Quantum SWITCH for Communication Enhancement

Siqi Yao

SCHOOL OF DATA SCIENCE

August 5, 2025

Table of Contents

Table of Contents

Intuition

- The quantum `SWITCH` is an operation that allows multiple quantum channels to act in a superposition of different orders, under the control of a **control qubit/order qubit**.
- **Classical order (fixed order):** Order of two channels $\mathcal{A}$ and $\mathcal{B}$ is fixed by $\mathcal{A} \circ \mathcal{B}(\rho)$ or $\mathcal{B} \circ \mathcal{A}(\rho)$.
- **Quantum `SWITCH` (indefinite order):** Let control qubit $|0\rangle$ controls $\mathcal{A} \circ \mathcal{B}$ and $|1\rangle$ controls $\mathcal{B} \circ \mathcal{A}$. If the control qubit is prepared as $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, then the entire process becomes a superposition of quantum channels, a.k.a **Indefinite Causal Order**.

Figure Illustration

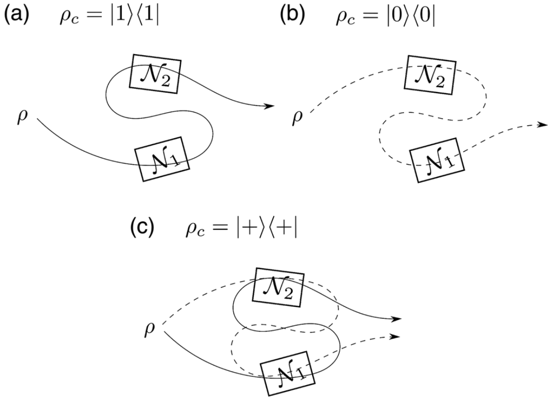

Figure 1: Superposition of Orders. Source: [Ebler et al., 2018]

Fixed Order Channels

- Suppose we have two quantum channels (CPTP maps) $\mathcal{E}$ and $\mathcal{F}$, whose action on quantum sate ρ can be expressed using the Kraus representation as:

$$\mathcal{E}(\rho) = \sum_i E_i \rho E_i^\dagger, \quad \mathcal{F}(\rho) = \sum_j F_j \rho F_j^\dagger \tag{1}$$

  where $\{E_i\}$ and $\{F_j\}$ are the Kraus operators satisfying $\sum_i E_i^\dagger E_i = \sum_j F_j^\dagger F_j = I$.
- Assume the channels are applied sequentially, giving rise to two possible orders:

$$\mathcal{F} \circ \mathcal{E}(\rho) = \sum_{j,i} F_j E_i \rho E_i^\dagger F_j^\dagger, \quad \mathcal{E} \circ \mathcal{F}(\rho) = \sum_{i,j} E_i F_j \rho F_j^\dagger E_i^\dagger. \tag{2}$$

- The order in which the channels are applied are fixed. In $\mathcal{F} \circ \mathcal{E}(\rho)$, $\mathcal{E}$ is applied first, while in $\mathcal{E} \circ \mathcal{F}(\rho)$ the opposite.

Indefinite Order Channels

- The quantum `SWITCH` is a higher-order quantum channel constructed from $\mathcal{E}, \mathcal{F}$ and an ancilla control qubit $|\omega\rangle$, defined as:

$$\mathcal{S}(\mathcal{E}, \mathcal{F}, |\omega\rangle)(\rho) = \sum_{i,j} K_{ij}(\rho \otimes \omega) K_{ij}^{\dagger} \tag{3}$$

  where $\omega = |\omega\rangle\langle\omega|$ and $\{K_{ij}\}$ are the Kraus operators:

$$K_{ij} = E_i F_j \otimes |0\rangle\langle 0| + F_j E_i \otimes |1\rangle\langle 1| \tag{4}$$

- The order in which the channels $\mathcal{E}$ and $\mathcal{F}$ act is determined by the state of the control qubit. If $|\omega\rangle = |0\rangle$, $\mathcal{F}$ is applied first, while if $|\omega\rangle = |1\rangle$ the opposite.
- If the control qubit is initially in a superposition state $\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$, then the quantum `SWITCH` creates a **superposition of orders**.
- The **Sender encodes** information into ρ and transmits it. The **Receiver** determines how to **decode** by the state of the control qubit. The control qubit is accessible only to the receiver.

Table of Contents

Holevo's Theorem: Setting

Suppose Alice transmits messages, in the form of density matrices, to Bob through the following procedure:

- Alice samples $X \in \Sigma \subseteq \{0,1\}^n$, where $X = x$ with probability $p(x)$.
- Alice sends $\sigma_X \in \mathbb{C}^{d \times d}$.
- Bob picks POVM's $\{E_y\}_{y \in \Gamma}$, where $\Gamma \subseteq \{0,1\}^n$.
- Bob measures σ_X, and receives output "$Y \in \Gamma$", where $Y = y$ given $X = x$ with probability $\mathrm{Tr}(E_y \sigma_x)$.
- Bob tries to infer X from Y. Note: X, Y are two **classical** probability distributions.

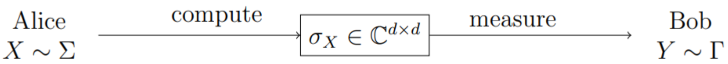

Figure 2: Communication Scheme. Source: [O'Donnell and Wright, 2015]

Analysis

- Bob sees the mixed state:

$$\begin{cases} \sigma_{x_1} \text{ with prob. } p(x_1), \\ \sigma_{x_2} \text{ with prob. } p(x_2), \\ \vdots \end{cases} \equiv \sum_{x\in\Sigma} p(x)\sigma_x =: \rho_B. \tag{5}$$

- Alice sees:

$$\begin{cases} |x_1\rangle \text{ with prob. } p(x_1), \\ |x_2\rangle \text{ with prob. } p(x_2), \\ \vdots \end{cases} \equiv \sum_{x\in\Sigma} p(x)|x\rangle\langle x| =: \rho_A. \tag{6}$$

- State of the joint mixed system is:

$$\rho := \sum_{x\in\Sigma} p(x)|x\rangle\langle x| \otimes \sigma_x. \tag{7}$$

- We wish to answer the question "how much does seeing one random variable tell me about the other". Solution is given by the mutual information.
- The **classical mutual information** $I(X;Y)$ between two random variables X and Y is

$$I(X;Y) = H(X) + H(Y) - H(X,Y), \tag{8}$$

where $H(\cdot)$ is the Shannon entropy. Mutual information represents the amount of information one learn about X from knowing Y, and vice-versa since it is symmetric in X and Y.

- The **accessible information** is defined as:

$$I_{\text{acc}}(\sigma, p) = \max_{\substack{\text{over all} \\ \text{POVMs} \\ \{E_y\}_{y \in \Gamma}}} I(X;Y). \tag{9}$$

This represents the best Bob can do given Alice's choice of the σ_x's and the distribution p.

- If ρ is the joint state of two quantum systems A and B, then the **quantum mutual information** is:

$$I(\rho_A; \rho_B) = H(\rho_A) + H(\rho_B) - H(\rho). \tag{10}$$

where $H(\cdot)$ is the von Neumann entropy. Note: If $\rho = \rho_A \otimes \rho_B$, then $I(\rho_A; \rho_B) = 0$ since $H(\rho_A \otimes \rho_B) = H(\rho_A) + H(\rho_B)$.

- The **Holevo information** is defined as:

$$\chi(\sigma, p) := I(\rho_A; \rho_B). \tag{11}$$

Statement

Theorem 1 (Holevo's Theorem/Holevo's Bound)

The accessible information is upper-bounded by the Holevo information:

$$I_{\mathrm{acc}}(\sigma, p) \leq \chi(\sigma, p). \tag{12}$$

Equivalent form of $\chi(\sigma, p)$:

$$\chi(\sigma, p) = \chi(\eta) := S\left(\sum_i p_i \rho_i\right) - \sum_i p_i S(\rho_i) \tag{13}$$

where $p_i = p(x_i), \rho_i = \sigma_{x_i}, \eta = \{(p_i, \rho_i)\}$ and $S(\cdot)$ denotes the von Neumann entropy. $\chi(\eta)$ is called the **Holevo information** or **Holevo χ quantity**.

We will prove (13).

Proof of (13)

Write out all the states:

$$\rho_A = \sum_i p_i |i\rangle\langle i|, \quad \rho_B = \sum_i p_i \rho_i, \quad \rho_{AB} = \sum_i p_i |i\rangle\langle i| \otimes \rho_i \tag{14}$$

Compute von Neumann entropy:

$$H(\rho_A) = H(\{p_i\}), \quad H(\rho_B) = S\left(\sum_i p_i \rho_i\right) \tag{15}$$

where $H(\{p_i\})$ is the Shannon entropy of $\{p_i\}$. Moreover we have:

$$H(\rho_{AB}) = H(\{p_i\}) + \sum_i p_i S(\rho_i) \tag{16}$$

Plug $H(\rho_A), H(\rho_B)$ and $H(\rho_{AB})$ into the definition of $I(\rho_A; \rho_B)$ will get the result.

Proof of (13) (cont'd)

Finally we will show that (16) holds. ρ_{AB} can be written as a block diagonal matrix:

$$\rho_{AB} = \begin{pmatrix} p_1\rho_1 & 0 & 0 & \cdots \\ 0 & p_2\rho_2 & 0 & \cdots \\ 0 & 0 & p_3\rho_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \tag{17}$$

Each ρ_i has its own eigenvalue r_{ik} (where k is the index of the eigenstate of ρ_i). The eigenvalues of ρ_{AB} are all $p_i r_{ik}$. So we have:

$$\begin{aligned} H(\rho_{AB}) &= -\sum_{i,k} p_i r_{ik} \log(p_i r_{ik}) && (18) \\ &= -\sum_{i,k} p_i r_{ik} (\log p_i + \log r_{ik}) && (19) \\ &= -\sum_i p_i \log p_i \left(\sum_k r_{ik} \right) - \sum_i p_i \sum_k r_{ik} \log r_{ik} && (20) \\ &= -\sum_i p_i \log p_i + \sum_i p_i S(\rho_i) && (21) \\ &= H(\{p_i\}) + \sum_i p_i S(\rho_i) && (22) \end{aligned}$$

where $\sum_k r_{ik} = 1$ (each density matrix's trace is 1) and $-\sum_k r_{ik} \log r_{ik} = S(\rho_i)$. $\square$

HSW Theorem

In classical information theory, the **channel capacity** C is defined as:

$$C = \max_{p(x)} I(X;Y) \tag{23}$$

where $p(x)$ is the probability distribution of the input X. Combining Holevo's theorem, we have the following result:

Theorem 2 (Holevo-Schumacher-Westmoreland Theorem)

Suppose quantum channel $\mathcal{E}$ has the Kraus representation $\mathcal{E}(\sigma) = \sum_j E_j \sigma E_j^\dagger$. $\rho := \sum_i p_i \rho_i$. Then the classical channel capacity of $\mathcal{E}$ is given by:

$$\chi(\mathcal{E}) = \max_{\{p_i, \rho_i\}} \left[S\left(\mathcal{E}(\rho)\right) - \sum_i p_i S(\mathcal{E}(\rho_i)) \right] \tag{24}$$

$$= \max_{\{p_i, \rho_i\}} \left[S\left(\sum_i p_i \sum_j E_j \rho_i E_j^\dagger \right) - \sum_i p_i S\left(\sum_j E_j \rho_i E_j^\dagger \right) \right] \tag{25}$$

where ρ_i is the input quantum state, p_i is the probability distribution of the quantum state and $S(\cdot)$ is the von Neumann entropy. $\chi(\mathcal{E})$ is called the **Holevo information**.

Table of Contents

Motivation for Quantum SWITCH

- In this section, we will focus on the communication of **classical** information.
- Any non-constant quantum channel $\mathcal{N}$ has positive Holevo information:
  - There exist at least two pure states $|\phi\rangle$ and $|\psi\rangle$, such that $\mathcal{N}(|\phi\rangle\langle\phi|) \neq \mathcal{N}(|\psi\rangle\langle\psi|)$.
  - Set $p(\phi) = p(\psi) = \frac{1}{2}$, then the joint state becomes

$$\sigma = \frac{1}{2}\mathcal{N}(|\phi\rangle\langle\phi|) + \frac{1}{2}\mathcal{N}(|\psi\rangle\langle\psi|) \tag{26}$$

  - So the Holevo information is

$$\chi(\mathcal{N}) = S(\sigma) - (\frac{1}{2}S(\mathcal{N}(|\phi\rangle\langle\phi|)) + \frac{1}{2}S(\mathcal{N}(|\psi\rangle\langle\psi|))) > 0 \tag{27}$$

    due to the concavity of $S(\cdot)$.
- The Holevo information of a constant channel, e.g. **completely depolarizing channel**, is zero, which makes it impossible to perform classical communication. However, this can be achieved with quantum SWITCH.

Kraus Operator for Depolarizing Channel

- A completely depolarizing channel $\mathcal{N}^D$ on a d-dimensional quantum system can be represented by d^2 orthogonal ($\mathrm{Tr}[U_i^\dagger U_j] = d \cdot \delta_{ij}]$) unitary operators U_i, such that its action on a state ρ is

$$\mathcal{N}^D(\rho) = \frac{1}{d^2} \sum_{i=1}^{d^2} U_i \rho U_i^\dagger = \mathrm{Tr}[\rho] \frac{I}{d}. \tag{28}$$

  Example for $d = 2$:
  $U_1 = \sqrt{2}|0\rangle\langle 0|, U_2 = \sqrt{2}|0\rangle\langle 1|, U_3 = \sqrt{2}|1\rangle\langle 0|, U_4 = \sqrt{2}|1\rangle\langle 1|.$

- Thus, according to (4), the overall quantum channel resulting from the quantum `SWITCH` of **two completely depolarizing channels** has the Kraus operator:

$$W_{ij} = \frac{1}{d^2} \left(U_i U_j \otimes |0\rangle\langle 0|_c + U_j U_i \otimes |1\rangle\langle 1|_c \right) \tag{29}$$

Quantum SWITCH for Depolarizing Channel

- Set control state to $\rho_c := |\psi_c\rangle\langle\psi_c|$, where $|\psi_c\rangle := \sqrt{p}|0\rangle + \sqrt{1-p}|1\rangle$.
- Set the input state to ρ, then the receiver will get the output state

$$\mathcal{S}(\mathcal{N}^D, \mathcal{N}^D, \rho_c)(\rho) = \frac{1}{d^4}\sum_{i,j}\Big(p|0\rangle\langle 0|_c \otimes U_i U_j \rho U_j^\dagger U_i^\dagger + (1-p)|1\rangle\langle 1|_c \otimes U_j U_i \rho U_i^\dagger U_j^\dagger \tag{30}$$

$$+ \sqrt{p(1-p)}|0\rangle\langle 1|_c \otimes U_i U_j \rho U_i^\dagger U_j^\dagger + \sqrt{p(1-p)}|1\rangle\langle 0|_c \otimes U_j U_i \rho U_j^\dagger U_i^\dagger\Big)$$

$$= p|0\rangle\langle 0|_c \otimes \frac{I}{d} + (1-p)|1\rangle\langle 1|_c \otimes \frac{I}{d} \tag{31}$$

$$+ \sqrt{p(1-p)}\frac{|0\rangle\langle 1|_c}{d^2} \otimes \sum_j \mathrm{Tr}[U_j\rho]\frac{U_j^\dagger}{d}$$

$$+ \sqrt{p(1-p)}\frac{|1\rangle\langle 0|_c}{d^2} \otimes \sum_j \mathrm{Tr}[\rho U_j^\dagger]\frac{U_j}{d}.$$

$$= (p|0\rangle\langle 0|_c + (1-p)|1\rangle\langle 1|_c) \otimes \frac{I}{d} + \sqrt{p(1-p)}(|0\rangle\langle 1|_c + |1\rangle\langle 0|_c) \otimes \frac{\rho}{d^2}. \tag{32}$$

Brief Explanation

- (30) results from (3).
- $\sum_j \mathrm{Tr}[U_j\rho]\frac{U_j^\dagger}{d}$ results from viewing $U_j\rho$ as a single state and applying (28).
- (32) follows from the fact that $\{U_j\}$ forms an orthonormal basis for the set of $d \times d$ matrices, i.e., $\rho = \sum_{j=1}^{d^2} \mathrm{Tr}[U_j\rho]\ \frac{U_j^\dagger}{d}$. This is related to the **Hilbert-Schmidt operator and Hilbert-Schmidt space**.
- The quantum `SWITCH` of two depolarizing channels depends on state ρ. Thus, **we can communicate classical information at a nonzero rate**.

Table of Contents

Extending to > 2 Channels

- When $N = 2$, completely depolarizing channels are unable to transmit quantum data, even when applying quantum `SWITCH`.
- We consider $N(\geq 3)$ completely depolarizing channels combined in a superposition of N causal orders related to each other by **cyclic permutations**.
- The intermediate nodes in Figure 3 (purple) are identity operations

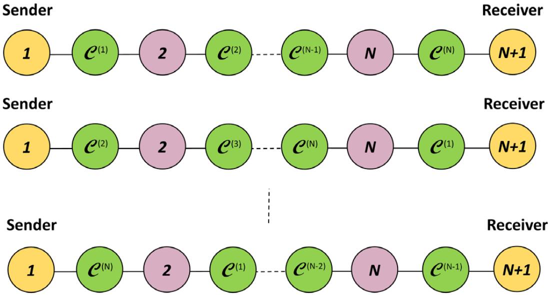

Figure 3: Cyclic Permutations. Source: [Chiribella et al., 2021a]

Settings

- Take $N = 3$ for simplicity, construct the quantum `SWITCH` $\mathcal{S}(\mathcal{E}, \mathcal{F}, \mathcal{G}, |\omega\rangle)(\rho)$ where:
  - $\mathcal{E}, \mathcal{F}, \mathcal{G}$ are three completely depolarizing channels.
  - $|\omega\rangle$ is the control **qutrit**.
- Set:
  - $\mathcal{E}(\rho) = \sum_i E_i \rho E_i^\dagger$
  - $\mathcal{F}(\rho) = \sum_j F_j \rho F_j^\dagger$
  - $\mathcal{G}(\rho) = \sum_k G_k \rho G_k^\dagger$
- Let the **orthonormal** basis of the control space be:
  - $|\pi_0\rangle$ denotes the order $\mathcal{E} \to \mathcal{F} \to \mathcal{G}$
  - $|\pi_1\rangle$ denotes the order $\mathcal{F} \to \mathcal{G} \to \mathcal{E}$
  - $|\pi_2\rangle$ denotes the order $\mathcal{G} \to \mathcal{E} \to \mathcal{F}$
- Let the control state be the equal superposition $|\omega\rangle = \frac{1}{\sqrt{3}} \left(|\pi_0\rangle + |\pi_1\rangle + |\pi_2\rangle\right)$ and let $\mathbf{S}$ denote the set of permutations. The density matrix becomes:

$$\omega = |\omega\rangle\langle\omega| = \frac{1}{3} \sum_{\pi, \pi' \in \mathbf{S}} |\pi\rangle\langle\pi'| \tag{33}$$

Kraus Operators

- $G_k F_j E_i := K_{ijk}^{(\pi_0)}, E_i G_k F_j := K_{ijk}^{(\pi_1)}, F_j E_i G_k := K_{ijk}^{(\pi_2)}$. Extending from (4) we have:

$$K_{ijk} = K_{ijk}^{(\pi_0)} \otimes |\pi_0\rangle\langle\pi_0| + K_{ijk}^{(\pi_1)} \otimes |\pi_1\rangle\langle\pi_1| + K_{ijk}^{(\pi_2)} \otimes |\pi_2\rangle\langle\pi_2| \tag{34}$$

$$= \sum_{\pi} K_{ijk}^{(\pi)} \otimes |\pi\rangle\langle\pi| \tag{35}$$

Quantum SWITCH

- So the channel becomes:

$$\mathcal{S} = \sum_{i,j,k} K_{ijk}(\rho \otimes \frac{1}{3}\sum_{\pi,\pi'} |\pi\rangle\langle\pi'|)K_{ijk}^{\dagger} \tag{36}$$

$$= \sum_{i,j,k}\sum_{\pi} K_{ijk}^{(\pi)} \otimes |\pi\rangle\langle\pi|(\rho \otimes \frac{1}{3}\sum_{\pi,\pi'} |\pi\rangle\langle\pi'|)\sum_{\pi'} K_{ijk}^{(\pi')\dagger} \otimes |\pi'\rangle\langle\pi'| \tag{37}$$

$$= \sum_{i,j,k}\sum_{\pi,\pi'} K_{ijk}^{(\pi)}\rho K_{ijk}^{(\pi')\dagger} \otimes \frac{1}{3}|\pi\rangle\langle\pi'| \tag{38}$$

$$= \sum_{\pi,\pi'} \mathcal{C}_{\pi\pi'}(\rho) \otimes \omega_{\pi,\pi'}\,|\pi\rangle\langle\pi'| := \mathcal{C}_{\text{eff}}(\rho) \tag{39}$$

  where $\mathcal{C}_{\pi\pi'}(\rho) := \sum_{i,j,k} K_{ijk}^{(\pi)}\rho K_{ijk}^{(\pi')\dagger}, \omega_{\pi,\pi'} = \frac{1}{3}$.
- Notation for general N channel setting: $C_{\pi\pi'}(\rho) = \sum_{s_1,\ldots,s_N} K_{(s_1,\ldots,s_N)}^{(\pi)}\,\rho\,K_{(s_1,\ldots,s_N)}^{(\pi')\dagger}$, where s_i's are indexes of Kraus operators corresponding to each channel.

Channel Construction

- Recall: $\omega = |\omega\rangle\langle\omega|$ where $|\omega\rangle = \frac{1}{\sqrt{N}} \sum_\pi |\pi\rangle$.
- [Chiribella et al., 2021b] showed that:

$$\mathcal{C}_{\pi\pi}(\rho) = \frac{I}{d} \quad \text{and} \quad \mathcal{C}_{\pi\pi'}(\rho) = \frac{\rho}{d^2} \quad \forall\, \pi \neq \pi' \tag{40}$$

- Plug into (39) yields:

$$\mathcal{C}_{\text{eff}}(\rho) = \sum_\pi \frac{I}{d} \otimes \frac{1}{N} |\pi\rangle\langle\pi| + \sum_{\pi\neq\pi'} \frac{\rho}{d^2} \otimes \frac{1}{N} |\pi\rangle\langle\pi'| \tag{41}$$

$$= \frac{I}{d} \otimes \frac{I}{N} + \frac{\rho}{Nd^2} \otimes \sum_{\pi\neq\pi'} |\pi\rangle\langle\pi'| \tag{42}$$

$$= \frac{I}{d} \otimes \frac{I}{N} + \frac{\rho}{Nd^2} \otimes (N|\omega\rangle\langle\omega| - I)\,, \tag{43}$$

where (43) comes from the relations $N|\omega\rangle\langle\omega| = \sum_{\pi,\pi'} |\pi\rangle\langle\pi'|$ and $I = \sum_\pi |\pi\rangle\langle\pi|$.

Channel as a Mixture of Channels

- Rearranging (43) we have:

$$\mathcal{C}_{\text{eff}}(\rho) = \mathcal{E}_0(\rho) \otimes (1-p)\rho_0 + \mathcal{E}_1(\rho) \otimes p\rho_1, \tag{44}$$

  where $\rho_0 := |\omega\rangle\langle\omega|$ and $\rho_1 := \frac{I-|\omega\rangle\langle\omega|}{N-1}, p := \frac{(N-1)(d^2-1)}{Nd^2}$. $\mathcal{E}_0, \mathcal{E}_0$ are quantum channels defined by

$$\mathcal{E}_0(\rho) := \frac{N-1}{N-1+d^2}\rho + \frac{d^2}{N-1+d^2}\frac{I}{d} \tag{45}$$

  and

$$\mathcal{E}_1(\rho) := \frac{d^2}{d^2-1}\frac{I}{d} - \frac{1}{d^2-1}\rho \tag{46}$$

- $\mathcal{C}_{\text{eff}}$ is a mixture of two channels $\mathcal{E}_0$ and $\mathcal{E}_1$. By measuring ρ_0 and ρ_1, it is possible to determine the occurrence of the channels $\mathcal{E}_0$ and $\mathcal{E}_1$.

Transmitting Quantum Data

- $\mathcal{E}_1$ is unable to transmit quantum data, see [Chiribella et al., 2021a].
- $\mathcal{E}_0$ is a depolarizing channel, with probability of depolarization equal to $\frac{d^2}{N+d^2-1}$.
- $\lim_{N\to\infty} \frac{d^2}{N+d^2-1} = 0 \Rightarrow \lim_{N\to\infty} \mathcal{E}_0(\rho) = \rho$. Therefore, as long as we wait until the control system is measured to be ρ_0, we achieve **nearly perfect quantum information transmission**.

Table of Contents

- Next, we discuss quantum SWITCH constructed by **general channels and permutations**. Let n denote the total number of channels, $m := |\mathbf{S}|$ denote the number of permutations in $\mathbf{S}$. n may not equal m.
- [Wu et al., 2025a] introduced the key quantity

$$\mathcal{P}_n = 1 - \frac{1}{m^2} \min_{\rho} \sum_{\pi,\pi' \in \mathbf{S}} \mathrm{Tr}(C_{\pi\pi'}(\rho)) \tag{47}$$

- $\mathcal{P}_n$ can be viewed as the maximum probability (as ρ varies over all input states) of obtaining the measurement outcome F_2 associated with the POVM $\{F_1 = |\omega\rangle\langle\omega| = \frac{1}{m}\sum_{i,j\in\mathbf{S}} |i\rangle\langle j|,\ F_2 = I - |\omega\rangle\langle\omega|\}$. Proof in next page.

Proof

Probability of obtaining outcome F_1 can be computed as:

$$\Pr[F_1] = \mathrm{Tr}\left[(I \otimes F_1)\,\mathcal{C}_{\mathrm{eff}}(\rho)\right]. \tag{48}$$

$$= \mathrm{Tr}\left[(I \otimes F_1)\sum_{\pi,\pi'} \mathcal{C}_{\pi\pi'}(\rho) \otimes \frac{1}{m}|\pi\rangle\langle\pi'|\right] \tag{49}$$

$$= \frac{1}{m}\sum_{\pi,\pi'} \mathrm{Tr}\left[\mathcal{C}_{\pi\pi'}(\rho) \otimes (F_1|\pi\rangle\langle\pi'|)\right] \tag{50}$$

$$= \frac{1}{m}\sum_{\pi,\pi'} \mathrm{Tr}(\mathcal{C}_{\pi\pi'}(\rho)) \cdot \mathrm{Tr}(F_1|\pi\rangle\langle\pi'|) \tag{51}$$

$$= \frac{1}{m^2}\sum_{\pi,\pi'} \mathrm{Tr}(\mathcal{C}_{\pi\pi'}(\rho)) \tag{52}$$

So probability of obtaining outcome F_2 is:

$$\Pr[F_2] = 1 - \Pr[F_1] = 1 - \frac{1}{m^2}\sum_{\pi,\pi'} \mathrm{Tr}(\mathcal{C}_{\pi\pi'}(\rho)) \tag{53}$$

So we have:

$$\mathcal{P}_n = \max_{\rho} \Pr[F_2] = 1 - \min_{\rho} \frac{1}{m^2}\sum_{\pi,\pi'} \mathrm{Tr}(\mathcal{C}_{\pi\pi'}(\rho)) \tag{54}$$

Proposition on S-invariance and Commutativity

Proposition 1

The quantity $\mathcal{P}_n = 0$ if and only if the Kraus operators $\{C_i\}$ are **S**-invariant, i.e.,

$$K^{(\pi)}_{(s_1,\ldots,s_n)} = K^{(\pi')}_{(s_1,\ldots,s_n)} \quad (55)$$

for all indices $s_1, \ldots, s_n$ and for all $\pi, \pi' \in \mathbf{S}$. That is, the Kraus operators are **independent of the causal order**. In particular, for $n = m = 2$, **S**-invariance is equivalent to commutativity: $E_i F_j = F_j E_i \quad \forall\, i, j$ for two sets of Kraus operators $\{E_i\}, \{F_j\}$.

Proof of Proposition 1

First we will show a simple fact that $\mathrm{Tr}(C_{\pi\pi}(\rho)) = 1 \quad \forall \pi \in \mathbf{S}$.
Recall:

$$C_{\pi\pi}(\rho) = \sum_{s_1,\dots,s_n} K^{(\pi)}_{(s_1,\dots,s_n)}\, \rho\, K^{(\pi)\dagger}_{(s_1,\dots,s_n)} \tag{56}$$

Here, $K^{(\pi)}_{(s_1,\dots,s_n)}$ is the composite Kraus operator under the permutation π:

$$K^{(\pi)}_{(s_1,\dots,s_n)} := A^{(\pi)}_{s_1,\dots,s_n} = A_{\pi(1),s_1} A_{\pi(2),s_2} \cdots A_{\pi(n),s_n}, \tag{57}$$

$$K^{(\pi)\dagger}_{(s_1,\dots,s_n)} := A^{(\pi)\dagger}_{s_1,\dots,s_n} = A^{\dagger}_{\pi(n),s_n} A^{\dagger}_{\pi(n-1),s_{n-1}} \cdots A^{\dagger}_{\pi(1),s_1} \tag{58}$$

where each A_{i,s_i} is the s_i-th Kraus operator of the i-th quantum channel.
Since each channel is a valid CPTP map, the composite channel must also be a CPTP map.
So we have:

$$\mathrm{Tr}(C_{\pi\pi}(\rho)) = 1 \quad \forall \pi \in \mathbf{S} \tag{59}$$

Proof of Proposition 1 (cont'd)

Let ρ be an arbitrary input state, and let $\pi, \pi' \in \mathbf{S}$ be arbitrary permutations in $\mathbf{S}$. We then have

$$\operatorname{Tr}(C_{\pi\pi'}(\rho)) = \sum_{s_1,\ldots,s_n} \operatorname{Tr}\left(K^{(\pi)}_{(s_1,\ldots,s_n)}\, \rho\, K^{(\pi')\dagger}_{(s_1,\ldots,s_n)}\right) \tag{60}$$

$$= \sum_{s_1,\ldots,s_n} \operatorname{Tr}\left(K^{(\pi)}_{(s_1,\ldots,s_n)} \sqrt{\rho} \left(K^{(\pi')}_{(s_1,\ldots,s_n)} \sqrt{\rho}\right)^\dagger\right) \tag{61}$$

$$\leq \frac{1}{2}\left[\operatorname{Tr}(C_{\pi\pi}(\rho)) + \operatorname{Tr}(C_{\pi'\pi'}(\rho))\right] = 1 \tag{62}$$

where the final equality follows from (59), and the first inequality follows from combining the Cauchy-Schwarz and AM-GM inequalities, which for arbitrary operators A and B yields

$$\operatorname{Tr}(AB^\dagger) \leq \sqrt{\operatorname{Tr}(AA^\dagger)\operatorname{Tr}(BB^\dagger)} \tag{63}$$

$$\leq \frac{1}{2}\left(\operatorname{Tr}(AA^\dagger) + \operatorname{Tr}(BB^\dagger)\right). \tag{64}$$

Explanations on the Cauchy-Schwarz Inequality

To apply Cauchy-Schwarz inequality, we define the **Hilbert-Schmidt inner product** as follows (this is also related to the Hilbert-Schmidt space):

$$\langle A, B \rangle := \mathrm{Tr}(AB^\dagger) \tag{65}$$

The norm induced by this inner product is:

$$\|A\|_{\mathrm{HS}} := \sqrt{\mathrm{Tr}(AA^\dagger)} \tag{66}$$

Applying the Cauchy-Schwarz inequality, we obtain:

$$|\mathrm{Tr}(AB^\dagger)| \leq \sqrt{\mathrm{Tr}(AA^\dagger)} \cdot \sqrt{\mathrm{Tr}(BB^\dagger)} \tag{67}$$

By setting $A_{(s_1,\ldots,s_n)} = K^{(\pi)}_{(s_1,\ldots,s_n)} \sqrt{\rho}, B_{(s_1,\ldots,s_n)} = K^{(\pi')}_{(s_1,\ldots,s_n)} \sqrt{\rho}$, and using Cauchy-Schwarz inequality for each pair of $A_{(s_1,\ldots,s_n)}, B_{(s_1,\ldots,s_n)}$, we get the inequality last page.

Conclusion and Remarks

Since the equality holds if and only if $A_{(s_1,\dots,s_n)} = B_{(s_1,\dots,s_n)} \quad \forall s_1,\dots,s_n$, it follows that

$$\mathrm{Tr}(C_{\pi\pi'}(\rho)) = 1, \tag{68}$$

if and only if

$$K^{(\pi)}_{(s_1,\dots,s_n)}\sqrt{\rho} = K^{(\pi')}_{(s_1,\dots,s_n)}\sqrt{\rho} \quad \forall s_1,\dots,s_n. \tag{69}$$

From the definition of $\mathcal{P}_n$, it follows that $\mathcal{P}_n = 0$ if and only if (69) holds for all states ρ, for all permutations $\pi, \pi' \in \mathbf{S}$, and for all indices $s_1,\dots,s_n$. By considering pure states, we see that $\mathcal{P}_n = 0$ if and only if

$$K^{(\pi)}_{(s_1,\dots,s_n)} = K^{(\pi')}_{(s_1,\dots,s_n)} \tag{70}$$

for all $s_1,\dots,s_n$, as desired. In particular, when $n = m = 2$, the permutation set $\mathbf{S}$ only contains two permutations $(1,2)$ and $(2,1)$, thus S-invariance implies

$$C^1_{s_1}C^2_{s_2} = C^2_{s_2}C^1_{s_1} \quad \text{for all } s_1, s_2, \tag{71}$$

i.e., the Kraus operators of the channels $\mathcal{C}^1$ and $\mathcal{C}^2$ pairwise commute.

Table of Contents

Notations

- Denote the quantum SWITCH channel $\mathcal{C}_{\text{eff}}(\rho)$ as $\mathcal{S}^n(\rho)$, where n is the number of channels.
- In the following, we will consider the case $\mathcal{C}^i = \mathcal{N}$ for $i = 1, \ldots, n$, where $\mathcal{N}$ is a fixed quantum channel, and let $\mathcal{N}^n$ denote the n-fold composition $\mathcal{N} \circ \cdots \circ \mathcal{N}$.

Tracing Out Control Recovers $\mathcal{N}^n$

- $\mathcal{N}^n$ can be obtained from $\mathcal{S}^n$ by tracing out the control system:

$$\mathcal{N}^n(\rho) = \mathcal{N} \circ \cdots \circ \mathcal{N}(\rho) = \sum_{s_1,\ldots,s_n} C_{s_n} \cdots C_{s_1} \rho C_{s_1}^\dagger \cdots C_{s_n}^\dagger \tag{72}$$

Take partial trace:

$$\begin{aligned}\mathrm{Tr}_C\left[\mathcal{S}^n(\rho)\right] &= \sum_{\pi,\pi'\in\mathbf{S}} C_{\pi\pi'}(\rho) \cdot \mathrm{Tr}\left(\omega_{\pi,\pi'}|\pi\rangle\langle\pi'|\right) && (73)\\ &= \sum_{\pi,\pi'\in\mathbf{S}} C_{\pi\pi'}(\rho) \cdot \frac{1}{m}\delta_{\pi\pi'} && (74)\\ &= \frac{1}{m}\sum_{\pi\in\mathbf{S}} C_{\pi\pi}(\rho) && (75)\end{aligned}$$

Clearly $C_{\pi\pi}(\rho) = \mathcal{N}^n(\rho)$ (since $C^i = \mathcal{N} \quad \forall i$), thus:

$$\mathrm{Tr}_C\left[\mathcal{S}^n(\rho)\right] = \mathcal{N}^n(\rho). \tag{76}$$

Quantifying Communication Enhancement

- For all capacity measure f satisfying the data-processing inequality $f(\mathcal{A} \circ \mathcal{B}) \leq f(\mathcal{B})$, for arbitrary quantum channels $\mathcal{A}$, $\mathcal{B}$, we obtain the **bottleneck inequality**:

$$f(\mathcal{N}^n) \leq f(\mathcal{S}^n). \tag{77}$$

  since taking partial trace is also a CPTP map.

- Capacity measures such as classical capacity, quantum capacity, Holevo information and coherent information all satisfy (77). Thus, we define the associated **causal gain** by

$$\delta_f = f(\mathcal{S}^n) - f(\mathcal{N}^n). \tag{78}$$

  δ_f is a direct measure of the communication enhancement of the channel $\mathcal{N}$ which is achieved by inputting n-copies of the channel into the quantum `SWITCH`.

Condition for $\delta_f = 0$

- $\mathcal{P}_n = 0$ if and only if $K^{(\pi)}_{(s_1,\ldots,s_n)} = K^{(\pi')}_{(s_1,\ldots,s_n)}$ for all $s_1, \ldots, s_n$. Hence, for a given subset of permutations $\mathbf{S}$, $\mathcal{P}_n = 0$ implies that $C_{\pi\pi'}(\rho) = C_{\pi\pi}(\rho) = C_{\pi'\pi'}(\rho) = C(\rho)$ for arbitrary permutations $\pi, \pi' \in \mathbf{S}$.
- Therefore $\mathcal{S}^n(\rho) = \mathcal{N}^n(\rho) \otimes \omega$, where ω is the control state and is independent of the input state ρ. Proof:

$$\begin{aligned}\mathcal{S}^n(\rho) &= \sum_{\pi,\pi'\in\mathbf{S}} C_{\pi\pi'}(\rho) \otimes \omega_{\pi,\pi'} |\pi\rangle\langle\pi'| && (79)\\ &= \sum_{\pi,\pi'\in\mathbf{S}} \mathcal{N}^n(\rho) \otimes \omega_{\pi,\pi'} |\pi\rangle\langle\pi'| && (80)\\ &= \mathcal{N}^n(\rho) \otimes \left(\sum_{\pi,\pi'\in\mathbf{S}} \omega_{\pi,\pi'} |\pi\rangle\langle\pi'| \right) && (81)\\ &= \mathcal{N}^n(\rho) \otimes \omega && (82)\end{aligned}$$

- Since ω is a constant, we have $f(\mathcal{S}^n) = f(\mathcal{N}^n)$, so that $\delta_f = 0$.

Main Conjecture

We have shown that $\mathcal{P}_n > 0$ is a necessary condition for positive causal gain, so it is natural to propose the following conjecture:

Conjecture

For all channels outside **a set of measure-zero**, the condition $\mathcal{P}_n > 0$ is necessary and sufficient for $\delta_f > 0$.

Table of Contents

Pauli Channel

The **Pauli channel** is a prototypical noise channel acting on a single qubit. It describes a probabilistic mixture of Pauli operations applied to the qubit:

$$\mathcal{E}(\rho) = \sum_{i=0}^{3} p_i \sigma_i \rho \sigma_i^\dagger, \quad \text{with } \sum_i p_i = 1 \tag{83}$$

where:

- ρ is the input quantum state (a density matrix),
- $\sigma_0 = I$, $\sigma_1 = X$, $\sigma_2 = Y$, $\sigma_3 = Z$ are the four Pauli operators,
- $p_i \geq 0$ is the probability of applying operation σ_i.

$$\sigma_0 = I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$
$$\sigma_2 = Y = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Setting

- In this section, we verify the conjecture in the space of all Pauli channels $\mathcal{N}(\rho) = \sum_{i=0}^{3} p_i \sigma_i \rho \sigma_i^\dagger$, and the permutation set $\mathbf{S}$ contains the identity permutation $(1, \dots, n)$ and its reversal $(n, \dots, 1)$.
- In such case, the quantum `SWITCH` places n copies of a Pauli channel $\mathcal{N}$ in a superposition of the forward and backward orders $\mathcal{N}_n \circ \cdots \circ \mathcal{N}_1$ and $\mathcal{N}_1 \circ \cdots \circ \mathcal{N}_n$ (where $\mathcal{N}_i = \mathcal{N}$ for all i).
- The forward and backward orders are indistinguishable when used individually. We show that when such orders are placed in a superposition via quantum `SWITCH`, an enhancement of classical capacity and coherent information occurs almost surely

Operational Meaning of $\mathcal{P}_n$

- Fix the control state to be $\omega = |+\rangle\langle+|$ with $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, then apply the projective measurement $\{F_1 = |\omega\rangle\langle\omega|, F_2 = I - |\omega\rangle\langle\omega|\}$.
- The probability of obtaining F_2 is independent of the initial state ρ (proof on next few pages), i.e., for all states ρ,

$$\mathrm{Tr}((I \otimes F_2)\mathcal{S}^n(\rho)) = \mathcal{P}_n. \tag{84}$$

Proof

Recall, probability of obtaining outcome F_2 is:

$$\Pr[F_2] = 1 - \frac{1}{m^2} \sum_{\pi, \pi'} \mathrm{Tr}(\mathcal{C}_{\pi\pi'}(\rho)) \tag{85}$$

For Pauli channel $\mathcal{N}(\rho) = \sum_{i=0}^{3} p_i \sigma_i \rho \sigma_i^\dagger$, its Kraus representation is:

$$\mathcal{N}(\rho) = \sum_i E_i \rho E_i^\dagger, \quad E_i = \sqrt{p_i}\sigma_i \tag{86}$$

For a quantum `SWITCH` of n channels, the composite Kraus operator is given by:

$$K^{(\pi)}_{(s_1,\ldots,s_n)} = \sqrt{p_{s_1} \cdots p_{s_n}} \cdot \sigma_{s_{\pi(1)}} \cdots \sigma_{s_{\pi(n)}}, \tag{87}$$

where π is the permutation (either forward or reverse).
Take trace:

$$\mathrm{Tr}(C_{\pi\pi'}(\rho)) = \sum_{s_1,\ldots,s_n} \mathrm{Tr}\left(\sigma_{s_{\pi(1)}} \cdots \sigma_{s_{\pi(n)}} \rho \sigma^\dagger_{s_{\pi'(n)}} \cdots \sigma^\dagger_{s_{\pi'(1)}}\right) \cdot p_{s_1} \cdots p_{s_n} \tag{88}$$

(89)

If $\pi = \pi'$:

$$\begin{aligned}
\mathrm{Tr}(C_{\pi\pi}(\rho)) &= \sum_{s_1,\ldots,s_n} \mathrm{Tr}\left(\sigma_{s_{\pi(1)}} \cdots \sigma_{s_{\pi(n)}} \rho \sigma^\dagger_{s_{\pi(n)}} \cdots \sigma^\dagger_{s_{\pi(1)}}\right) \cdot p_{s_1} \cdots p_{s_n} && (90)\\
&= \sum_{s_1,\ldots,s_n} \mathrm{Tr}\left(\rho \sigma^\dagger_{s_{\pi(n)}} \cdots \sigma^\dagger_{s_{\pi(1)}} \sigma_{s_{\pi(1)}} \cdots \sigma_{s_{\pi(n)}}\right) \cdot p_{s_1} \cdots p_{s_n} && (91)\\
&= \sum_{s_1,\ldots,s_n} \mathrm{Tr}\left(\rho\right) \cdot p_{s_1} \cdots p_{s_n} && (92)\\
&= \sum_{s_1,\ldots,s_n} p_{s_1} \cdots p_{s_n} && (93)\\
&= \left(\sum_{s_1} p_{s_1}\right) \cdots \left(\sum_{s_n} p_{s_n}\right) = 1 \cdots 1 = 1. && (94)
\end{aligned}$$

Proof (cont'd)

If $\pi \neq \pi'$, take $n = 2$ as example, $\pi = (1, 2), \pi' = (2, 1)$:

$$\begin{aligned}
\mathrm{Tr}(C_{\pi\pi'}(\rho)) &= \sum_{s_1, s_2} \mathrm{Tr}(\sigma_{s_1}\sigma_{s_2}\rho\sigma_{s_1}^{\dagger}\sigma_{s_2}^{\dagger}) \cdot p_{s_1}p_{s_2} && (95)\\
&= \sum_{s_1, s_2} \mathrm{Tr}(\sigma_{s_1}\sigma_{s_2}\rho\sigma_{s_1}\sigma_{s_2}) \cdot p_{s_1}p_{s_2} && (96)\\
&= \sum_{s_1, s_2} \mathrm{Tr}(\rho\sigma_{s_1}\sigma_{s_2}\sigma_{s_1}\sigma_{s_2}) \cdot p_{s_1}p_{s_2} && (97)
\end{aligned}$$

Case 1, $s_1 = s_2$:

$$\sigma_{s_1}\sigma_{s_1}\sigma_{s_1}\sigma_{s_1} = I \tag{98}$$

Case 2, $s_1 \neq s_2$:

- If $i \neq j$ and $i, j \neq 0$, use the relationship $\sigma_i\sigma_j = -\sigma_j\sigma_i$:

$$\sigma_{s_1}\sigma_{s_2}\sigma_{s_1}\sigma_{s_2} = -\sigma_{s_1}^2\sigma_{s_2}^2 = -I \tag{99}$$

- If $s_1 = 0$ or $s_2 = 0$ (i.e., $\sigma_0 = I$), the product is I.

In all cases, the outcome is independent of ρ.

Structural Characterization of $\mathcal{P}_n = 0$

[Wu et al., 2025b] proved the following proposition that characterizes the condition for $\mathcal{P}_n = 0$:

Proposition 2

The quantity $\mathcal{P}_n$ is zero if and only if:

1. n is even and the Kraus operators of $\mathcal{N}$ are commutative.
2. n is odd and the number of Kraus operators of $\mathcal{N}$ is not more than 2.

Degradable Channels

- For general n, $\mathcal{P}_n = 0$ only if the **Choi rank** of $\mathcal{N} \leq 2$, and such Pauli channels are shown to be either **degradable or antidegradable**.
- A quantum channel has a Stinespring representation:

$$\mathcal{N}(\rho) = \mathrm{Tr}_E\, U_{\mathcal{N}}(\rho) \tag{100}$$

where E is the environment, $\mathcal{N} : \mathcal{H}_A \to \mathcal{H}_B, U_{\mathcal{N}} : \mathcal{H}_A \to \mathcal{H}_B \otimes \mathcal{H}_E$.
- The complementary channel $\mathcal{N}^c : \mathcal{H}_A \to \mathcal{H}_E$ is defined by

$$\mathcal{N}^c(\rho) = \mathrm{Tr}_B\, U_{\mathcal{N}}(\rho). \tag{101}$$

- A channel $\mathcal{N}$ is degradable when there exists a CPTP map $\mathcal{T} : \mathcal{H}_B \to \mathcal{H}_E$ such that:

$$\mathcal{N}^c = \mathcal{T} \circ \mathcal{N}. \tag{102}$$

- Anti-degradable channel: a channel whose complement is degradable, i.e. there exists a CPTP map $\mathcal{S} : \mathcal{H}_E \to \mathcal{H}_B$ such that $\mathcal{N} = \mathcal{S} \circ \mathcal{N}^c$.

$\mathcal{P}_n = 0$ Lies on a Measure-Zero Set

- The set of Pauli channels is parametrized by the three-dimensional simplex of probability vectors in $\mathbb{R}^4$, as shown in Figure 4
- [Wu et al., 2025b] also showed that the subset satisfying $\mathcal{P}_n = 0$ lies in the edges of this simplex, which is a **measure-zero set**.
- Thus, $\mathcal{P}_n > 0$ almost surely.

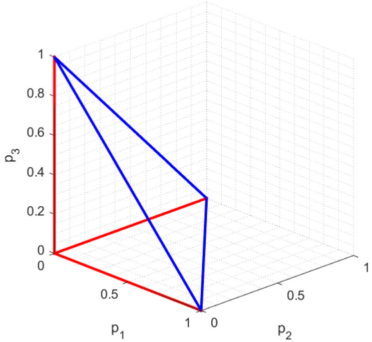

Figure 4: Simplex Representing Pauli Channels. Source: [Wu et al., 2025a]

Explicit Decomposition of $\mathcal{S}^n$ via $\mathcal{P}_n$

- Now, we establish the connection between $\mathcal{P}_n$ and δ_f when f is either the classical capacity C or the coherent information I_c
- [Wu et al., 2025b] proved the following expression for $\mathcal{S}^n$:

$$\mathcal{S}^n(\rho) = (1 - \mathcal{P}_n)\Phi_+(\rho) \otimes \omega + \mathcal{P}_n \Phi_-(\rho) \otimes (I - \omega), \tag{103}$$

where $\omega = |+\rangle\langle+|$ and $\Phi_\pm$ are two Pauli channels.

- Take partial trace we have:

$$\begin{aligned} \mathrm{Tr}_C[\mathcal{S}^n(\rho)] &= (1 - \mathcal{P}_n)\Phi_+(\rho) \cdot \mathrm{Tr}(\omega) + \mathcal{P}_n \Phi_-(\rho) \cdot \mathrm{Tr}(1 - \omega) && (104) \\ &= (1 - \mathcal{P}_n)\Phi_+(\rho) + \mathcal{P}_n \Phi_-(\rho) = \mathcal{N}^n(\rho) && (105) \end{aligned}$$

where we used the fact that $\mathrm{Tr}(\omega) = \mathrm{Tr}(1 - \omega) = 1$. Thus,

$$\mathcal{N}^n = (1 - \mathcal{P}_n)\Phi_+ + \mathcal{P}_n \Phi_- \tag{106}$$

Classical Capacity of $\mathcal{S}^n$

[Wu et al., 2025b] proved that the classical capacity C of $\mathcal{S}^n$ takes a similar form:

Theorem 3

Let $\Phi_\pm$ be as in (103). Then

$$C(\mathcal{S}^n) = (1 - \mathcal{P}_n)C(\Phi_+) + \mathcal{P}_n C(\Phi_-). \tag{107}$$

Deriving Classical Causal Gain for Qubit Unital Channels

- We now use Theorem 3 to derive an explicit expression for the **classical causal gain** δ_C.
- Qubit unital channel are channels that satisfy the following conditions:
  - **Action on Qubit**: The input and output quantum states are density matrices in a two-dimensional Hilbert space (i.e., 2×2 matrices).
  - **Unital Property**: The channel maps the identity matrix I to itself, i.e.,

    $$\mathcal{M}(I) = I.$$

    For a single qubit, I is the 2×2 identity matrix.

Classical Capacity of Qubit Unital Channels

- Classical capacity of qubit unital Channels can be computed via the formula

$$C(\mathcal{M}) = \chi(\mathcal{M}) = 1 - H^{\min}(\mathcal{M}) \tag{108}$$

  where $\chi(\mathcal{M})$ is the Holevo information, $H^{\min}(M) = \min_\rho H(\mathcal{M}(\rho))$ is the minimum output von Neumann entropy of the channel. More general result can be found in [Müller-Hermes, 2021].

- For qubit unital Channels, $\mathcal{M}$'s eigenvalues α_i (satisfying $\mathcal{M}(A) = \alpha_i A$) determine the minimum output entropy:

$$H^{\min}(\mathcal{M}) = h(\alpha), \quad \alpha = \max_i |\alpha_i|, \tag{109}$$

  where $h(x)$ is the binary entropy function $H_b\left(\frac{1+x}{2}\right) = -\frac{1+x}{2}\log\frac{1+x}{2} - \frac{1-x}{2}\log\frac{1-x}{2}$, with domain $[-1, 1]$.

Range of Eigenvalues of Pauli Channels

- Pauli channel $\mathcal{N}(\rho) = \sum_{i=0}^{3} p_i \sigma_i \rho \sigma_i^\dagger$ is Pauli diagonal:

$$\mathcal{N}(\sigma_j) = \sum_{i=0}^{3} p_i \sigma_i \sigma_j \sigma_i = \lambda_j \sigma_j, \quad \text{for } j = 0, 1, 2, 3 \tag{110}$$

  where λ_j is the jth eigenvalue.
- Apply the following relationship:

$$\sigma_i \sigma_j \sigma_i = \begin{cases} \sigma_j, & \text{if } \sigma_i \text{ commutes with } \sigma_j, \\ -\sigma_j, & \text{if } \sigma_i \text{ anti-commutes with } \sigma_j. \end{cases} \tag{111}$$

- Define $\chi_{ij} = \pm 1$ to represent $\sigma_i \sigma_j \sigma_i = \chi_{ij} \sigma_j$, we have:

$$\mathcal{N}(\sigma_j) = \left(\sum_{i=0}^{3} p_i \chi_{ij} \right) \sigma_j \tag{112}$$

Range of Eigenvalues of Pauli Channels (cont'd)

- Finally we have the eigenvalues:

$$\lambda_j = \sum_{i=0}^{3} p_i \cdot \chi_{ij}, \quad \text{where } \chi_{ij} = \begin{cases} 1 & \text{if } [\sigma_i, \sigma_j] = 0 \\ -1 & \text{if } \{\sigma_i, \sigma_j\} = 0 \end{cases} \tag{113}$$

- For $\sigma_0 = I : \mathcal{N}(I) = I \Rightarrow \lambda_0 = 1$.
- For $\sigma_1 = X$:

$$\sigma_i X \sigma_i = \begin{cases} X & i = 0, 1 \\ -X & i = 2, 3 \end{cases} \Rightarrow \lambda_1 = p_0 + p_1 - p_2 - p_3 \tag{114}$$

- For $\sigma_2 = Y$:

$$\sigma_i Y \sigma_i = \begin{cases} Y & i = 0, 2 \\ -Y & i = 1, 3 \end{cases} \Rightarrow \lambda_2 = p_0 + p_2 - p_1 - p_3 \tag{115}$$

- For $\sigma_3 = Z$:

$$\sigma_i Z \sigma_i = \begin{cases} Z & i = 0, 3 \\ -Z & i = 1, 2 \end{cases} \Rightarrow \lambda_3 = p_0 + p_3 - p_1 - p_2 \tag{116}$$

Range of Eigenvalues of Pauli Channels (cont'd)

- Take λ_1 for example. Since $\sum_{i=0}^{3} p_i = 1$, we have $\lambda_1 = 2(p_0 + p_1) - 1$. Thus:

$$-1 \leq \lambda_1 \leq 1 \tag{117}$$

where $\lambda_1 = 1$ iff $p_0 + p_1 = 1$, $\lambda_1 = -1$ iff $p_2 + p_3 = 1$.

- Similarly, we have $-1 \leq \lambda_2 \leq 1, -1 \leq \lambda_3 \leq 1$. As a result:

$$\lambda_j \in [-1, 1] \quad \forall j \tag{118}$$

- Set γ, μ and ν to be the maximum of absolute values of eigenvalues of $\mathcal{N}^n$, Φ_+ and Φ_- respectively.
- Assume the eigenvalues of Φ_+ are $\{\alpha_i\}$ and those of Φ_- are $\{\beta_i\}$. Then the eigenvalues of $\mathcal{N}^n$ are given by:

$$\lambda_i = (1 - \mathcal{P}_n)\alpha_i + \mathcal{P}_n\beta_i. \tag{119}$$

- γ satisfies:

$$\gamma = \max_i |\lambda_i| \leq (1 - \mathcal{P}_n) \max_i |\alpha_i| + \mathcal{P}_n \max_i |\beta_i| \leq 1. \tag{120}$$

Explicit Formula for Classical Causal Gain

- The classical capacity of the effective channel is

$$C(\mathcal{S}^n) = (1-\mathcal{P}_n)C(\Phi_+) + \mathcal{P}_n C(\Phi_-) \tag{121}$$
$$= (1-\mathcal{P}_n)(1-h(\mu)) + \mathcal{P}_n(1-h(\nu))) \tag{122}$$
$$= 1-(1-\mathcal{P}_n)h(\mu) - \mathcal{P}_n h(\nu) \tag{123}$$

- The classical causal gain is then

$$\delta_C = C(\mathcal{S}^n) - C(\mathcal{N}^n) \tag{124}$$
$$= C(\mathcal{S}^n) - (1-h(\gamma))) \tag{125}$$
$$= h(\gamma) - (1-\mathcal{P}_n)h(\mu) - \mathcal{P}_n h(\nu) \tag{126}$$

Sufficiency Proof of the Conjecture

- We will prove the sufficiency part of the conjecture ($\mathcal{P}_n > 0 \Rightarrow \delta_C > 0$) holds for Pauli channels except the completely depolarizing channel with forward and backward orders.
- Combining (120), and the concavity and monotonicity of $h(x)$, we have:

$$h(\gamma) \geq h\left((1-\mathcal{P}_n)\mu + \mathcal{P}_n\nu\right) \geq (1-\mathcal{P}_n)h(\mu) + \mathcal{P}_n h(\nu). \qquad (127)$$

Therefore,

$$\delta_C = h(\gamma) - (1-\mathcal{P}_n)h(\mu) - \mathcal{P}_n h(\nu) \geq 0. \qquad (128)$$

The equality holds iff $\mu = \nu$.

Important Note

- When computing $\alpha = \max_i |\alpha_i|$ for Φ_+ and Φ_-, we have ruled out the trivial eigenvalue $\lambda_0 = 1$.
- I is an **indistinguishable** input state: the completely mixed state carries no information.
- The true contributors to information transmission are the non-trivial actions of operators like X, Y, Z.
- Therefore, the eigenvalues actually used for $\alpha = \max_i |\alpha_i|$ are those corresponding to $\sigma_1, \sigma_2, \sigma_3$.

Coherent Information

- Next, we will discuss the communication of **coherent information**.
- For a channel $\mathcal{N}$ and an input state ρ, the coherent information is defined as:

$$I_c(\rho, \mathcal{N}) = S(\mathcal{N}(\rho)) - S(\mathcal{N}^c(\rho)), \tag{129}$$

  where $S(\cdot)$ is the von Neumann entropy and $\mathcal{N}^c$ is the complementary channel of $\mathcal{N}$.
- $S(\mathcal{N}(\rho))$: The entropy of the output state, quantifying the information obtained by the receiver.
- $S(\mathcal{N}^c(\rho))$: The entropy of the environment state, quantifying the amount of quantum information leaked to the environment.
- Quantum information transmission is enabled when $I_c > 0$.

Quantum Causal Gain in Coherent Information

- For a Pauli channel $\mathcal{N}$, the coherent information attains a maximum on the completely mixed state $\frac{I}{2}$, which is called the **hashing bound**:

$$I_c\left(\frac{I}{2},\mathcal{N}\right) = S\left(\mathcal{N}\left(\frac{I}{2}\right)\right) - S\left(\mathcal{N}^c\left(\frac{I}{2}\right)\right) = 1 - H(\vec{p}) \tag{130}$$

where $\vec{p} = (p_0, p_1, p_2, p_3)$ is the probability vector associated with the Pauli channel N.

- Since $\Phi_\pm$ are both Pauli channels, it follows from (103) that the coherent information of $\mathcal{S}^n$ also attains a maximum on $\frac{I}{2}$:

$$I_c(\mathcal{S}^n) = 1 - (1 - \mathcal{P}_n)H(\vec{s}) - \mathcal{P}_n H(\vec{t}) \tag{131}$$

- So we have the causal gain:

$$\delta_I = H(\vec{q}) - (1 - \mathcal{P}_n)H(\vec{s}) - \mathcal{P}_n H(\vec{t}) \tag{132}$$

- Similar to the case of δ_c, it is then straightforward to deduce that $\mathcal{P}_n > 0$ is a sufficient condition for $\delta_I > 0$

Necessary and Sufficient Condition for Causal Gain

Based on previous results, [Wu et al., 2025b] proved the following theorem:

Theorem 4

Let f be the classical capacity or coherent information, and let δ_f be the causal gain associated with forward and backward orders. Then for all Pauli channels (except the completely depolarizing channel when n is odd), the condition $\mathcal{P}_n > 0 \iff \delta_f > 0$ holds.

References I

Chiribella, G., Wilson, M., and Chau, H. (2021a).
Quantum and classical data transmission through completely depolarizing channels in a superposition of cyclic orders.
Physical review letters, 127(19):190502.

Chiribella, G., Wilson, M., and Chau, H. (2021b).
Supplemental material for quantum and classical data transmission through completely depolarizing channels in a superposition of cyclic orders.
http://link.aps.org/supplemental/10.1103/PhysRevLett.127.190502.

Cubitt, T. S., Ruskai, M. B., and Smith, G. (2008).
The structure of degradable quantum channels.
Journal of Mathematical Physics, 49(10).

References II

Das, D. and Bandyopadhyay, S. (2022).
Quantum communication using a quantum switch of quantum switches.
Proceedings of the Royal Society A, 478(2266):20220231.

Dong, Q., Quintino, M. T., Soeda, A., and Murao, M. (2023).
The quantum switch is uniquely defined by its action on unitary operations.
Quantum, 7:1169.

Ebler, D., Salek, S., and Chiribella, G. (2018).
Enhanced communication with the assistance of indefinite causal order.
Physical review letters, 120(12):120502.

References III

Li, Y. (2010).
Degradable quantum channels.
https://quantum.phys.cmu.edu/QIP/DegChannels.pdf.
Department of Physics, Carnegie-Mellon University, Pittsburgh, PA.

Mendl, C. (2007).
Unital quantum channels.
http://christian.mendl.net/science/talks/ringberg%202007.pdf.
Presentation, 11 December 2007.

References IV

Müller-Hermes, A. (2021).
Lecture 13: Additivity problems.
https://www.uio.no/studier/emner/matnat/math/MAT4430/v22/lecture-notes/lecture13.pdf.
Lecture notes for the course Quantum Information Theory (MAT4430), University of Oslo, Spring 2021.

O'Donnell, R. and Wright, J. (2015).
Lecture 18: Quantum information theory and holevo's bound.
https://www.cs.cmu.edu/~odonnell/quantum15/.
Course 15-859BB: Quantum Computation and Information, Carnegie Mellon University.

References V

Quantum Zeitgeist (2024).
Quantum switch characterized for first time, paving way for advanced quantum technology.
https://quantumzeitgeist.com/quantum-switch-characterized-for-first-time-paving-way-for-

Rana, F. (2024).
Quantum information theory.
https://courses.cit.cornell.edu/ece531/Lectures/Lectures.htm.
ECE 5310 Quantum Optics, Slides 16.

Wikipedia contributors (2025a).
Hilbert–Schmidt operator — Wikipedia, The Free Encyclopedia.
https://en.wikipedia.org/wiki/HilbertâĂŞSchmidt_operator.

References VI

Wikipedia contributors (2025b).
Qutrit — Wikipedia, The Free Encyclopedia.
https://en.wikipedia.org/wiki/Qutrit.

Wolf, M. M. and Pérez-García, D. (2007).
Quantum capacities of channels with small environment.
Physical Review A—Atomic, Molecular, and Optical Physics, 75(1):012303.

Wu, Z., Fullwood, J., Ma, Z., Zhou, S.-Q., Zhao, Q., and Chiribella, G. (2025a).
General communication enhancement via the quantum switch.
Physical Review A, 111(1):012605.

References VII

Wu, Z., Fullwood, J., Ma, Z., Zhou, S.-Q., Zhao, Q., and Chiribella, G. (2025b).
Supplemental material for general communication enhancement via the quantum switch.
http://link.aps.org/supplemental/10.1103/PhysRevA.111.012605.

Thank you!
Any questions?