

## Lecture 16

First a note about weak joint and conditional typicality:

As part of the definition of joint typicality, we require both that the sample entropy  $\bar{H}(x^n, y^n)$  be within  $\delta$  of  $H(X, Y)$  and that  $x^n$  and  $y^n$  both be marginally typical. The former condition is not strong enough to imply the latter.

Similarly, we only consider  $y^n | x^n$  to be conditionally typical if  $x^n$  is typical.

As we will see, with strong typicality we don't need the extra condition.

②

## Strong typicality

Defn A type is a function  $t(x)$  for  $x \in X$  such that the  $t(x)$  are a set of frequencies:

$$\sum_{x \in X} t(x) = 1, \quad t(x) \geq 0 \quad \forall x \in X.$$

The type  $t_{x^n}(x)$  of a sequence  $x^n$  is the empirical distribution:

$$t_{x^n}(x) = \frac{1}{n} \underbrace{N(x|x^n)}_{\substack{\text{\# of times } x \text{ occurs in } x^n}}$$

Defn The strongly typical set  $T_\delta^{X^n}$  is the set of all sequences whose type has a maximum deviation of  $\delta$  from  $p_X(x)$ :

$$T_\delta^{X^n} = \left\{ x^n : |t_{x^n}(x) - p_X(x)| \leq \delta \quad \forall x \in X \right. \\ \left. \text{and } t_{x^n}(x) = 0 \text{ if } p_X(x) = 0 \right\}$$

This extra condition is necessary so that no typical sequence can include a symbol that has probability zero.

# Properties of $T_{\delta}^{\Sigma^n}$ :

(3)

i. Unit probability

$$\forall \epsilon > 0 \quad \Pr\{x^n \in T_{\delta}^{\Sigma^n}\} \geq 1 - \epsilon \text{ for large enough } n.$$

2. Exponentially small cardinality

$$|T_{\delta}^{\Sigma^n}| \leq 2^{n(H(\bar{x}) + c\delta)} \quad \text{for a positive const. } c.$$

$$\text{lower bound: } |T_{\delta}^{\Sigma^n}| \geq (1-\epsilon) 2^{n(H(\bar{x}) - c\delta)} \text{ for large enough } n.$$

3. Equipartition

$$2^{-n(H(\bar{x}) + c\delta)} \leq P_{\Sigma^n}(x^n) \leq 2^{-n(H(\bar{x}) - c\delta)}$$

for all  $x^n \in T_{\delta}^{\Sigma^n}$ .

Proof of 3:

$$P_{\Sigma^n}(x^n) = \prod_{x \in X^+} P_{\Sigma}(x)^{N(x|x^n)}$$

$$\Rightarrow \frac{1}{n} \log P_{\Sigma^n}(x^n) = \sum_{x \in X^+} \frac{1}{n} N(x|x^n) \log P_{\Sigma}(x)$$

Because  $x^n$  is strongly typical,

$$P_{\Sigma^n}(x) - \delta \leq \frac{1}{n} N(x|x^n) \leq P_{\Sigma^n}(x) + \delta$$

$$\Rightarrow H(\bar{x}) - c\delta \leq \frac{1}{n} \log P_{\Sigma^n}(x^n) \leq H(\bar{x}) + c\delta$$

$$\text{where } c = \sum_{x \in X^+} -\log(P_{\Sigma}(x)).$$

$X^+$  = symbols w/p > 0

Given the proof of 3, the proofs of 1 & 2  
 are just like the weakly typical case. ④

□

Defn A typical type  $t(x)$  satisfies

$$|t(x) - p_x(x)| \leq \delta \quad \forall x \in X$$

and  $t(x) = 0$  if  $p_x(x) = 0$ .

The set of all  $\delta$ -typical types is  $\mathcal{T}_\delta$ .

Defn A typical type class is ~~is~~ the set of <sup>all</sup> typical sequences with the same typical type  $t(x)$ .

$$T_t^{X^n} = \{x^n : t_{x^n}(x) = t(x) \quad \forall x \in X, \\ t(x) \text{ typical}\}.$$

The total number of types for sequences of length  $n$  is  $\binom{n+|X|-1}{|X|-1}$ , but most of them are not typical. (For  $|X|=2$  this is  $n+1$ .)  $T_\delta^{X^n}$  is the union of all typical type classes.

5

Surprisingly, the cardinality of a typical type class is not much smaller than all of  $T_{\delta}^{\mathbb{X}^n}$ :

$$|T_{\delta}^{\mathbb{X}^n}| \geq \frac{1}{(n+1)^{|X|}} 2^n (H(\mathbb{X}) - \eta(|X| \delta))$$

for sufficiently large  $n$ , where  $\eta(q)$  is a function that  $\rightarrow 0$  as  $q \rightarrow 0$ .

No time for the proof, but it is in the book. As  $n \rightarrow \infty$  and  $\delta \rightarrow 0$

$$|T_{\delta}^{\mathbb{X}^n}| \rightarrow \mathbb{R} 2^n H(\mathbb{X}).$$


---

Strong joint typicality

We can define a type  $t_{x^n y^n}(x, y) = \frac{1}{n} N(x, y | x^n, y^n)$  and a strongly typical set  $T_{\delta}^{\mathbb{X}^n \mathbb{Y}^n}$  just as above.

In this case, strong joint typicality implies strong marginal typicality, as is not hard to see.

The 3 properties hold for this set as well with  $H(\mathbb{X})$  replaced by  $H(\mathbb{X}, \mathbb{Y})$ .

## Strong Conditional Typicality

Let  $x^n$  be a (strongly) typical sequence drawn from the marginal of the i.i.d. distribution  $p_{\mathcal{X}^n}(x^n)$ .

We want to define a strongly, conditionally typical set  $T_s^{y^n|x^n}$ . Imagine lining up the symbols  $\{y_j\}$  from  $y^n$  with their corresponding  $\{x_j\}$  from  $x^n$ :

$x_1 x_2 x_3 \dots x_n \leftarrow$  All symbols drawn from  $\mathcal{X}^+$

$y_1 y_2 y_3 \dots y_n$

We can partition  $y^n$  into  $|\mathcal{X}^+|$  subsets  $S_x$ ,

$S_x = \text{ordered set } \{y_j : x_j = x\}$ .

Each of these subsets looks like a sequence drawn from a fixed prob. dist.  ~~$p_{\mathcal{Y}^n|x}(y|x)$~~ .

So each of these should be typical with high prob in the limit where  $n$  becomes large, but

for finite  $n$  they may deviate.

~~we compare~~  
~~so rather than comparing the distribution of~~  
~~conditional~~  
 ~~$\{y_j\}$  to the empirical distribution~~

$$\hat{f}_{y^n|x^n}(y|x) = \frac{1}{n}$$

So we compare the conditional empirical distribution

$$T_{y^n|x^n}(y|x) = \frac{T_{x^n y^n}(x,y)}{T_{x^n}(x)}$$

to the a priori distribution for each  $x \in X^+$ :

$$\bar{T}_S^{Y^n|X^n} = \left\{ Y^n : \forall (x,y) \in X \times Y \right.$$

$$\left. |N(x,y|x^n, y^n) - p(y|x)N(x|x^n)| \leq n\delta \right\}$$

and  $N(x,y|x^n, y^n) = 0$  if  $p(y|x) = 0$

where we also assume  $x^n \in T_S^{X^n}$ .

Properties

1. Unit prob:  $\forall \epsilon > 0 \quad \Pr\{Y^n \in \bar{T}_S^{Y^n|X^n}\} \geq 1 - \epsilon$  For large enough  $n$ .

2. Exp. ly small cardinality

$$|\bar{T}_S^{Y^n|X^n}| \leq 2^{n(H(Y|X) + c(\delta + \delta'))}$$

$$|\bar{T}_S^{Y^n|X^n}| \geq (1 - \epsilon) 2^{n(H(Y|X) - c(\delta + \delta'))} \quad \text{for large enough } n$$

3. Equipartition

$$2^{-n(H(Y|X) + c(\delta + \delta'))} \leq P_{Y^n|X^n}(y^n|x^n) \leq 2^{-n(H(Y|X) - c(\delta + \delta'))}$$

## Strong Q Typicality

⑧

We can define a strongly typical subspace in an exactly analogous way to the case of weak typicality:

$$T_S^{\bar{X}^n} = \text{span}\{|x^n\rangle : x^n \in T_S^{\bar{X}^n}\}$$

where the state from the Q source is  $p^{\otimes n}$  and  $P = \sum_x p_{\bar{x}}(x) |x\rangle\langle x|$ . The projector onto  $T_S^{\bar{X}^n}$  is  $\Pi_{T_S^{\bar{X}^n}}$ . The properties follow instantly from the classical properties:

$$1. \text{Tr}\{\Pi_{T_S^{\bar{X}^n}} P^{\otimes n}\} \geq 1 - \epsilon \quad \text{for large enough } n \\ \forall \epsilon > 0.$$

$$2. \text{Tr}\{\Pi_{T_S^{\bar{X}^n}}\} \leq 2^{n(H(\bar{X}) + c\delta)} \\ \text{Tr}\{\Pi_{T_S^{\bar{X}^n}}\} \geq (1 - \epsilon) 2^{n(H(\bar{X}) - c\delta)} \quad \text{for large enough } n$$

$$3. \Pi_{T_S^{\bar{X}^n}} 2^{-n(H(\bar{X}) + c\delta)} \leq \Pi P^{\otimes n} \Pi \leq \Pi 2^{-n(H(\bar{X}) - c\delta)}$$

## Joint Q Typicality

Once again, we can define a typical subspace for a bipartite source  $P_{AB}^{\otimes n}$ ; but we cannot necessarily also apply the requirement that A & B are separately typical. This can be done for certain states, such as classical states.

# Conditional Q Typicality

We can define this for classical-quantum states

$$\rho_{XB} = \sum_x p_S(x) |x\rangle\langle x| \otimes \rho_x \rightarrow \rho_{XB}^{\otimes n}$$

This is an important case, because a common situation is when we condition on a measurement outcome  $x$ , or an input variable  $x$ .

A complication is that the states  $\rho_x$  may have different eigenstates for different  $x$ 's:

$$\rho_x = \sum_y p_{Y|x} (y|x) |y_x\rangle\langle y_x|$$

So we have to use the right basis for each subsystem, depending on the corresponding symbol  $x_j$ :

$$T_S^{Y^n|x^n} = \text{span} \left\{ \bigotimes_{j=1}^n |(y_{x_j})_j\rangle : y^n = y_1 \dots y_n \in T_S^{Y|x^n} \right\}$$

The usual properties apply, as shown in the text book.