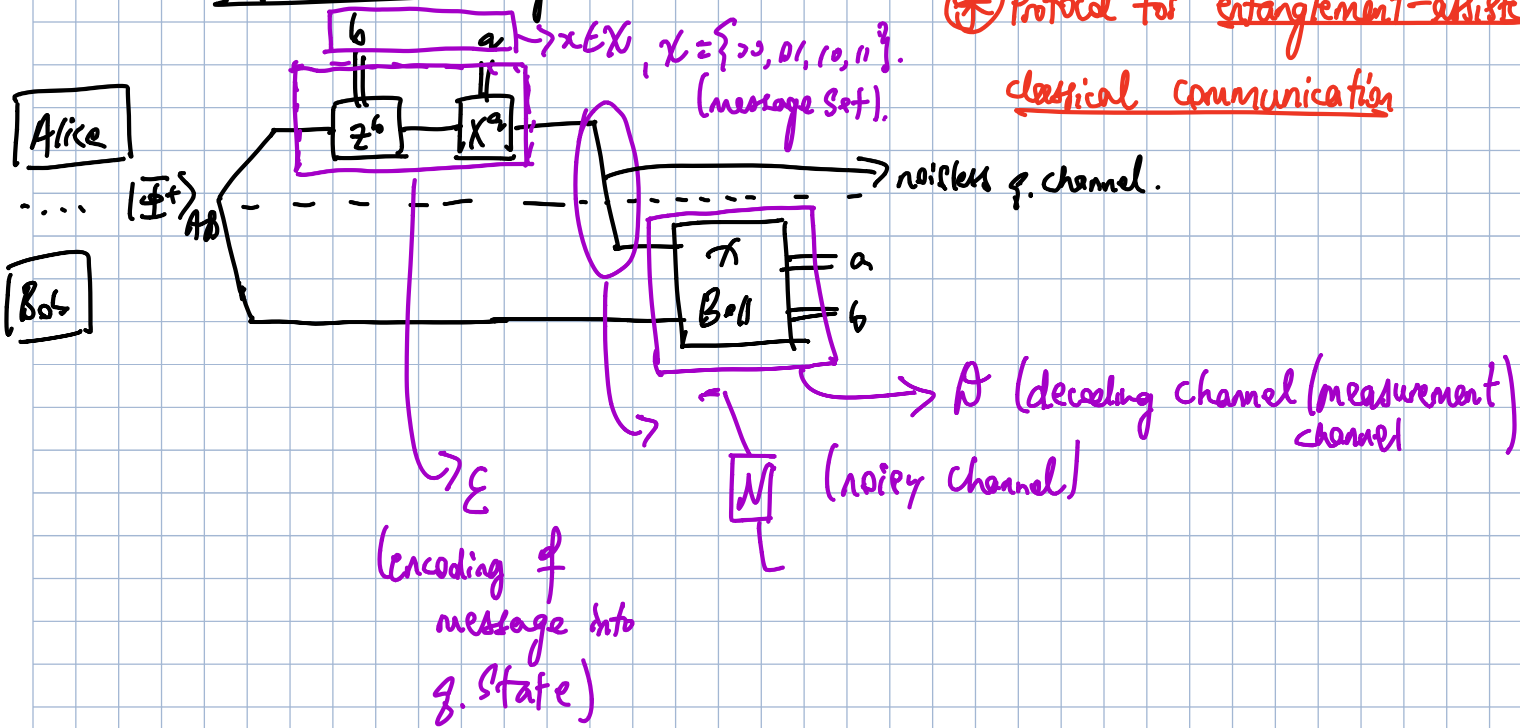
Recall Superdense coding

(*) Protocol for entanglement-assisted classical communication

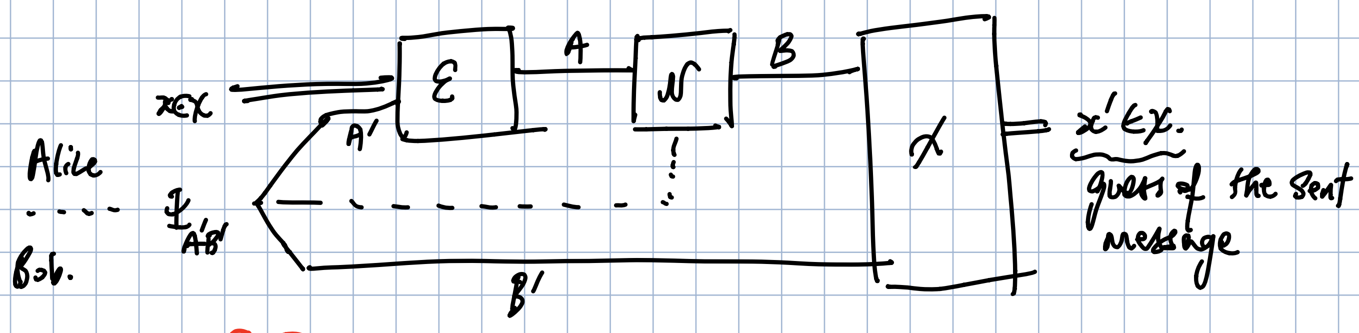


General Classical Communication.

- Messages $\mathcal{X} = \{x_1, x_2, \dots, x_M\}$ to be sent by Alice ($\log_2|\mathcal{X}|$ bits of inf.)
- Prior $p: \mathcal{X} \to [0,1]$ prob. distribution. $\to$ take uniform : $p(x) = \frac{1}{|\mathcal{X}|}$ $\forall x \in \mathcal{X}$.
- Given :
  1. Q. channel $\mathcal{N}_{A\to B}$ from Alice to Bob (with many uses).
  2. Local encoding and decoding.
  3. Entanglement-assistance

  These elements combined are called a code for ent.-assisted classical communication

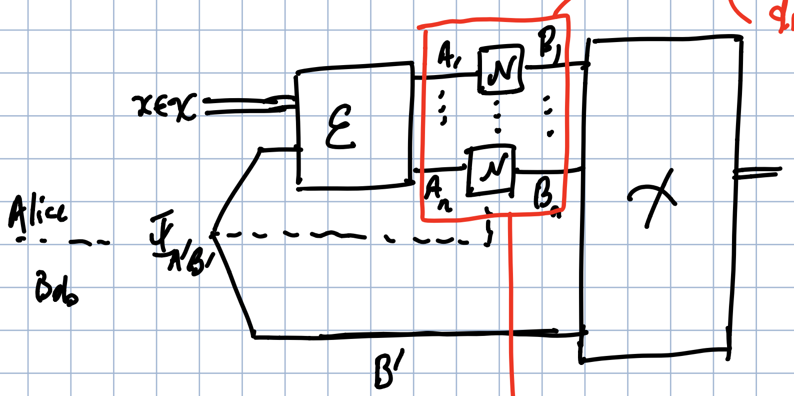
- Protocol : "One-shot" setting : (one channel use only).



(*) In the one-shot setting, we allow for some error. In general, it is not possible to get zero with just one channel use.

Asymptotic Setting : unlimited # of channel uses.

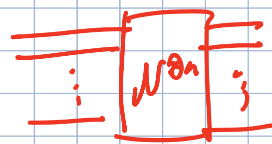
(Do not necessarily need n physically distinct of. channels here)



$x' \in \mathcal{X}$. (guess for the message sent)

⊛ In this setting, we want error to be zero, so we allow an unlimited # of channel uses in order to make this happen.

Mathematically equivalent to $\mathcal{N}^{\otimes n}$:



⊛ So n channel uses is equivalent to one use of $\mathcal{N}^{\otimes n}$

⊛ **Main Goal** : perfectly send the classical message from Alice to Bob.

$$|\Phi\rangle_{X_A X_A' A'B'} = \sum_{x \in \mathcal{X}} \sqrt{\frac{1}{|\mathcal{X}|}} |x\rangle_{X_A} \otimes |x\rangle_{X_A'} \otimes |\Phi\rangle_{A'B'}$$

$\downarrow$ measure – picks a message

$$\overline{\Phi}_{X_A X_A'} = \sum_{x \in \mathcal{X}} \frac{1}{|\mathcal{X}|} |x\rangle\langle x|_{X_A} \otimes |x\rangle\langle x|_{X_A'} \otimes \Phi_{A'B'}$$

$\downarrow$ encode

$$\rho_{X_A A} = \sum_{x \in \mathcal{X}} \frac{1}{|\mathcal{X}|} |x\rangle\langle x|_{X_A} \otimes \underbrace{\mathcal{E}_{X_A' A' \to A}(|x\rangle\langle x|_{X_A'} \otimes \Phi_{A'B'})}_{\tau^x_{AB'}}$$

$\downarrow$ Send through channel

$$\rho_{X_A B} = \sum_{x \in \mathcal{X}} \frac{1}{|\mathcal{X}|} |x\rangle\langle x|_{X_A} \otimes \rho^x_{BB'}, \quad \rho^x_{BB'} = \mathcal{N}_{A \to B}(\tau^x_{AB'})$$

decode : $\{M^{x}_{BB'}\}_{x\in\mathcal{X}}$

$$\boxed{\omega_{X_A X_B} = \mathcal{M}_{BB'\to X_B}\left(\rho_{X_A B}\right) = \sum_{x,x'\in\mathcal{X}} \frac{1}{|\mathcal{X}|} \mathrm{Tr}\left[M^{x'}_{BB'}\rho^{x}_{BB'}\right] |x,x'\rangle\langle x,x'|_{X_A X_B}}$$

Figure of merit : Registers X_A + X_B should have the same value

$$\Pi_{X_A X_B} = \sum_{x\in\mathcal{X}} |x\rangle\langle x|_{X_A} \otimes |x\rangle\langle x|_{X_B} \quad \text{(projector corresponding to } X_A = X_B\text{).}$$

$$\Rightarrow \boxed{p_{err}(\mathcal{E},\mathcal{D},\Phi_{AB'};\mathcal{N}) = 1 - \mathrm{Tr}\left[\Pi_{X_A X_B}\,\omega_{X_A X_B}\right]}$$ (probability that X_A + X_B do not have the same value.)

Main Questions :

① What is the smallest possible value of p^{*}_{err} over all codes $\mathcal{E},\mathcal{D},\Phi_{AB'}$? (given $\mathcal{X}$)

$$\varepsilon_E(|\mathcal{X}|;\mathcal{N}) = \min_{(\Phi,\mathcal{E},\mathcal{D})}\left\{p_{err}(\mathcal{E},\mathcal{D},\Phi;\mathcal{N}) : \mathcal{E}_{X_A A'\to A},\ \mathcal{D}_{BB'\to X_B}\ \text{CPTP},\ d_{X_A} = d_{X_B} = |\mathcal{X}|\right\}$$

Main focus.

② Given tolerable error ε, what is the largest value of $\log_2|\mathcal{X}|$?

$$C^{\varepsilon}_E(\mathcal{N}) = \max_{(|\mathcal{X}|,\mathcal{E},\mathcal{D},\Phi)}\left\{\log_2|\mathcal{X}| : p_{err}(\mathcal{E},\mathcal{D},\Phi;\mathcal{N}) \le \varepsilon\right\}$$

(*) In the asymptotic setting, what is the largest value of $\frac{1}{n}\log_2|\mathcal{X}|$ such that $\varepsilon\to 0$ as $n\to\infty$ (n: # of channel uses)?

$$\text{Rate} \equiv R = \frac{\log_2|\mathcal{X}|}{n} \Rightarrow |\mathcal{X}| = 2^{nR}$$

bits per channel use

Entanglement-assisted Classical Capacity of $\mathcal{N}$:

$$\boxed{C_E(\mathcal{N}) = \max\left\{R : \lim_{n\to\infty}\varepsilon_E\left(2^{n(R-\delta)};\mathcal{N}^{\otimes n}\right) = 0\ \ \forall\,\delta>0\right\}}$$

"Highest achievable rate"

Alternate Formula: $C_E(\mathcal{N}) = \lim_{\varepsilon\to 0} \liminf_{n\to\infty} \frac{1}{n} C_E^{\varepsilon}(\mathcal{N}^{\otimes n})$

$$= \lim_{\varepsilon\to 0} \sup_{n\in\mathbb{N}} \inf\left\{ \frac{1}{m} C_E^{\varepsilon}(\mathcal{N}^{\otimes m}) : m \geq n \right\}$$

$$\boxed{\begin{aligned} &\textbf{Main Result:} \quad C_E(\mathcal{N}) = I(\mathcal{N}) = \sup_{\psi_{RA}} \inf_{\sigma_B} D(\mathcal{N}_{A'\to B}(\psi_{AA'}) \| \psi_A \otimes \sigma_B) \\ &= \sup_{\psi_{AA'}} I(A;B)_\omega \quad \left(\omega_{AB} = \mathcal{N}_{A'\to B}(\psi_{AA'})\right) \end{aligned}}$$

Direct q. analogue of Shannon's noisy channel coding theorem:

$C(\mathcal{N}) = \max_{p_X} I(X;Y)$

$$D(\rho\|\sigma) = \begin{cases} \mathrm{Tr}[\rho\log\rho - \rho\log\sigma] & \text{if } \mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma) \\ +\infty & \text{otherwise.} \end{cases}$$

Quantum Relative entropy (generalization of KL-divergence for classical prob. distributions: $D(p\|q) = \sum_{x\in\mathcal{X}} p(x)\log\left(\frac{p(x)}{q(x)}\right)$)

$$I(A;B)_\rho = D(\rho_{AB} \| \rho_A \otimes \rho_B)$$
$$= \inf_{\sigma_B} D(\rho_{AB} \| \rho_A \otimes \sigma_B) \text{ is called } \underline{\text{mutual inf.}}$$

(Intuitively: how close the given state is to a product state)

$H(A)_\rho \equiv H(\rho) = -\mathrm{Tr}[\rho\log\rho] = -D(\rho\|\mathbb{1})$ ("von Neumann Entropy")

$$I(A;B)_\rho = H(A)_\rho + H(B)_\rho - H(AB)_\rho$$

Properties of Q. Relative Entropy:

① Faithful: $D(\rho\|\sigma) = 0$ iff $\rho = \sigma$.

② If $\rho \leq \sigma$, then $D(\rho\|\sigma) \leq 0$.

③ $D(\rho_1 \otimes \rho_2 \| \sigma_1 \otimes \sigma_2) = D(\rho_1\|\sigma_1) + D(\rho_2\|\sigma_2)$.

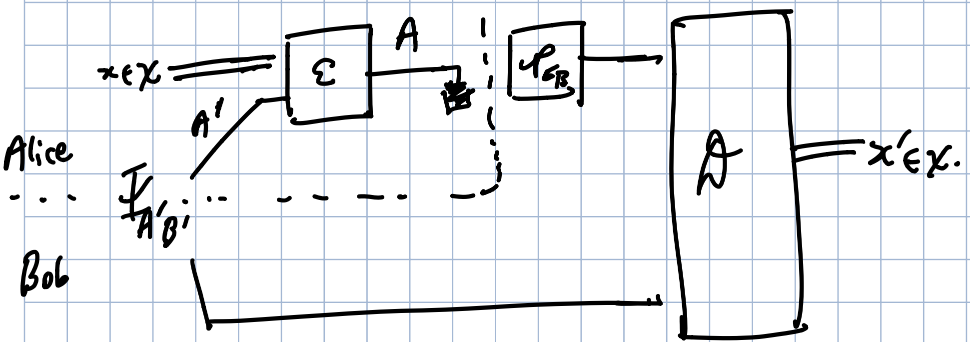
④ $D(\mathcal{N}(\rho) \| \mathcal{N}(\sigma)) \leq D(\rho\|\sigma)$ $\forall$ q. channels $\mathcal{N}$ — Data-processing inequality.

Prove in two parts

① $C_E(\mathcal{N}) \leq I(\mathcal{N})$ ⟶ main focus

② $C_E(\mathcal{N}) \geq I(\mathcal{N})$ — requires an explicit protocol construction.

↳ Key idea: Compare the protocol over $\mathcal{N}$ with the protocol over a useless channel (i.e. replacement channel). Intuitively, the useless channel is the one that "cuts the line".



$\tilde{\mathcal{N}}_{A\to B} = \varphi_{\sigma_B} \circ \mathrm{Tr}_A$

$\tilde{\mathcal{N}}_{A\to B}(\rho_A) = \sigma_B \mathrm{Tr}[\rho_A]$

↳ all info. about message is lost.

From above:

$$\rho_{X_A A} = \sum_{x\in\mathcal{X}} \frac{1}{|\mathcal{X}|} |x\rangle\langle x|_{X_A} \otimes \underbrace{\mathcal{E}_{X_A' A' \to A}(|x\rangle\langle x|_{X_A'} \otimes \Phi_{A'B'})}_{\tau^x_{AB'}}$$

$$= \frac{1}{|\mathcal{X}|} \sum_{x\in\mathcal{X}} |x\rangle\langle x|_{X_A} \otimes \tau^x_{AB'}$$

$$\tilde{\mathcal{N}}_{A\to B}(\tau^x_{AB'}) = \sigma_B \otimes \mathrm{Tr}_A[\tau^x_{AB'}] = \sigma_B \otimes \Phi_{B'} \quad (\Phi_{B'} = \mathrm{Tr}_A[\Phi_{AB'}])$$

⇒ Final state: $\tilde{\omega}_{X_A X_B} = \frac{1}{|\mathcal{X}|} \sum_{x\in\mathcal{X}} |x\rangle\langle x|_{X_A} \otimes \sum_{x'\in\mathcal{X}} \mathrm{Tr}[M^{x'}_{BB'}(\sigma_B \otimes \Phi_{B'})]\, |x'\rangle\langle x'|_{X_B}$

$$= \frac{1}{|\mathcal{X}|} \sum_{x,x'\in\mathcal{X}} \mathrm{Tr}[M^{x'}_{BB'}(\sigma_B \otimes \Phi_{B'})]\, |x,x'\rangle\langle x,x'|_{X_A X_B}$$

$$= \frac{1}{|\mathcal{X}|} \sum_{x\in\mathcal{X}} |x\rangle\langle x|_{X_A} \otimes \sum_{x'\in\mathcal{X}} \mathrm{Tr}[M^{x'}_{BB'}(\sigma_B \otimes \Phi_{B'})]\, |x'\rangle\langle x'|_{X_B}$$

$$= \frac{1}{|X|} \mathbb{1}_{X_A} \otimes \tilde{\omega}_{X_B} = \omega_{X_A} \otimes \tilde{\omega}_{X_B}! \quad \tilde{\omega}_{X_B} = \mathcal{D}_{BB' \to X_B}(\sigma_B \otimes \Psi_{B'})$$

Now, <u>comparator test</u>: $\mathrm{Tr}[\Pi_{X_A X_B} \tilde{\omega}_{X_A X_B}]$

$$= \frac{1}{|X|} \mathrm{Tr}[\Pi_{X_A X_B} (\mathbb{1}_{X_A} \otimes \tilde{\omega}_{X_B})]$$

$$= \frac{1}{|X|} \mathrm{Tr}[\underbrace{\mathrm{Tr}_{X_A}[\Pi_{X_A X_B}]}_{\mathbb{1}_{X_B}} \tilde{\omega}_{X_B}]$$

$$= \frac{1}{|X|} \underbrace{\mathrm{Tr}[\tilde{\omega}_{X_B}]}_{1}$$

$$= \frac{1}{|X|} \Rightarrow \log_2 |X| = -\log_2 \mathrm{Tr}[\Pi_{X_A X_B} (\omega_{X_A} \otimes \tilde{\omega}_{X_B})]$$

For the actual channel $\mathcal{N}_{A \to B}$, assume that $p_{err}(\mathcal{E}, \mathcal{D}, \Psi; \mathcal{N}) \leq \varepsilon$

$$\Rightarrow \mathrm{Tr}[\Pi_{X_A X_B} \omega_{X_A X_B}] \geq 1 - \varepsilon$$

(Protocol is an arbitrary element in the optimization for $C_E^{\varepsilon}(\mathcal{N})$.)

Then $\log_2 |X| = -\log_2 \mathrm{Tr}[\Pi_{X_A X_B} (\omega_{X_A} \otimes \tilde{\omega}_{X_B})]$

$$\leq \max_{\Lambda} \left\{ -\log_2 \mathrm{Tr}[\Lambda (\omega_{X_A} \otimes \tilde{\omega}_{X_B})] : 0 \leq \Lambda \leq \mathbb{1}, \ \mathrm{Tr}[\Lambda \omega_{X_A X_B}] \geq 1 - \varepsilon \right\}$$

$$\equiv D_H^{\varepsilon}(\omega_{X_A X_B} \| \omega_{X_A} \otimes \tilde{\omega}_{X_B})$$ $\longrightarrow$ Hypothesis testing relative entropy.

$$D_H^{\varepsilon}(\rho \| \sigma) = \max_{\Lambda} \{ -\log_2 \mathrm{Tr}[\Lambda \sigma] : 0 \leq \Lambda \leq \mathbb{1}, \ \mathrm{Tr}[\Lambda \rho] \geq 1 - \varepsilon \}$$

<u>(can be computed via Semi-definite programming (SDP))</u>

$$\Rightarrow \log_2 |X| \leq D_H^{\varepsilon}(\omega_{X_A X_B} \| \omega_{X_A} \otimes \tilde{\omega}_{X_B})$$ $\longrightarrow$ depends on the protocol! Get a bound that depends only on $\mathcal{N}$!

$$\omega_{X_A X_B} = (\mathcal{D}_{BB' \to X_B} \circ \mathcal{N}_{A \to B} \circ \mathcal{E}_{X_A' A' \to A})(\overline{\Phi}_{X_A X_A'} \otimes \Psi_{A' B'})$$

$$\Rightarrow \omega_{X_A} = \frac{1}{|\mathcal{X}|} \mathbb{1}_{X_A}$$

$$= D_H^{\varepsilon}\left(\mathcal{D}_{BB'\to X_B}(\theta_{X_A BB'}) \,\|\, \mathcal{D}_{BB'\to X_B}(\omega_{X_A} \otimes \sigma_B \otimes \Phi_{B'})\right)$$

$$\leq D_H^{\varepsilon}(\theta_{X_A BB'} \,\|\, \omega_{X_A} \otimes \sigma_B \otimes \Phi_{B'})$$

$\left(D_H^{\varepsilon}(\mathcal{N}(\rho)\|\mathcal{N}(\sigma)) \leq D_H^{\varepsilon}(\rho\|\sigma) \right.$ Data-processing inequality$\left.\right)$.

$$= D_H^{\varepsilon}(\mathcal{N}_{A\to B}(\rho_{X_A B' A}) \,\|\, \omega_{X_A} \otimes \Phi_{B'} \otimes \sigma_B)$$

$$\rho_{X_A B' A} = \mathcal{E}_{X_A' A' \to A}(\overline{\Phi}_{X_A X_A'} \otimes \Phi_{A'B'})$$

$$\mathrm{Tr}_A[\rho_{X_A B' A}] = \underbrace{\overline{\Phi}_{X_A}}_{\frac{1}{|\mathcal{X}|}\mathbb{1}_{X_A} = \omega_{X_A}} \otimes \Phi_{B'}$$

$$= D_H^{\varepsilon}(\mathcal{N}_{A\to B}(\rho_{RA}) \,\|\, \rho_R \otimes \sigma_B). \quad (R \equiv X_A B')$$

σ_B above was arbitrary $\Rightarrow$ optimize over all states to get the tightest upper bound.

$$\Rightarrow \log_2|\mathcal{X}| \leq \inf_{\sigma_B} D_H^{\varepsilon}(\mathcal{N}_{A\to B}(\rho_{RA}) \,\|\, \rho_R \otimes \sigma_B)$$

Now optimize over all states ρ_{RA} — like optimizing over all encoding channels.

$$\Rightarrow \log_2|\mathcal{X}| \leq \sup_{\rho_{RA}} \inf_{\sigma_B} D_H^{\varepsilon}(\mathcal{N}_{A\to B}(\rho_{RA}) \,\|\, \rho_R \otimes \sigma_B) = I_H^{\varepsilon}(\mathcal{N}).$$

$\Rightarrow$ Suffices to optimize over pure states ψ_{RA}, $d_R = d_A$

$\Rightarrow \log_2|\mathcal{X}| \leq I_H^{\varepsilon}(\mathcal{N})$ — Hypothesis testing mutual information.

$$\Rightarrow \boxed{C_E^{\varepsilon}(\mathcal{N}) \leq I_H^{\varepsilon}(\mathcal{N})}$$

Now, can get upper bound on $C_E(\mathcal{N})$. Important fact:

$$\lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}D_H^{\varepsilon}(\rho^{\otimes n}\|\sigma^{\otimes n}) = D(\rho\|\sigma).$$ (Quantum Stein's Lemma).

$$D(\rho\|\sigma) = \begin{cases} \mathrm{Tr}[\rho\log\rho - \rho\log\sigma] & \text{if } \mathrm{supp}(\rho)\subseteq\mathrm{supp}(\sigma) \\ +\infty & \text{otherwise.}\end{cases}$$

Quantum Relative entropy (generalization of KL-divergence for classical prob. distributions: $D(p\|q) = \sum_{x\in\mathcal{X}} p(x)\log\left(\frac{p(x)}{q(x)}\right)$)

$$\Rightarrow C_E(\mathcal{N}) = \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}C_E^{\varepsilon}(\mathcal{N}^{\otimes n})$$

$$\leq \lim_{\varepsilon\to 0}\lim_{n\to\infty}\frac{1}{n}I_H^{\varepsilon}(\mathcal{N}^{\otimes n})$$

$$\leq I(\mathcal{N}). \quad \Rightarrow \boxed{C_E(\mathcal{N})\leq I(\mathcal{N}).}$$

$$I(A;B)_\rho = D(\rho_{AB}\,\|\,\rho_A\otimes\rho_B)$$
$$= \inf_{\sigma_B} D(\rho_{AB}\|\rho_A\otimes\sigma_B)$$ is called mutual inf.

(Intuitively: how close the given state is to a product state)

$H(A)_\rho \equiv H(\rho) = -\mathrm{Tr}[\rho\log\rho] = -D(\rho\|\mathbb{1})$ ("von Neumann Entropy")

$$I(A;B)_\rho = H(A)_\rho + H(B)_\rho - H(AB)_\rho$$

Also, coherent information: $I(A\rangle B)_\rho = H(B)_\rho - H(AB)_\rho$

Properties of Q. Relative Entropy:

① Faithful: $D(\rho\|\sigma)=0$ iff $\rho=\sigma$.

② If $\rho\leq\sigma$, then $D(\rho\|\sigma)\leq 0$.

③ $D(\rho_1\otimes\rho_2\|\sigma_1\otimes\sigma_2)=D(\rho_1\|\sigma_1)+D(\rho_2\|\sigma_2)$.

$\boxed{\text{④ } D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))\leq D(\rho\|\sigma) \ \forall \text{ q-channels } \mathcal{N} \text{ — Data-processing inequality.}}$