From classical to quantum compression

Last week: Shannon entropy $H(p)$ = optimal rate for compressing a classical data source described by $p \in P(\Sigma)$.

Today: $\rho \in D(\mathcal{H})$? Let's first define entropy + then discuss compression.

Von Neumann Entropy of $\rho \in D(\mathcal{H})$:

as usual: $\rho = \sum_i p_i |e_i\rangle\langle e_i|$

$f(\rho) := \sum_i f(p_i) |e_i\rangle\langle e_i|$

$$H(\rho) = H(p) = -\sum_i p(i) \log p(i) = -\mathrm{tr}[\rho \cdot \log \rho]$$

where p vector of eigenvalues of ρ (repeated acc. to multiplicity).

Properties: * $H(\rho) = H(V\rho V^\dagger)$ for any isometry V

* $0 \leq H(\rho) \leq \log \mathrm{rk}(\rho) \leq \log d$, where $d := \dim \mathcal{H}$

* $=0$ iff pure, $= \log d$ iff $\rho = \frac{I}{d}$ (maximally mixed)

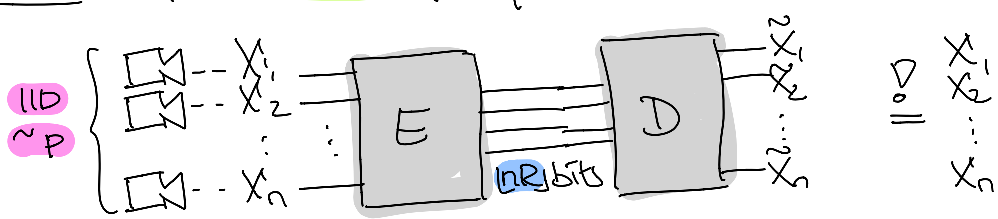
* Continuous Pf: $\rho \approx \sigma \Rightarrow$ sorted eigenvalues of $\rho \approx$ those of σ (□)

* Concave, i.e. $H(\sum_x p_x \rho_x) \geq \sum_x p_x H(\rho_x)$ HW

How to define the task of q. compression?

Motivation: Classical compression and correlations

Recall: (n,R,δ)-code for $p \in P(\Sigma)$:



with probability $\geq 1-\delta$

i.e. $\left\{ \begin{array}{l} E: \Sigma^n \to \{0,1\}^{\lfloor nR \rfloor} \\ D: \{0,1\}^{\lfloor nR \rfloor} \to \Sigma^n \end{array} \right\}$ s.th. $\Pr(\tilde{X}^n \neq X^n) \leq \delta$

Shannon: $H(p)$ is optimal rate → last week for precise statement

How about if $X_1, \dots, X_n \overset{\text{IID}}{\sim} p$, but correlated to another random variable Y?

e.g. $Y = X_1$ or $Y = X_1 \oplus \dots \oplus X_n$ or even $Y = X^n$

$D(E(X^n))$

Let P_{X^nY}, $P_{\tilde{X}^nY}$ joint distribution of (X^n, Y) and of $(\tilde{X}^n, Y)$, respectively

Lem: (D,E) is an (n,R,δ)-code iff $T(P_{X^nY}, P_{\tilde{X}^nY}) \le \delta$ for all RV's (X^n, Y) s.th. $X^n \overset{\text{IID}}{\sim} p$

↑ trace dist of probability distr's

$T(q,\tilde{q}) := \frac{1}{2}\sum_z |q(z) - \tilde{q}(z)|$

Sketch: $(\Rightarrow)$ $T(P_{X^nY}, P_{\tilde{X}^nY})$

$\le \Pr((X^n,Y) \neq (\tilde{X}^n, Y)) = \Pr(X^n \neq \tilde{X}^n) \le \delta$

→ PRACTICE

$(\Leftarrow)$ Choose $Y = X^n$! Then:

$$\Pr(D(E(X^n)) \neq X^n) = \big|\Pr(\tilde{X}^n \neq Y) - \underbrace{\Pr(X^n \neq Y)}_{=0}\big| \le \delta \qquad (\square)$$

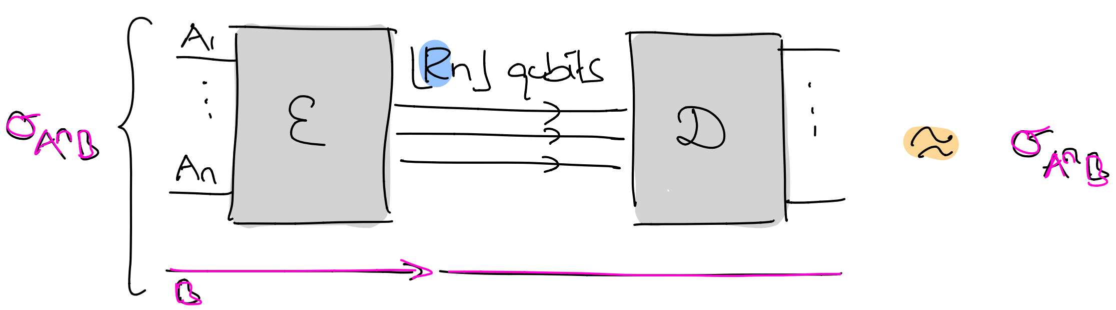
same

Thus: Correlations are preserved & this characterizes a reliable compression protocol.

We will take this as the def'n in the q. case!

Quantum compression

Def: (n,R,δ)-quantum code for $\rho \in D(\mathcal{H}_A)$: quantum channels
$$\mathcal{E} \in C(\mathcal{H}_A^{\otimes n}, (\mathbb{C}^2)^{\otimes \lfloor Rn \rfloor}) \quad \text{and} \quad \mathcal{D} \in C((\mathbb{C}^2)^{\otimes \lfloor Rn \rfloor}, \mathcal{H}_A^{\otimes n})$$
s.th.
$$F((\mathcal{D} \circ \mathcal{E} \otimes \mathcal{I}_B)[\sigma_{A^nB}], \sigma_{A^nB}) \ge 1 - \delta$$
for all $\mathcal{H}_B$, $\sigma_{A^nB} \in D(\mathcal{H}_A^{\otimes n} \otimes \mathcal{H}_B)$ with $\sigma_{A^n} = \rho^{\otimes n}$



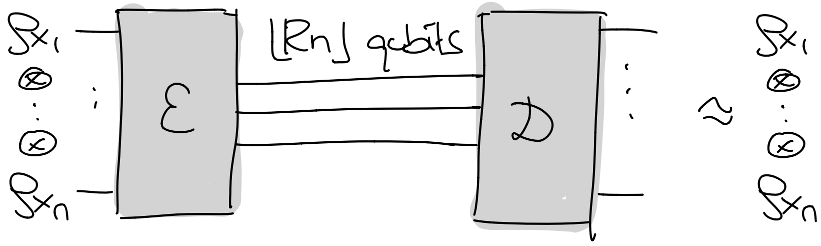
Same idea as $T \le \delta$, except that we use fidelity

Plan: ① Relate to "ordinary" compression. ② Simplify definition
③ Prove analog of Shannon's theorem.

Surprising definition! How is this related to compression of a q. source?

emits $\rho_x \in D(\mathcal{H}_A)$ w/ probability $p(x)$, $x \in \Sigma$ $\overset{\text{IID}}{\Longrightarrow}$ emits $\rho_{x_1} \otimes \cdots \otimes \rho_{x_n}$ w/ probability $p(x_1)\cdots p(x_n)$

$\boxed{\text{WANT:}}$



$\approx$ on average or even w.h.p.

Choose (n, R, δ) – q. code for average output state $\boxed{\rho = \sum_x p(x) \rho_x}$

Then:

$$\sum_{x^n} p(x_1)\cdots p(x_n)\, F(\rho_{x_1}\otimes\cdots\otimes\rho_{x_n}, \mathcal{D}(\mathcal{E}(\rho_{x_1}\otimes\cdots\otimes\rho_{x_n}))) \geq 1-\delta$$

Pf: Define $\sigma_{A^nY} = \sum_{x^n} p(x_1)\cdots p(x_n)\, \rho_{x_1}\otimes\cdots\otimes\rho_{x_n} \otimes |x^n\rangle\langle x^n|_Y \in D(\mathcal{H}_A^{\otimes n}\otimes\mathcal{H}_Y)$, $\mathcal{H}_Y = \mathbb{C}^{\Sigma^n}$

$\Rightarrow$ LHS $\overset{\text{HW 5}}{=} F(\sigma_{A^nY}, (\mathcal{D}\circ\mathcal{E}\otimes\mathcal{I}_Y)[\sigma_{A^nY}]) \geq 1-\delta$ $\square$

Thus: Can compress ANY source w/ average output state ρ! $\rightarrow$ $\boxed{\text{HW}}$

$\boxed{\text{WARNING:}}$ In general:
- $|\Sigma| \neq \dim \mathcal{H}_A$
- ρ_x NOT pairwise orthogonal,
- $p(x)$ NOT eigenvalues of ρ

Next, let's simplify the definition of a q. code...

> **Def:** Channel fidelity of $T_A \in C(\mathcal{H}_A, \mathcal{H}_A)$ and $\rho_A \in D(\mathcal{H}_A)$:
>
> $$F(\Phi_A, \rho_A) = \inf\{F(\sigma_{AB}, (T_A\otimes\mathcal{I}_B)[\sigma_{AB}]) : \mathcal{H}_B, \sigma_{AB} \text{ s.th. } \sigma_A = \rho_A\}$$

$\hookrightarrow$ Condition in (n,R,δ)–code can be written as $F(\mathcal{D}\circ\mathcal{E}, \rho^{\otimes n}) \geq 1-\delta$

> **Lem:** Let $|\Psi_{AB}\rangle \in \mathcal{H}_A\otimes\mathcal{H}_B$ be ANY purification of ρ_A. Then:
>
> $$F(T_A, \rho_A) = F(|\Psi_{AB}\rangle\langle\Psi_{AB}|, (T_A\otimes\mathcal{I}_B)[|\Psi_{AB}\rangle\langle\Psi_{AB}|])$$

$\leftarrow$ no more "for all σ_{AB}" !

Pf: Monotonicity $\leadsto$ σ_{AB} pure. Isometric invariance $\leadsto$ Any purification is ok. $\square$

$\boxed{\text{Cor: If } T_A[M_A]=\sum_i X_i M_A X_i^\dagger \text{ Kraus repr: } F(T_A, \rho_A)^2=\sum_i |\mathrm{tr}[X_i \rho_A]|^2}$

Pf: Using the previous lemma:

$$F(T_A, \rho_A)^2 = \langle \Psi_{AB}| \sum_i (X_i \otimes I_B)|\Psi_{AB}\rangle\langle\Psi_{AB}|(X_i^\dagger \otimes I_B)|\Psi_{AB}\rangle$$

Fidelity with one state pure

$$= \sum_i |\underbrace{\langle \Psi_{AB}|X_i \otimes I_B|\Psi_{AB}\rangle}_{= \mathrm{tr}[|\Psi\rangle\langle\Psi|(X_i \otimes I_B)]}|^2 = \sum_i |\mathrm{tr}[\rho_A X_i]|^2 \qquad \square$$

Schumacher's theorem & Typical Subspaces

Schumacher's Theorem: Let $0<\delta<1$:

① If $R > H(\rho)$: $\exists n_0 : \forall n \geq n_0 : \exists (n, R, \delta)$-q code

② If $R < H(\rho)$: $\exists n_0 : \forall n \geq n_0 : \nexists (n, R, \delta)$-q code

key tool:

Typical subspace: $S_{n,\varepsilon}(\rho) = \mathrm{Span}\{|e_{y_1}\rangle \otimes \dots \otimes |e_{y_n}\rangle : y^n \in T_{n,\varepsilon}(q)\}$

where $\rho = \sum_y q_y |e_y\rangle\langle e_y|$ eigendecomposition

$\Pi_{n,\varepsilon}$ = orthog. projection onto $S_{n,\varepsilon}(\rho)$ "typical projector"

Properties. ⓪ Eigenvalues of $\Pi_{n,\varepsilon}\rho^{\otimes n}\Pi_{n,\varepsilon}$ are in $2^{-n(H(\rho)\pm\varepsilon)}$

① $\dim S_{n,\varepsilon}(\rho) = |T_{n,\varepsilon}(q)| \leq 2^{n(H(\rho)+\varepsilon)}$

② $\underbrace{\mathrm{tr}[\Pi_{n,\varepsilon}\rho^{\otimes n}]}_{= \sum_{y^n \in T_{n,\varepsilon}(q)} q(y_1)\cdots q(y_n)} \longrightarrow 1$

Pf: This follows directly from the corresponding properties of $T_{n,\varepsilon}(q)$. $\square$

Next week, we will use this to prove Schumacher's theorem.

→ PRACTICE & LECTURE NOTES for a sneak peak!