# Final Recitation – Information Theory, Application for Quantum Fault Tolerance.

David Ponarovsky

June 20, 2025

#### Introduction

- ▶ Brief overview of the topic
- ► Importance and relevance
- Objectives of the presentation

# **Key Points**

- ▶ Main point 1
- ► Main point 2
- ► Main point 3

#### Claim

Let Y be a bit given by moving X trough BSC(p), Then there is  $\gamma_p < 1$  such :

$$1 - H(Y) \le \gamma \left(1 - H(X)\right)$$

Denote by  $\delta$  the parameter for which X distributed as  $\sim Bin(\frac{1+\delta}{2})$ . First notice that:

$$\Pr(Y = 1) = \frac{1+\delta}{2}(1-p) + \frac{1-\delta}{2}p = \frac{1+\delta-2\delta p}{2}$$

So 
$$Y \sim \text{Bin}(\frac{1-\delta(1-2p)}{2})$$
, Or  $\delta \mapsto 1-2p\delta$ .

Now expand 1 - H(X) to it's Taylor Seryias at  $\delta$  gives:

$$1 - H(X) = 1 - \frac{1}{2} \left( (1 + \delta) \log \left( \frac{1 + \delta}{2} \right) + (1 - \delta) \log \left( \frac{1 - \delta}{2} \right) \right)$$

$$= -\frac{1}{2} \left( (1 + \delta) \log \left( \frac{1 + \delta}{2} \right) + (1 - \delta) \log \left( \frac{1 - \delta}{2} \right) \right)$$

$$= -\frac{1}{2} \cdot (1 + \delta) \sum_{i=1}^{\infty} \frac{(-1)^{n+1} \delta^n}{n} + (1 - \delta) \sum_{i=1}^{\infty} \frac{(-1)^{n+1} (-\delta)^n}{n}$$

$$= -\frac{1}{2} \cdot \sum_{i=1}^{\infty} 2 \frac{\delta^{2n}}{2n} - \sum_{i=1}^{\infty} 2 \frac{\delta^{2n}}{2n-1}$$

$$= \sum_{i=1}^{\infty} \frac{\delta^{2n}}{2n(2n-1)}$$

Denote the above by  $K(\delta)$ 

Now, observes that:

$$1 - H(Y) = K(2p\delta) = \sum_{i=1}^{\infty} \frac{(2p\delta)^{2n}}{2n(2n-1)}$$
  
 
$$\leq (1 - 2p)^2 K(\delta) = (1 - 2p)^2 (1 - H(X))$$

And notice that since p<1 we have  $\gamma<1$ , noitce also that inequlity is symmetric to  $p\mapsto 1-p$ , in paritcular the entropy is not increase if either p=0 or p=1.

#### Claim

Let  $Y = (Y_1, Y_2, ..., Y_m)$  be a bit given by moving each of  $X_i \in X = (X_1, X_2, ..., X_m)$  trough BSC(p). Then:

$$m - H(Y) \le \gamma (m - H(X))$$

$$m - H(Y_1, Y_2, ..., Y_m) = m - \sum_{i} H(Y_i | Y_1, Y_2, ..., Y_{i-1})$$

$$\leq m - \sum_{i} H(Y_i | X_1, X_2, ..., X_{i-1})$$

$$\leq \sum_{i} 1 - H(Y_i | X_1, X_2, ..., X_{i-1})$$

$$\leq \sum_{i} \gamma (1 - H(X_i | X_1, X_2, ..., X_{i-1}))$$

$$\leq \gamma \sum_{i} (1 - H(X_i | X_1, X_2, ..., X_{i-1}))$$

$$= \gamma (m - H(X))$$

#### Claim

Denote b  $X = (X_1, X_2, ..., X_m)$  and  $Y = (Y_1, Y_2, ..., Y_m)$  the input and the output distrubtions of reversible p-noisy computation at widith m (bits) and depth d. Then, there:

$$m - H(Y) \le \gamma^d (m - H(X))$$

In particular, for  $d = \Omega(\log m)$  we have  $H(Y) \to m$ .

#### Claim

Let  $\rho_1$  be a reduce density matrix of  $\rho$  Then:

$$-\mathsf{Tr}\ \rho\log\left(\rho_1\otimes I\right)=S(\rho_1)$$

First consider the case in which  $\rho$  is a tensor of  $\rho_1$  namely  $\rho=\rho_1\otimes\rho_2$ , Then clearly  $\rho$  and  $\log\rho_1\otimes I$  commute. Denote by  $\lambda_1,...\lambda_n$  and  $\mu_1,...\mu_m$  the eigen values of  $\rho_1$  and  $\rho_2$ . So the trace equals:

$$\sum \lambda_i \mu_j \log(\lambda_i \cdot 1) = \left(\sum \mu_j\right) \left(\sum_i \lambda_i \log \lambda_i\right)$$
  
=  $(\operatorname{Tr} \ 
ho_2) \sum_i \lambda_i \log \lambda_i = -S(
ho_1)$ 

Let's use the notation  $\sum_{A_k} \rho|_{A_k}$  to denote the sum over all the reduced matrices over k qubits.

#### Claim

Let  $\rho$  be a density matrix over n qubits then:

$$\binom{n}{k}^{-1} \sum_{A_k} I(\rho|_{A_k}) \leq \frac{k}{n} I(\rho)$$

Let  $\rho_1$  be  $\rho^{\otimes k\binom{n}{k}}$  and let  $\rho_2 = \left(\prod_{A_k} \rho|_{A_k}\right)^{\otimes n}$ .

$$0 \ge S(\rho_2|\rho_1) = \operatorname{Tr} \left(\rho_1 \left(\log \rho_1 - \log \rho_2\right)\right) = -S(\rho_1) - \operatorname{Tr} \left(\rho_1 \log \rho_2\right)$$
$$= -k \binom{n}{k} S(\rho) - \sum_{A_k} \operatorname{Tr} \left(\rho_1 \log \left(\rho|A_k\right)^n \otimes I^n\right)$$

Now observes that  $\rho|_{A_K}^{\otimes n}$  is a reduced density matrix of  $\rho_1$ . So we get:

$$0 \leq -k \binom{n}{k} S(\rho) - \sum_{A_k} n S(\rho|_{A_k})$$
$$\Rightarrow \sum_{A_k} I(\rho|_{A_k}) \leq \frac{k}{n} \binom{n}{k} I(\rho)$$

#### Claim

Let  $\rho$  be a density matrix of n qubits. Let each qubit be replaced with independent probability  $\rho$  by a fully mixed qubit denoted by v, to give the density matrix  $\sigma$ . Then  $I(\sigma) \leq (1-\rho) \, I(\rho)$ .

Let us write:

$$\sigma = \sum_{k=1}^{n} \sum_{A_k} p^{n-k} (1-p)^k \rho|_{A_k} \otimes v^{n-k}$$

By the concavity of the entropy (convexity of I), We have:

$$I(\sigma) \leq \sum_{k=1}^{n} \sum_{A_{k}} p^{n-k} (1-p)^{k} [I(\rho|_{A_{k}}) + (n-k)I(v)]$$

$$= \sum_{k=1}^{n} \sum_{A_{k}} p^{n-k} (1-p)^{k} I(\rho|_{A_{k}})$$

$$\leq \sum_{k=1}^{n} \sum_{A_{k}} p^{n-k} (1-p)^{k} \frac{k}{n} \binom{n}{k} I(\rho)$$

$$= (1-p) I(\rho)$$

Zeros Distillation. Consider the unitary majority gate over 3 qubits, which on the computational basis sets the last 3rd bit to be the majority, and acts as follows:

$$M |0,0,1\rangle \mapsto |1,1,0\rangle$$
  
 $M |1,1,0\rangle \mapsto |0,0,1\rangle$ 

And acts trivially on every other configuration.

#### Claim

Consider the noisy zero  $ho=\left(1-\frac{p}{2}\right)\ket{0}\bra{0}+\frac{p}{2}\ket{1}\bra{1}$ . Then:

$$\mathbf{Tr}_{\ [1,2]} M \rho^3 = \left(1 - \frac{p}{2} - \frac{p^2}{4} + \frac{p^3}{8}\right) |0\rangle \langle 0| +,,$$

$$\begin{split} \rho^{\otimes 3} &= \left(1 - \frac{p}{2}\right)^{3} |000\rangle \langle 000| \\ &+ \left(1 - \frac{p}{2}\right)^{2} \frac{p}{2} \left(|001\rangle \langle 001| + |010\rangle \langle 010| + |100\rangle \langle 100|\right) \\ &+ \left(1 - \frac{p}{2}\right) \left(\frac{p}{2}\right)^{2} \left(|110\rangle \langle 110| + |101\rangle \langle 101| + |011\rangle \langle 011|\right) \\ &+ \left(\frac{p}{2}\right)^{3} |111\rangle \langle 111| \\ M\rho^{\otimes 3} &= " \\ &+ \left(1 - \frac{p}{2}\right)^{2} \frac{p}{2} \left(|110\rangle \langle 110| + "\right) \\ &+ \left(1 - \frac{p}{2}\right) \left(\frac{p}{2}\right)^{2} \left(|001\rangle \langle 001| + "\right) \\ &+ " \end{split}$$