— Again because of the random coding, $x^{(n)}(w)$ and $x^{(n)}(1)$ are <u>independently occurring</u> and thus $y^{(n)}$ (associated with $w=1$) and $x^{(n)}(w)$ when $w \neq 1$. Therefore,

$$\boxed{P^{av}_{rob}(E_w) \le 2^{-n(I(X;Y)-3\varepsilon)}, \quad w \neq 1}$$

— Finally, choosing n large enough, one gets: $\boxed{\begin{aligned} P^{av}_{rob}(\mathcal{E}) &\le \varepsilon + (2^{nR}-1)\,2^{-n(I(X;Y)-3\varepsilon)} \\ &\le 2\varepsilon \Leftarrow R < I(X;Y) - 3\varepsilon \end{aligned}}$

— Choose $\pi = \{p(x)\}_{x\in\mathcal{X}}$ such that $I(X;Y) = C$: $\boxed{P^{av}_{rob}(\mathcal{E}) \le 2\varepsilon \text{ when } R < C - 3\varepsilon}$

— $\boxed{\exists\ \mathcal{E}^* \text{ such that } \frac{1}{2^{nR}} \sum_{w\in\mathcal{W}} d^{(n)}_w(\mathcal{E}^*) \le 2\varepsilon}$, otherwise $P^{av}_{rob}(\mathcal{E}) > 2\varepsilon$

— Let $\mathcal{W} = \mathcal{W}_1 \cup \mathcal{W}_2$ where $w \in \mathcal{W}_2 \Rightarrow d^{(n)}_w(\mathcal{E}^*) > d^{(n)}_{\bar{w}}(\mathcal{E}^*)\ \forall \bar{w} \in \mathcal{W}_1$

$$\#(\mathcal{W}_2) = \frac{\#(\mathcal{W})}{2}$$

Then, $\boxed{d^{(n)}_w(\mathcal{E}^*) \le 4\varepsilon}\ \forall w \in \mathcal{W}_1$; otherwise $2\varepsilon \ge \frac{1}{2^{nR}} \sum_{w\in\mathcal{W}_2} d^{(n)}_w(\mathcal{E}^*) > \frac{4\varepsilon}{2^{nR}} \frac{2^{nR}}{2} = 2\varepsilon$

Concluding, we know there exists a code $\mathcal{E}^*$ with $2^{nR-1} = 2^{n(R-\frac{1}{n})}$ with $R-\frac{1}{n} < C-3\varepsilon$ such that the maximum error $d^{(n)} \xrightarrow[n]{} 0$.

Proof of Part 2)

- Let W be the uniformly distributed random variables with values in $\mathcal{W}$:

$\boxed{P(w) = \frac{1}{2^{nR}}}$. Let $\hat{W}$ be the random variable with values in $\mathcal{W}$ which is the outcome of the decoding procedure g : $\boxed{\hat{W} = g(Y^{(n)})}$

Then, the probability of error is $\boxed{P^{(n)}_{err} = \mathrm{Prob}(\hat{W} \neq W)}$

- Set $\boxed{E = \begin{cases} 1 & \text{if } \hat{W} \neq W \\ 0 & \text{if } \hat{W} = W \end{cases}}$

- Fano's inequality :
$$\begin{aligned} H(E, W \mid Y^{(n)}) &= H(E, W, Y^{(n)}) - H(Y^{(n)}) \\ &= H(E, W, Y^{(n)}) - H(W, Y^{(n)}) + H(W, Y^{(n)}) - H(Y^{(n)}) \\ &= H(E \mid W, Y^{(n)}) + H(W \mid Y^{(n)}) \end{aligned}$$

$$H(E \mid W, Y^{(n)}) + H(W \mid Y^{(n)}) = H(W \mid E, Y^{(n)}) + H(E \mid Y^{(n)})$$

$\boxed{H(E \mid W, Y^{(n)}) = 0}$ since W and $Y^{(n)}$ determine wheter $E = 0$ or $E = 1$

$$H(W \mid Y^{(n)}) = P(E=0)\, H(W \mid E=0, Y^{(n)}) + P(E=1)\, H(W \mid E=1) + H(E \mid Y^{(n)})$$

$\boxed{H(W \mid E=0, Y^{(n)}) = 0}$ since in absence of errors $(E=0)$ knowing $Y^{(n)}$ determine W

$P(E=1) = P_{err}^{(n)} \Longrightarrow H(W \mid Y^{(n)}) = P_{err}^{(n)} H(W \mid E=1) + H(E \mid Y^{(n)})$

$\boxed{H(E \mid Y^{(n)}) \leq H(E) \leq 1}$ and $\boxed{H(W \mid E=1) \leq \log(2^{nR}-1) \leq nR}$

Then, $\boxed{H(W \mid Y^{(n)}) \leq 1 + nR\, P_{err}^{(n)}}$ (Fano's inequality)

$$- \boxed{I(X^{(n)}; Y^{(n)})} \overset{\text{(Lemma 2.3.1)}}{=} H(Y^{(n)}) - \sum_{j=1}^{m} H(Y_j \mid X^{(n)}, Y^{(j-1)}) \overset{\text{(memoryless channel)}}{=} H(Y^{(n)}) - \sum_{j=1}^{m} H(Y_j \mid X_j)$$

$$\boxed{\leq} \sum_{j=1}^{m} \left(H(Y_j) - H(Y_j \mid X_j) \right) = \boxed{\sum_{j=1}^{m} I(X_j; Y_j)} \Longrightarrow \boxed{I(X^{(n)}; Y^{(n)}) \leq nC}$$

– If the maximal probability of error $d^{(n)} \xrightarrow[n]{} 0$ so does $P^{(n)}_{err}$:

$$\boxed{P^{(n)}_{err} = \frac{1}{2^{nR}} \sum_{w \in \mathcal{W}} d^{(n)}_w \xrightarrow[n]{} 0}$$

used in Fano's inequality

– Since W is uniformly distributed: $\boxed{P^{(n)}_{err} = \text{Prob}(\hat{W} \neq W)}$ and

$$\boxed{nR = H(W) = H(W \mid Y^{(n)}) + I(W; Y^{(n)})}$$

– $W \longrightarrow X^{(n)}(W) \longrightarrow Y^{(n)}$ is a <u>Markov chain</u> (see Definition 2.3.4.)

as well as (see Example 2.3.1, 2) $Y^{(n)} \longrightarrow X^{(n)}(W) \longrightarrow W$.

Then, the data processing inequality (Proposition 2.3.1) gives

$$\boxed{I(X^{(n)}; Y^{(n)}) \geq I(W; Y^{(n)})} \Longrightarrow \boxed{I(W; Y^{(n)}) \leq nC}$$

– $H(W \mid Y^{(n)}) \leq 1 + P^{(n)}_{err}\, nR \Longrightarrow \boxed{R \leq \frac{1}{n} + P^{(n)}_{err} R + C \xrightarrow[n \to +\infty]{} C}$

– $\boxed{P^{(n)}_{err} \geq 1 - \frac{1}{nR} - \frac{C}{R}} \Longrightarrow$ the error probability is <u>bounded away from 0</u> $\boxed{\text{if } R > C}$

2.4. Quantum channel coding

A quantum channel is any completely positive map and trace-preserving $\Phi : M_s(\mathbb{C}) \longrightarrow M_r(\mathbb{C})$ acting on a sender quantum state $\rho \in M_s(\mathbb{C})$ and yielding an output state $\Phi[\rho]$:

$$\xrightarrow[\text{input state}]{\rho} \boxed{\text{CPTP } \Phi} \xrightarrow[\text{output state}]{\Phi[\rho]}$$

$\boxed{\text{Definition 2.4.1}}$ A quantum channel Φ is said memoryless if using the channel n times successively is described by the n-fold tensor product

$$\Phi^{(n)} := \underbrace{\Phi \otimes \Phi \otimes \cdots \otimes \Phi}_{n \text{ times}} \; : \; M_s(\mathbb{C})^{\otimes n} \longrightarrow M_r^{\otimes n}(\mathbb{C})$$

$\boxed{\text{Example 2.4.1}}$ Depolarizing channel : $\Phi : M_2(\mathbb{C}) \to M_2(\mathbb{C})$

$$\Phi[\rho] = (1-p)\rho + \frac{p}{3}\left(\sigma_x \rho \sigma_x + \sigma_y \rho \sigma_y + \sigma_z \rho \sigma_z\right) \quad , \quad 0 \leq p \leq 1 .$$

$$\Phi'[\rho] = \frac{1}{2}\left(\sigma_0 \rho \sigma_0 + \sigma_x \rho \sigma_x + \sigma_y \rho \sigma_y + \sigma_z \rho \sigma_z\right) \quad , \quad \sigma_0 = \mathbb{1} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

$$\Phi'[\sigma_0] = 2\sigma_0 \quad ; \quad \Phi'[\sigma_x] = \frac{\sigma_x + \sigma_x - \sigma_x - \sigma_x}{2} = 0 \quad ; \quad \Phi'[\sigma_y] = \Phi'[\sigma_z] = 0$$

$$\Phi'[\rho] = \mathrm{Tr}(\rho)\sigma_0 \quad \Rightarrow \quad \boxed{\Phi[\rho]} = (1-p)\rho + \frac{2}{3}p\frac{\mathbb{1}}{2} - \frac{p}{3}\rho \boxed{= \left(1 - \frac{4}{3}p\right)\rho + \frac{2p}{3}\frac{\mathbb{1}}{2}}$$

extendible to n-qubits

$$\mathrm{Sp}(\Phi[\rho]) = \left\{ \frac{p}{3} + \left(1 - \frac{4}{3}p\right) r_1 \, , \, \frac{p}{3} + \left(1 - \frac{4}{3}p\right) r_2 \right\}$$

$$\Phi[\rho] = \sum_{\alpha=0}^{z} E_\alpha \rho E_\alpha^\dagger \quad ; \quad E_0 = \sqrt{1-p}\,\sigma_0 \, ; \quad E_x = \sqrt{\frac{p}{3}}\,\sigma_x \quad ; \quad E_y = \sqrt{\frac{p}{3}}\,\sigma_y \quad ; \quad E_z = \sqrt{\frac{p}{3}}\,\sigma_z$$

$$\sum_{\alpha=0}^{3} E_\alpha^\dagger E_\alpha = 1 - p + \frac{p}{3} \times 3 = 1$$

$\boxed{\text{Example 2.4.2}}$ Amplitude damping channel: spontaneous decay from an excited state.

$\Phi : M_2(\mathbb{C}) \to M_2(\mathbb{C})$

$$\Phi[\rho] = A_1 \rho A_1^+ + A_2 \rho A_2^+ \quad ; \quad A_1 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix} \quad ; \quad A_2 = \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix} \qquad 0 \le p \le 1$$

$$A_1^+ A_1 = A_1^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1-p \end{pmatrix} \; ; \quad A_2^+ A_2 = \begin{pmatrix} 0 & 0 \\ \sqrt{p} & 0 \end{pmatrix}\begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & p \end{pmatrix} \longrightarrow \boxed{A_1^+ A_1 + A_2^+ A_2 = \mathbb{1}}$$

$$\begin{cases} |0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \\ |1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix} \end{cases} : \quad \boxed{\begin{matrix} A_1|0\rangle = |0\rangle \\ A_1|1\rangle = \sqrt{1-p}\,|1\rangle \end{matrix}} \; , \quad \boxed{\begin{matrix} A_2|0\rangle = 0 \\ A_2|1\rangle = \sqrt{p}\,|0\rangle \end{matrix}}$$

$|0\rangle$ ground state

$|1\rangle$ excited state

$$\Phi[\rho] = \Phi\left[\begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix}\right] = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix}\begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix} + \begin{pmatrix} 0 & \sqrt{p} \\ 0 & 0 \end{pmatrix}\begin{pmatrix} \rho_{00} & \rho_{01} \\ \rho_{10} & \rho_{11} \end{pmatrix}\begin{pmatrix} 0 & 0 \\ \sqrt{p} & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \rho_{00} & \rho_{01} \\ \sqrt{1-p}\,\rho_{10} & \sqrt{1-p}\,\rho_{11} \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-p} \end{pmatrix} + \begin{pmatrix} \sqrt{p}\,\rho_{10} & \sqrt{p}\,\rho_{11} \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 0 \\ \sqrt{p} & 0 \end{pmatrix}$$

$$= \begin{pmatrix} \rho_{00} & \sqrt{1-p}\,\rho_{01} \\ \sqrt{1-p}\,\rho_{10} & (1-p)\,\rho_{11} \end{pmatrix} + \begin{pmatrix} p\,\rho_{11} & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} \rho_{00} + p\,\rho_{11} & \sqrt{1-p}\,\rho_{01} \\ \sqrt{1-p}\,\rho_{10} & (1-p)\,\rho_{11} \end{pmatrix}$$

Exercise 2.4.1 show that $\lim_{n} \Phi^n[\rho] = |0\rangle\langle 0| \quad \forall\, \rho \in M_2(\mathbb{C})$

$$\Phi[\rho] = \begin{pmatrix} \rho_{00} + p\rho_{11} & \sqrt{1-p}\,\rho_{01} \\ \sqrt{1-p}\,\rho_{10} & (1-p)\rho_{11} \end{pmatrix} = \begin{pmatrix} 1-(1-p)\rho_{11} & \sqrt{1-p}\,\rho_{01} \\ \sqrt{1-p}\,\rho_{10} & (1-p)\rho_{11} \end{pmatrix}$$

$$\Phi^2[\rho] = \Phi[\Phi[\rho]] = \Phi\left[\begin{pmatrix} 1-(1-p)\rho_{11} & \sqrt{1-p}\,\rho_{01} \\ \sqrt{1-p}\,\rho_{10} & (1-p)\rho_{11} \end{pmatrix}\right] = \begin{pmatrix} 1-(1-p)^2\rho_{11} & (\sqrt{1-p})^2\rho_{01} \\ (\sqrt{1-p})^2\rho_{01} & (1-p)^2\rho_{11} \end{pmatrix}$$

$\vdots$

$$\Phi^n[\rho] = \begin{pmatrix} 1-(1-p)^n\rho_{11} & +(1-p)^{\frac{n}{2}}\rho_{01} \\ (1-p)^{\frac{n}{2}}\rho_{10} & (1-p)^n\rho_{11} \end{pmatrix} \xrightarrow[n]{} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = |0\rangle\langle 0|.$$

– **Holevo's bound**

Alice encodes classical letters $x \in \{1,2,\dots,M\} = \mathcal{X}$ into quantum states $\rho_x \in M_n(\mathbb{C})$ with probabilities $p_x \geq 0$, $\sum_{x=1}^{M} p_x = 1$; Alice sends the states ρ_x to Bob through a noiseless channel.

Bob receives the states ρ_x and tries to retrieve the encoded information by measuring on them a POVM $\{E_y\}_{y\in\mathcal{Y}}$: $\boxed{\rho_x \longmapsto \sum_{y\in\mathcal{Y}} E_y\,\rho\,E_y^\dagger \,,\ \sum_y E_y^\dagger E_y = 1}$

Remark: X with values $x \in \mathcal{X}$ and Y with values in the POVM index set $\mathcal{Y}$ are random variables.

Definition 2.4.2 The maximal accessible information obtainable by Bob about X is given by:

$$\boxed{I_{acc}(X) = \max_{\{E_y\}_{y \in \mathcal{Y}}} I(X;Y)}$$

where $I(X;Y)$ is the mutual information of X and Y, while the maximum is computed over all possible POVM's available to Bob.

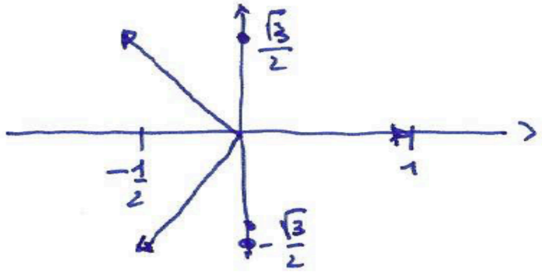
Remark: $I(X;Y) = \sum_{\substack{x \in \mathcal{X} \\ y \in \mathcal{Y}}} p(x,y) \log \frac{p(x,y)}{p(x)\, p(y)}$

1. conditional probability of getting y upon receiving x: $\boxed{p(y|x) = \mathrm{Tr}(\rho_x E_y^\dagger E_y)}$;
2. joint probability of getting x and y: $\boxed{p(x,y) = p_x \mathrm{Tr}(\rho_x E_y^\dagger E_y)}$;
3. probability of the output y: $\boxed{p(y) = \sum_{x \in \mathcal{X}} p(x,y) = \sum_{\mathcal{X} \ni x} p_x \mathrm{Tr}(\rho_x E_y^\dagger E_y) = \mathrm{Tr}(\rho E_y^\dagger E_y)}$ where $\rho = \sum_{x \in \mathcal{X}} p_x \rho_x$.

$\boxed{\text{Example } 2.4.3}$ $\quad \mathcal{X} = \{1, 2, 3\}$; $|\psi_1\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, $|\psi_2\rangle = \frac{1}{2}\begin{pmatrix} -1 \\ \sqrt{3} \end{pmatrix}$, $|\psi_3\rangle = \frac{1}{2}\begin{pmatrix} -1 \\ -\sqrt{3} \end{pmatrix}$

Encoding:

$$\begin{cases} 1 \longrightarrow P_1 = |\psi_1\rangle\langle\psi_1| = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \; ; \; p_1 = \frac{1}{3} \\ 2 \longrightarrow P_2 = |\psi_2\rangle\langle\psi_2| = \frac{1}{4}\begin{pmatrix} 1 & -\sqrt{3} \\ -\sqrt{3} & 3 \end{pmatrix} \; ; \; p_2 = \frac{1}{3} \\ 3 \longrightarrow P_3 = |\psi_3\rangle\langle\psi_3| = \frac{1}{4}\begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & 3 \end{pmatrix} \; ; \; p_3 = \frac{1}{3} \end{cases}$$



$$\boxed{\langle\psi_i|\psi_j\rangle = -\frac{1}{2}, \; i \neq j} \qquad \boxed{\rho = \frac{1}{3}(P_1 + P_2 + P_3) = \frac{1}{2}}$$

Optimal POVM: $\boxed{E_i = \sqrt{\frac{2}{3}}\, Q_i \;, \quad Q_i = 1 - P_i} \Longrightarrow \sum_{i=1}^{3} E_i^\dagger E_i = \frac{2}{3}\sum_{i=1}^{3}(1 - P_i) = 2 - 1 = 1$

$$\boxed{p(j|i) =} \; \mathrm{Tr}(P_i E_j^\dagger E_j) = \frac{2}{3}\mathrm{Tr}(P_i Q_j) = \boxed{\begin{cases} 0 & i = j \\ \frac{2}{3}(1 - |\langle\psi_i|\psi_j\rangle|^2) = \frac{2}{3}\left(1 - \frac{1}{4}\right) = \frac{1}{2} \end{cases}}$$

$$p(i,j) = \frac{1}{3}\, p(j|i) \Longrightarrow \boxed{p(j) = \frac{1}{3} \times \frac{1}{2} \times 2 = \frac{1}{3}}$$

$$\boxed{I(X;Y) =} \; H(Y) - H(Y|X) = \log 3 - \frac{1}{3} \times 3 \; \boxed{= \log 3 - 1 < 1 = S(\rho)}$$

Theorem 2.4.1 Holevo's bound

$$\boxed{I_{acc}(X) \leq S(\rho) - \sum_{x\in\mathcal{X}} p_x S(\rho_x)}, \quad \text{where} \quad \rho = \sum_{x\in\mathcal{X}} p_x \rho_x \in M_d(\mathbb{C})$$

Proof: $$I(X:Y) = \sum_{x\in\mathcal{X}, y\in\mathcal{Y}} p(x,y) \log \frac{p(x,y)}{p(x)p(y)}$$

$$= \sum_{x\in\mathcal{X}, y\in\mathcal{Y}} p_x \operatorname{Tr}(\rho_x E_y^\dagger E_y) \left(\log p_x + \log\left(\operatorname{Tr}(\rho_x E_y^\dagger E_y)\right) - \log p_x - \log\left(\operatorname{Tr}(\rho E_y^\dagger E_y)\right)\right)$$

$$= \sum_{x\in\mathcal{X}, y\in\mathcal{Y}} p_x \operatorname{Tr}(\rho_x E_y^\dagger E_y) \left(\log \operatorname{Tr}(\rho_x E_y^\dagger E_y) - \log\left(\operatorname{Tr}(\rho E_y^\dagger E_y)\right)\right)$$

Let $\{|y\rangle\}_{y=1}^{|\mathcal{Y}|}$ be an orthonormal basis in $\mathbb{C}^{|\mathcal{Y}|}$ and define the linear map

$\gamma : M_d(\mathbb{C}) \longrightarrow M_{|\mathcal{Y}|}(\mathbb{C})$ such that $\boxed{\gamma[\rho] = \sum_{y\in\mathcal{Y}} \sum_{j=1}^{d} |y\rangle\langle j| E_y \rho E_y^\dagger |j\rangle\langle y|}$

- γ is completely positive : $\gamma[\rho] = \sum_\alpha A_\alpha \rho A_\alpha^\dagger$, $A_\alpha = |y\rangle\langle j| E_y : \mathbb{C}^d \to \mathbb{C}^{|\mathcal{Y}|}$
- γ is trace-preserving : $\operatorname{Tr}(\gamma[\rho]) = \sum_{y\in\mathcal{Y}} \sum_{j=1}^{d} \langle j| E_y \rho E_y^\dagger |j\rangle = \sum_{y\in\mathcal{Y}} \operatorname{Tr}(\rho E_y^\dagger E_y) = \operatorname{Tr}\rho$
- $\gamma[\rho] = \sum_{y\in\mathcal{Y}} \operatorname{Tr}(\rho E_y^\dagger E_y) |y\rangle\langle y|$

- Relative Entropy: $S(\rho_1;\rho_2) = \mathrm{Tr}\big(\rho_1(\log\rho_1 - \log\rho_2)\big)$

$$\gamma[\rho] \text{ diagonal } \Rightarrow S(\gamma[\rho_1],\gamma[\rho_2]) = \sum_{y\in\mathcal{Y}} \mathrm{Tr}(\rho_1 E_y^\dagger E_y)\Big(\log\big(\mathrm{Tr}(\rho_1 E_y^\dagger E_y)\big) - \log\big(\mathrm{Tr}(\rho_2 E_y^\dagger E_y)\big)\Big)$$

$$\Downarrow$$

$$I(X;Y) = \sum_{x\in\mathcal{X}} p_x\, S(\gamma[\rho_x],\gamma[\rho])$$

- Monotonicity of the relative entropy under CPTP maps yields:

$$\boxed{I(X;Y) \le \sum_{x\in\mathcal{X}} p_x\, S(\rho_x,\rho) = S(\rho) - \sum_{x\in\mathcal{X}} p_x\, S(\rho_x)}$$

- The upper bound is independent of the POVM $\{E_y\}_{y\in\mathcal{Y}}$; therefore,

$$I_{acc}(X) := \max_{\{E_y\}_{y\in\mathcal{Y}}} I(X;Y) \le S(\rho) - \sum_{x\in\mathcal{X}} p_x\, S(\rho_x)$$

Definition 2.4.3 $\boxed{\chi := S(\rho) - \sum_{x\in\mathcal{X}} p_x\, S(\rho_x)}$ is called <u>Holevo χ-quantity.</u>

Remark

1. $\rho = \sum_{x \in \mathcal{X}} p_x \rho_x \implies \boxed{\begin{array}{l} S(\rho) \leq \sum_{x \in \mathcal{X}} p_x S(\rho_x) + H(X) \\ H(X) = - \sum_{x \in \mathcal{X}} p_x \log p_x \end{array}} \implies I_{acc}(X) \leq H(X)$

By quantum encodings of n classical bits one cannot transmit more than n bits of information.

2. From Fano's inequality: $H(X|Y) \leq H(E) + p(E=1) \log(|\mathcal{X}|-1)$

$$E = \begin{cases} 1 & \text{if } X \neq f(X) = Y \\ 0 & \text{if } X = f(X) = Y \end{cases}$$

Then, $H(E) + p(E=1)\log(|\mathcal{X}|-1) \geq H(X|Y) = H(X) - I(X;Y) \geq H(X) - \chi$

$$\boxed{H(E) \geq H(X) - \chi - p(E=1)\log(|\mathcal{X}|-1)}$$

provides bounds on how well Bob may infer the value of X.

$\boxed{\text{Example 2.4.4}}$ $\mathcal{X} = \{0, 1\}$; $|\mathcal{X}| = 2$

$$\begin{cases} 0 \longrightarrow P_0 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} (1\,0) = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} ; \quad p_0 = \frac{1}{2} \\ 1 \longrightarrow P_1 = \begin{pmatrix} \cos\vartheta \\ \sin\vartheta \end{pmatrix} (\cos\vartheta \ \sin\vartheta) = \begin{pmatrix} \cos^2\vartheta & \cos\vartheta\sin\vartheta \\ \cos\vartheta\sin\vartheta & \sin^2\vartheta \end{pmatrix} ; \quad p_1 = \frac{1}{2} \end{cases}$$

$$\rho = \frac{1}{2}(P_0 + P_1) = \frac{1}{2} \begin{pmatrix} 1 + \cos^2\vartheta & \cos\vartheta\sin\vartheta \\ \cos\vartheta\sin\vartheta & \sin^2\vartheta \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 1 + r_3 & r_1 - i r_2 \\ r_1 + i r_2 & 1 - r_3 \end{pmatrix}$$

$$\bar{r} = (r_1, r_2, r_3) = (\cos\vartheta\sin\vartheta, 0, \cos^2\vartheta) ; \quad r = \|\bar{r}\| = \cos\vartheta , \quad 0 \le \vartheta \le \pi/2$$

$$\rho = \frac{1+r}{2} \frac{1 + \bar{r}/r \cdot \bar{\sigma}}{2} + \frac{1-r}{2} \frac{1 - \bar{r}/r \cdot \bar{\sigma}}{2} \Longrightarrow Sp(\rho) = \left\{ \frac{1+\cos\vartheta}{2}, \frac{1-\cos\vartheta}{2} \right\}$$

$$\boxed{\chi = h\left(\frac{1+\cos\vartheta}{2}\right) = -\frac{1+\cos\vartheta}{2} \log \frac{1+\cos\vartheta}{2} - \frac{1-\cos\vartheta}{2} \log \frac{1-\cos\vartheta}{2}} \Longleftarrow \boxed{\chi = S(\rho)}$$

$$\boxed{H(\varepsilon) \ge 1 - \chi} \Longleftarrow \boxed{H(X) = h(1/2) = 1 , \quad \log(|\mathcal{X}| - 1) = \log 1 = 0}$$

- $\vartheta = \frac{\pi}{2} \Longrightarrow P_0 \perp P_1 \quad \not\Rightarrow \quad H(\varepsilon) \ge 0$
- $\vartheta \to 0 \Longrightarrow P_1 \to P_0 \quad \not\Rightarrow \quad H(\varepsilon) \ge 1 - \varepsilon$