QIC 820 CO781 / 486 CS 867 F23

Part 3 lecture 2 : von Neumann entropy
Quantum data compression
Entanglement concentration & dilution
Entropy of entanglement

Def: Let $\rho \in D(X)$ with spectral decomp

$$\rho = \sum_{v=1}^{d} p(v)\, |e_v\rangle\langle e_v| \qquad (d = \dim(X))$$

This induces a rv V on sample space $\{1,2,\dots d\}$ with distribution $p(v)$.

① The von Neumann entropy of ρ, denoted $S(\rho)$, is $H(V)$.

② Fix $n \in \mathbb{N}$, $\delta > 0$. Let $T_{n,\delta}$ be typical set for n iid draws of V.

For $\vec{v} = v_1 v_2 \dots v_n$, let

$$|e_{\vec{v}}\rangle = |e_{v_1}\rangle \otimes |e_{v_2}\rangle \otimes \dots \otimes |e_{v_n}\rangle \in X^{\otimes n}$$

ⓐ The δ-typical space of $\rho^{\otimes n}$, $S_{n,\delta} := \text{span}\{ |e_{\vec{v}}\rangle : \vec{v} \in T_{n,\delta} \}$

ⓑ Let $\Pi_{n,\delta} = \sum_{\vec{v} \in T_{n,\delta}} |e_{\vec{v}}\rangle\langle e_{\vec{v}}|$ (projection onto $S_{n,\delta}$)

Obs:

① $\dim(S_{n,\delta}) = |T_{n,\delta}| \overset{\text{AEP}}{\leq} 2^{n(H(V)+\delta)} = 2^{n(S(\rho)+\delta)}$

② $$\mathrm{Tr}(\rho^{\otimes n}\, \Pi_{n,\delta}) = \mathrm{Tr}\left[\sum_{\substack{\text{all } \vec{v}' \\ \uparrow \\ n\text{-tuple}}} p(\vec{v}')\, |e_{\vec{v}'}\rangle\langle e_{\vec{v}'}| \right]\left[\sum_{\vec{v} \in T_{n,\delta}} |e_{\vec{v}}\rangle\langle e_{\vec{v}}| \right]$$

$$= \sum_{\vec{v} \in T_{n,\delta}} p(\vec{v}) \geq 1 - \varepsilon \qquad (\text{if } n \geq n_0)$$

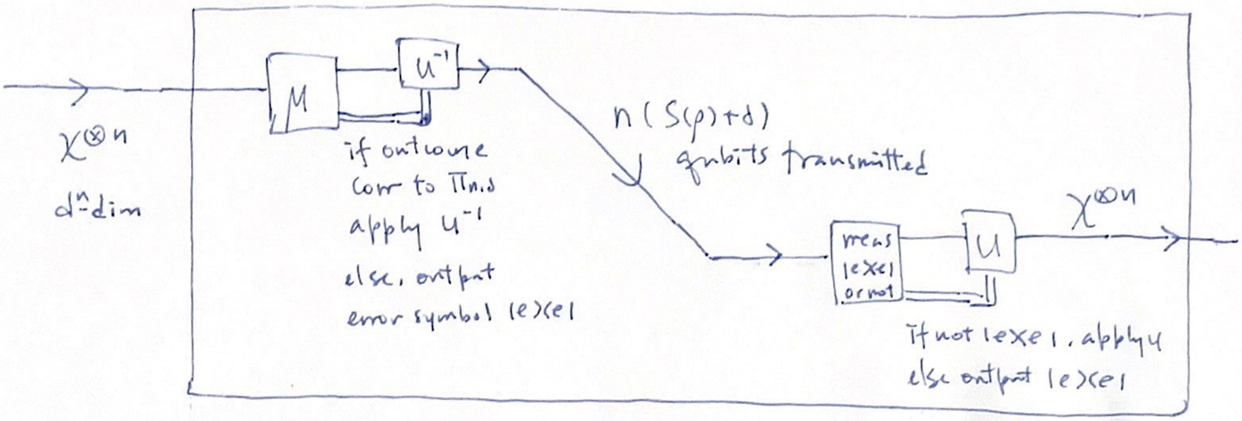
Next, we consider the Quantum analogue of transmitting only typical sequences and discuss 2 applications.

- $\rho, \delta, \varepsilon, n \geq n_0, T_{n,\delta}, S_{n,\delta}, \Pi_{n,\delta}$ defined as before.

$\mathcal{M}$ = binary meas with POVM $\{\Pi_{n,\delta}, I - \Pi_{n,\delta}\} \subseteq \mathrm{Pos}(\mathcal{X}^{\otimes n})$

Isometric bijections : $\mathbb{C}^{2^{n(S(\rho)+\delta)}} \sim (\mathbb{C}^2)^{\otimes n(S(\rho)+\delta)} \underset{\mathcal{U}^{-1}}{\overset{\mathcal{U}}{\rightleftarrows}} S_{n,\delta}$

The "transmit the typical space" (TTS) protocol : Υ :



$$\forall Y \in L(\mathcal{X}^{\otimes n}),\ \tilde{\Upsilon}(Y) = \Pi_{n,\delta} Y \Pi_{n,\delta} \otimes |0\rangle\langle 0| \otimes |e\rangle\langle e|$$
$$+ |e\rangle\langle e| \otimes |1\rangle\langle 1| \otimes (\mathbb{1} - \Pi_{n,\delta}) Y (\mathbb{1} - \Pi_{n,\delta})$$
$$\Upsilon(Y) = \mathrm{tr}_{2,3}\, \tilde{\Upsilon}(Y).$$

NB If input $= \rho^{\otimes n} = \sum_{\text{all } \vec{v}} p(\vec{v})\, |e_{\vec{v}}\rangle\langle e_{\vec{v}}|$,

$$\text{output } \Upsilon(\rho^{\otimes n}) = \sum_{\vec{v} \in T_{n,\delta}} p(\vec{v})\, |e_{\vec{v}}\rangle\langle e_{\vec{v}}| + \underbrace{(1 - p(T_{n,\delta}))}_{\leq \varepsilon \text{ by AEP ①}} |e\rangle\langle e|$$

$$\therefore\ \| \rho^{\otimes n} - \Upsilon(\rho^{\otimes n}) \|_1 \leq \varepsilon.$$

NB: Not the most interesting task to send $\rho^{\otimes n}$ (see will why soon).

Quantum source / ensemble:

Let X be a r.v. with sample space Ω, $Pr(X=x) = p(x)$.

Let A be CES, $\forall x$, $\rho_x \in D(A)$ q state labeled by x.

Consider the process:

① Sample X, obtain $x \in \Omega$ wp $p(x)$, store x in sys R

② Prepare ρ_x in sys A

Resulting state: $$\Lambda = \sum_x p(x) \; \underset{\substack{\text{classical}\\\text{random}\\\text{outcome}}}{|x\rangle\langle x|_R} \otimes \underset{\substack{\text{induced}\\\text{quantum}\\\text{random}\\\text{outcome}}}{\rho_{x\,A}}$$

Terminology:

- Above process is called 1 draw of the ensemble $\mathcal{E} = \{p(x), \rho_x\}_{x\in\Omega}$
- Average state of $\mathcal{E}$ is $\rho = \sum_x p(x)\,\rho_x = \operatorname{tr}_R \Lambda$

eg. B92, $\Omega = \{0,1\}$, $p(0) = p(1) = \frac{1}{2}$

$$\rho_0 = |0\rangle\langle 0|, \quad \rho_1 = |+\rangle\langle +|, \quad |+\rangle = \tfrac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$$

$$\Lambda = \tfrac{1}{2}\,|0\rangle\langle 0|_R \otimes |0\rangle\langle 0|_A + \tfrac{1}{2}\,|1\rangle\langle 1|_R \otimes |+\rangle\langle +|_A$$

$$\rho = \tfrac{1}{2}\left(|0\rangle\langle 0|_A + |+\rangle\langle +|_A\right) = \begin{bmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} \end{bmatrix}$$

- iid Q source: repeat the above process, say, n times:

① Sample X n times iid, obtain $\vec{x} = x_1 x_2 \ldots x_n$ wp $p(\vec{x}) = p(x_1) \ldots p(x_n)$

② Prepare $\rho_{\vec{x}} = \rho_{x_1} \otimes \rho_{x_2} \otimes \ldots \otimes \rho_{x_n}$

Resulting state: $$\sum_{\vec{x}} p(\vec{x}) \underbrace{|x_1\rangle\langle x_1| \cdots |x_n\rangle\langle x_n|}_{\text{in } R_1 R_2 \cdots R_n} \otimes \underbrace{\rho_{x_1} \otimes \rho_{x_2} \otimes \cdots \otimes \rho_{x_n}}_{\text{in } A_1 A_2 \cdots A_n} = \Lambda^{\otimes n}$$

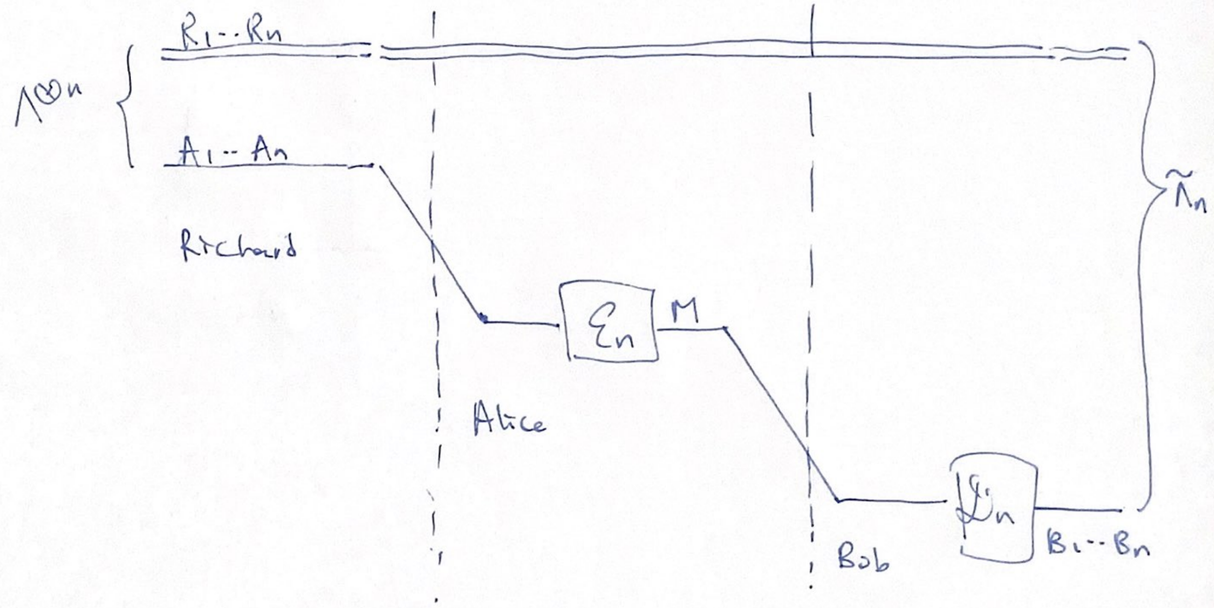
Quantum data compression: Notation as in page ③

① Referee Richard prepares $\Lambda^{\otimes n}$ and gives $A_1 \ldots A_n$ to sender Alice

② Alice encodes $A_1 \cdots A_n$ in nr qubits (system M)

③ Alice transmits M to Bob

④ Bob decodes M to $B_1 B_2 \cdots B_n \sim A_1 A_2 \cdots A_n$.



Goal: min r with constraint $\| \Lambda^{\otimes n} - \tilde{\Lambda}_n \|_1 \to 0$ as $n \to \infty$.

Above model: blind compression, Alice doesn't know $x_1 x_2 \ldots x_n$

Alternative model: Richard is Alice, who knows $x_1 \ldots x_n$, visible compression.

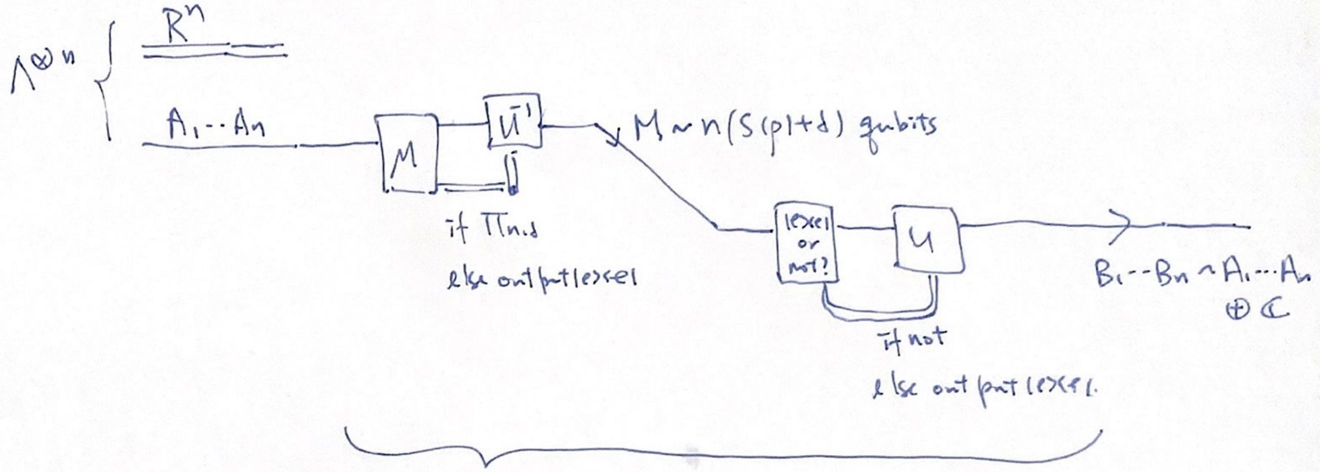
Thm: min $r = S(\rho)$ if ρ_x pure ($\rho_x = |\psi_x\rangle\langle\psi_x|$) $\forall x \in \Omega$.

- Direct coding theorem: Schumacher compression
- Strong converse: Winter (partial converse, weak converse by Schumacher 95, Jozsa-Schumacher 94, Barnum 96, Horodecki 98)

NB: spec decomp of $\rho = \sum_v p(v) |e_v\rangle\langle e_v|$. Note $p(v) \neq p(x)$.

Direct coding thm:

Schumacher compression: Apply TTS for $\rho^{\otimes n}$ (as defined on pages ①②)



TTS Υ

$$\Upsilon(Y) = \Pi_{n,\delta} Y \Pi_{n,\delta} + |e\rangle\langle e| \cdot \mathrm{tr}\left[\left(\underset{\substack{\uparrow \\ \text{on } A^{\otimes n}}}{\mathbb{1}} - \Pi_{n,\delta}\right) Y\right]$$

NB: while reduced state is $\rho^{\otimes n}$ on $A_1 \cdots A_n$, the goal of QDC is <u>NOT</u> transmission of $\rho^{\otimes n}$ (Bob knows ρ, can prepare $\rho^{\otimes n}$ himself). He needs to output $\rho_{x_1} \otimes \cdots \otimes \rho_{x_n}$ if Richard has $|x_1\rangle\langle x_1| \otimes \cdots |x_n\rangle\langle x_n|$, ie correlation with $R_1 \cdots R_n$ has to be preserved. Thus we must think of preserving $\Lambda^{\otimes n}$.

Claim: $\| \Lambda^{\otimes n} - I_{R_1 \cdots R_n} \otimes T(\Lambda^{\otimes n}) \|_1 \leq 2\sqrt{2}\sqrt{\varepsilon}$.

Pf: Consider the following purifications of $\Lambda^{\otimes n}$ and $I_{R_1 \cdots R_n} \otimes T(\Lambda^{\otimes n})$.

Let $$|u\rangle = \sum_{\vec{x}} \sqrt{q(\vec{x})} \underbrace{\left(|x_1\rangle \cdots |x_n\rangle\right)}_{Z_1} \underbrace{\left(|x_1\rangle \cdots |x_n\rangle\right)}_{R_1 \cdots R_n} \otimes \underbrace{|\psi_{x_1}\rangle |\psi_{x_2}\rangle \cdots |\psi_{x_n}\rangle}_{\tilde{A}} \otimes \underbrace{|0\rangle}_{Z_2} \otimes \underbrace{|e\rangle}_{Z_3}$$

$$|v\rangle = \sum_{\vec{x}} \sqrt{q(\vec{x})} \underbrace{\left(|x_1\rangle \cdots |x_n\rangle\right)}_{Z_1} \underbrace{\left(|x_1\rangle \cdots |x_n\rangle\right)}_{R_1 \cdots R_n} \otimes \left[\underbrace{\left(\Pi_{n,\delta}\right)\left(|\psi_{x_1}\rangle \cdots |\psi_{x_n}\rangle\right)}_{\tilde{A}} \otimes \underbrace{|0\rangle}_{Z_2} \otimes \underbrace{|e\rangle}_{Z_3} + \underbrace{|e\rangle}_{\tilde{A}} \otimes \underbrace{|1\rangle}_{Z_2} \otimes \underbrace{\left(\mathbb{1} - \Pi_{n,\delta}\right)\left(|\psi_{x_1}\rangle \cdots |\psi_{x_n}\rangle\right)}_{Z_3} \right]$$

where $\tilde{A} \cong A^{\otimes n} \oplus \mathbb{C} \cong Z_3$

$Z_1 \cong R_1 \cdots R_n$, $Z_2 = \mathbb{C}^{\{0,1\}}$.

Then $\operatorname{tr}_{Z_1 Z_2 Z_3} |u\rangle\langle u| = \Lambda^{\otimes n}$

$\operatorname{tr}_{Z_1 Z_2 Z_3} |v\rangle\langle v| = I_{R_1 \cdots R_n} \otimes T(\Lambda^{\otimes n})$

$$\langle u|v\rangle = \sum_{\substack{\vec{x} \\ \parallel \\ x_1 x_2 \cdots x_n}} q(\vec{x}) \langle \psi_{x_1}| \langle \psi_{x_2}| \cdots \langle \psi_{x_n}| \Pi_{n,\delta} |\psi_{x_1}\rangle |\psi_{x_2}\rangle \cdots |\psi_{x_n}\rangle$$

$$= \operatorname{tr} q(x_1) \cdots q(x_n) |\psi_{x_1}\rangle\langle\psi_{x_1}| \otimes |\psi_{x_2}\rangle\langle\psi_{x_2}| \otimes \cdots \otimes |\psi_{x_n}\rangle\langle\psi_{x_n}| \, \Pi_{n,\delta}$$

$$= \operatorname{tr} \rho^{\otimes n} \, \Pi_{n,\delta} \quad !!!$$

$$\geq 1 - \varepsilon$$

Apply $\| uu^* - vv^* \|_1 = 2\sqrt{1 - |\langle u, v\rangle|^2}$ (P46 in proof of Fuchs - vande Graaf ineq)

$$\| |u\rangle\langle u| - |v\rangle\langle v| \|_1 = 2\sqrt{1 - \langle u, v\rangle}\sqrt{1 + \langle u, v\rangle} \leq 2\sqrt{\varepsilon}\sqrt{2}$$

$$\therefore \| \Lambda^{\otimes n} - I_{R_1 \cdots R_n} \otimes T(\Lambda^{\otimes n}) \|_1 \underset{\substack{\uparrow \\ \text{mono of } \|\cdot\|_1}}{\leq} \| |u\rangle\langle u| - |v\rangle\langle v| \|_1 \leq 2\sqrt{2}\sqrt{\varepsilon}.$$

∴ The rate $r = S(\rho)$ can be attained.

ie as $n \to \infty$, $\varepsilon \to 0$, $\delta \to 0$, $r = S(\rho) + \delta \to S(\rho)$.

Converse:

This requires results from LN 2011 Lecture 10 and 11.

① Thm 10.5 (Fannes inequality)

$\mathcal{X}$: CES, $\forall\, \rho, \xi \in D(\mathcal{X})$ s.t. $\|\rho - \xi\|_1 \leq 1/e$

$$|S(\rho) - S(\xi)| \leq \log(\dim(\mathcal{X}))\, \|\rho - \xi\|_1 + \frac{1}{\ln 2}\, \eta(\|\rho - \xi\|_1)$$

where $\eta(x) = \begin{cases} -x \log x & \text{if } x > 0 \\ 0 & x = 0 \end{cases}$

② For $P, Q \in Pd(\mathcal{X})$, define the q relative entropy of P with Q as:

$$S(P \| Q) = \operatorname{Tr} P(\log P - \log Q)$$

③ Thm 10.6

Let $\rho, \xi \in D(\mathcal{X}) \cap Pd(\mathcal{X})$. Then $S(\rho \| \xi) \geq \frac{1}{2 \ln 2} \|\rho - \xi\|_{\underset{\uparrow}{2}}^2 \geq 0$

This is 1 for Pinsker's Ineq

④ Cor 11.8 (Monotonicity of QRE)

$\forall\, \Phi \in C(\mathcal{X}, \mathcal{Y})$, $\forall\, \rho, \xi \in D(\mathcal{X})$. $S(\Phi(\rho) \| \Phi(\xi)) \leq S(\rho \| \xi)$

⑤ If $\rho = \sum_k p_k \rho_k$, then $\sum_k p_k S(\rho_k \| \rho)$

$$= \sum_k p_k \operatorname{Tr} \rho_k (\log \rho_k - \log \rho)$$

$$= \sum_k p_k (-S(\rho_k)) + S(\rho)$$

⑥ $\max_{\rho \in D(\mathcal{X})} S(\rho) = \log(\dim(\mathcal{X}))$

Pf (converse, version by A. Winter 99)

For any $\mathcal{E}_n$, $\mathcal{D}_n$, suppose $\| \Lambda^{\otimes n} - \mathcal{I}_{R_1 \cdots R_n} \otimes (\mathcal{D}_n \circ \mathcal{E}_n)(\Lambda^{\otimes n}) \|_1 \leq \lambda \ll 1$

$$\left.\begin{array}{l} \Lambda^{\otimes n} = \sum_{\vec{x}} q(\vec{x}) \underset{\substack{\| \\ |x_1\rangle\langle x_1| \otimes |x_2\rangle\langle x_2| \otimes \cdots \otimes |x_n\rangle\langle x_n|}}{|\vec{x}\rangle\langle\vec{x}|} \otimes \rho_{\vec{x}} = \rho_{x_1} \otimes \rho_{x_2} \otimes \cdots \otimes \rho_{x_n} \\ \\ \mathcal{I}_{R_1 \cdots R_n} \otimes (\mathcal{D}_n \circ \mathcal{E}_n)(\Lambda^{\otimes n}) = \sum_{\vec{x}} q(\vec{x}) |\vec{x}\rangle\langle\vec{x}| \otimes (\mathcal{D}_n \circ \mathcal{E}_n)(\rho_{\vec{x}}) \end{array}\right\} \text{both block diagonal}$$

block structure is unchanged by composit.

$\therefore \| \Lambda^{\otimes n} - \mathcal{I}_{R_1 \cdots R_n} \otimes (\mathcal{D}_n \circ \mathcal{E}_n)(\Lambda^{\otimes n}) \|_1$

$= \sum_{\vec{x}} q(\vec{x}) \| \rho_{\vec{x}} - \mathcal{D}_n \circ \mathcal{E}_n (\rho_{\vec{x}}) \|_1$

$$\left(\left\| \begin{array}{cc} \boxed{A} & \boxed{0} \\ \boxed{0} & \boxed{B} \end{array} - \begin{array}{cc} \boxed{C} & \boxed{0} \\ \boxed{U} & \boxed{D} \end{array} \right\|_1 = \|A-C\|_1 + \|B-D\|_1 \right)$$

$\leq \lambda$ — $(**)$ upp bdd on the average
can be lifted to looser upp bdd on most. (Markov ineq)

Let $G \subseteq \Omega^{\otimes n}$, $G = \{ \vec{x} \in \Omega^{\otimes n} : \| \rho_{\vec{x}} - \mathcal{D}_n \circ \mathcal{E}_n (\rho_{\vec{x}}) \|_1 \leq \sqrt{\lambda} \}$

$B = \Omega^{\otimes n} \setminus G$

Then $\underset{\substack{\uparrow \\ \text{def of } B}}{\Pr(B) \cdot \sqrt{\lambda} \leq} \sum_{\vec{x} \in B} q(\vec{x}) \| \rho_{\vec{x}} - \mathcal{D}_n \circ \mathcal{E}_n(\vec{x}) \|_1 \underset{\substack{\uparrow \\ \text{from } (**)}}{\leq} \lambda$

$\Pr(B) \leq \sqrt{\lambda}$, $\Pr(G) \geq 1 - \sqrt{\lambda}$

Bad set has
$\boxed{\text{low prob}}$

Good set has
$\boxed{\text{low error}}$

Recall $\rho^{\otimes n} = \sum_{\vec{x}} p(\vec{x})\, \rho_{\vec{x}}$.

By mono of QRE, $\mathcal{D}_n$ TCP,

$\forall \vec{x} \in \Omega^{\otimes n}$, $S(\mathcal{E}_n(\rho_{\vec{x}}) \| \mathcal{E}_n(\rho^{\otimes n})) \geq S(\mathcal{D}_n \circ \mathcal{E}_n(\rho_{\vec{x}}) \| \mathcal{D}_n \circ \mathcal{E}_n(\rho^{\otimes n}))$

Average over $\vec{x} \in \Omega^{\otimes n}$:

$$\sum_{\vec{x}} p(\vec{x})\, S(\mathcal{E}_n(\rho_{\vec{x}}) \| \mathcal{E}_n(\rho^{\otimes n})) \geq \sum_{\vec{x}} p(\vec{x})\, S(\mathcal{D}_n \circ \mathcal{E}_n(\rho_{\vec{x}}) \| \mathcal{D}_n \circ \mathcal{E}_n(\rho^{\otimes n}))$$

Apply Item ⑤ from page ⑦ on each side:

$$-\sum_{\vec{x}} p(\vec{x})\, S(\mathcal{E}_n(\rho_{\vec{x}})) + S(\mathcal{E}_n(\rho^{\otimes n})) \overset{(***)}{\geq} -\sum_{\vec{x}} p(\vec{x})\, S(\mathcal{D}_n \circ \mathcal{E}_n(\rho_{\vec{x}})) + S(\mathcal{D}_n \circ \mathcal{E}_n(\rho^{\otimes n}))$$

$$\overset{\#\text{ qubits sent}}{nr} \overset{\text{item ⑥ on page ⑦}}{\geq} S(\mathcal{E}_n(\rho^{\otimes n}))$$

$$\overset{(***)}{\geq} -\sum_{\vec{x}} p(\vec{x})\, S(\mathcal{D}_n \circ \mathcal{E}_n(\rho_{\vec{x}})) + S(\mathcal{D}_n \circ \mathcal{E}_n(\rho^{\otimes n}))$$

$$= \underbrace{-\sum_{\vec{x} \in B} p(\vec{x})\, S(\mathcal{D}_n \circ \mathcal{E}_n(\rho_{\vec{x}}))}_{\text{bound prob}} - \sum_{\vec{x} \in G} p(\vec{x})\, S(\underbrace{\mathcal{D}_n \circ \mathcal{E}_n(\rho_{\vec{x}})}_{\sqrt{\lambda}\text{-close to } \rho_{\vec{x}}}) + S(\underbrace{\mathcal{D}_n \circ \mathcal{E}_n(\rho^{\otimes n})}_{\lambda\text{-close to } \rho_{\vec{x}}})$$

Fannes ineq twice

$$\geq -\underset{①}{\Pr(B)} \cdot \underbrace{n \log(\dim A)}_{\text{worst case entropy}} - \sum_{\vec{x} \in G} p(\vec{x}) \left(\underset{⑤}{S(\rho_{\vec{x}})} + \underset{②}{\log(\dim A^{\otimes n})} \cdot \sqrt{\lambda} + \frac{\eta(\sqrt{\lambda})}{\ln 2} \right)$$

$$+ \underset{④}{S(\rho^{\otimes n})} - \underset{③}{\log(\dim A^{\otimes n})}\, \lambda - \frac{\eta(\lambda)}{\ln 2}$$

$$\geq n \left[\log(\dim A) \left(\underset{①②}{-2\sqrt{\lambda}} \underset{③}{-\lambda} \right) + \underset{④}{S(\rho)} \right] \underbrace{- \sum_{\vec{x} \in G} p(\vec{x})\, \underset{⑤}{S(\rho_{\vec{x}})}}_{\overset{\shortparallel}{0} \text{ for pure } \rho_{\vec{x}}} + \text{const}$$

i.e. $r \geq S(\rho)$ as $n \to \infty$, $\lambda \to 0$

This gives $S(\rho)$ an <u>operational</u> meaning: the best rate for blind compression of pure state ensembles, providing a proper Q analogue to Shannon entropy.

The entire quantum info theory is founded on this notion of entropy, and we will investigate more in the rest of the term.

But first, more thoughts on Q data compression...

(I) We can tighten the converse bound:

$$-\sum_{\vec{x}\in G} p(\vec{x})\, S(\rho_{\vec{x}}) \overset{①}{\geq} -\sum_{\text{all } \vec{x}} p(\vec{x})\, S(\rho_{\vec{x}})$$

$$\overset{②}{=} -\sum_{x_1}\sum_{x_2}\cdots\sum_{x_n} p(x_1)\cdots p(x_n)\left(S(\rho_{x_1})+\dots+S(\rho_{x_n})\right)$$

$$= -n\sum_{x\in\Omega} p(x)\, S(\rho_x)$$

① = non neg of vN entropy $S(\sigma) \geq 0 \quad \forall\, \sigma \in D(\cdots)$

$$\begin{gathered} S(\sigma) \\ \| \\ \underset{\ominus}{-}\,\mathrm{tr}\,\underset{(\geq 0)}{\sigma}\,\underbrace{\log\sigma}_{(\leq 0)} \end{gathered}$$

②: Additivity of vN entropy for product states $S(\underbrace{\sigma_1\otimes\sigma_2}_{\substack{\text{spectrum is product of}\\ \text{the spectrum of } \sigma_1 \text{ \& that of } \sigma_2}}) = S(\sigma_1)+S(\sigma_2)$

Shannon entropy on indep $X_1\ X_2$

$H(X_1 X_2) = H(X_1) + H(X_2)$

∴ for general ρ_x,

$$nr \geq n\Big[\log(\dim A)(-2\sqrt{\lambda}-\lambda) + \underbrace{\big[S(\rho) - \sum_x p(x)\, S(\rho_x)\big]}_{\text{lower bound for mixed states, Holevo info for ensemble.}}\Big] + \text{const}$$

(II) If ρ_x is mixed, we can decompose $\rho_x = \sum_y t_y |\psi_{xy}\rangle\langle\psi_{xy}|$ for each x

and consider new ensemble $\mathcal{E} = \{ p_x \cdot t_y , |\psi_{xy}\rangle\langle\psi_{xy}| \}$

with the same average state ρ.

$\therefore$ TIS / Schumacher compression works for mixed state ensembles

<u>but</u> the rate $S(\rho)$ differs from the lower bound $S(\rho) - \sum_x q(x) S(\rho_x)$.

(III) Actual best rate for mixed state compression was found by Koashi & Imoto $\approx$ 2001

Idea: given a set of states ρ_x, there is a basis in which

$$\forall x, \quad \rho_x = \sum_c t_c |c\rangle\langle c| \otimes \sigma_{cx} \otimes \gamma_c$$

↑ classical part ↑ send through Q or classical channel.

↑ Q part ↑ Schumacher compression

↑ Redundant part not depending on x (do not send)

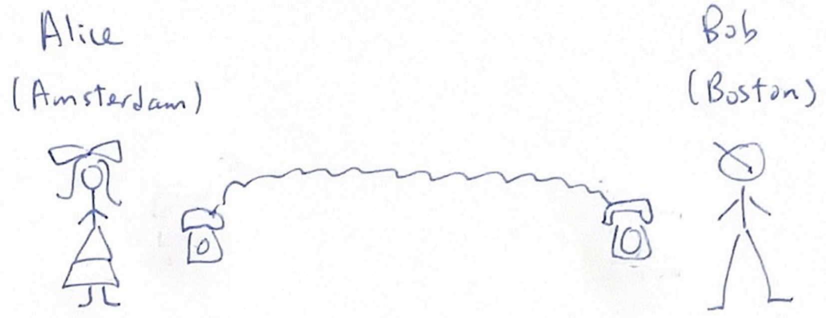
(IV) Visible compression ....

If ρ_x pure, Schumacher compression is optimal

If ρ_x mixed, Alice's knowledge of $x_1 \dots x_n$ can be used to lower the rate

(A4 ??)

Entanglement concentration & dilution BBPS 9511030



"LOCC"

- Each can perform local unitaries, and can communicate classical data both ways.

An ebit: $\frac{1}{\sqrt{2}}\left(|0\rangle_A |0\rangle_B + |1\rangle_A |1\rangle_B \right)$

m ebits: $$\frac{1}{\sqrt{2^m}}\left(|0\rangle|0\rangle + |1\rangle|1\rangle\right)^{\otimes m} = \frac{1}{\sqrt{2^m}}\left(\sum_{x \in \{0,1\}^n} \underset{\substack{\uparrow \\ \text{in } \mathbb{C}^{2\otimes m}}}{|x\rangle} \underset{\substack{\nwarrow \\ \otimes\, \mathbb{C}^{2\otimes m}}}{|x\rangle} \right)$$

Arbitrary bipartite pure state in Schmidt decomposition:

$$|\psi\rangle = \sum_{v=1}^{d} \sqrt{p_v}\, |a_v\rangle_A |b_v\rangle_B \quad , \quad A \sim B \sim \mathbb{C}^d .$$

Since local unitaries are free WLOG

$$|\psi\rangle = \sum_{v=1}^{d} \sqrt{p_v}\, |v\rangle_A |v\rangle_B$$

$$|\psi\rangle^{\otimes n} = \sum_{\substack{\vec{v} \\ \| \\ v_1 \cdots v_n}} \sqrt{p_{\vec{v}}}\ \underset{\substack{\uparrow \\ \text{in } A_1 \otimes A_2 \cdots \otimes A_n}}{|\vec{v}\rangle}\ \underset{\substack{\nwarrow \\ B_1 \otimes B_2 \otimes \cdots \otimes B_n}}{|\vec{v}\rangle}$$

2 related Qns in bipartite LOCC:

(a) How many ebits are needed to create approx of $|\psi\rangle^{\otimes n}$?

(b) How many "approx ebits" can be extracted from $|\psi\rangle^{\otimes n}$?

$$\begin{tikzcd}
\begin{matrix} n r_1 = m_1 \\ n r_2 = m_2 \end{matrix} \text{ ebits} \arrow[r, shift left, "\text{entanglement dilution}"] & {|\psi\rangle^{\otimes n}} \arrow[l, shift left, "\text{entanglement concentration}"] \\
\text{" max ent states "} \arrow[u]
\end{tikzcd}$$

- Nice surprise = $\min r_1 \approx \max r_2 \approx S(\rho)$, $\rho = \mathrm{Tr}_A |\psi\rangle\langle\psi| = \mathrm{tr}_B |\psi\rangle\langle\psi|$.

NB= Both Alice & Bob have a description of $|\psi\rangle$.

Remarks:-

① For large n, this allows us to view $S(\rho)$ as the "amount of entanglement," measured in ebits, contained in each copy of $|\psi\rangle$.

Furthermore, the "exchange rate" between any 2 bipartite pure states is simply the ratio of their entanglements (instead of some complex function of $|\psi_1\rangle, |\psi_2\rangle$):

$$\begin{tikzcd}
{|\psi_1\rangle^{\otimes n r_1}_{\quad\ \ r_2}} \arrow[r, shift left] \arrow[dr, shift left] & {|\psi_2\rangle^{\otimes n}} \arrow[l, shift left] \arrow[d, shift left] \\
& n r S(\rho_1) \text{ ebits} \arrow[ul, shift left] \arrow[u, shift left]
\end{tikzcd}$$

$$r \approx \min r_1 \approx \max r_2 \approx \frac{S(\rho_2)}{S(\rho_1)}, \quad \rho_i = \mathrm{tr}_A(|\psi_i\rangle\langle\psi_i|)$$

② Assuming LOCC cannot create entanglement, the rate $r = S(\rho)$ is thus optimal for both dilution & concentration (clearly $r_1 \geq r_2$ so optimal if $r_1 = r_2$).

Protocol for dilution:

① Alice locally prepares $|\psi\rangle^{\otimes n}$ on $\underbrace{A_1 A_2 \cdots A_n \tilde{A}_1 \tilde{A}_2 \cdots \tilde{A}_n}$

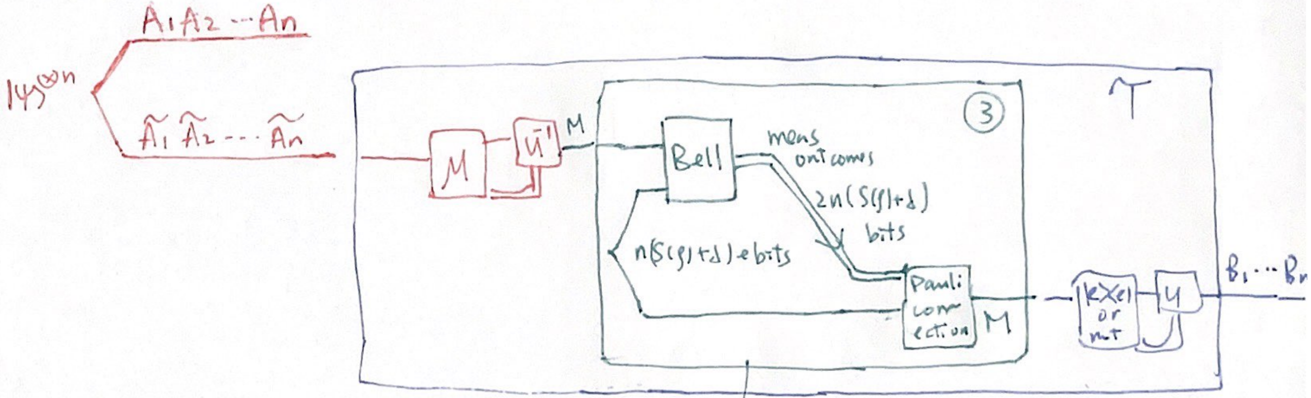
in $|\psi\rangle\ |\psi\rangle \cdots |\psi\rangle$

② Alice applies $\boxed{TTS}$ to $\tilde{A}_1 \tilde{A}_2 \cdots \tilde{A}_n$, output is $B_1 \cdots B_n$, with a twist ....

③ Suppose to send $n(S(\rho)+\delta)$ qubits to Bob

Instead Alice teleports $n(S(\rho)+\delta)$ qubits to Bob

consuming $n(S(\rho)+\delta)$ ebits and sending $2n(S(\rho)+\delta)$ c. bits.



nonlocal resources (ebits, c. comm)
only used in ③.

Note teleportation achieves perfect quantum communication & preserves all correlations.

∴ suffices to check TTS works.

$$|\psi\rangle^{\otimes n} = \sum_{\vec{v}} \sqrt{p(\vec{v})}\, |\vec{v}\rangle |\vec{v}\rangle, \quad \text{purified by } |u\rangle = \sum_{\vec{v}} \sqrt{p(\vec{v})}\, |\vec{v}\rangle |\vec{v}\rangle \otimes |0\rangle \otimes |e\rangle$$

$$|\omega\rangle = \underset{\text{all}}{\sum_{\vec{v}}} \sqrt{p(\vec{v})}\, \underset{A_1 \cdots A_n}{|\vec{v}\rangle} \otimes \Big[\underset{\tilde{A}_1 \cdots \tilde{A}_n}{(\Pi_{n,\delta} |\vec{v}\rangle)} \otimes |0\rangle \otimes |e\rangle + |e\rangle \otimes |1\rangle \otimes (\mathbb{1} - \Pi_{n,\delta}) |\vec{v}\rangle \Big]$$

defined by $\rho^{\otimes n}$, $\rho = \sum_{v} p_v |v\rangle\langle v|$.

• NB $|\omega\rangle$ purifies output on $A_1 \cdots A_n$
$B_1 \cdots B_n$

Again. $\langle u|w\rangle \geq 1-\varepsilon$

$\therefore \| |u\rangle\langle u| - |w\rangle\langle w| \|_1 \leq 2\sqrt{2}\sqrt{\varepsilon}$

$\therefore \| |\psi\rangle\langle\psi|^{\otimes n} - I\otimes\Upsilon(|\psi\rangle\langle\psi|^{\otimes n}) \|_1 \leq 2\sqrt{2}\sqrt{\varepsilon}$.

<u>Protocol for concentation:</u>

- First TTS does <u>NOT</u> work.

(max entangled)

$\sum_{\vec{v}\in T_{n,\delta}} \sqrt{p(\vec{v})}\,|\vec{v}\rangle|\vec{v}\rangle$ has $\underline{\underline{\text{low}}}$ fidelity with $\frac{1}{\sqrt{|T_{n,\delta}|}}\sum_{\vec{v}\in T_{n,\delta}} |\vec{v}\rangle|\vec{v}\rangle$

$\uparrow$

$$p(\vec{v}) \in \left[2^{-n(S(\rho)\underset{\uparrow}{+\delta})},\ 2^{-n(S(\rho)\underset{\uparrow}{-\delta})}\right]$$ ranges too much ....

can be substantial.

- Instead, Alice and Bob independently meas the <u>type class</u>.

Def: Let V has sample space $\{1,2,\ldots,d\}$.

Let $\vec{v} = v_1 v_2 \ldots v_n$ be the outcome for n iid draws of V

for $k=1,2,\ldots,d$, let n_k = # v_i's that are equal to k.

Then $\vec{v}$ is in the type class $T_{(n_1,n_2,\ldots,n_d)}$.

eg. $n=20$, $d=6$ (20 throws of a dice).

outcome $443\underset{\triangle}{2}65\underset{=}{1}\underset{=}{1}\ 5643\underset{=}{1}4\ 6\underset{\triangle}{2}\underset{\triangle}{2}\ \underset{\triangle}{2}46$

$n_1=3,\ n_2=4,\ n_3=2,\ n_4=5,\ n_5=2,\ n_6=4$

$\therefore$ outcome in type class $T_{(3,4,2,5,2,4)}$.

Obs: $n_1+\cdots+n_d=n$. Outcomes in one type class are all exactly equiprob.

eg. prob(outcome) = $p_1^3\, p_2^4\, p_3^2\, p_4^5\, p_5^2\, p_6^4$.

If initial state of Alice & Bob =

$$|\psi\rangle^{\otimes n} = \sum_{\vec{v}\ (\text{all})} p(\vec{v})\ \underset{\text{on } A_1\cdots A_n}{|\vec{v}\rangle}\underset{\text{on } B_1\cdots B_n}{|\vec{v}\rangle}$$

each meas the type class (Alice meas on $A_1 \cdots A_n$, Bob on $B_1 \cdots B_n$)

post meas state for outcome $T_{(n_1,n_2,\dots,n_d)}$:

$$\frac{1}{\sqrt{|T_{(n_1,\dots,n_d)}|}} \sum_{\vec{v} \in T_{(n_1,n_2,\dots,n_d)}} |\vec{v}\rangle|\vec{v}\rangle$$

Note that for large n, any $T_{(n_1,n_2,\dots,n_d)}$ very unlikely,

but Alice & Bob will get *some* outcome; *and* a max ent state of

Schmidt rank $|T_{(n_1,n_2,\dots,n_d)}|$, giving $\log|T_{(n_1,n_2,\dots,n_d)}|$ ebits via local unitaries

Claim: $\underset{\substack{\uparrow \\ \text{over outcomes} \\ T_{(n_1,n_2,\dots,n_d)}}}{\mathbb{E}} \log|T_{(n_1,n_2,\dots,n_d)}| \approx nS(\rho) - o(n)$.

Pf: maybe A4 ?! maybe not ....

Intuition: only $(n+1)^d$ type classes but at least $(1-\varepsilon)\, 2^{n(S(\rho)-\delta)}$ typical sequences.

∴ The probable type classes cannot have too few elements....

(1) Operational meaning:
for bipartite pure entangled state in many copies

- LOCC conversion is approx reversible
- single "currency" (ebit) of entanglement

This gives meaning to the quantity $S(\mathrm{tr}_B |\psi\rangle\langle\psi|)$ as "the amount of entanglement" in the state $|\psi\rangle$.

(2) Even better, above holds even if CC is charged

- concentration requires no CC
- dilution requires $\Theta(\sqrt{n})$ cbits

achievability: Lo-Popescu, necessity: Harrow-Lo-Hayden-Winter

(3) For bipartite pure state single copy, (1)-(2) don't hold.

LOCC conversion: Lo-Popescu 9703038
Nielsen 9811053 (majorization)

(4) For bipartite mixed state, many copies

Task (a) has no name (?). # ebits per copy required:
entanglement of formation Bennett DiVincenzo Smolin Wootters 96
regularized to "entanglement cost" Hayden (M) Horodecki Terhal 0008134
and the two can be different
Shor 0305035, Hastings 0809.3972

restricting to vanishing CC, # ebits per copy is called the entanglement of purification. Terhal Horodecki Leung DiVincenzo 0202044

Task (b) is called entanglement distillation
Bennett DiVincenzo Smolin Wootters 96

ebits extracted per copy: distillable entanglement

Mix state has "noise" ... to be removed by distillation.

Distillation (with 1- or 2-way CC) is mathematically equivalent to noisy channel coding for sending quantum data through noisy quantum channels (+crypto apps)!

Also, distillable entanglement can be strictly smaller than entanglement cost for some state (e.g., "bound entangled states" have 0 distillable entanglement but positive entanglement cost). M, P, R Horodecki 9801069

So, no single entanglement measure for mixed state, and LOCC conversion can be irreversible.

(5) For 3 or more parties, pure state, large # of copies

no comparable conversion theory
many types of incomparable entanglement

Bennett Popescu Rohrlich Smolin Thapliyal 9908073

Entanglement spread:

Def: for bipartite state $|\Psi\rangle_{AB}$, $\rho = \mathrm{tr}_B |\Psi\rangle\langle\Psi|$,

- its entanglement spread is defined as

$$\Delta(|\Psi\rangle) = \log(\mathrm{rank}(\rho)) - \log \frac{1}{\|\rho\|_\infty}$$

- its ϵ perturbed entanglement spread is defined as

$$\Delta_\epsilon(|\Psi\rangle) = \min_{\substack{P:\ \text{projectors} \\ \mathrm{tr}(\rho P) \geq 1-\epsilon}} \Delta(P \otimes I |\Psi\rangle)$$

e.g., $\Delta = 0$ iff all nonzero Schmidt coeffs are equal.

The transformation: $|\phi\rangle \rightarrow |\tilde{\Psi}\rangle$

s.t. fidelity of $|\Psi\rangle, |\tilde{\Psi}\rangle \geq 1-\epsilon$

requires $C \geq \Delta_{(4\epsilon)^{\frac{1}{8}}}(|\Psi\rangle) - \Delta_0(|\phi\rangle) + 2\log(1-(4\epsilon)^{\frac{1}{8}})$ cbits.

i.e., increase in spread must be "paid for" by classical communication.

For entanglement dilution:

$$\underset{\substack{\uparrow \\ \text{ebits} \\ \text{0 spread}}}{|\Phi\rangle^{\otimes nr}} \longrightarrow \underset{\substack{\uparrow \\ O(n)\ \text{spread} \\ \Theta(\sqrt{n})\ \epsilon\ \text{perturbed spread} \\ \text{lower bound of cbits needed}}}{|\Psi\rangle^{\otimes n}}$$