

Lecture 15

(i)

We can define a typical subspace for a bipartite state P_{AB} , just as for any state:

$$P_{AB} = \sum_z p_z(z) |z\rangle\langle z|.$$

However the states $\{|z\rangle\}$ will generally be entangled. Classically, a jointly typical sequence (x^n, y^n) will also have x^n and y^n be typical for the marginal distributions. But that need not be the case for quantum states; the typical projectors $\Pi_s^{x^n}$ and $\Pi_s^{y^n}$ for P_A and P_B need not commute with Π_s^z for P_{AB} . So joint typicality is not generically well defined in the Q case.

For classical states

$$P_{AB} = \sum_{x,y} p_{xy}(x,y) |x\rangle_A\langle x| \otimes |y\rangle_B\langle y|$$

joint typicality is well defined, and they inherit the properties of jointly typical sequences.

Conditional Q-Typicality

A case of some importance in QIT are so-called "classical-quantum states":

$$\rho_{XB} = \sum_x p_X(x) |x\rangle\langle x| \otimes \rho_x$$

For states of this kind (which can result from, e.g., measuring half of a bipartite state), we can define conditionally typical subspaces:

~~ρ_X~~ $\rho_X = \sum_y p_{Y|X}(y|x) |y_x\rangle\langle y_x|$
spectral
decomposition

The orthonormal bases $\{|y_x\rangle\}$ are different, in general for different values of x . So

$$\rho_{XB}^{\otimes n} = \sum_{x^n} p_{X^n}(x^n) |x^n\rangle\langle x^n| \otimes \sum_{y^n} p_{Y|X}(y^n|x^n) |y_{x^n}\rangle\langle y_{x^n}|$$

$$= \sum_{x^n} p_{X^n}(x^n) |x^n\rangle\langle x^n| \otimes \rho_{X^n}^{\otimes n} |y_{x^n}\rangle\langle y_{x^n}|$$

$$\text{where } |x^n\rangle \equiv |x_1\rangle |x_2\rangle \dots \otimes |x_n\rangle$$

$$\text{and } |y_{x^n}\rangle = |(y_1)_{x_1}\rangle |(y_2)_{x_2}\rangle \dots |(y_n)_{x_n}\rangle.$$

Putting this all together:

Defn The weakly conditional subspace $T_S^{B^n|X^n}$ is

$$T_S^{B^n|X^n} = \overline{\text{span}} \left\{ |Y_{X^n}^n\rangle \mid |\bar{H}(Y^n|X^n) - H(B|X)| \leq \delta \right\}$$

Here, $H(B|X) = \sum_x P_X(x) H(P_x)$ and

$$\bar{H}(Y^n|X^n) = -\frac{1}{n} \log P_{Y^n|X^n}(Y^n|X^n)$$

$$P_{Y^n|X^n}(Y^n|X^n) = \prod_{j=1}^n P_{Y|X}(Y_j|X_j).$$

The conditionally typical projector onto $T_S^{B^n|X^n}$

is $\Pi_S^{B^n|X^n} = \sum_{Y^n \in T_S^{B^n|X^n}} |Y_{X^n}^n\rangle \langle Y_{X^n}^n|$.

The properties of these subspaces are similar to typical subspaces, but with a subtle distinction:

1. Unit prob.

$$\forall \epsilon > 0 \quad \mathbb{E}_{X^n} \left[\text{Tr} \left\{ \Pi_S^{B^n|X^n} P_{X^n}^{\otimes n} \right\} \right] \geq 1 - \epsilon \quad \text{for large enough } n.$$

2. Exponentially small dimension

$$\text{Tr} \left\{ \Pi_S^{B^n|X^n} \right\} \leq 2^{n(H(B|X) + \delta)}$$

$$\mathbb{E}_{X^n} \left[\text{Tr} \left\{ \Pi_S^{B^n|X^n} \right\} \right] \geq (1 - \epsilon) 2^{n(H(B|X) - \delta)} \quad \text{for large enough } n$$

3. Equipartition

④

$$2^{-n(H(B|\bar{X})+\delta)} \prod_{\gamma}^{B^n|X^n} \leq \prod_{\gamma}^{B^n|X^n} P_{X^n} \prod_{\gamma}^{B^n|X^n} \leq 2^{-n(H(B|\bar{X})-\delta)} \prod_{\gamma}^{B^n|X^n}$$

The difference in the conditional case is in properties 1 & 2, where we take expectations of the sequences \bar{X}^n .

Schumacher Compression

We now have the tools for the first asymptotic QIT result: Schumacher compression:

Thm Let ρ_A be the density matrix of a quantum source. Then $H(\rho)$ is the smallest achievable rate for Q data compression.

Proof: This is two-fold—a direct coding theorem (achievability) and a converse theorem (optimality).

Direct coding Spectrally decompose P :

$$\rho_A = \sum_z p_z(z) |z\rangle\langle z|.$$

There will be a set $T_S^{2^n}$ of typical sequences z^n , ⑤
 that give us a typical subspace $T_S^{2^n}$ with
 typical projector $\Pi_S^{2^n}$.

Here is the procedure:

1. Measure $P_A^{2^n}$ with the projectors $\{\Pi_S^{2^n}, I - \Pi_S^{2^n}\}$.
 If we get the second result the compression fails.
 For any $\epsilon > 0$ we can make $P_{\text{fail}} < \epsilon$ for large enough n

2. $T_S^{2^n}$ is spanned by states $|z^n\rangle$ where $z^n \in T_S^{2^n}$.
 The size of this set is $\leq 2^{n(H(A)+S)}$. Define a function
 $f(z^n) \in \{0, 1\}^{n(H(A)+S)}$ that maps the typical sequences
 z^n to binary strings of length $n(H(z) + S)$.

3. If I succeeded, apply an isometry

$$U = \sum_{z^n \in T_S^{2^n}} |f(z^n)\rangle \langle z^n|$$

to the state. Send the resulting $\otimes^{n(H(A)+S)}$ qubits
 through a noiseless channel to Bob.

(If I failed, send a standard "failed" string)

4. Bob applies the inverse of the isometry H , his state

$\rightarrow \Pi_S^{2^n} P_A^{\otimes n} \Pi_S^{2^n}$; we know that

$$\|P_A^{\otimes n} - \Pi_S^{2^n} P_A^{\otimes n} \Pi_S^{2^n}\|_1 \leq 2\sqrt{\epsilon},$$

and its fidelity is therefore $\geq 1 - 2\sqrt{\epsilon}$.

By making n large, both the failure prob ϵ
 and the trace distance $2\sqrt{\epsilon} \rightarrow 0$, and the
 rate approaches $H(A)$ □

(6)

Converse Theorem

Suppose Alice wants to reversibly compress n copies of the state σ_A with error prob $< \epsilon$.

First purify $\sigma_A \rightarrow \Phi_{AR}$ $\sigma_A^{\otimes n} \rightarrow \Phi_{AR}^{\otimes n} \equiv \Phi_{A^n R^n}$.

An $CQ(n, R+S, \epsilon)$ compression code compresses n copies with a rate $R+S$ and error prob ϵ .

$$A^n \xrightarrow[\text{compress}]{} W \xrightarrow[\text{recover}]{} \hat{A}^n$$

Suppose the final state is $\omega_{A^n R^n}$. We want

$$\|\omega_{A^n R^n} - \Phi_{A^n R^n}\|_1 \leq \epsilon.$$

We will use the data processing inequality.

$$\begin{aligned} 2^n R &= \log(2^{nR}) + \log(2^{nR}) && \text{based on dim.} \\ &\geq |H(W)_\omega| + |H(W|R^n)_\omega| && \text{properties of I} \\ &\geq |H(W)_\omega - H(W|R^n)_\omega| && \text{Def. of } I(W; R^n) \\ &= I(W; R^n)_\omega \end{aligned}$$

$$\geq I(\hat{A}^n; R^n)_\omega \quad \text{... data processing}$$

$$\geq I(\hat{A}^n; R^n)_\phi - n\epsilon' \quad \text{... Alicki-Fannes} \quad \epsilon' = G(R) + H(\epsilon)/n$$

$$= I(A^n; R^n)_\phi - n\epsilon'$$

$$= H(A^n)_{\phi} + H(R^n)_{\phi} - H(A^n R^n)_{\phi} - n\epsilon' \quad \leftarrow \text{defn of } I$$

$$= 2H(A^n)_{\phi} - n\epsilon' \quad \leftarrow \text{because } \phi \text{ is pure.}$$

Putting this all together we get

$$2nR \geq \underbrace{2H(A^n)_{\phi}}_{nH(A)_\phi} - n\epsilon'$$

$$\Rightarrow R \geq H(A)_\phi - \epsilon'. \quad \square$$

For example, if our source is

$$\left\{ \left(\frac{1}{2}, |0\rangle \right), \left(\frac{1}{2}, |+\rangle \right) \right\}$$

then $P_A = \frac{1}{2}|0\rangle\langle 0| + \frac{1}{2}|+\rangle\langle +| = \cos^2 \frac{\pi}{8}|+\rangle\langle +| + \sin^2 \frac{\pi}{8}|-\rangle\langle -|$

$$|+\rangle = \cos \frac{\pi}{8}|0\rangle + \sin \frac{\pi}{8}|1\rangle$$

$$|-\rangle = \sin \frac{\pi}{8}|0\rangle - \cos \frac{\pi}{8}|1\rangle$$

So we can compress by an amount

$$h(\cos^2 \pi/8) \approx 0.6009.$$

The compressability arises because $|0\rangle$ and $|+\rangle$ are not distinguishable.