{N}_+^d$, let $D_{\overline{n}} = [0, n_1 - 1] \times \cdots [0, n_d - 1] \cap \mathbb{Z}^d$. Then

$$\lim_{\overline{n} \to \infty} \frac{1}{n_1 \cdots n_d} H(\alpha^{D_{\overline{n}}}) = \inf_{\overline{n} \in \mathbb{N}^d} \frac{1}{n_1 \cdots n_d} H(\alpha^{D_{\overline{n}}})$$

the limit exists (independently of how $\overline{n} \to \infty$).

Proof:

Defn: A function $f : \mathbb{N}_+^d \to \mathbb{R}$ is *subadditive* if for all $k = 1, \ldots, d$ and all i, j,

$$f(n_1, \ldots, n_{k-1}, i+j, n_{k+1}, \ldots n_d) \leq f(n_1, \ldots, n_{k-1}, i, n_{k+1}, \ldots n_d)$$
$$+ f(n_1, \ldots, n_{k-1}, j, n_{k+1}, \ldots n_d)$$

The theorem follows immediately from the following two lemmas:

Lemma 1: The function $f(\overline{n}) := H(\alpha^{D_{\overline{n}}})$ is subadditive.

Lemma 2 (sometimes called Fekete's lemma): For any subadditive function f,

$$\lim_{\overline{n} \to \infty} \frac{1}{n_1 \cdots n_d} f(\overline{n}) = \inf_{\overline{n} \in \mathbb{N}^d} \frac{1}{n_1 \cdots n_d} f(\overline{n})$$

Proof of Lemma 1: $(d = 2)$:

$\alpha^{D_{n_1, i+j}} = \alpha^{D_{n_1, i}} \vee \alpha^{D_{n_1, j} + (0,i)} = \alpha^{D_{n_1, i}} \vee T^{-(0,i)}(\alpha^{D_{n_1, j}})$.

So,

$$H(\alpha^{D_{n_1,i+j}}) \le H(\alpha^{D_{n_1,i}})+H(T^{-(0,i)}(\alpha^{D_{n_1,j}})) = H(\alpha^{D_{n_1,i}})+H(\alpha^{D_{n_1,j}})$$

Similarly,

$$H(\alpha^{D_{i+j,n_2}}) \le H(\alpha^{D_{i,n_2}}) + H(\alpha^{D_{j,n_2}})$$

□

Proof of Lemma 2: Idea for $d = 1$:
$f(rs) \le rf(s)$ and so $f(rs)/(rs) \le f(s)/s$.
Proof for $(d = 2)$:
Fix $(t_1, t_2) \in \mathbb{N}_+^2$. For $(n_1, n_2) \in \mathbb{N}_+^2$, write

$$n_1 = q_1 t_1 + r_1,\ n_2 = q_2 t_2 + r_2,\ 0 \le r_1 < t_1,\ 0 \le r_2 < t_2$$

Then

$$f(n_1, n_2) \le f(q_1 t_1, n_2) + f(r_1, n_2) \le q_1 f(t_1, n_2) + f(r_1, n_2)$$
$$\le q_1(f(t_1, q_2 t_2) + f(t_1, r_2)) + f(r_1, q_2 t_2) + f(r_1, r_2)$$
$$\le q_1 q_2 f(t_1, t_2) + q_1 f(t_1, r_2) + q_2 f(r_1, t_2) + f(r_1, r_2)$$

Then

$$\frac{1}{n_1 n_2} f(n_1, n_2) \le \frac{q_1}{n_1}\frac{q_2}{n_2} f(t_1, t_2) + \frac{q_1}{n_1}\frac{1}{n_2} f(t_1, r_2)$$
$$+\frac{1}{n_1}\frac{q_2}{n_2} f(r_1, t_2) + \frac{1}{n_1 n_2} f(r_1, r_2)$$

Since $q_1/n_1 \to 1/t_1$ as $n_1 \to \infty$ and $q_2/n_2 \to 1/t_2$ as $n_2 \to \infty$

$$\limsup_{\overline{n}\to\infty} \frac{1}{n_1 n_2} f(n_1, n_2) \le \frac{1}{t_1 t_2} f(t_1, t_2)$$

Since this holds for all (t_1, t_2),

$$\liminf_{\overline{n}\to\infty} \frac{1}{n_1 n_2} f(n_1, n_2) \le \limsup_{\overline{n}\to\infty} \frac{1}{n_1 n_2} f(n_1, n_2) \le \inf_{\overline{n}\in\mathbb{N}^d} \frac{1}{n_1 n_2} f(n_1, n_2)$$

□

For a finite alphabet F let $M_+ := F^{\mathbb{Z}^d_+}$.

The law of a stationary random field $X_{i_1,\ldots,i_d}$ defines a Borel probability measure μ^+ on M_+ and the $\mathbb{Z}^d_+$ shift action $\mathcal{T}_+$ on $(M_+, \mathcal{A}_+, \mu_+)$ is an MP semigroup action.

Similarly, letting $M := F^{\mathbb{Z}^d}$, the same stationary process defines a Borel probability measure μ on M and the $\mathbb{Z}^d$ shift action $\mathcal{T}$ on $(M, \mathcal{A}, \mu)$ is an MP group action.

For $a \in F$, let

$$A^a_+ := \{x \in M_+ : x_{\overline{0}} = a\}$$

and $\alpha_+ := \{A^a_+ : a \in F\}$. Similarly, let

$$A^a := \{x \in M : x_{\overline{0}} = a\}$$

and $\alpha := \{A^a : a \in F\}$.

Fact: $h(\mathcal{T}, \alpha) = h(\mathcal{T}_+, \alpha_+)$

Proof: the definitions give the same result:

$H(\alpha^{D_n}) = H(\alpha_+^{D_n})$ is the joint distribution of the random variable $X_{\overline{i}} : \overline{i} \in D_n$.

In particular for the iid process iid$\overline{p}$,

$$h(\mathcal{T}, \alpha) = h(\mathcal{T}, \alpha_+) = H(\overline{p})$$

because

$$H(\alpha^{D_{\overline{n}}}) = n^d H(\alpha) = n^d H(\overline{p}).$$

For any d and $\overline{p} = (1/2, 1/2)$, t hen $h(\mathcal{T}, \alpha) = \log 2$; for $\overline{p} = (1/3, 1/3, 1/3)$, then $h(\mathcal{T}, \alpha) = \log 3$.

We often write $h(X) := h(\mathcal{T}, \alpha)$.

In fact, every MP $\mathbb{Z}^d_+$ action has a corresponding MP $\mathbb{Z}^d$ action that is minimal in a categorical (universal mapping property) way;

the two actions share all of their ergodic properties: ergodicity, mixing, entropy.

Natural extension: Given a $\mathbb{Z}_+^d$ action $\mathcal{T}_+$ on $(M_+, \mathcal{A}_+, \mu_+)$, define

$$M = \{x = (x_{\bar{i}}) \in M_+^{\mathbb{Z}^d} : \bar{i} \in \mathbb{Z}^d, x_{\bar{i}} \in M_+, T^{\bar{j}}(x_{\bar{i}}) = x_{\bar{i}+\bar{j}} \ \ \forall \bar{j}\}$$

Define $S^{\bar{i}}(x)_{\bar{j}} = x_{\bar{i}+\bar{j}}$.

There is a measure μ on the product σ-algebra $\mathcal{A}$ of $M_+^{\mathbb{Z}^d}$ s.t. § is an MP $\mathbb{Z}^d$ action on $(M, \mathcal{A}, \mu)$ and the map

$$\pi : M \to M_+, \ x \mapsto x_{\bar{0}}$$

is MP and for all $\bar{i} \in \mathbb{Z}_+^d$, $\pi \circ S^{\bar{i}} = T^{\bar{i}} \circ \pi$

AND

π is minimal in the sense that for any other $(M', \mathcal{A}', \mu')$, $\mathbb{Z}^d$ action $(\S')^{\bar{i}}$ and MP $\pi' : M' \to M_+$ s.t for all $\bar{i} \in \mathbb{Z}_+^d$, $\pi' \circ (S')^{\bar{i}} = T^{\bar{i}} \circ \pi$, the map π' factors through π.

$(M, \mathcal{A}, \mu, S, \pi)$ is called the *natural extension* of $\mathcal{T}$. See section 2.5 of Keller for more on this.

The $\mathbb{Z}^d$ action is better than the $\mathbb{Z}_+^d$ action!

Lecture 20:

Recall some of the entropy properties.

Recall:

Let $\mathcal{T} := (\bar{i} \mapsto T^{\bar{i}})$ be a MP $\mathbb{Z}_+^d$ action. Let α be a finite measble. partition. For a finite $G \subset \mathbb{Z}_+^d$, let

$$\alpha^G := \vee_{g \in G} T^{-g} \alpha$$

Let $D_n = [0, n-1]^d \cap \mathbb{Z}^d$.

Defn: For an MP $\mathbb{Z}_+^d$ action $\mathcal{T}$ and (finite, measble.) partition α,

$$h(\mathcal{T}, \alpha) := \lim_{n \to \infty} (1/n^d) H(\alpha^{D_n})$$

The limit exists.

This defines $h(\mathcal{T}, \alpha)$ for an MP $\mathbb{Z}^d$ action $\mathcal{T}$.

For any $G \subset \mathbb{Z}^d$, let

$$\alpha^G := \vee_{g \in G} T^{-g} \alpha$$

if G is infinite, we take α^G to mean the σ-algebra generated α^G.

Let $\prec$ denote lexicographic order on $\mathbb{Z}^d$: $\bar{i} = (i_1, \ldots, i_d) \prec \bar{j} = (i_1, \ldots, i_d)$ if there exists $1 \le k \le d$ s.t. $i_k < j_k$ and $i_m = j_m$ for all $m > k$ (look for the "last" disagreement; this is really anti-lexicographic).

For $\bar{j} \in \mathbb{Z}^d$, let

$$\mathcal{P}^-(\bar{j}) := \{\bar{i} \in \mathbb{Z}^d : \ \bar{i} \prec \bar{j}\}$$

the lexicographic past of $\bar{j}$, and $\mathcal{P}^- := \mathcal{P}^-(\bar{0})$

Theorem Past: Let $\mathcal{T}$ be a $\mathbb{Z}^d$ action.

$$h(\mathcal{T}, \alpha) = H(\bar{0} | \ \mathcal{P}^-).$$

(the entropy of the present conditioned on the past)

Note: this yields a result for $\mathbb{Z}_+^d$ actions via the natural extension.

Proof: $(d = 2)$: Decompose

$$H(D_n) = \sum_{\bar{j} \in D_n} H(\bar{j} | \mathcal{P}^-(\bar{j}) \cap D_n)$$

Let $B_m := [-m, m]^d$.

By continuity of entropy, $H(\bar{0}|\mathcal{P}^- \cap B_m) \downarrow H(\bar{0}|\mathcal{P}^-)$. So, given $\epsilon > 0$, there exists m s.t.

$$H(\bar{0}|\mathcal{P}^-) \leq H(\bar{0}|\mathcal{P}^- \cap B_m) \leq H(\bar{0}|\mathcal{P}^-) + \epsilon$$

Let

$$S_{m,n} = \{\bar{j} \in D_n : \mathcal{P}^- \cap B_m + \ \bar{j} \subseteq \mathcal{P}^-(\bar{j}) \cap D_n\}$$

Then $|S_{m,n}| \geq n^2 - 3nm$.

For each $\bar{j} \in S_{m,n}$,

$$H(\bar{0}|\mathcal{P}^-) = H(\bar{j}|\mathcal{P}^-(\bar{j})) \leq H(\bar{j}|\mathcal{P}^-(\bar{j}) \cap D_n) \leq H(\bar{j}|\mathcal{P}^- \cap B_m + \bar{j})$$
$$= H(\bar{0}|\mathcal{P}^- \cap B_m) \leq H(\bar{0}|\mathcal{P}^-) + \epsilon.$$

Thus,

$$|\frac{H(D_n)}{n^2} - H(\bar{0}|\mathcal{P}^-)| \leq \frac{\sum_{\bar{j} \in S_{m,n}} |(H(\bar{j}|\mathcal{P}^-(\bar{j}) \cap D_n) - H(\bar{0}|\mathcal{P}^-))|}{n^2}$$
$$+ \frac{\sum_{\bar{j} \in D_n \setminus S_{m,n}} (H(\bar{j}|\mathcal{P}^-(\bar{j}) \cap D_n))}{n^2} < \epsilon + \frac{3mn \log |\alpha|}{n^2} < 2\epsilon$$

for sufficiently large n. □

Prop 1: Let $\mathcal{T}$ be an MP $\mathbb{Z}_+^d$ action and α a (finite, measble.) partition. Then for all m,

$$h(T, \alpha) = h(T, \alpha^{D_m})$$

Proof:

$$h(\mathcal{T}, \alpha^{D_m}) = \lim_{n \to \infty} (1/n^d) H((\alpha^{D_m})^{D_n})$$

$$= \lim_{n\to\infty} (1/n^d) H(\alpha^{D_{n+m}}) == \lim_{n\to\infty} (\frac{(n+m)^d}{n^d})(\frac{1}{(n+m)^d}) H(\alpha^{D_{n+m}}) = h(\mathcal{T}, \alpha)$$

Prop 2: Let $\mathcal{T}$ be an MP $\mathbb{Z}_+^d$ action and α, β be (finite, measble.) partitions. Then

$$h(\mathcal{T}, \beta) \leq h(\mathcal{T}, \alpha) + H(\beta|\alpha)$$

Proof:

$$H(\beta^{D_n}) \leq H(((\beta^{D_n}) \vee (\alpha^{D_n}))$$
$$= H(\alpha^{D_n}) + H(\beta^{D_n}|\alpha^{D_n})$$

$$\text{2nd term } \leq \sum_{\bar{j}\in D_n} H(\beta^{\bar{j}} \mid \vee_{\bar{i}\in D_n} \alpha^{\bar{i}}) \leq \sum_{\bar{j}\in D_n} H(\beta^{\bar{j}} \mid \alpha^{\bar{j}}) = n^d H(\beta|\alpha)$$

Thus,

$$(1/n^d) H(\beta^{D_n}) \leq (1/n^d) H(\alpha^{D_n}) + H(\beta \mid \alpha).$$

Let $n \to \infty$. $\square$