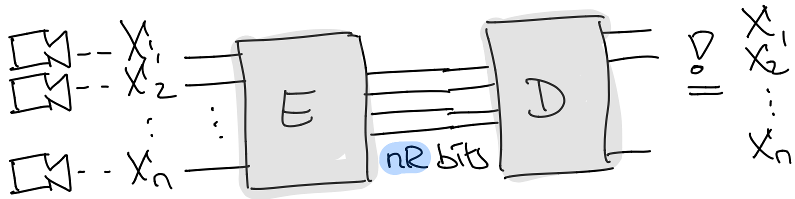
From classical to quantum compression

Last week: Compressing a classical data source described by $p \in \mathcal{P}(\Sigma)$ using (n,R,δ)-code:



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. $\displaystyle\sum_{x : D(E(x))=x} p(x_1)\cdots p(x_n) \geq 1-\delta.$ i.i.d.

Shannon: $H(p) = -\sum_x p(x) \log p(x)$ is optimal rate ⇝ last week for precise stmt

---

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

---

e.g. $Y = X_1 \oplus \dots \oplus X_n$

Let $q, \tilde{q} \in \mathcal{P}(\Sigma^n \times \Gamma)$ distribution of (X,Y) and $(D(E(X)),Y)$, respectively.

Then:

$$\|q-\tilde{q}\|_1 := \sum_{x_1,\dots,x_n,y} |q(x,y) - \tilde{q}(x,y)| \leq 2\delta \qquad (\ast)$$

Fact: (D,E) defines an (n,R,δ)-code for p iff $(\ast)$ holds for any joint distr. of RVs (X,Y) s.th. $X_1,\dots,X_n \overset{\text{IID}}{\sim} p$.

Sketch: Set $Z=(X,Y)$ & $\tilde{Z} = (D(E(X)),Y)$. Then:

$(\Rightarrow)$ $\|q-\tilde{q}\|_1 = \|p_Z - p_{\tilde{Z}}\|_1 \leq 2\cdot \Pr(Z \neq \tilde{Z}) = 2 \cdot \Pr(X \neq D(E(X))) \leq 2\delta$

$(\Leftarrow)$ Choose $Y=X$. Then:

$$\Pr(D(E(X)) \neq X) = |\Pr(D(E(X)) \neq Y) - \underbrace{\Pr(X \neq Y)}_{=0}| \leq \tfrac{1}{2} 2\delta = \delta. \qquad \square$$

⇝ $\boxed{\text{ESET}}$

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

How do these two types of compression translate to q. states?

Let's start with 2nd. Will see that stronger in the q. world...!

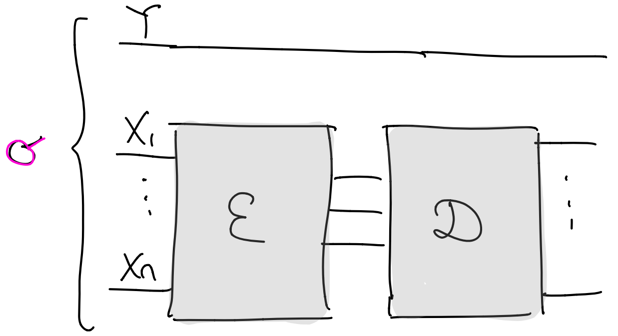
$p \in \mathcal{P}(\Sigma)$?⇝ $\rho \in D(\mathcal{X})$

Quantum compression

(n,R,δ) – quantum code for $\rho\in D(\mathcal{X})$: q. channels $\mathcal{E}\in C(\mathcal{X}^{\otimes n}, (\mathbb{C}^2)^{\otimes\lfloor Rn\rfloor})$, $\mathcal{D}\in C((\mathbb{C}^2)^{\otimes\lfloor Rn\rfloor}, \mathcal{X}^{\otimes n})$ s.th.

$$F((\mathcal{D}\circ\mathcal{E}\otimes\mathcal{I}_Y)[\sigma],\sigma)\geq 1-\delta \qquad \boxed{\otimes}$$

for all Y, $\sigma\in D(\mathcal{X}^{\otimes n}\otimes Y)$ with $\mathrm{tr}_Y[\sigma]=\rho^{\otimes n}$.

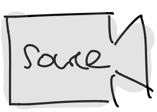


$\approx \sigma$ Compare $\otimes$ above

Here we used the fidelity $F(\sigma,\gamma)=\|\sqrt{\sigma}\sqrt{\gamma}\|_1=\mathrm{tr}\sqrt{\sqrt{\sigma}\gamma\sqrt{\sigma}}$ from L2.

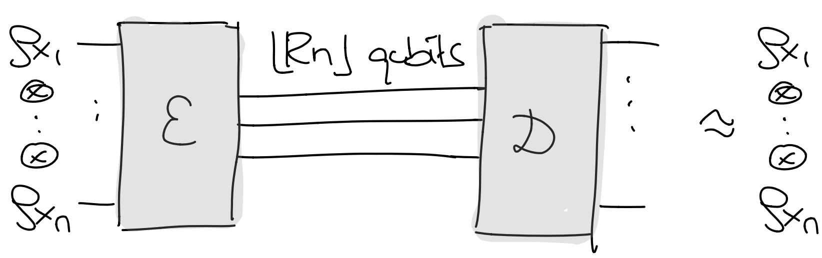
Plan: ① Relate to "ordinary" compression. ② Simplify def'n. ③ Prove analog of Shannon's theorem.

How does this relate to "ordinary" compression of a q. data source?



emits $\rho_x\in D(\mathcal{X})$ w/ probability p_x $(x\in\Sigma)$ $\Longrightarrow$ emits $\rho_{x_1}\otimes\cdots\otimes\rho_{x_n}$ w/ probability $p_{x_1}\cdots p_{x_n}$

Goal:



with high probability

Choose (n,R,δ)-code for average output state $\rho=\sum_x p_x\rho_x$. Then:

$$\sum_x p(x_1)\cdots p(x_n)\, F(\mathcal{D}[\mathcal{E}[\rho_{x_1}\otimes\cdots\otimes\rho_{x_n}]],\rho_{x_1}\otimes\cdots\otimes\rho_{x_n})$$

$$\overset{\text{ESET}}{=} F(\sigma,(\mathcal{D}\circ\mathcal{E}\otimes\mathcal{I}_Y)[\sigma])\geq 1-\delta$$

where $\sigma = \sum_x p(x_1)\cdots p(x_n)\, \rho_{x_1}\otimes\cdots\otimes\rho_{x_n}\otimes|x_1\cdots x_n\rangle\langle x_1\cdots x_n| \in D(\mathcal{X}^{\otimes n}\otimes\mathcal{Y}^{\otimes n})$ where $\mathcal{Y}=\mathbb{C}^{\Sigma}$

WARNING: In general: $|\Sigma| \neq \dim \mathcal{X}$, (ρ_x) NOT pairwise orthogonal, (p_x) NOT eigenvalues of ρ.

Fidelity & Channels

Recall:

* $F(|\psi\rangle\langle\psi|, \gamma) = \sqrt{|\langle\psi|\gamma|\psi\rangle|}$
* monotonicity: $F(\mathrm{tr}_Y[\rho], \mathrm{tr}_Y[\sigma]) \geq F(\rho,\sigma)$
* unitary invariance: $F(\rho,\sigma) = F(U\rho U^*, U\sigma U^*)$
* multiplicativity: $F(\rho\otimes\tilde\rho, \sigma\otimes\tilde\sigma) = F(\rho,\sigma)\,F(\tilde\rho,\tilde\sigma)$

→ PSET for further properties

Channel fidelity: Given $T\in C(\mathcal{X},\mathcal{X})$ and $\rho\in D(\mathcal{X})$, define

$$F(T,\rho) = \inf\{F((T\otimes I_Y)[\sigma],\sigma) : \mathcal{Y},\ \sigma\in D(\mathcal{X}\otimes\mathcal{Y}) \text{ s.th. } \mathrm{tr}_Y[\sigma]=\rho\}$$

↳ Condition ① can be written as $F(D\circ E, \rho) \geq 1-\delta$. How to calculate?

Lem: Let $|\psi\rangle\in\mathcal{X}\otimes\mathcal{Y}$ be ANY purification of ρ. Then:

$$F(T,\rho) = F(|\psi\rangle\langle\psi|, (T\otimes I_Y)[|\psi\rangle\langle\psi|])$$

← no optimization!

Pf: Monot ⇝ pure states. Addit. ⇝ same space. Unitary invariance ⇝ QED. □

Lem: If $\{A_x\}$ Kraus ops for T: $F(T,\rho)^2 = \sum_x |\mathrm{tr}[A_x\rho]|^2$

Pf: Let $|\psi\rangle\in\mathcal{X}\otimes\mathcal{Y}$ purification. Then:

$$F(T,\rho)^2 = \langle\psi|(T\otimes I_Y)[|\psi\rangle\langle\psi|]\,|\psi\rangle = \sum_x |\langle\psi|A_x\otimes I_Y|\psi\rangle|^2$$

fidelity w/ one pure state

$$= \sum_x |\mathrm{tr}[|\psi\rangle\langle\psi|(A_x\otimes I_Y)]|^2 = \sum_x |\mathrm{tr}[\rho A_x]|^2.$$

□

Schumacher's theorem

Q. entropy of $\rho \in D(\mathcal{H})$: as usual: $\rho = \sum_y q_y |e_y\rangle\langle e_y|$

$f(\rho) := \sum_y f(q_y) |e_y\rangle\langle e_y|$

$$H(\rho) = H(q) = -\sum_y q_y \log q_y = -\mathrm{tr}[\rho \cdot \log \rho]$$

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

also known as "von Neumann entropy"

Thm (Schumacher): Let $\delta \in (0,1)$.

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

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

entropy = "optimal" rate

Main tool: Typical subspaces for $\rho = \sum_y q_y |e_y\rangle\langle e_y|$ (eigendecomposition):

$$S_{n,\varepsilon}(\rho) = \mathrm{span}\, \{ |e_{y_1}\rangle \otimes \cdots \otimes |e_{y_n}\rangle : y = (y_1, \ldots, y_n) \in T_{n,\varepsilon}(q) \}$$

Properties:

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

② $\mathrm{tr}[\Pi_{n,\varepsilon} \rho^{\otimes n}] \longrightarrow 1$ where $\Pi_{n,\varepsilon}$ = orthog. projection onto $S_{n,\varepsilon}(\rho)$

follow from properties of $T_{n,\varepsilon}(q)$

Proof of Schumacher's thm, part (A): Choose $\varepsilon = \frac{R - H(\rho)}{2}$. Then:

$$n(H(\rho) + \varepsilon) = n(R - \varepsilon) \leq \lfloor nR \rfloor \quad \text{if } n \geq \tfrac{1}{\varepsilon}.$$

Then, $\exists$ injection $E_n : T_{n,\varepsilon}(q) \to \{0,1\}^{\lfloor nR \rfloor}$. Define

$$V_n := \sum_{y \in T_{n,\varepsilon}(q)} |E_n(y)\rangle (\langle e_{y_1}| \otimes \cdots \otimes \langle e_{y_n}|)$$

partial isometry $S_{n,\varepsilon} \hookrightarrow \ldots$

Note: $V_n^* V_n = \Pi_{n,\varepsilon}$. Define

Interpretation: measure $\{\Pi, I - \Pi\}$. if Π, map into $\lfloor nR \rfloor$ qubits. else send α.

$$\mathcal{E}_n(\sigma) := V_n \sigma V_n^* + \mathrm{tr}[(I - \Pi_{n,\varepsilon})\sigma] \cdot \alpha$$

arbitrary states

$$\mathcal{D}_n(\sigma) := V_n^* \sigma V_n + \mathrm{tr}[(I - V_n V_n^\dagger)\sigma] \cdot \beta$$

similar.

$\mathcal{D}_n\mathcal{E}_n$ has kraus operators $\{V_n^\dagger V_n = \Pi_n, \ldots\}$

$\overset{\text{Lem}}{\Longrightarrow}$ $F(\mathcal{D}_n\mathcal{E}_n, \rho^{\otimes n}) \geq |\mathrm{tr}[\Pi_n \rho^{\otimes n}] \longrightarrow 1.$ $\square$

Proof of part (B): Crucial fact: If P is an orthogonal projection of rank $\leq 2^{nR}$ then, for $\varepsilon = (H(\rho) - R)/2$,

$$\mathrm{tr}[P\rho^{\otimes n}] = \mathrm{tr}[P \cdot \Pi_{n,\varepsilon} \rho^{\otimes n} \Pi_{n,\varepsilon}] + \mathrm{tr}[P(I - \Pi_{n,\varepsilon})\rho^{\otimes n}(\ )]$$

commute

$$\leq 2^{-\varepsilon n} + (1 - \mathrm{tr}[\Pi_{n,\varepsilon}\rho^{\otimes n}]) \longrightarrow 0 \quad \text{uniformly in choice of } P$$

Now choose kraus op's for $\mathcal{E}_n, \mathcal{D}_n$:

$$A_i \in L(\mathcal{H}^{\otimes n}, (\mathbb{C}^2)^{\otimes \lfloor Rn \rfloor}) \ \& \ B_j \in L((\mathbb{C}^2)^{\otimes \lfloor Rn \rfloor}, \mathcal{H}^{\otimes n})$$

$\hookrightarrow$ $\{C_k\} = \{A_i B_j\}$ Kraus op's have rank $\leq 2^{Rn}$

Choose $P_k :=$ orthog. projection onto range of C_k.

$$\Longrightarrow F(\mathcal{D}_n\mathcal{E}_n, \rho^{\otimes n})^2 = \sum_k |\mathrm{tr}[C_k \rho^{\otimes n}]|^2 = \sum_k |\mathrm{tr}[P_k C_k \rho^{\otimes n}]|^2$$

$$= \sum_k |\mathrm{tr}[C_k \sqrt{\rho^{\otimes n}} \sqrt{\rho^{\otimes n}} P_k]|^2$$

$$\leq \sum_k \underbrace{\mathrm{tr}[C_k \rho^{\otimes n} C_k^*]}_{\text{prob. dist.}} \cdot \underbrace{\mathrm{tr}[P_k \rho^{\otimes n}]}_{\to 0 \text{ uniformly in } k} \longrightarrow 0. \quad \square$$