

Quantum SWITCH for Communication Enhancement

Siqi Yao

SCHOOL OF DATA SCIENCE

August 7, 2025

Table of Contents

- 1 Quantum SWITCH
 - Brief Introduction
- 2 Quantum SWITCH for Capacity Enhancement of Specific Channels
 - Quantum Information Theory Preliminaries
 - Completely Depolarizing Channels-Classical Information
 - Completely Depolarizing Channels-Quantum Information
- 3 Quantum SWITCH for Capacity Enhancement of Generic Channels
 - Key Quantity and Properties
 - Conjecture on Conditions of Communication Enhancement
 - Verification of Conjecture on Pauli Channels
 - Verification of Conjecture on Qudit Depolarizing Channels
- 4 References

Table of Contents

- 1 Quantum SWITCH
 - Brief Introduction
- 2 Quantum SWITCH for Capacity Enhancement of Specific Channels
 - Quantum Information Theory Preliminaries
 - Completely Depolarizing Channels-Classical Information
 - Completely Depolarizing Channels-Quantum Information
- 3 Quantum SWITCH for Capacity Enhancement of Generic Channels
 - Key Quantity and Properties
 - Conjecture on Conditions of Communication Enhancement
 - Verification of Conjecture on Pauli Channels
 - Verification of Conjecture on Qudit Depolarizing Channels
- 4 References

- 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

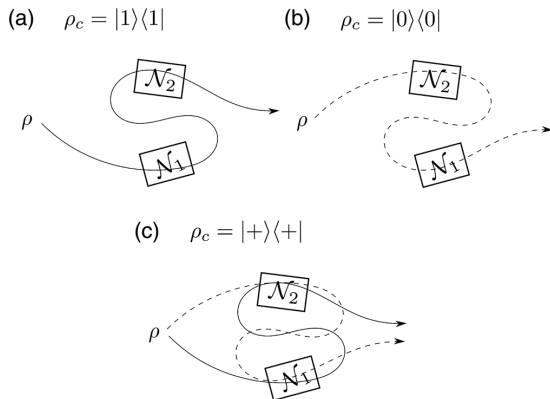


Figure 1: Superposition of Orders. Source: [Ebler et al., 2018]

Fixed Order Channels

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

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

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:

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

- 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:

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

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

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

- 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

- 1 Quantum SWITCH
 - Brief Introduction
- 2 Quantum SWITCH for Capacity Enhancement of Specific Channels
 - Quantum Information Theory Preliminaries
 - Completely Depolarizing Channels-Classical Information
 - Completely Depolarizing Channels-Quantum Information
- 3 Quantum SWITCH for Capacity Enhancement of Generic Channels
 - Key Quantity and Properties
 - Conjecture on Conditions of Communication Enhancement
 - Verification of Conjecture on Pauli Channels
 - Verification of Conjecture on Qudit Depolarizing Channels
- 4 References

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 $\text{Tr}(E_y \sigma_x)$.
- Bob tries to infer X from Y . Note: X, Y are two **classical** probability distributions.

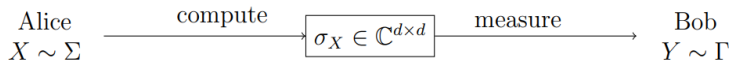


Figure 2: Communication Scheme. Source: [O'Donnell and Wright, 2015]

- Bob sees the mixed state:

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

- Alice sees:

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

- State of the joint mixed system is:

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

- 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

$$I(X; Y) = H(X) + H(Y) - H(X, Y), \quad (8)$$

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:

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

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:

$$I(\rho_A; \rho_B) = H(\rho_A) + H(\rho_B) - H(\rho). \quad (10)$$

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:

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

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

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

$$I_{\text{acc}}(\sigma, p) \leq \chi(\sigma, p). \quad (12)$$

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

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

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:

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

Compute von Neumann entropy:

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

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

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

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:

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

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:

$$H(\rho_{AB}) = - \sum_{i,k} p_i r_{ik} \log(p_i r_{ik}) \quad (18)$$

$$= - \sum_{i,k} p_i r_{ik} (\log p_i + \log r_{ik}) \quad (19)$$

$$= - \sum_i p_i \log p_i \left(\sum_k r_{ik} \right) - \sum_i p_i \sum_k r_{ik} \log r_{ik} \quad (20)$$

$$= - \sum_i p_i \log p_i + \sum_i p_i S(\rho_i) \quad (21)$$

$$= H(\{p_i\}) + \sum_i p_i S(\rho_i) \quad (22)$$

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

HSW Theorem

The classical channel capacity C of a quantum channel is given by the HSW theorem:

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 the following regularized expression:

$$C(\mathcal{E}) = \lim_{n \rightarrow \infty} \frac{1}{n} \chi^*(\mathcal{E}^{\otimes n}) \quad (23)$$

where $\chi^*(\mathcal{E})$ is the **maximum Holevo information**:

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

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

ρ_i is the input quantum state, p_i is the probability distribution of the quantum state.

- Unlike the classical case, $C(\mathcal{E})$ is not given by a single-letter formula involving a single copy of the \mathcal{E} . Instead it is given by a **regularization of the maximum Holevo information**, which involves many copies of \mathcal{E} . In general, it is unclear how to compute such a formula, and a simpler expression for $C(\mathcal{E})$ is not yet known.
- $\mathcal{E}^{\otimes k} : \mathcal{B}(\mathcal{H}_A^{\otimes k}) \rightarrow \mathcal{B}(\mathcal{H}_B^{\otimes k})$ means composing k independent copies of \mathcal{E} in parallel. For any input density operator $\rho^{(k)} \in \mathcal{B}(\mathcal{H}_A^{\otimes k})$, the output is:

$$\mathcal{E}^{\otimes k}(\rho^{(k)}) = (\mathcal{E} \otimes \mathcal{E} \otimes \cdots \otimes \mathcal{E})(\rho^{(k)}) \quad (26)$$

Table of Contents

- 1 Quantum SWITCH
 - Brief Introduction
- 2 Quantum SWITCH for Capacity Enhancement of Specific Channels
 - Quantum Information Theory Preliminaries
 - Completely Depolarizing Channels-Classical Information
 - Completely Depolarizing Channels-Quantum Information
- 3 Quantum SWITCH for Capacity Enhancement of Generic Channels
 - Key Quantity and Properties
 - Conjecture on Conditions of Communication Enhancement
 - Verification of Conjecture on Pauli Channels
 - Verification of Conjecture on Qudit Depolarizing Channels
- 4 References

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

$$\sigma = \frac{1}{2}\mathcal{N}(|\phi\rangle\langle\phi|) + \frac{1}{2}\mathcal{N}(|\psi\rangle\langle\psi|) \quad (27)$$

- So the Holevo information is

$$\chi(\mathcal{N}) = S(\sigma) - \left(\frac{1}{2}S(\mathcal{N}(|\phi\rangle\langle\phi|)) + \frac{1}{2}S(\mathcal{N}(|\psi\rangle\langle\psi|))\right) > 0 \quad (28)$$

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 ($\text{Tr}[U_i^\dagger U_j] = d \cdot \delta_{ij}$) unitary operators U_i , such that its action on a state ρ is

$$\mathcal{N}^D(\rho) = \frac{1}{d^2} \sum_{i=1}^{d^2} U_i \rho U_i^\dagger = \text{Tr}[\rho] \frac{I}{d}. \quad (29)$$

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:

$$W_{ij} = \frac{1}{d^2} (U_i U_j \otimes |0\rangle\langle 0|_c + U_j U_i \otimes |1\rangle\langle 1|_c) \quad (30)$$

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

$$\mathcal{S}(\mathcal{N}^D, \mathcal{N}^D, \rho_c)(\rho) = \frac{1}{d^4} \sum_{i,j} \left(p|0\rangle\langle 0|_c \otimes U_i U_j \rho U_j^\dagger U_i^\dagger + (1-p)|1\rangle\langle 1|_c \otimes U_j U_i \rho U_i^\dagger U_j^\dagger \right. \\ \left. + \sqrt{p(1-p)}|0\rangle\langle 1|_c \otimes U_i U_j \rho U_i^\dagger U_j^\dagger + \sqrt{p(1-p)}|1\rangle\langle 0|_c \otimes U_j U_i \rho U_j^\dagger U_i^\dagger \right) \quad (31)$$

$$= p|0\rangle\langle 0|_c \otimes \frac{I}{d} + (1-p)|1\rangle\langle 1|_c \otimes \frac{I}{d} \quad (32)$$

$$+ \sqrt{p(1-p)} \frac{|0\rangle\langle 1|_c}{d^2} \otimes \sum_j \text{Tr}[U_j \rho] \frac{U_j^\dagger}{d} \\ + \sqrt{p(1-p)} \frac{|1\rangle\langle 0|_c}{d^2} \otimes \sum_j \text{Tr}[\rho U_j^\dagger] \frac{U_j}{d}. \\ = (p|0\rangle\langle 0|_c + (1-p)|1\rangle\langle 1|_c) \otimes \frac{I}{d} + \sqrt{p(1-p)}(|0\rangle\langle 1|_c + |1\rangle\langle 0|_c) \otimes \frac{\rho}{d^2}. \quad (33)$$

- (31) results from (3).
- $\sum_j \text{Tr}[U_j \rho] \frac{U_j^\dagger}{d}$ results from viewing $U_j \rho$ as a single state and applying (29).
- (33) follows from the fact that $\{U_j\}$ forms an orthonormal basis for the set of $d \times d$ matrices, i.e., $\rho = \sum_{j=1}^{d^2} \text{Tr}[U_j \rho] \frac{U_j^\dagger}{d}$. 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

- 1 Quantum SWITCH
 - Brief Introduction
- 2 Quantum SWITCH for Capacity Enhancement of Specific Channels
 - Quantum Information Theory Preliminaries
 - Completely Depolarizing Channels-Classical Information
 - Completely Depolarizing Channels-Quantum Information
- 3 Quantum SWITCH for Capacity Enhancement of Generic Channels
 - Key Quantity and Properties
 - Conjecture on Conditions of Communication Enhancement
 - Verification of Conjecture on Pauli Channels
 - Verification of Conjecture on Qudit Depolarizing Channels
- 4 References

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

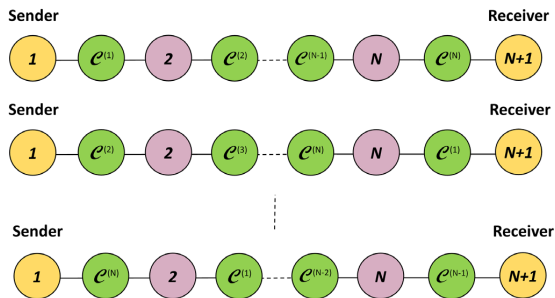


Figure 3: Cyclic Permutations. Source: [Chiribella et al., 2021a]

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 $\mathcal{E} \rightarrow \mathcal{F} \rightarrow \mathcal{G}$
 - $|\pi_1\rangle$ denotes the order $\mathcal{F} \rightarrow \mathcal{G} \rightarrow \mathcal{E}$
 - $|\pi_2\rangle$ denotes the order $\mathcal{G} \rightarrow \mathcal{E} \rightarrow \mathcal{F}$
- Let the control state be the equal superposition $|\omega\rangle = \frac{1}{\sqrt{3}} (|\pi_0\rangle + |\pi_1\rangle + |\pi_2\rangle)$ and let \mathbf{S} denote the set of permutations. The density matrix becomes:

$$\omega = |\omega\rangle\langle\omega| = \frac{1}{3} \sum_{\pi, \pi' \in \mathbf{S}} |\pi\rangle\langle\pi'| \quad (34)$$

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

$$K_{ijk} = K_{ijk}^{(\pi_0)} \otimes |\pi_0\rangle\langle\pi_0| + K_{ijk}^{(\pi_1)} \otimes |\pi_1\rangle\langle\pi_1| + K_{ijk}^{(\pi_2)} \otimes |\pi_2\rangle\langle\pi_2| \quad (35)$$

$$= \sum_{\pi} K_{ijk}^{(\pi)} \otimes |\pi\rangle\langle\pi| \quad (36)$$

- So the channel becomes:

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

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

$$= \sum_{i,j,k} \sum_{\pi,\pi'} K_{ijk}^{(\pi)} \rho K_{ijk}^{(\pi')\dagger} \otimes \frac{1}{3} |\pi\rangle\langle\pi'| \quad (39)$$

$$= \sum_{\pi,\pi'} \mathcal{C}_{\pi\pi'}(\rho) \otimes \omega_{\pi,\pi'} |\pi\rangle\langle\pi'| := \mathcal{C}_{\text{eff}}(\rho) \quad (40)$$

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:

$\mathcal{C}_{\pi\pi'}(\rho) = \sum_{s_1,\dots,s_N} K_{(s_1,\dots,s_N)}^{(\pi)} \rho K_{(s_1,\dots,s_N)}^{(\pi')\dagger}$, 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:

$$\mathcal{C}_{\pi\pi}(\rho) = \frac{I}{d} \quad \text{and} \quad \mathcal{C}_{\pi\pi'}(\rho) = \frac{\rho}{d^2} \quad \forall \pi \neq \pi' \quad (41)$$

- Plug into (40) yields:

$$\mathcal{C}_{\text{eff}}(\rho) = \sum_{\pi} \frac{I}{d} \otimes \frac{1}{N} |\pi\rangle\langle\pi| + \sum_{\pi \neq \pi'} \frac{\rho}{d^2} \otimes \frac{1}{N} |\pi\rangle\langle\pi'| \quad (42)$$

$$= \frac{I}{d} \otimes \frac{I}{N} + \frac{\rho}{Nd^2} \otimes \sum_{\pi \neq \pi'} |\pi\rangle\langle\pi'| \quad (43)$$

$$= \frac{I}{d} \otimes \frac{I}{N} + \frac{\rho}{Nd^2} \otimes (N|\omega\rangle\langle\omega| - I), \quad (44)$$

where (44) 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 (44) we have:

$$\mathcal{C}_{\text{eff}}(\rho) = \mathcal{E}_0(\rho) \otimes (1-p)\rho_0 + \mathcal{E}_1(\rho) \otimes p\rho_1, \quad (45)$$

where $\rho_0 := |\omega\rangle\langle\omega|$ and $\rho_1 := \frac{I - |\omega\rangle\langle\omega|}{N-1}$, $p := \frac{(N-1)(d^2-1)}{Nd^2}$. $\mathcal{E}_0, \mathcal{E}_1$ are quantum channels defined by

$$\mathcal{E}_0(\rho) := \frac{N-1}{N-1+d^2}\rho + \frac{d^2}{N-1+d^2}\frac{I}{d} \quad (46)$$

and

$$\mathcal{E}_1(\rho) := \frac{d^2}{d^2-1}\frac{I}{d} - \frac{1}{d^2-1}\rho \quad (47)$$

- \mathcal{C}_{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 .

- \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 $\frac{d^2}{N+d^2-1}$.
- $\lim_{N \rightarrow \infty} \frac{d^2}{N+d^2-1} = 0 \Rightarrow \lim_{N \rightarrow \infty} \mathcal{E}_0(\rho) = \rho$. 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

- 1 Quantum SWITCH
 - Brief Introduction
- 2 Quantum SWITCH for Capacity Enhancement of Specific Channels
 - Quantum Information Theory Preliminaries
 - Completely Depolarizing Channels-Classical Information
 - Completely Depolarizing Channels-Quantum Information
- 3 Quantum SWITCH for Capacity Enhancement of Generic Channels
 - Key Quantity and Properties
 - Conjecture on Conditions of Communication Enhancement
 - Verification of Conjecture on Pauli Channels
 - Verification of Conjecture on Qudit Depolarizing Channels
- 4 References

- 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

$$\mathcal{P}_n = 1 - \frac{1}{m^2} \min_{\rho} \sum_{\pi, \pi' \in \mathbf{S}} \text{Tr}(C_{\pi\pi'}(\rho)) \quad (48)$$

- \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.

Probability of obtaining outcome F_1 can be computed as:

$$\Pr[F_1] = \text{Tr}[(I \otimes F_1) \mathcal{C}_{\text{eff}}(\rho)]. \quad (49)$$

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

$$= \frac{1}{m} \sum_{\pi, \pi'} \text{Tr} [\mathcal{C}_{\pi\pi'}(\rho) \otimes (F_1 |\pi\rangle\langle\pi'|)] \quad (51)$$

$$= \frac{1}{m} \sum_{\pi, \pi'} \text{Tr}(\mathcal{C}_{\pi\pi'}(\rho)) \cdot \text{Tr}(F_1 |\pi\rangle\langle\pi'|) \quad (52)$$

$$= \frac{1}{m^2} \sum_{\pi, \pi'} \text{Tr}(\mathcal{C}_{\pi\pi'}(\rho)) \quad (53)$$

So probability of obtaining outcome F_2 is:

$$\Pr[F_2] = 1 - \Pr[F_1] = 1 - \frac{1}{m^2} \sum_{\pi, \pi'} \text{Tr}(\mathcal{C}_{\pi\pi'}(\rho)) \quad (54)$$

So we have:

$$\mathcal{P}_n = \max_{\rho} \Pr[F_2] = 1 - \min_{\rho} \frac{1}{m^2} \sum_{\pi, \pi'} \text{Tr}(\mathcal{C}_{\pi\pi'}(\rho)) \quad (55)$$

Proposition 1

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

$$K_{(s_1, \dots, s_n)}^{(\pi)} = K_{(s_1, \dots, s_n)}^{(\pi')} \quad (56)$$

for all indices s_1, \dots, 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 $\text{Tr}(C_{\pi\pi}(\rho)) = 1 \quad \forall \pi \in \mathbf{S}$.

Recall:

$$C_{\pi\pi}(\rho) = \sum_{s_1, \dots, s_n} K_{(s_1, \dots, s_n)}^{(\pi)} \rho K_{(s_1, \dots, s_n)}^{(\pi)\dagger} \quad (57)$$

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

$$K_{(s_1, \dots, s_n)}^{(\pi)} := A_{s_1, \dots, s_n}^{(\pi)} = A_{\pi(1), s_1} A_{\pi(2), s_2} \cdots A_{\pi(n), s_n}, \quad (58)$$

$$K_{(s_1, \dots, s_n)}^{(\pi)\dagger} := A_{s_1, \dots, s_n}^{(\pi)\dagger} = A_{\pi(n), s_n}^\dagger A_{\pi(n-1), s_{n-1}}^\dagger \cdots A_{\pi(1), s_1}^\dagger \quad (59)$$

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:

$$\text{Tr}(C_{\pi\pi}(\rho)) = 1 \quad \forall \pi \in \mathbf{S} \quad (60)$$

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

$$\mathrm{Tr}(C_{\pi\pi'}(\rho)) = \sum_{s_1, \dots, s_n} \mathrm{Tr} \left(K_{(s_1, \dots, s_n)}^{(\pi)} \rho K_{(s_1, \dots, s_n)}^{(\pi')\dagger} \right) \quad (61)$$

$$= \sum_{s_1, \dots, s_n} \mathrm{Tr} \left(K_{(s_1, \dots, s_n)}^{(\pi)} \sqrt{\rho} \left(K_{(s_1, \dots, s_n)}^{(\pi')} \sqrt{\rho} \right)^\dagger \right) \quad (62)$$

$$\leq \frac{1}{2} [\mathrm{Tr}(C_{\pi\pi}(\rho)) + \mathrm{Tr}(C_{\pi'\pi'}(\rho))] = 1 \quad (63)$$

where the final equality follows from (60), and the first inequality follows from combining the Cauchy-Schwarz and AM-GM inequalities, which for arbitrary operators A and B yields

$$\mathrm{Tr}(AB^\dagger) \leq \sqrt{\mathrm{Tr}(AA^\dagger)\mathrm{Tr}(BB^\dagger)} \quad (64)$$

$$\leq \frac{1}{2} \left(\mathrm{Tr}(AA^\dagger) + \mathrm{Tr}(BB^\dagger) \right). \quad (65)$$

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):

$$\langle A, B \rangle := \text{Tr}(AB^\dagger) \quad (66)$$

The norm induced by this inner product is:

$$\|A\|_{\text{HS}} := \sqrt{\text{Tr}(AA^\dagger)} \quad (67)$$

Applying the Cauchy-Schwarz inequality, we obtain:

$$|\text{Tr}(AB^\dagger)| \leq \sqrt{\text{Tr}(AA^\dagger)} \cdot \sqrt{\text{Tr}(BB^\dagger)} \quad (68)$$

By setting $A_{(s_1, \dots, s_n)} = K_{(s_1, \dots, s_n)}^{(\pi)} \sqrt{\rho}$, $B_{(s_1, \dots, s_n)} = K_{(s_1, \dots, s_n)}^{(\pi')} \sqrt{\rho}$, and using Cauchy-Schwarz inequality for each pair of $A_{(s_1, \dots, s_n)}$, $B_{(s_1, \dots, s_n)}$, 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

$$\text{Tr}(C_{\pi\pi'}(\rho)) = 1, \quad (69)$$

if and only if

$$K_{(s_1, \dots, s_n)}^{(\pi)} \sqrt{\rho} = K_{(s_1, \dots, s_n)}^{(\pi')} \sqrt{\rho} \quad \forall s_1, \dots, s_n. \quad (70)$$

From the definition of \mathcal{P}_n , it follows that $\mathcal{P}_n = 0$ if and only if (70) 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

$$K_{(s_1, \dots, s_n)}^{(\pi)} = K_{(s_1, \dots, s_n)}^{(\pi')} \quad (71)$$

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

$$C_{s_1}^1 C_{s_2}^2 = C_{s_2}^2 C_{s_1}^1 \quad \text{for all } s_1, s_2, \quad (72)$$

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

Table of Contents

- 1 Quantum SWITCH
 - Brief Introduction
- 2 Quantum SWITCH for Capacity Enhancement of Specific Channels
 - Quantum Information Theory Preliminaries
 - Completely Depolarizing Channels-Classical Information
 - Completely Depolarizing Channels-Quantum Information
- 3 Quantum SWITCH for Capacity Enhancement of Generic Channels
 - Key Quantity and Properties
 - Conjecture on Conditions of Communication Enhancement
 - Verification of Conjecture on Pauli Channels
 - Verification of Conjecture on Qudit Depolarizing Channels
- 4 References

- 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 $i = 1, \dots, n$, where \mathcal{N} is a fixed quantum channel, and let \mathcal{N}^n denote the n -fold composition $\mathcal{N} \circ \dots \circ \mathcal{N}$.

Tracing Out Control Recovers \mathcal{N}^n

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

$$\mathcal{N}^n(\rho) = \mathcal{N} \circ \dots \circ \mathcal{N}(\rho) = \sum_{s_1, \dots, s_n} C_{s_n} \dots C_{s_1} \rho C_{s_1}^\dagger \dots C_{s_n}^\dagger \quad (73)$$

Take partial trace:

$$\text{Tr}_C [\mathcal{S}^n(\rho)] = \sum_{\pi, \pi' \in \mathbf{S}} C_{\pi\pi'}(\rho) \cdot \text{Tr} (\omega_{\pi, \pi'} |\pi\rangle \langle \pi'|) \quad (74)$$

$$= \sum_{\pi, \pi' \in \mathbf{S}} C_{\pi\pi'}(\rho) \cdot \frac{1}{m} \delta_{\pi\pi'} \quad (75)$$

$$= \frac{1}{m} \sum_{\pi \in \mathbf{S}} C_{\pi\pi}(\rho) \quad (76)$$

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

$$\text{Tr}_C [\mathcal{S}^n(\rho)] = \mathcal{N}^n(\rho). \quad (77)$$

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**:

$$f(\mathcal{N}^n) \leq f(\mathcal{S}^n). \quad (78)$$

since taking partial trace is also a CPTP map.

- Capacity measures such as classical capacity, quantum capacity, Holevo information and coherent information all satisfy (78). Thus, we define the associated **causal gain** by

$$\delta_f = f(\mathcal{S}^n) - f(\mathcal{N}^n). \quad (79)$$

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.

Sufficient Condition for $\delta_f = 0$

- $\mathcal{P}_n = 0$ if and only if $K_{(s_1, \dots, s_n)}^{(\pi)} = K_{(s_1, \dots, s_n)}^{(\pi')}$ for all s_1, \dots, 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:

$$\mathcal{S}^n(\rho) = \sum_{\pi, \pi' \in \mathbf{S}} C_{\pi\pi'}(\rho) \otimes \omega_{\pi, \pi'} |\pi\rangle \langle \pi'| \quad (80)$$

$$= \sum_{\pi, \pi' \in \mathbf{S}} \mathcal{N}^n(\rho) \otimes \omega_{\pi, \pi'} |\pi\rangle \langle \pi'| \quad (81)$$

$$= \mathcal{N}^n(\rho) \otimes \left(\sum_{\pi, \pi' \in \mathbf{S}} \omega_{\pi, \pi'} |\pi\rangle \langle \pi'| \right) \quad (82)$$

$$= \mathcal{N}^n(\rho) \otimes \omega \quad (83)$$

- 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

- 1 Quantum SWITCH
 - Brief Introduction
- 2 Quantum SWITCH for Capacity Enhancement of Specific Channels
 - Quantum Information Theory Preliminaries
 - Completely Depolarizing Channels-Classical Information
 - Completely Depolarizing Channels-Quantum Information
- 3 Quantum SWITCH for Capacity Enhancement of Generic Channels
 - Key Quantity and Properties
 - Conjecture on Conditions of Communication Enhancement
 - Verification of Conjecture on Pauli Channels
 - Verification of Conjecture on Qudit Depolarizing Channels
- 4 References

The **Pauli channel** describes a probabilistic mixture of Pauli operations applied to a single qubit (density matrix) ρ :

$$\mathcal{N}(\rho) = \sum_{i=0}^3 p_i \sigma_i \rho \sigma_i^\dagger, \quad \text{with} \quad \sum_i p_i = 1 \quad (84)$$

where:

- $\sigma_0 = I$, $\sigma_1 = X$, $\sigma_2 = Y$, $\sigma_3 = Z$ are the four Pauli operators,
- $p_i \geq 0$ is the probability of applying operation σ_i .
- Completely depolarizing channel is a special case of Pauli channel:
 $\mathcal{N}(\rho) = \frac{1}{4} (I\rho I + X\rho X + Y\rho Y + Z\rho Z) = \frac{I}{2}.$

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

- 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 $\mathcal{N}_n \circ \dots \circ \mathcal{N}_1$ and $\mathcal{N}_1 \circ \dots \circ \mathcal{N}_n$ (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

- 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 ,

$$\text{Tr}((I \otimes F_2)\mathcal{S}^n(\rho)) = \mathcal{P}_n. \quad (85)$$

Recall, probability of obtaining outcome F_2 is:

$$\Pr[F_2] = 1 - \frac{1}{m^2} \sum_{\pi, \pi'} \text{Tr}(\mathcal{C}_{\pi\pi'}(\rho)) \quad (86)$$

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

$$\mathcal{N}(\rho) = \sum_i E_i \rho E_i^\dagger, \quad E_i = \sqrt{p_i} \sigma_i \quad (87)$$

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

$$K_{(s_1, \dots, s_n)}^{(\pi)} = \sqrt{p_{s_1} \cdots p_{s_n}} \cdot \sigma_{s_{\pi(1)}} \cdots \sigma_{s_{\pi(n)}}, \quad (88)$$

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

Take trace:

$$\text{Tr}(\mathcal{C}_{\pi\pi'}(\rho)) = \sum_{s_1, \dots, s_n} \text{Tr} \left(\sigma_{s_{\pi(1)}} \cdots \sigma_{s_{\pi(n)}} \rho \sigma_{s_{\pi'(n)}}^\dagger \cdots \sigma_{s_{\pi'(1)}}^\dagger \right) \cdot p_{s_1} \cdots p_{s_n} \quad (89)$$

$$(90)$$

Proof (cont'd)

If $\pi = \pi'$:

$$\text{Tr}(C_{\pi\pi}(\rho)) = \sum_{s_1, \dots, s_n} \text{Tr} \left(\sigma_{s_{\pi(1)}} \cdots \sigma_{s_{\pi(n)}} \rho \sigma_{s_{\pi(n)}}^\dagger \cdots \sigma_{s_{\pi(1)}}^\dagger \right) \cdot p_{s_1} \cdots p_{s_n} \quad (91)$$

$$= \sum_{s_1, \dots, s_n} \text{Tr} \left(\rho \sigma_{s_{\pi(n)}}^\dagger \cdots \sigma_{s_{\pi(1)}}^\dagger \sigma_{s_{\pi(1)}} \cdots \sigma_{s_{\pi(n)}} \right) \cdot p_{s_1} \cdots p_{s_n} \quad (92)$$

$$= \sum_{s_1, \dots, s_n} \text{Tr}(\rho) \cdot p_{s_1} \cdots p_{s_n} \quad (93)$$

$$= \sum_{s_1, \dots, s_n} p_{s_1} \cdots p_{s_n} \quad (94)$$

$$= \left(\sum_{s_1} p_{s_1} \right) \cdots \left(\sum_{s_n} p_{s_n} \right) = 1 \cdots 1 = 1. \quad (95)$$

Proof (cont'd)

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

$$\text{Tr}(C_{\pi\pi'}(\rho)) = \sum_{s_1, s_2} \text{Tr}(\sigma_{s_1} \sigma_{s_2} \rho \sigma_{s_1}^\dagger \sigma_{s_2}^\dagger) \cdot p_{s_1} p_{s_2} \quad (96)$$

$$= \sum_{s_1, s_2} \text{Tr}(\sigma_{s_1} \sigma_{s_2} \rho \sigma_{s_1} \sigma_{s_2}) \cdot p_{s_1} p_{s_2} \quad (97)$$

$$= \sum_{s_1, s_2} \text{Tr}(\rho \sigma_{s_1} \sigma_{s_2} \sigma_{s_1} \sigma_{s_2}) \cdot p_{s_1} p_{s_2} \quad (98)$$

Case 1, $s_1 = s_2$:

$$\sigma_{s_1} \sigma_{s_1} \sigma_{s_1} \sigma_{s_1} = I \quad (99)$$

Case 2, $s_1 \neq s_2$:

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

$$\sigma_{s_1} \sigma_{s_2} \sigma_{s_1} \sigma_{s_2} = -\sigma_{s_1}^2 \sigma_{s_2}^2 = -I \quad (100)$$

- 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:

$$\mathcal{N}(\rho) = \text{Tr}_E U_{\mathcal{N}}(\rho) \quad (101)$$

where E is the environment, $\mathcal{N} : \mathcal{H}_A \rightarrow \mathcal{H}_B$, $U_{\mathcal{N}} : \mathcal{H}_A \rightarrow \mathcal{H}_B \otimes \mathcal{H}_E$.

- The complementary channel $\mathcal{N}^c : \mathcal{H}_A \rightarrow \mathcal{H}_E$ is defined by

$$\mathcal{N}^c(\rho) = \text{Tr}_B U_{\mathcal{N}}(\rho). \quad (102)$$

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

$$\mathcal{N}^c = \mathcal{T} \circ \mathcal{N}. \quad (103)$$

- Anti-degradable channel: a channel whose complement is degradable, i.e. there exists a CPTP map $\mathcal{S} : \mathcal{H}_E \rightarrow \mathcal{H}_B$ 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.

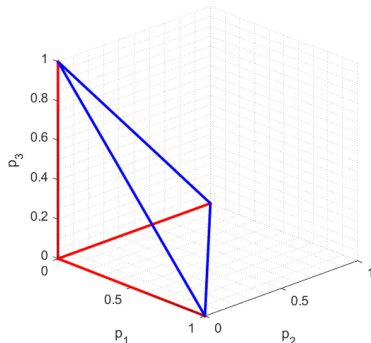


Figure 4: Simplex Representing Pauli Channels. Source: [Wu et al., 2025a]

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 :

$$\mathcal{S}^n(\rho) = (1 - \mathcal{P}_n)\Phi_+(\rho) \otimes \omega + \mathcal{P}_n\Phi_-(\rho) \otimes (I - \omega), \quad (104)$$

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

- Take partial trace we have:

$$\text{Tr}_C[\mathcal{S}^n(\rho)] = (1 - \mathcal{P}_n)\Phi_+(\rho) \cdot \text{Tr}(\omega) + \mathcal{P}_n\Phi_-(\rho) \cdot \text{Tr}(1 - \omega) \quad (105)$$

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

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

$$\mathcal{N}^n = (1 - \mathcal{P}_n)\Phi_+ + \mathcal{P}_n\Phi_- \quad (107)$$

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

Theorem 3

Let Φ_{\pm} be as in (104). Then

$$C(\mathcal{S}^n) = (1 - \mathcal{P}_n)C(\Phi_+) + \mathcal{P}_n C(\Phi_-). \quad (108)$$

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.

- **Examples:** Pauli channels, completely depolarizing channel, unital channel.

Classical Capacity of Qubit Unital Channels

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

$$C(\mathcal{M}) = \chi(\mathcal{M}) = 1 - H^{\min}(\mathcal{M}) \quad (109)$$

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

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

$$H^{\min}(\mathcal{M}) = h(\alpha), \quad \alpha = \max_i |\alpha_i|, \quad (110)$$

where $h(x)$ is the binary entropy function

$$H_b\left(\frac{1+x}{2}\right) = -\frac{1+x}{2} \log \frac{1+x}{2} - \frac{1-x}{2} \log \frac{1-x}{2}, \text{ with domain } [-1, 1].$$

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:

$$\mathcal{N}(\sigma_j) = \sum_{i=0}^3 p_i \sigma_i \sigma_j \sigma_i = \lambda_j \sigma_j, \quad \text{for } j = 0, 1, 2, 3 \quad (111)$$

where λ_j is the j th eigenvalue.

- Apply the following relationship:

$$\sigma_i \sigma_j \sigma_i = \begin{cases} \sigma_j, & \text{if } \sigma_i \text{ commutes with } \sigma_j, \\ -\sigma_j, & \text{if } \sigma_i \text{ anti-commutes with } \sigma_j. \end{cases} \quad (112)$$

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

$$\mathcal{N}(\sigma_j) = \left(\sum_{i=0}^3 p_i \chi_{ij} \right) \sigma_j \quad (113)$$

Range of Eigenvalues of Pauli Channels (cont'd)

- Finally we have the eigenvalues:

$$\lambda_j = \sum_{i=0}^3 p_i \cdot \chi_{ij}, \quad \text{where } \chi_{ij} = \begin{cases} 1 & \text{if } [\sigma_i, \sigma_j] = 0 \\ -1 & \text{if } \{\sigma_i, \sigma_j\} = 0 \end{cases} \quad (114)$$

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

$$\sigma_i X \sigma_i = \begin{cases} X & i = 0, 1 \\ -X & i = 2, 3 \end{cases} \Rightarrow \lambda_1 = p_0 + p_1 - p_2 - p_3 \quad (115)$$

- For $\sigma_2 = Y$:

$$\sigma_i Y \sigma_i = \begin{cases} Y & i = 0, 2 \\ -Y & i = 1, 3 \end{cases} \Rightarrow \lambda_2 = p_0 + p_2 - p_1 - p_3 \quad (116)$$

- For $\sigma_3 = Z$:

$$\sigma_i Z \sigma_i = \begin{cases} Z & i = 0, 3 \\ -Z & i = 1, 2 \end{cases} \Rightarrow \lambda_3 = p_0 + p_3 - p_1 - p_2 \quad (117)$$

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:

$$-1 \leq \lambda_1 \leq 1 \quad (118)$$

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:

$$\lambda_j \in [-1, 1] \quad \forall j \quad (119)$$

- 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:

$$\lambda_i = (1 - \mathcal{P}_n)\alpha_i + \mathcal{P}_n\beta_i. \quad (120)$$

- γ satisfies:

$$\gamma = \max_i |\lambda_i| \leq (1 - \mathcal{P}_n) \max_i |\alpha_i| + \mathcal{P}_n \max_i |\beta_i| \leq 1. \quad (121)$$

Explicit Formula for Classical Causal Gain

- The classical capacity of the effective channel is

$$C(\mathcal{S}^n) = (1 - \mathcal{P}_n)C(\Phi_+) + \mathcal{P}_n C(\Phi_-) \quad (122)$$

$$= (1 - \mathcal{P}_n)(1 - h(\mu)) + \mathcal{P}_n(1 - h(\nu)) \quad (123)$$

$$= 1 - (1 - \mathcal{P}_n)h(\mu) - \mathcal{P}_n h(\nu) \quad (124)$$

- The classical causal gain is then

$$\delta_C = C(\mathcal{S}^n) - C(\mathcal{N}^n) \quad (125)$$

$$= C(\mathcal{S}^n) - (1 - h(\gamma)) \quad (126)$$

$$= h(\gamma) - (1 - \mathcal{P}_n)h(\mu) - \mathcal{P}_n h(\nu) \quad (127)$$

- Note: The composition channel $\mathcal{N} = \mathcal{N} \circ \dots \circ \mathcal{N}$ is still a Pauli channel, due to the (anti-)commutativity of Pauli operators.

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 (121), and the concavity and monotonicity of $h(x)$, we have:

$$h(\gamma) \geq h((1 - \mathcal{P}_n)\mu + \mathcal{P}_n\nu) \geq (1 - \mathcal{P}_n)h(\mu) + \mathcal{P}_nh(\nu). \quad (128)$$

Therefore,

$$\delta_C = h(\gamma) - (1 - \mathcal{P}_n)h(\mu) - \mathcal{P}_nh(\nu) \geq 0. \quad (129)$$

- The equality holds iff $\mu = \nu$. E.g., for completely depolarizing channel, $p_0 = \cdots = p_3 = \frac{1}{4}$, $\lambda_1 = \cdots = \lambda_3 = 0$. Thus, $\mu = \nu = 0$.

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 $\sigma_1, \sigma_2, \sigma_3$.

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

$$I_c(\rho, \mathcal{N}) = S(\mathcal{N}(\rho)) - S(\mathcal{N}^c(\rho)), \quad (130)$$

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**:

$$I_c\left(\frac{I}{2}, \mathcal{N}\right) = S\left(\mathcal{N}\left(\frac{I}{2}\right)\right) - S\left(\mathcal{N}^c\left(\frac{I}{2}\right)\right) = 1 - H(\vec{p}) \quad (131)$$

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

- Since Φ_{\pm} are both Pauli channels, it follows from (104) that the coherent information of \mathcal{S}^n also attains a maximum on $\frac{I}{2}$:

$$I_c(\mathcal{S}^n) = 1 - (1 - \mathcal{P}_n)H(\vec{s}) - \mathcal{P}_n H(\vec{t}) \quad (132)$$

- So we have the causal gain:

$$\delta_I = H(\vec{q}) - (1 - \mathcal{P}_n)H(\vec{s}) - \mathcal{P}_n H(\vec{t}) \quad (133)$$

- 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.

Table of Contents

- 1 Quantum SWITCH
 - Brief Introduction
- 2 Quantum SWITCH for Capacity Enhancement of Specific Channels
 - Quantum Information Theory Preliminaries
 - Completely Depolarizing Channels-Classical Information
 - Completely Depolarizing Channels-Quantum Information
- 3 Quantum SWITCH for Capacity Enhancement of Generic Channels
 - Key Quantity and Properties
 - Conjecture on Conditions of Communication Enhancement
 - Verification of Conjecture on Pauli Channels
 - Verification of Conjecture on Qudit Depolarizing Channels
- 4 References

Qudit Depolarizing Channels

A qudit depolarizing channel \mathcal{D}_p^d is given by:

$$\mathcal{D}_p^d(\rho) = (1 - p)\rho + p \operatorname{Tr}(\rho) \frac{I}{d} \quad (134)$$

where $p \in (0, 1)$ and $d > 1$.

[Wu et al., 2025b] proved the following result:

Theorem 5

Let δ_C be the classical causal gain with respect to forward and backward orders. Then for all qudit depolarizing channels (except the completely depolarizing channel), the condition $\mathcal{P}_n > 0 \iff \delta_C > 0$ holds.



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).



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.



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:](http://christian.mendl.net/science/talks/ringberg%202007.pdf)

[//christian.mendl.net/science/talks/ringberg%202007.pdf](http://christian.mendl.net/science/talks/ringberg%202007.pdf).

Presentation, 11 December 2007.



Müller-Hermes, A. (2021a).

Lecture 12: The holevo-schumacher-westmoreland theorem.

<https://www.uio.no/studier/emner/matnat/math/MAT4430/v22/lecture-notes/lecture12.pdf>.

Lecture notes for the course Quantum Information Theory (MAT4430), University of Oslo, Spring 2021.



Müller-Hermes, A. (2021b).

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.



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:](https://courses.cit.cornell.edu/ece531/Lectures/Lectures.htm)

[//courses.cit.cornell.edu/ece531/Lectures/Lectures.htm](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/Hilbert–Schmidt_operator.



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.



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:](http://link.aps.org/supplemental/10.1103/PhysRevA.111.012605)

[//link.aps.org/supplemental/10.1103/PhysRevA.111.012605](http://link.aps.org/supplemental/10.1103/PhysRevA.111.012605).

Thank you!
Any questions?