Quantum Channels: Transition from Classical to Quantum Channels

Himanshu Singh Raagav Ramakrishnan Shreyas Sinha Yash Seri

November 21, 2024

Contents

1 Introduction

Quantum channels, fundamental in the study of quantum information theory, describe how quantum states evolve under various types of noise and interaction. These channels serve as models for real-world scenarios where quantum systems are exposed to external environments, resulting in state transformations that are often non-ideal. Quantum channels are central to understanding and developing quantum communication, computation, and cryptography, as they capture the limitations and potential of transferring quantum information through noisy or imperfect media.

In quantum information theory, channels are typically categorized based on their capacities to transmit information. Just as in classical information theory, where the Shannon capacity limits the rate at which information can be reliably transmitted, quantum channels are analyzed through various capacity measures, each capturing different facets of information transfer. The primary capacities of interest include **quantum capacity**, **classical capacity**, and **private capacity**, each reflecting the channel's ability to transmit quantum information, classical information, or private (secure) information, respectively. These capacities vary depending on the nature and severity of the noise introduced by the channel, making it essential to understand the characteristics and limitations of each type.

One of the distinctive aspects of quantum channels is that their capacities are not always additive, a property that contrasts with classical channels. The superadditivity phenomenon, where multiple channel uses jointly yield a higher capacity than the sum of individual capacities, adds a layer of complexity to the analysis. For example, entanglement-assisted strategies can enhance both classical and quantum transmission rates, exploiting the channel's underlying quantum mechanics to achieve greater efficiency. This superadditivity highlights the advantage of leveraging quantum properties like entanglement and coherence for information transmission.

Our term paper explores the nature of quantum channels, and delves into the mathematical frameworks used to quantify their capacities. We discuss the fundamental concepts underlying quantum channel capacities, analyze the theoretical limits of information transmission, and cement theses ideas by taking a suitable example.

Before we begin our discussion on channel capacities though, we need to first define some basic concepts in Quantum Information Theory. Since Quantum Systems work in a fundamentally different way compared to their classical counterparts, we must define new quantities and constructs to take these into account.

1.1 Trace

It is helpful for us to introduce the **Trace** of a square operator. It is defined as:

$$Tr[A] \equiv \sum_i \langle i|A|i\rangle$$

where A is an operator in the Hilbert space $\mathscr{H}$ and $\{|i\rangle\}$ is a complete orthonormal basis in $\mathscr{H}$. We note the following properties about the Trace:

- Trace is Linear: $Tr[aA + bB] = aTr[A] + bTr[B]$ where a and b are scalars
- Trace is independent of the choice of the basis, as long as it is complete and orthonormal
- Trace is Cyclic: $Tr[ABC] = Tr[CAB]$

1.2 States

In quantum mechanics, the state of a system encapsulates all the information needed to describe it and predict the outcomes of measurements. For a simple, isolated quantum system, the state is typically represented as a **state vector** (or **ket**) $|\psi\rangle$ in a complex Hilbert space $\mathscr{H}$. This vector describes a **pure state**, where the system is in a specific, well-defined quantum state. Mathematically, pure states satisfy $\langle\psi|\psi\rangle = 1$, indicating that the vector is normalized. For example, in a two-level quantum system (qubit), a general pure state is represented as $|\psi\rangle = \alpha|0\rangle + \beta|1\rangle$, where α and β are complex coefficients satisfying $|\alpha|^2 + |\beta|^2 = 1$.

It is important to note that the trace of the square of a pure state is 1, or $Tr[|\psi\rangle^2] = 1$.

However, in many real-world scenarios, systems are not isolated and can exist in **mixed states**, where they are probabilistically distributed over multiple possible states. Mixed states are described by **density operators**, which generalize the concept of quantum states to incorporate probabilistic mixtures and interactions with external environments. Trace of squares of mixed states is not necessarily 1. $\rho = 0.5|0\rangle\langle 0| + 0.5|1\rangle\langle 1|$ is an example of such a mixed state.

In quantum information theory, we define a quantity called **purity** to judge if a state is pure or not:

Definition 1.1 (Purity)**.** *The purity of a state $P[\rho]$ is defined as:*

$$P[\rho] = Tr[\rho^\dagger \rho] = Tr[\rho^2]$$

$P[\rho] = 1$ if ρ is pure and $P[\rho] < 1$ if ρ is NOT pure.

1.3 Density Operators

As described above, the **density operator** (or **density matrix**) is a mathematical tool that describes the state of a quantum system. Unlike a pure state, which represents a system with complete certainty about its state vector, the density operator allows for the representation of mixed states, where a system may exist in a probabilistic mixture of several possible states. This flexibility makes density operators essential for describing real-world quantum systems, particularly when dealing with noise, entanglement, and decoherence effects in quantum channels. The density operator for a system encodes all observable properties of the system.

Mathematically, a density operator ρ corresponding to the ensemble $\mathscr{E} \equiv \{p_X(x), |\psi\rangle\}_{x\in\mathscr{X}}$ is defined as:

$$\rho \equiv \sum_{x\in\mathscr{X}} p_X(x)|\psi\rangle\langle\psi|$$

This definition of density operators can also be interpreted as the state of the system. One more way of understanding the density operator is to think of it as the expected state:

$$\rho = \mathbb{E}_X\{|\psi\rangle\langle\psi|\}$$

The definition gives us a wide picture of the Density operator. As a consequence of the above, the Density operator has certain properties that it always satisfies. These are summarized here as follows:

- Density operator has unit trace: $Tr[\rho] = 1$
- Density operator is Hermitian: $\rho^\dagger = \rho$
- Density operators are positive semi-definite, meaning $\langle\varphi|\rho|\varphi\rangle \geq 0 \forall\varphi$ which may be written as $\rho >= 0$

It is worth mentioning that every ensemble has a unique density operator, but the opposite is not necessarily true, and the same density operator could correspond to multiple ensembles.

1.4 Environment and Reference Systems

In studying quantum systems, we often introduce reference systems and environments to clarify how entanglement, decoherence, and information transfer occur. A reference system R is often used as an ancillary(additional) system that helps track information about another system A. Reference systems are useful for defining quantities like coherent information and entanglement. By examining how the information in A relates to the reference R, we gain insights into the entanglement properties and capacities of quantum channels or systems.

In realistic scenarios, no quantum system is completely isolated. An environment E represents everything outside the system of interest, which may interact with it and cause decoherence—the process by which a quantum system loses its coherence and behaves more classically. When analyzing quantum systems, we model the system and its environment as a combined entity as mixed states that allow us to study how noise and information loss occur in quantum channels.

A reference system is often modelled as a **Product state**:

Definition 1.2 (Product State)**.** *The Tensor Product of a state with another state (often the environment or a reference system).*

$$\rho_{AB} = \rho_A \otimes \rho_B$$

The product state models all the interactions between the system and the environment. Besides modelling environment interaction, product states can also model interaction between two systems and give information about the correlation between them.

Definition 1.3 (Bipartite states)**.** *A state is said to be Bipartite if it is a Tensor Product of two different states.*

This brings us to the related definition of separable states and entanglement:

Definition 1.4 (Separable states)**.** *A bipartite state σ_{AB} is said to be Separable if it can be written in the form:*

$$\sigma_{AB} = \sum_x p_X(x)|\psi_x\rangle\langle\psi_x|_A \otimes |\phi_x\rangle\langle\phi_x|_B$$

for some probability distribution $p_X(x)$ and sets $\{|\psi_x\rangle_A\}$ and $\{|\phi_x\rangle_B\}$ of pure states.

Definition 1.5 (Entanglement)**.** *A state that is not separable is said to be entangled.*

One more tool that is often used in quantum information theory is the Partial Trace. It is defined as follows:

Definition 1.6 (Partial Trace)**.** *Let X_{AB} be a square operator acting on the tensor product Hilbert Space $\mathscr{H}_A \otimes \mathscr{H}_B$ and let $\{|l\rangle_B\}$ be an orthonormal basis for $\mathscr{H}_B$. Then, the partial trace over B is:*

$$Tr_B[X_{AB}] \equiv \sum_l (I_A \otimes \langle l|_B) X_{AB} (I_A \otimes |l\rangle_B)$$

The partial trace effectively ”traces out” one part of the system, leaving us with the ”local” density operator to work with.

Now that we have defined the environment, we can go on to describe how a channel process a state. Although the exact details of a channel will be described below, we are right now interested in describing how the channel is used to process information.

This usually takes place is 3 steps. As seen in figure 1, we usually deal with classical information that needs to be converted to its quantum analogue first. This conversion takes place through an ”encoder”,

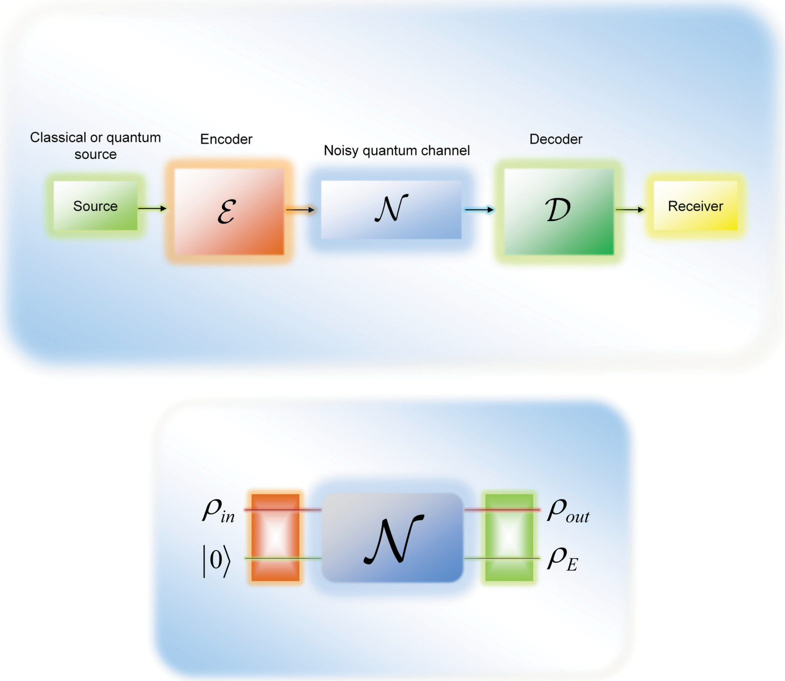


Figure 1: (Top) Illustration of the 3 step process picture of quantum communication, where classical information is encoded as quantum information, then evolved through quantum channels, then decoded back to classical information - the encoder and decoder can both be modelled as quantum channels themselves. (Bottom) Illustration of how a reference system (in this case it is known to be $|0\rangle$) may be considered along with the main system when processing through a channel. Images borrowed from [1]

in a process called state preparation. This prepared state can then be evolved through various quantum channels to get our final result, which is usually quantum information that needs to be converted to some classical form for analysis through some ”decoder”. This ”decoder” is more often than not some measurement through some instrument. It is interesting to note that the process of measurement itself can be modelled as a quantum channel! It is convenient, however, to model measurements as POVMs (Positive Operator Valuled Measures). We define these as follows:

Definition 1.7 (POVM)**.** *A Positive Operator Valued Measure (POVM) is a set $\{\Lambda_j\}_j$ of operators that satisfy non-negativity and completeness:*

$$\forall j : \Lambda_j \geq 0$$
$$\sum_j \Lambda_j = I$$

and the probability of obtaining outcome j is:

$$\langle\psi|\Lambda_j|\psi\rangle$$

1.5 Channels as CPTP maps

Now we have established most things required to begin our major discussion on Quantum Channels. Any Quantum channel is a model of a physical transformation. Hence, there are certain rules that such a channel must follow in order to make physical sense.

Let us take $\mathscr{D}_A$ and $\mathscr{D}_B$ to be the set all density operators in Hilbert spaces $\mathscr{H}_A$ and $\mathscr{H}_B$ respectively. Our channel must be a map from $\mathscr{D}_A$ to $\mathscr{D}_B$. We use $\mathscr{L}(\mathscr{H})$ to denote the set of all square linear operators on $\mathscr{H}$. Let us call our channel $\mathscr{N}$ For the channel to make physical sense, we must apply 3 restrictions on it:

Linearity It makes sense to have linearity in quantum channels. Although the physical basis for this condition is beyond the scope of this short review, if channels were not linear, the results of many experiments would not make sense. We may take this as an axiom for now.

Hence, a quantum channel must satisfy:

$$\mathscr{N}(\alpha X_A + \beta Y_A) = \alpha \mathscr{N}(X_A) + \beta \mathscr{N}(Y_A)$$

Complete Positivity Since $\mathscr{N}$ maps density operator to density operator, the output of this map must be positive semi-definite whenever the input is positive semi-definite as all density operators are always positive semi-definite. For this we define positivity:

Definition 1.8 (Positivity)**.** *A linear map $\mathscr{M} : \mathscr{L}(\mathscr{H}_A) \to \mathscr{L}(\mathscr{H}_B)$ is said to be positive if $\mathscr{M}(X_A)$ is positive semi-definite for all positive semi-definite X_A.*

However, this condition must be extended a little bit. Since we are often working with product states with reference systems and our channel may apply to only one of the states in the product, the above definition may prove to be lacking. To remedy this, we define complete positivity:

Definition 1.9 (Complete Positivity)**.** *A linear map $\mathscr{M} : \mathscr{L}(\mathscr{H}_A) \to \mathscr{L}(\mathscr{H}_B)$ is said to be completely positive if $id_R \otimes \mathscr{M}$ is positive map for a reference system R of arbitrary size.*

We require our channel map to be completely positive.

Trace Preserving Lastly, since (again) our channel maps density operators to density operators, the input and output both must have Trace 1. In other words, the trace of the output must be the same as the trace of the input. This is defined as:

Definition 1.10 (Trace Preservation)**.** *A map $\mathscr{N}$ is trace preserving if $Tr[X_A] = Tr[\mathscr{N}(X_A)]$ for all input $X_A \in \mathscr{L}(\mathscr{H}_A)$*

With the above conditions defined, we may now define a quantum channel:

Definition 1.11 (Quantum Channel)**.** *A quantum channel is a linear, completely positive, trace preserving map, corresponding to a quantum physical evolution.*

1.6 Choi-Kraus Representation and the Choi operator

The above definition (CPTP maps) of the quantum channel leads to a certain representation that turns out to be incredibly useful and insightful. We summarize this representation as the Choi-Kraus theorem:

Theorem 1 (Choi-Kraus Theorem)**.** *Any map $\mathscr{N}_{A\to B} : \mathscr{L}(\mathscr{H}_A) \to \mathscr{L}(\mathscr{H}_B)$ (where $\mathscr{L}(\mathscr{H}_A)$ is the space of all Linear operators on $\mathscr{H}_A$) is Linear, Completely Positive and Trace Preserving if and only if it has a Choi-Kraus decomposition as follows:*

$$\mathscr{N}_{A\to B}(X_A) = \sum_{l=0}^{d-1} V_l X_A V_l^\dagger$$

where $X_A \in \mathscr{L}(\mathscr{H}_A)$*,* $V_l \in \mathscr{L}(\mathscr{H}_\mathscr{A}, \mathscr{H}_\mathscr{B})$ $\forall l \in \{0,\ldots,d-1\}$ *with* $d < dim(\mathscr{H}_A)dim(\mathscr{H}_B)$ *and*

$$\sum_{l=0}^{d-1} V_l^\dagger V_l = I_A$$

Before we provide an outline of the proof, we define Choi operators:

Definition 1.12 (Choi Operators)**.** *Let* $\mathscr{H}_R$ *and* $\mathscr{H}_A$ *be isomorphic Hilbert spaces, and let* $\{|i\rangle_R\}$ *and* $\{|i\rangle_A\}$ *be orthonormal bases for* $\mathscr{H}_R$ *and* $\mathscr{H}_A$*, respectively. Let* $\mathscr{H}_B$ *be some other Hilbert space, and let* $\mathscr{N} : \mathscr{L}(\mathscr{H}_A) \to \mathscr{L}(\mathscr{H}_B)$ *be a linear map (written also as* $N_{A\to B}$*). The Choi operator corresponding to* $N_{A\to B}$ *and the bases* $\{|i\rangle_R\}$ *and* $\{|i\rangle_A\}$ *is defined as the following operator:*

$$(id_R \otimes \mathscr{N}_{A\to B})(|\Gamma\rangle\langle\Gamma|_{RA}) = \sum_{i,j=0}^{d_A-1} |i\rangle\langle j|_R \otimes \mathscr{N}_{A\to B}(|i\rangle\langle j|_A)$$

where $d_A \equiv dim(\mathscr{H}_A)$ *and* $|\Gamma\rangle_{RA}$ *is the unnormalized maximally entangled vector:*

$$|\Gamma\rangle_{RA} \equiv \sum_{i=0}^{d_A-1} |i\rangle_R \otimes |i\rangle_A$$

The Choi operator is a profound and very useful tool. The very first thing that we may state (without proof) is that the positive semi-definiteness of the Choi operator implies the complete positivity of the channel map that it corresponds to, and vice versa. This is not the only information about the channel that the Choi operator carries though. It encodes in itself the entire behavior of the channel. It is in essence a ”finger print” of the channel.

The proof of the Choi-Kraus theorem is very long and beyond the scope of this very short review, but it uses the Choi operator and various tricks to decompose the channel.

1.7 POVM Measurements

A **Positive Operator-Valued Measure (POVM)** is a quantum measurement described by positive semi-definite operators on a Hilbert space. POVMs generalize projection-valued measures (PVMs), allowing for a broader class of quantum measurements.

In analogy, while a PVM describes pure state measurements, a POVM is to a *mixed state*, often arising from subsystems of larger quantum systems or effects of noisy environments. POVMs are extensively used in quantum information theory.

1.7.1 Definition

Let $\mathscr{H}$ denote a Hilbert space and (X,M) a measurable space. A POVM is a function F defined on M whose values are positive bounded self-adjoint operators on $\mathscr{H}$, satisfying:

$$F(X) = I_{\mathscr{H}}, \quad \langle\psi|F(E)|\psi\rangle \geq 0,$$

where $\langle\psi|F(E)|\psi\rangle$ gives the probability of event E for state $|\psi\rangle$.

In finite dimensions, a POVM is a set of positive semi-definite Hermitian matrices $\{F_i\}$ satisfying:

$$\sum_i F_i = I,$$

where I is the identity matrix.

1.7.2 Properties

1. **Probability Measure:** For a quantum state ρ, the probability of measurement outcome i is:

$$\text{Prob}(i) = \text{Tr}(\rho F_i).$$

If ρ is pure ($\rho = |\psi\rangle\langle\psi|$), this reduces to:

$$\text{Prob}(i) = \langle\psi|F_i|\psi\rangle.$$

2. **Comparison with PVMs:** - In PVMs, measurement operators $\{P_i\}$ are orthogonal projectors ($P_iP_j = \delta_{ij}P_i$). - POVMs relax this restriction, enabling non-orthogonal measurement operators and greater flexibility.

1.7.3 Applications

POVMs are widely used in quantum information and communication:

- **Quantum State Discrimination:** Optimal strategies for distinguishing non-orthogonal quantum states.
- **Quantum Cryptography:** Analyze eavesdropping and maximize information extraction.
- **Quantum Communication:** Determine classical information capacity of quantum channels.

POVMs generalize quantum measurements, extending beyond projective measurements, and play a crucial role in understanding and utilizing quantum systems.

2 Quantum Entropy, Accessible Information and Holevo Bound

2.1 Quantum Entropy

Quantum (Von Neumann) entropy is the measure of disorder of a quantum state, which can be calculated as:

$$S(\rho) = -\mathrm{Tr}(\rho \log \rho),$$

where Tr is the trace of the density matrix. It can also be written as:

$$S(\rho) = -\sum_i \lambda_i \log \lambda_i,$$

where λ_i is an eigenvalue of the density matrix ρ.

For example, a quantum system with the mixed state:

$$\rho = \begin{pmatrix} 0.8 & 0 \\ 0 & 0.2 \end{pmatrix}$$

has eigenvalues 0.8 and 0.2 (since the quantum state is diagonal, the eigenvalues are simply the diagonal elements). The von Neumann entropy of this system is:

$$S(\rho) = -\sum_i \lambda_i \log \lambda_i = -(0.8 \log 0.8 + 0.2 \log 0.2) = 0.3974.$$

2.1.1 Properties of Quantum Entropy

Non-Negativity The quantum entropy $H(\rho)$ is non-negative for any density operator ρ:

$$H(\rho) \geq 0.$$

Proof: The proof of this follows from the non-negativity of Shannon entropy.

Minimum Value The minimum value of $H(\rho)$ is zero, which occurs when the density operator ρ is in its pure state.

Proof: When the eigenvalues of the density operator are distributed such that all the probability mass is concentrated on one eigenvector (and the other eigenvalues are zero), the density operator has a rank of one. This corresponds to its pure state.

Maximum Value The maximum value of $H(\rho)$ is $\log d$, where d is the dimension of the system. This occurs when the density operator is in its maximally mixed state.

Proof: The proof of this property is the same as in the classical case. A maximally mixed state distributes the probability mass equally among all eigenvalues, achieving the maximum entropy.

Concavity Let $\rho_x \in$ D(H) and let $p_x(x)$ be the probability distribution. The entropy is concave in the density operator:

$$H(\rho) \geq \sum_x p_X(x) H(\rho_x),$$

where $\rho \equiv \sum_x p_x(x)\rho_x$

2.1.2 Joint Quantum Entropy

The joint quantum entropy H(AB)$_\rho$ of the density operator $\rho_{AB} \in \mathscr{D}(\mathscr{H}_{\mathscr{A}} \otimes \mathscr{H}_{\mathscr{B}})$ for a bipartite system AB follows naturally from the definition of quantum entropy

$$H(AB)_\rho \equiv -Tr\{\rho_{AB} \log \rho_{AB}\},$$

2.1.3 Conditional Quantum Entropy

Let $\rho_{AB} \in D(H_A \otimes H_B)$. The conditional quantum entropy $H(A|B)_\rho$ of ρ_{AB} is equal to the difference of the joint quantum entropy $H(AB)_\rho$ and marginal entropy $H(B)_\rho$

$$H(A|B)_\rho \equiv H(AB)_\rho - H(B)_\rho,$$

Conditioning does not increase entropy Consider a bipartite quantum state ρ_{AB}. Then the following applies for marginal entropy $H(A)_\rho$ and conditional quantum entropy $H(A|B)_\rho$:

$$H(A)_\rho \geq H(A|B)_\rho$$

The proof for this property is the same as the classical one

2.1.4 Quantum Mutual Information

the quantum mutual information of a bipartite state $\rho_{AB} \in \mathscr{D}(\mathscr{H}_{\mathscr{A}} \otimes \mathscr{H}_{\mathscr{B}})$ is defined as follows:

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

The following relations hold for quantum mutual information the same as for classical mutual information:

$$I(A;B)_\rho = H(A)_\rho - H(A|B)_\rho.$$
$$= H(B)_\rho - H(B|A)_\rho$$

These formulae lead to the following relations between quantum mutual information and coherent information:

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

Non-negativity of Quantum Mutual Information The quantum mutual information of any bipartite quantum state ρ_{AB} is non-negative.

$$I(A;B)_\rho \geq 0$$

2.1.5 Conditional Quantum Mutual Information

We define the conditional quantum mutual information $I(A;B|C)_\rho$ of any tripartite state $\rho_{ABC} \in \mathscr{D}(\mathscr{H}_{\mathscr{A}} \otimes \mathscr{H}_{\mathscr{B}} \otimes \mathscr{H}_{\mathscr{C}})$ similarly to how we did in the classical one:

$$I(A;B|C)_\rho \equiv H(A|C)_\rho + H(B|C)_\rho - H(AB|C)_\rho$$

Chain rule The CQMI obeys a chain rule:

$$I(A;BC)_\rho = I(A;B)_\rho + I(A;C|B)_\rho$$

The proof for this follows the classical one

Non-Negativity if CQMI Let $\rho_{ABC} \equiv \mathscr{D}(\mathscr{H}_{\mathscr{A}} \otimes \mathscr{H}_{\mathscr{B}} \otimes \mathscr{H}_{\mathscr{C}}))$.Then the CQMI is non-negative:

$$I(A;B|C)_\rho \geq 0.$$

2.1.6 Quantum Relative Entropy

The quantum relative entropy $D(\rho||\sigma)$ between a density operator $\rho \in D(H)$ and a positive semi-definite operator $\sigma \in L(H)$ is defined as follows:

$$D(\rho\|\sigma) = Tr\{\rho[log\rho - log\sigma]\}$$

if the following condition is met:

$$supp(\rho) \subseteq supp(\sigma)$$

Else it is defined as +∞

Monotonicity of Quantum Relative Entropy Let $\rho \in \mathscr{D}(\mathscr{H}), \sigma \in \mathscr{L}(\mathscr{H})$ and $\mathscr{N} : \mathscr{L}(\mathscr{H}) \to \mathscr{L}(\mathscr{H}\prime)$ be a quantum channel. The quantum relative entropy can only decrease or stay the same if we apply the same quantum channel N to ρ and σ:

$$D(\rho||\sigma) \geq D(N(\rho)||N(\sigma))$$

This implies non-negativity of quantum relative entropy in some cases

Non-negativity Let $\rho \in \mathscr{D}(\mathscr{H})$, and let $\sigma \in \mathscr{L}(\mathscr{H})$ be positive semi-definite such that $Tr\{\rho\} \leq 1$. Then the quantum relative entropy $D(\rho||\sigma)$ is non-negative:

$$D(\rho||\sigma) \geq 0$$

and $D(\rho||\sigma) = 0$ if and only if $\rho = \sigma$

Proof: The first part of this proof follows from the monotonicity of relativity

$$D(\rho||\sigma) \geq D(Tr\{\rho\}||Tr\{\sigma\} = Tr\{\rho\}log\left(\frac{Tr\{\rho\}}{Tr\{\sigma\}}\right) \geq 0$$

If $\rho = \sigma$, then we get that $D(\rho||\sigma)$ = 0.This implies that the inequality above is saturated and thus Tr{ρ} = Tr{σ} = 1. Let M be an arbitrary measurement channel. From the monotonicity of relative entropy we can conclude that $D(M(\rho)||M(\sigma)$ = 0. The equality condition for non-negative entropy gives us $M(\rho) = M(\sigma)$. Since this equality holds for any possible measurement channel, we can conclude that $\rho = \sigma$

Property Let $\rho \in \mathscr{D}(\mathscr{H})$ and $\sigma, \sigma' \in \mathscr{L}(\mathscr{H})$ be positive semi-definite. Suppose that $\sigma \leq \sigma'$. Then

$$D(\rho\|\sigma') \leq D(\rho\|\sigma).$$

Proof: The assumption that $\sigma \leq \sigma'$ is equivalent to $\sigma' - \sigma$ being positive semi-definite. Then the following operator is positive semi-definite:

$$\sigma \otimes |0\rangle\langle 0|_X + (\sigma' - \sigma) \otimes |1\rangle\langle 1|_X,$$

and as a consequence

$$D(\rho\|\sigma) = D(\rho \otimes |0\rangle\langle 0|_X \| \sigma \otimes |0\rangle\langle 0|_X + (\sigma' - \sigma) \otimes |1\rangle\langle 1|_X), \tag{11.141}$$

which follows by a direct calculation (essentially the same reasoning as that used to solve Exercise 11.8.8). By monotonicity of quantum relative entropy (Theorem 11.8.1), the quantum relative entropy does not increase after discarding the system X, so that

$$D(\rho \otimes |0\rangle\langle 0|_X \| \sigma \otimes |0\rangle\langle 0|_X + (\sigma' - \sigma) \otimes |1\rangle\langle 1|_X) \geq D(\rho\|[\sigma + (\sigma' - \sigma)]) = D(\rho\|\sigma'),$$

concluding the proof.

2.2 Accessible Information

The Quantum no-cloning theorem establishes that quantum states cannot be perfectly cloned. This implies that there is an upper bound to the information that can be extracted from a quantum channel. The upper bound of information that is extractable from a quantum system is called the Holevo bound.

In classical systems, information is encoded deterministically, such as binary bits. In principle, all the encoded information can be accessed without limitations. In Quantum systems, however, information is encoded in quantum states, which can be superpositions or mixed states. This limits the amount of information that can be extracted from the system.

Let a sender(Alice) encode a classical variable X into a quantum state ρ_x. The receiver(Bob) attempts to decode X by measuring the quantum state. The accessible information is defined as the maximum classical mutual information between X and the output decoded Y:

$$I_{accessible} = \max I(X;Y)$$

Where the maximum is taken over all possible measurements that can be performed, represented by POVMs.

2.3 Holevo Bound

The Holevo bound is a fundamental result in quantum information theory that quantifies the maxiumum amount of classical information that can be extracted from a quantum system.

$$\chi = S(\rho) - \sum_i p_i S(\rho_i).$$

The proof for this is as follows:

Consider a quantum system that encodes classical information, which can be represented by a distribution with probability p_x for each input x. This allows us to define the classical state ρ_x as:

$$\rho_X := \sum_x p_x |x\rangle\langle x|,$$

(where $|x\rangle$ (ket x) represents a quantum state and $\langle x|$ (bra x) represents the Hermitian conjugate of $|x\rangle$.)

Since each input is mapped to a quantum state ρ_x, the combined state can be written as:

$$\rho_{XQ} := \sum_x p_x |x\rangle\langle x| \otimes \rho_x,$$

where ρ_x represents the density matrix of the quantum system.

The received combined state is represented as:

$$\rho := \mathrm{Tr}_X(\rho_{XQ}) = \sum_x p_x \rho_x,$$

where Tr_X represents the trace of ρ_{XQ}.

To bound the maximum information obtainable, we need to bound the mutual information $I(X:Y)$ with $I(X:Q)$, where X is the input, Y is the output, and Q is the quantum state. From the monotonicity of quantum mutual information, quantum mutual information does not increase under quantum operations. Hence, we get:

$$I(X:Q'Y) \leq I(X:Q).$$

Similarly,

$$I(X:Y) \leq I(X:Q'Y).$$

From these two inequalities, we get:

$$I(X:Y) \leq I(X:Q).$$

Simplifying $S(X:Q)$, we have:

$$I(X:Q) = S(X) + S(Q) - S(XQ),$$

$$= S(X) + S(\rho) - \mathrm{Tr}(\rho_{XQ} \log \rho_{XQ}),$$

$$= S(\rho) - \sum_x p_x S(\rho_x).$$

Thus, we arrive at the Holevo bound.

3 Schumacher’s Quantum Noiseless Channel Coding Theorem

3.1 Channel Compression

3.1.1 What is Channel Compression?

Channel compression is the process of encoding information in a way that minimizes the required resources, like **bits** in Classical Channels or **qubits** in Quantum Channels. In this way, we can preserve the integrity of the original message during transmission across the channel.

3.1.2 Why Channel Compression?

- **Efficiency:** All communication channels, whether classical or quantum, have finite capacities. Compressing information ensures that these capacities are utilized effectively without wastage.
- **Adaptability to Noise:** By compressing and encoding the message efficiently, we can cut-down the effects of noise present in the channel, improving reliability of transmission in communication systems.

3.1.3 In the Classical Context

In classical information theory, channel compression minimizes the number of bits needed to transmit a message over a channel without loss. These compression techniques are based on Shannon’s Source Coding Theorem, which aim to achieve rates close to the source entropy $H(X)$, where X represents the random variable describing the source in question.

3.1.4 In the Quantum Context

Schumacher’s Compression Theorem provides a way to compress quantum information down to the von Neumann entropy, indicating the minimum qubits needed to achieve high-fidelity state reconstruction. Quantum (communication) systems are inherently fragile and resource-intensive. We reduce the physical and computational resources needed to transmit relevant quantum information.

3.2 Classical Compression - Shannon’s Source Coding Theorem

Shannon’s Source Coding Theorem establishes the minimum average number of bits needed to represent information from a classical source without loss, which is directly related to the entropy of the source.

For a discrete memoryless source X, with alphabet $\mathscr{X}$, and probability distribution $P(X)$, the minimum rate R (in bits per symbol) at which information can be transmitted without loss is given by *Shannon entropy* $H(X)$:

$$H(X) = -\sum_{x\in\mathscr{X}} P(x)\log_2 P(x)$$

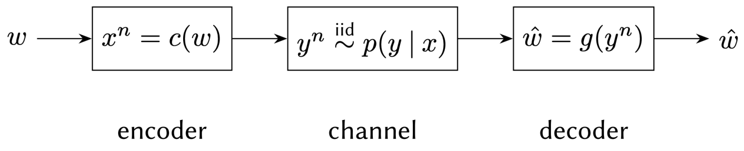
This entropy $H(X)$, represents the average information per symbol.

3.3 Proof of Shannon’s Source Coding Theorem

We consider the steps below, and draw a parallel with the proof of Schumacher’s Compression

1. **The Source Representation:** A discrete memoryless source X generates n-length sequences x^n with probabilities $P(x^n) = \prod_{i=1}^{n} P(x_i)$.

2. **Asymptotic Equipartition Property (AEP):** As $n \to \infty$, most sequences will fall into the **typical set** $A_\varepsilon^{(n)}$. We consider its *size* $2^{nH(X)}$ and each sequence is *equally likely*.
3. **Compression:** Typical sequences are encoded using $nH(X)$ bits.
4. **Achievability and Converse:**
   - Rates $R > H(X)$: Lossless compression is possible.
   - Rates $R < H(X)$: There **will be** information loss.



$p(y \mid x)$ probability of transmitted symbol x being receved as y.

Figure 2: Shannon's Noisy Channel Theorem [2]

3.4 Important quantities to consider for Schumacher's Compression

- **Entropy Measures:** For Shannon Entropy $H(X)$, its analog is $S(\rho)$.
- **Typicality:** We consider typical subspaces drawn from $\mathscr{H}$.
- **Dimension of Representation:** Schumacher's typical subspace has dimension $2^{nS(\rho)}$.
- **Compression:** Schumacher compresses quantum states by projecting them onto the typical subspace using an **isometry** $V_{n,\varepsilon}$.

3.5 Quantum Compression - Schumacher's Compression Theorem

1. Consider a quantum source producing states described by a density matrix ρ, associated with an ensemble $\{p_i, |\psi_i\rangle\}$ of pure states and probabilities. This acts on a *Complex Euclidean Space* (Hilbert space), with dimension $d = \dim(\mathscr{H})$.
2. Direct Part (Achievability): For any rate $R > S(\rho)$, there exists a sequence of (n, m, δ)-compression schemes that compress $\rho^{\otimes n}$ into a subspace of dimension 2^{nR} with perfect fidelity as $n \to \infty$.
$$\lim_{n\to\infty} F(D \circ E, \rho^{\otimes n}) = 1,$$
where $F(D \circ E, \rho^{\otimes n})$ is the channel fidelity.
3. Converse Part (Optimality): If there exists a sequence of (n, m, δ)-compression schemes, that compress $\rho^{\otimes n}$ into subspaces of dimension 2^{nR} with perfect fidelity, then the rate R must satisfy $R \geq S(\rho)$.
$$\lim_{n\to\infty} F(D \circ E, \rho^{\otimes n}) = 1,$$
then $R \geq S(\rho)$.

3.6 Typical Subspaces

They capture the most probable sequences emitted by the quantum source.

3.6.1 Definition

An ε-typical sequence $(i_1, i_2, \ldots, i_n)$ satisfies:

$$\left| -\frac{1}{n} \log p_{i_1 i_2 \ldots i_n} - S(\rho) \right| < \varepsilon,$$

where $p_{i_1 i_2 \ldots i_n} = p_{i_1} p_{i_2} \cdots p_{i_n}$, and $S(\rho)$ is the von Neumann entropy of ρ:

$$S(\rho) = -\sum_{i=1}^{d} p_i \log p_i.$$

The ε-typical subspace $\mathscr{H}_{\text{typ}}$ of $\mathscr{H}^{\otimes n}$ is the *span* of tensor product states corresponding to ε-typical sequences of eigenstates $|\psi_i\rangle$:

$$\mathscr{H}_{\text{typ}} = \text{span}\left\{ |\psi_{i_1}\rangle \otimes |\psi_{i_2}\rangle \otimes \cdots \otimes |\psi_{i_n}\rangle \,|\, (i_1, i_2, \ldots, i_n) \in T_{n,\varepsilon}(p) \right\},$$

where $T_{n,\varepsilon}(p)$ denotes the set of ε-typical sequences.

3.6.2 Properties

1. **High Probability:** The projection of $\rho^{\otimes n}$ onto $\mathscr{H}_{\text{typ}}$ captures most of the probability mass:

   $$\text{Tr}[\Pi_{n,\varepsilon} \rho^{\otimes n}] \geq 1 - \delta_n,$$

   only when $\delta_n \to 0$ as $n \to \infty$, and $\Pi_{n,\varepsilon}$ is the projector onto $\mathscr{H}_{\text{typ}}$.

2. **Dimension Bound:** The dimension $D_{\text{typ}} = \text{Tr}[\Pi_{n,\varepsilon}]$ of $\mathscr{H}_{\text{typ}}$ satisfies:

   $$D_{\text{typ}} \leq 2^{n(S(\rho)+\varepsilon)}.$$

   For sufficiently large n we have:

   $$D_{\text{typ}} \geq (1 - \delta_n) 2^{n(S(\rho)-\varepsilon)}.$$

3. **Projected State in the Typical Subspace:** When $\rho^{\otimes n}$ is projected onto the typical subspace $\mathscr{H}_{\text{typ}}$, it is confined within the subspace:

   $$2^{-n(S(\rho)+\varepsilon)} \Pi_{n,\varepsilon} \leq \Pi_{n,\varepsilon} \rho^{\otimes n} \Pi_{n,\varepsilon} \leq 2^{-n(S(\rho)-\varepsilon)} \Pi_{n,\varepsilon},$$
   $$\implies \rho^{\otimes n} \text{ is approximately uniform within } \mathscr{H}_{\text{typ}}.$$

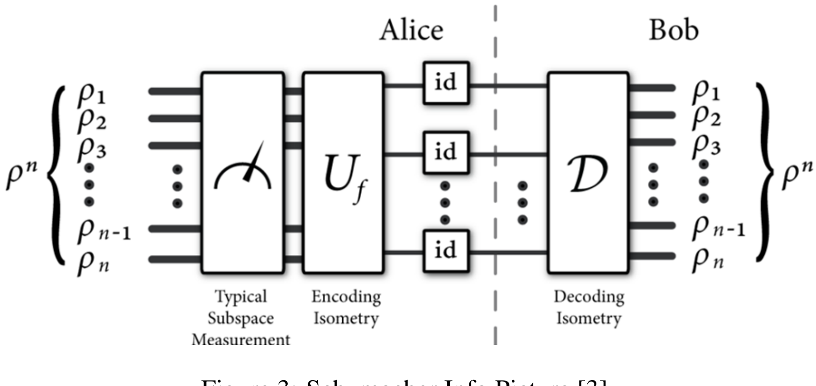


Figure 3: Schumacher Info Picture [3]

3.7 Proof of Schumacher's Compression Theorem

3.7.1 Direct Part: Achievability of Compression Rates

We aim to construct a compression scheme that achieves *asymptotically* perfect fidelity for any rate $R > S(\rho)$.

Step 1: Typical Projector $\Pi_{n,\varepsilon}$ Consider the typical projector $\Pi_{n,\varepsilon}$. It's defined as:

$$\Pi_{n,\varepsilon} = \sum_{(i_1,\ldots,i_n)\in T_{n,\varepsilon}(p)} |\psi_{i_1} \otimes \cdots \otimes \psi_{i_n}\rangle\langle\psi_{i_1} \otimes \cdots \otimes \psi_{i_n}|,$$

where $T_{n,\varepsilon}(p)$ is the set of ε-typical strings of length n, and is based on $\{p_i\}$ over the basis states $\{|\psi_i\rangle\}$ of ρ.
The typical projector maps states into the typical subspace $\mathscr{H}_{\text{typ}}$, which has the following properties:

- **Dimension Bound:** The dimension of the typical subspace is approximately:

$$\dim(\mathscr{H}_{\text{typ}}) \leq 2^{n(S(\rho)+\varepsilon)}.$$

- **Probability Concentration:** The typical subspace captures almost all the probability mass of $\rho^{\otimes n}$ for sufficiently large n.

$$\text{Tr}[\Pi_{n,\varepsilon}\rho^{\otimes n}] \to 1 \quad \text{as } n \to \infty.$$

- **Entropy Bound:** For any state in the typical subspace, the eigenvalues of $\rho^{\otimes n}$ satisfy:

$$2^{-n(S(\rho)+\varepsilon)} \leq \langle\psi|\rho^{\otimes n}|\psi\rangle \leq 2^{-n(S(\rho)-\varepsilon)},$$

where $|\psi\rangle \in \mathscr{H}_{\text{typ}}$.

This subspace represents where most of the information of $\rho^{\otimes n}$ is present.

Step 2: Encoding and Decoding Let $m_n = \lceil nR \rceil$, where $R > S(\rho)$. For sufficiently large n, the dimension of the compressed Hilbert space satisfies:

$$2^{m_n} > 2^{n(S(\rho)+\varepsilon)}.$$

Thus, the typical subspace can be mapped onto a smaller lower-dimensional Hilbert space $(\mathbb{C}^2)^{\otimes m_n}$.

- **Encoding Channel:** The encoding map $E_n : \mathscr{H}^{\otimes n} \to (\mathbb{C}^2)^{\otimes m_n}$ is defined as:

$$E_n(\rho^{\otimes n}) = \begin{cases} V_{n,\varepsilon}\rho^{\otimes n}V_{n,\varepsilon}^{\dagger}, & \text{if } \rho^{\otimes n} \in \mathscr{H}_{\text{typ}}, \\ |0\rangle\langle 0|, & \text{otherwise.} \end{cases}$$

Here, $V_{n,\varepsilon}$ is an *isometry* that maps $\mathscr{H}_{\text{typ}}$ into $(\mathbb{C}^2)^{\otimes m_n}$. And any state outside the typical subspace is mapped to a fixed state $|0\rangle \in (\mathbb{C}^2)^{\otimes m_n}$.

- **Decoding Channel:** The decoding map $D_n : (\mathbb{C}^2)^{\otimes m_n} \to \mathscr{H}^{\otimes n}$ is defined as:

$$D_n(Y) = V_{n,\varepsilon}^{\dagger} Y V_{n,\varepsilon}.$$

This will reconstruct the original state from its compressed representation in the typical subspace, while states outside the typical subspace are mapped to their nearest approximation within $\mathscr{H}_{\text{typ}}$.

Step 3: Fidelity Analysis To analyze the fidelity of the compression scheme, consider the combined operation $D_n \circ E_n$, represented by its Kraus operators $\{A_l^{(n)}\}$. The fidelity between the original state $\rho^{\otimes n}$ and the reconstructed state is given by:

$$F(D_n \circ E_n, \rho^{\otimes n}) = \left(\sum_{l=1}^{L_n} \left| \mathrm{Tr}(A_l^{(n)} \rho^{\otimes n}) \right| \right)^2 .$$

Using the Cauchy-Schwarz inequality, we can bound the term $\left| \mathrm{Tr}(A_l^{(n)} \rho^{\otimes n}) \right|$ as:

$$\left| \mathrm{Tr}(A_l^{(n)} \rho^{\otimes n}) \right|^2 \leq \mathrm{Tr}[\Pi_l^{(n)} \rho^{\otimes n}] \cdot q_l^{(n)},$$

where:

$$q_l^{(n)} = \mathrm{Tr}[A_l^{(n)} A_l^{(n)\dagger}] \quad \text{and} \quad \Pi_l^{(n)} = A_l^{(n)\dagger} A_l^{(n)}.$$

Substituting this back, we bound the fidelity as:

$$F(D_n \circ E_n, \rho^{\otimes n}) \leq \sum_{l=1}^{L_n} q_l^{(n)} \cdot \mathrm{Tr}[\Pi_l^{(n)} \rho^{\otimes n}].$$

Since $D_n \circ E_n$ is a quantum channel, the total probability weights satisfy:

$$\sum_{l=1}^{L_n} q_l^{(n)} = 1.$$

For sufficiently large n, the probability mass outside the typical subspace $\mathscr{H}_{\mathrm{typ}}$ becomes negligible.

$$\mathrm{Tr}[(I - \Pi_{n,\varepsilon})\rho^{\otimes n}] \to 0 \quad \text{as } n \to \infty.$$

Thus, the contribution to the fidelity from outside the typical subspace vanishes, ensuring:

$$F(D_n \circ E_n, \rho^{\otimes n}) \to 1 \quad \text{as } n \to \infty.$$

Step 4: Achievable Compression Rate The dimension of the typical subspace $\mathscr{H}_{\mathrm{typ}}$ is bounded by:

$$D_{\mathrm{typ}} \leq 2^{n(S(\rho)+\varepsilon)},$$

To encode the typical subspace, the number of qubits required is:

$$m_n = \lceil \log_2 D_{\mathrm{typ}} \rceil \leq n(S(\rho) + \varepsilon)$$

The compression rate R, defined as the number of qubits per copy of ρ, is:

$$R = \lim_{n\to\infty} \frac{m_n}{n}$$

And finally, from the bound on m_n, it follows that:

$$R \leq S(\rho) + \varepsilon$$

By choosing $\varepsilon > 0$ arbitrarily small, we ensure that R can approach $S(\rho)$. Therefore, for any $R > S(\rho)$, there exists a sufficiently small $\varepsilon > 0$ such that compression is achievable with asymptotically perfect fidelity.

3.7.2 Converse Part

For rates $R < S(\rho)$, fidelity approaches zero. Assume that there's a sequence of compression schemes with (n_k, m_k, δ_k) satisfying:

$$R = \lim_{k\to\infty} \frac{m_k}{n_k} < S(\rho) \quad \text{and} \quad \lim_{k\to\infty} \delta_k = 0.$$

Using Kraus decomposition for the channel, fidelity can be bounded:

$$F(D_k \circ E_k, \rho^{\otimes n_k}) \leq \sqrt{\sum_{l=1}^{L_k} q_l \mathrm{Tr}[\Pi_l \rho^{\otimes n_k}]},$$

where q_l are probabilities associated with the Kraus operators and Π_l are projections with dimensions constrained by the compression rate.

The typical subspace of $\rho^{\otimes n_k}$, which captures almost all of the probability mass, has dimension approximately $2^{n_k S(\rho)}$. When $R < S(\rho)$:

- The compressed space has dimension $2^{n_k R}$.
- The typical subspace has dimension $2^{n_k S(\rho)}$.

Since $R < S(\rho)$, the compressed space ($2^{n_k R}$) is exponentially smaller than the typical subspace ($2^{n_k S(\rho)}$). This mismatch ensures that the projection operators Π_l cover only a negligible fraction of the typical subspace.

For $R < S(\rho)$, it follows that:

$$\begin{aligned} \mathrm{Tr}[\Pi_l \rho^{\otimes n_k}] \to 0 \quad \text{as } k \to \infty, \\ \implies F(D_k \circ E_k, \rho^{\otimes n_k}) \to 0. \end{aligned}$$

4 Classical Capacity of a Quantum Channel

4.1 Definition of Classical Capacity

The classical capacity $C(\mathscr{N})$ of a quantum channel $\mathscr{N}$ quantifies the maximum amount of classical information it can transmit with an arbitrarily low error as the number of channel uses goes to infinity. It's measured in *bits per channel use*

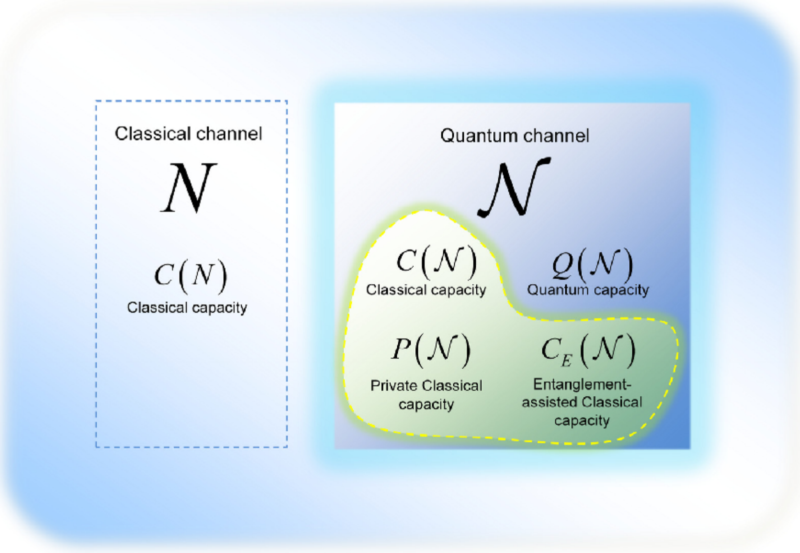


Figure 4: Properties of the Quantum Channel [4]

Practical Example: Consider a noisy quantum channel, such as a fiber optic cable carrying photons. Each photon represents a *quantum state*, which carries the classical information like *0*, *1*, or a *superposiition* of values. In the real world, however, these cables introduce noise.

- **Photon Loss:** Some photons fail to reach the receiver.
- **Depolarization:** The state of a photon changes unpredictably.
- **Environmental interference:** Vibrations, temperature changes, or impurities in the fiber can alter the transmitted states.

4.2 Formula for Classical Capacity

The classical capacity is determined using the Holevo-Schumacher-Westmoreland (HSW) theorem, which relates it to the **Holevo quantity**. Mathematically, it is given by:

$$C(\mathscr{N}) = \lim_{n\to\infty} \frac{1}{n}\chi\left(\mathscr{N}^{\otimes n}\right),$$

where:

- n: The number of channel uses.
- $\mathscr{N}^{\otimes n}$: The channel applied in parallel over n uses.
- $\chi(\mathscr{N})$: The Holevo quantity, an upper bound on the information that can be transmitted.

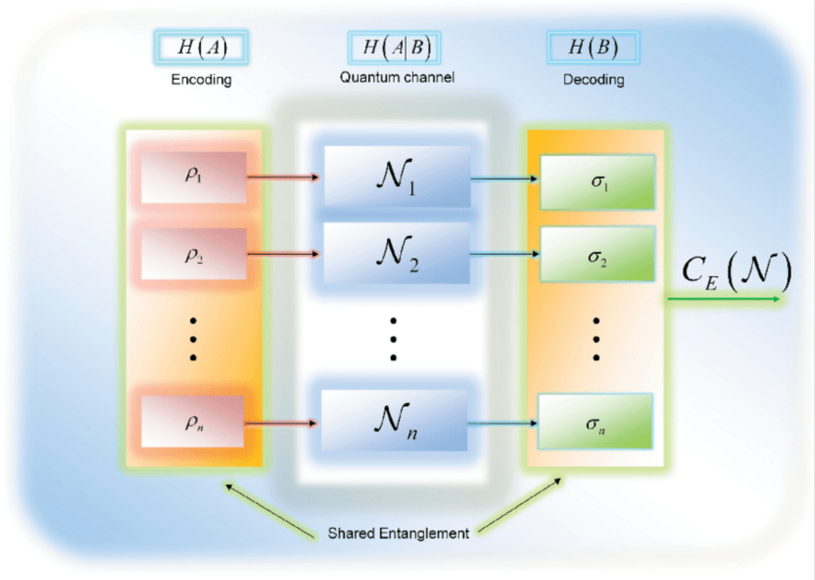


Figure 5: Classical Capacity Illustration [5]

4.2.1 The Holevo Quantity

The Holevo quantity χ quantifies the classical information extractable from quantum states. For a quantum channel $\mathscr{N}$ and a classical ensemble of quantum states $\{p_i, \rho_i\}$, it is defined as:

$$\chi(\mathscr{N}, \{p_i, \rho_i\}) = S\left(\sum_i p_i \mathscr{N}(\rho_i)\right) - \sum_i p_i S(\mathscr{N}(\rho_i)),$$

where:

- $\{p_i, \rho_i\}$: The ensemble of quantum states ρ_i sent with probabilities p_i.
- $S(\rho) = -\text{Tr}[\rho \log \rho]$: The von Neumann entropy.
- $\mathscr{N}(\rho_i)$: The output of the quantum channel for input state ρ_i.

4.3 Types of Quantum Channels

Quantum channels are often noisy, and can degrade the transmitted states. Understanding the noise characteristics is crucial for determining capacity.

Common Types of Quantum Channels:

- **Depolarizing Channel:** Randomly replaces a qubit with a completely mixed state.
- **Dephasing Channel:** Introduces noise that destroys superposition by affecting phase information.
- **Amplitude Damping Channel:** Models energy loss, such as photon leakage in optical fibers.

4.4 Computing Classical Capacity

Determining $C(\mathscr{N})$ involves:

1. **Step 1:** Evaluate the Holevo quantity $\chi(\mathscr{N})$ for a single channel use. This requires finding the optimal input ensemble $\{p_i, \rho_i\}$ that maximizes χ.

2. **Step 2:** Extend to n-parallel channel uses, taking advantage of collective strategies to enhance transmission. Regularization over n ensures the true capacity is approached.

4.4.1 Special Cases and Examples

- **Perfect Channel:** For an ideal quantum channel with no noise, the classical capacity equals the Shannon entropy of the input states.
- **Completely Depolarizing Channel:** For a channel that randomizes input states entirely, $C(\mathcal{N}) = 0$, as no meaningful classical information can be transmitted.

4.4.2 Practical Implications

1. **Communication Systems:** Quantum channels enable classical data transmission through optical fibers and satellites. Noise mitigation and error correction are essential to maximize capacity.
2. **Quantum Cryptography:** Classical capacity helps design secure communication protocols, where channel noise complicates eavesdropping but reduces capacity.
3. **Quantum Networks:** Classical capacity optimizes hybrid quantum-classical protocols in distributed systems.

4.5 Summary

The classical capacity $C(\mathcal{N})$ measures the maximum reliable classical information transmission over a quantum channel. While the Holevo quantity provides theoretical limits, practical systems use approximations and error correction to approach these limits.

5 Private Capacity of Quantum Channels

5.1 Private Capacity of a Wiretap Channel

Consider a three-user communication scenario involving Alice, Bob, and Eve, where Alice wishes to send message to Bob while ensuring privacy from Eve. The communication channel $\mathcal{N}$, called the wiretap channel, is defined by the conditional probability distribution $p_{Y,Z|X}(y,z|x)$. Here, X is the input random variable Alice controls, Y is Bob's received output, and Z is Eve's received output.

The information throughput may be expressed as $I(U;Y) - I(U;Z)$, where U is an auxiliary random variable chosen by Alice to optimize the input distribution $p_{U,X}(u,x)$. The expression thus captures the difference in the understanding between the sender and the receiver, and the sender and the eavesdropper. The private capacity $P(\mathcal{N})$ of a classical wiretap channel $\mathcal{N} = p_{Y,Z|X}$ is defined as the maximum achievable information throughput.

Definition 5.1 (Private Capacity of a Wiretap Channel)**.** *The private information $P(\mathcal{N})$ of a classical wiretap channel $\mathcal{N} = p_{Y,Z|X}$ is defined as:*

$$P(\mathcal{N}) = \max_{p_{U,X}(u,x)} [I(U;Y) - I(U;Z)]$$

We would like to examine the properties of this defined measure, and compare it against our intuitive understanding of channel capacity. This includes non-negativity – is the capacity always positive, additivity – does n usage of the channel result in n times the throughput, and asymptotic capacity – is the average throughput constant in the long run.

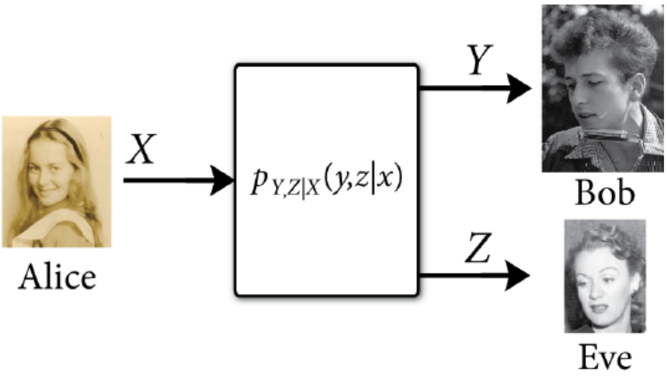


Figure 6: Model of a classical wiretap channel [6]. Alice would like to send messages to Bob while keeping them private from Eve.

5.2 Properties of Private Capacity of a Wiretap Channel

Theorem 2 (Non-negativity)**.** *The private capacity $P(\mathcal{N})$ of a wiretap channel $\mathcal{N}$ is non-negative.*

$$P(\mathcal{N}) \geq 0$$

Proof. By setting the joint density $p_{U,X}(u,x)$ to the degenerate distribution $\delta_{u,u_0}\delta_{x,x_0}$ for specific values u_0 and x_0, both mutual information terms $I(U;Y)$ and $I(U;Z)$ become zero, resulting in a difference of zero. Since $P(\mathcal{N})$ involves a maximization over all possible distributions $p_{U,X}(u,x)$, the value of $P(\mathcal{N})$ cannot be negative. □

Theorem 3 (Additivity)**.** *Let $\mathscr{N}_i$ represent a classical wiretap channel given by $p_{Y_i,Z_i|X_i}$. The private capacity of the combined classical wiretap channel $\mathscr{N}_1 \otimes \mathscr{N}_2$ is given by the sum of the individual private capacities of $\mathscr{N}_1$ and $\mathscr{N}_2$.*

$$P(\mathscr{N}_1 \otimes \mathscr{N}_2) = P(\mathscr{N}_1) + P(\mathscr{N}_2)$$

Proof. The inequality $P(\mathscr{N}_1 \otimes \mathscr{N}_2) \geq P(\mathscr{N}_1) + P(\mathscr{N}_2)$ follows from the fact that the two channels could be given by uncorrelated probability distributions. Since the maximization is over all possible probability distributions, the capacity would thus be atleast the throughput obtained with the uncorrelated probability distributions, which is equal to the sum of the individual probability distributions.

In order to prove the other direction, we consider any probability distribution $p_{U,X_1,X_2}(u,x_1,x_2)$ for $P(\mathscr{N}_1 \otimes \mathscr{N}_2)$, where X_1, X_2 are the input random variables transmitted to the two channels. We can write the combined channel capacity as follows.

$$\begin{aligned} P(\mathscr{N}_1 \otimes \mathscr{N}_2) &= I(U;Y_1Y_2) - I(U;Z_1Z_2) \\ &= I(U;Y_1) + I(U;Y_2|Y_1) - I(U;Z_2) - I(U;Z_1|Z_2) \\ &= I(U;Y_1|Z_2) - I(U;Z_1|Z_2) + I(U;Y_2|Y_1) - I(U;Z_2|Y_1) \end{aligned}$$

The first equality follows from the chain rule for mutual information, while the second follows from the following identity and some rearrangement of terms.

$$I(U;Y_1) - I(U;Z_2) = I(U;Y_1|Z_2) - I(U;Z_2|Y_1)$$

Consider the first two terms from the expanded form of combined channel capacity.

$$\begin{aligned} I(U;Y_1|Z_2) - I(U;Z_1|Z_2) &= \sum_{z_2} p_{Z_2}(z_2)[I(U;Y_1|Z_2 = z_2) - I(U;Z_1|Z_2 = z_2)] \\ &\leq \max_{z_2}[I(U;Y_1|Z_2 = z_2) - I(U;Z_1|Z_2 = z_2)] \\ &= P(\mathscr{N}_1) \end{aligned}$$

The last inequality holds because $U|Z_2$ can be viewed as an auxiliary random variable, resulting in the same form as observed in the definition of private capacity. Similar argument yields the following relation.

$$I(U;Y_2|Y_1) - I(U;Z_2|Y_1) \leq P(\mathscr{N}_2)$$

Combining the two relations proves the other side of the inequality. The two inequalities together establish the additivity relation.

$$P(\mathscr{N}_1 \otimes \mathscr{N}_2) = P(\mathscr{N}_1) + P(\mathscr{N}_2)$$

□

Theorem 4 (Equivalence of asymptotic and single use capacity)**.** *The asymptotic private capacity of a wiretap channel $\mathscr{N}$ is equal to its private capacity.*

$$\lim_{n\to\infty} \frac{1}{n} P(\mathscr{N}^{\otimes n}) = P(\mathscr{N})$$

Proof. Consider the following inductive hypothesis for any positive value of k. It trivially holds for $k = 1$.

$$P(\mathscr{N}^{\otimes k}) = \sum_{i=1}^{k} P(\mathscr{N})$$

$$= kP(\mathcal{N}_i)$$

Using the additivity relation and the inductive hypothesis, we can prove that the hypothesis holds for $k+1$.

$$\begin{aligned} P(\mathcal{N}^{\otimes k} \otimes \mathcal{N}_{k+1}) &= P(\mathcal{N}^{\otimes k}) + P(\mathcal{N}_{k+1}) \\ &= \sum_{i=1}^{k} P(\mathcal{N}_i) + P(\mathcal{N}_{k+1}) \\ &= \sum_{i=1}^{k+1} P(\mathcal{N}_i) \\ &= (k+1)P(\mathcal{N}) \end{aligned}$$

Moving the constant factor to the LHS and applying limit gives us the desired relation.

$$\lim_{n\to\infty} \frac{1}{n} P(\mathcal{N}^{\otimes n}) = P(\mathcal{N})$$

□

5.3 Private Capacity of a Quantum Channel

Consider a quantum channel $\mathcal{N}$ given by the operator $\mathcal{U}^{\mathcal{N}}_{A'\to BE}$. In the private communication model, Alice would like to establish classical correlations with Bob, without the environment of the channel having access to the correlations. Let the following be the density matrix of the classical-quantum state she transmits to the channel.

$$\rho_{XA'} = \sum_x p_X(x) |x\rangle\langle x|_X \otimes \rho^x_{A'}$$

The private capacity $P(\mathcal{N})$ of a quantum channel $\mathcal{N}$ is defined similar to the private capacity of a wiretap channel. Here, we will replace the classical mutual information with quantum mutual information, and the maximization will be over all the classical quantum states described earlier.

Definition 5.2 (Private Capacity of a Quantum Channel)**.** *The private information $P(\mathcal{N})$ of a quantum channel $\mathcal{N}$ is defined as:*

$$P(\mathcal{N}) = \max_{\rho_{XA'}} \left[I(X;B)_\rho - I(X;E)_\rho \right]$$

where the term being maximized is called the private information of the state.

We shall see that this definition of private capacity too satisifies non-negativity, but unlike the wiretap channel it violates the additivity relation.

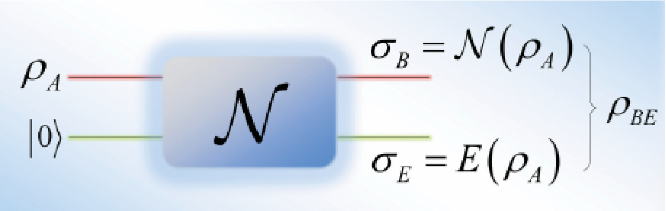


Figure 7: Private communication through a quantum channel [1].

5.4 Properties of Private Capacity of a Quantum Channel

Theorem 5 (Non-negativity)**.** *The private capacity $P(\mathscr{N})$ of a quantum channel $\mathscr{N}$ is non-negative.*

$$P(\mathscr{N}) \geq 0$$

Proof. Similar to the proof in the classical wiretap channel case, we can consider a degenerate input $\rho_{XA'}$ of the form $|0\rangle\langle 0|_X \otimes \psi_{A'}$, where $\psi_{A'}$ is a pure state. The private information of this state is zero. Since the maximization is over all possible input states, the private capacity of the channel is guaranteed to be atleast zero, or non-negative. □

Theorem 6 (Non-additivity)**.** *Let $\mathscr{N}_1$ and $\mathscr{N}_2$ represent two quantum channels. The private capacity of the combined quantum channel $\mathscr{N}_1 \otimes \mathscr{N}_2$ is not equal to the sum of the individual private capacities of $\mathscr{N}_1$ and $\mathscr{N}_2$.*

$$P(\mathscr{N}_1 \otimes \mathscr{N}_2) \neq P(\mathscr{N}_1) + P(\mathscr{N}_2)$$

Proof. We will construct a new channel $T_k^{\mathscr{N}}$, on the basis of an arbitrary channel $\mathscr{N}$, such that $C(T_k^{\mathscr{N}})$ is close to $C(\mathscr{N})$.

$$C(\mathscr{N}) \leq C(T_k^{\mathscr{N}}) \leq C(\mathscr{N}) + \delta(k)$$

Further, consider a 50% erasure channel $\mathscr{A}$. The quantum capacity of the combined channel $Q(T_k^N \otimes \mathscr{A})$ satisfies the following relation with the entanglement-assisted capacity $Q_E(\mathscr{N})$ [7].

$$Q(T_k^{\mathscr{N}} \otimes \mathscr{A}) \geq Q_E(\mathscr{N})$$

For channels $\mathscr{N}$ such that $Q_E(\mathscr{N}) > C(\mathscr{N})$, we can combine the two results as follows.

$$Q(T_k^{\mathscr{N}} \otimes \mathscr{A}) \geq Q_E(\mathscr{N}) > C(\mathscr{N}) \geq C(T_k^{\mathscr{N}})$$

As we shall prove in the subsequent section, the classical, private and quantum capacities of a quantum channel $\mathscr{N}$ are related as $C(\mathscr{N}) \geq P(\mathscr{N}) \geq Q(\mathscr{N})$. We can use this fact to write the following.

$$P(T_k^{\mathscr{N}} \otimes \mathscr{A}) \geq Q(T_k^{\mathscr{N}} \otimes \mathscr{A}) \geq Q_E(\mathscr{N}) > C(\mathscr{N}) \geq C(T_k^{\mathscr{N}}) \geq P(T_k^{\mathscr{N}})$$

Since erasure channels have zero private and quantum capacities, when $p \geq 1/2$, $P(\mathscr{A}) = Q(\mathscr{A}) = 0$ [8]. This gives us the following result, which violates the possibility of private capacity being an additive measure.

$$P(T_k^{\mathscr{N}} \otimes \mathscr{A}) > P(T_k^{\mathscr{N}}) + P(\mathscr{A})$$

Thus, the private capacity $P(\mathscr{N})$ of a quantum channel $\mathscr{N}$ is non-additive. □

6 Quantum Capacity of Quantum Channels

6.1 Coherent Information and Quantum Capacity

We now concern ourselves with the transmission of quantum states. Recall that both classical and private communication models concerned themselves with reliable transmission of classical bits. This was typically modelled with orthogonal basis states. However, transmission of arbitrary quantum states involve dealing with more complex phenomenon such as entanglement.

Suppose Alice prepares a pure state $\phi_{AA'}$ and inputs the system A' to a quantum channel $\mathscr{N}_{A'\to B}$. This transmission gives us the bipartite state $\rho_{AB} = \mathscr{N}_{A'\to B}(\phi_{AA'})$. An intuitive way of measuring the information throughput is to measure how uncertain the receiver is of the input given the received quantum state. This aligns with the definiton of conditional entropy. Since we are concerned with reducing the conditional entropy, we can define an analogous measure that can be maximized instead.

Definition 6.1 (Coherent Information of a Quantum State)**.** *The coherent information of a bipartite state $\rho_{AB} \in \mathscr{D}(\mathscr{H}_A \otimes \mathscr{H}_B)$ is defined as:*

$$I(A\rangle B)_\rho = S(B)_\rho - S(AB)_\rho$$

Definition 6.2 (Quantum Capacity of a Quantum Channel)**.** *The quantum information $Q(\mathscr{N})$ of a quantum channel $\mathscr{N}$ is defined as:*

$$\begin{aligned} Q(\mathscr{N}) &= \max_{\phi_{AA'}} \left[I(A\rangle B)_\rho \right] \\ &= \max_{\phi_{AA'}} \left[S(B)_\rho - S(AB)_\rho \right] \end{aligned}$$

An equivalent way of writing the quantum capacity is as follows, where $|\psi\rangle_{ABE} = U^{\mathscr{N}}_{A'\to BE}|\phi\rangle_{AA'}$.

$$Q(\mathscr{N}) = \max_{\phi_{AA'}} \left[S(B)_\psi - S(E)_\psi \right]$$

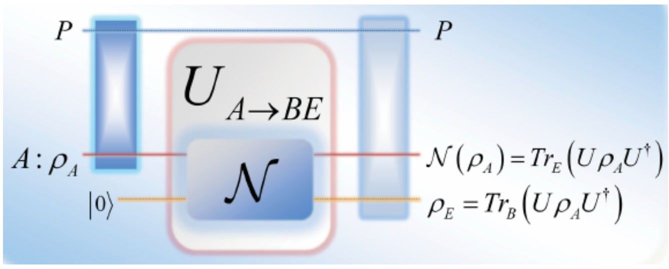

P
P
$U_{A\to BE}$
$A : \rho_A$
$|0\rangle$
$\mathcal{N}$
$\mathcal{N}(\rho_A) = Tr_E(U\rho_A U^\dagger)$
$\rho_E = Tr_B(U\rho_A U^\dagger)$

Figure 8: Quantum communication through a quantum channel [1]. The channel is represented as the unitary transformation. P refers to the reference system or the purification state. The outputs received by Bob (first) and the environment (second) can be computed using the partial trace operator.

We shall see that this definition of quantum capacity satisifies non-negativity, but violates the additivity relation similar to private capacity. We would then examine a special class of quantum channels that satisfies the additive relation for both private and quantum capacity. Moreover, their private and quantum capacities happen to be always equal.

6.2 Properties of Quantum Capacity

Theorem 7 (Non-negativity)**.** *The quantum capacity $Q(\mathscr{N})$ of a quantum channel $\mathscr{N}$ is non-negative.*

$$Q(\mathscr{N}) \geq 0$$

Proof. Similar to the non-negativity proofs earlier, we would construct a state that achieves zero coherent information. Since the quantum capacity is the maximum over all such states, it is guaranteed to be non-negative.

Consider the input state $\phi_{AA'}$ to be a product state of the form $\psi_A \otimes \Phi_{A'}$, where the state A is pure. We can evaluate its coherent information as follows.

$$\begin{aligned} I(A\rangle B)_{\psi\otimes\mathscr{N}(\Phi)} &= S(B)_{\mathscr{N}(\Phi)} - S(AB)_{\psi\otimes\mathscr{N}(\Phi)} \\ &= S(B)_{\mathscr{N}(\Phi)} - S(A)_{\psi} - S(B)_{\mathscr{N}(\Phi)} \\ &= -S(A)_{\psi} \\ &= 0 \end{aligned}$$

The first equality follows from the definition of coherent information. The second makes use of the fact that AB is a product state. The third cancels out common terms, while the fourth makes use of the fact that A is a pure state. □

Theorem 8 (Non-additivity)**.** *Let $\mathscr{N}_1$ and $\mathscr{N}_2$ represent two quantum channels. The quantum capacity of the combined quantum channel $\mathscr{N}_1 \otimes \mathscr{N}_2$ is not equal to the sum of the individual quantum capacities of $\mathscr{N}_1$ and $\mathscr{N}_2$.*

$$Q(\mathscr{N}_1 \otimes \mathscr{N}_2) \neq Q(\mathscr{N}_1) + Q(\mathscr{N}_2)$$

Proof. Consider a private Horodecki channel $\mathscr{N}_H$ and a symmetric channel of unbounded dimension $\mathscr{A}$. These channels are known to have the following properties [9] [10] [11].

$$P(\mathscr{N}_H) > 0, Q(\mathscr{N}_H) = 0, Q(\mathscr{A}) = 0$$

Further, we will utilize the following relation between the private capacity and the assisted capacity [12].

$$\frac{1}{2}P(N_H) \leq Q_{\mathscr{A}}(N_H)$$

Since $P(\mathscr{N}_H) > 0$, it follows that the assisted capacity $Q_{\mathscr{A}}(\mathscr{N}_H) > 0$. Further, since $\mathscr{A}$ is a symmetric channel of unbounded dimension, the quantum capacity of the joint channel can be related with the assisted capacity as follows [11].

$$Q(\mathscr{N}_H \times \mathscr{A}) = Q_{\mathscr{A}}(\mathscr{N}_H) > 0$$

Finally, note that $Q(\mathscr{N}_H) + Q(\mathscr{A}) = 0$. We can thus conclude the following relation establishing non-additivity.

$$Q(\mathscr{N}_H \times \mathscr{A}) \neq Q(\mathscr{N}_H) + Q(\mathscr{A})$$

□

Theorem 9 (Relation with other capacities)**.** *The single use capacities of a quantum channel $\mathscr{N}$ are related to each other as:*

$$Q(\mathscr{N}) \leq P(\mathscr{N}) \leq C(\mathscr{N})$$

Proof. The reason why $P(\mathcal{N}) \leq C(\mathcal{N})$ holds is easy to see. In both classical and private communication models, our objective is to transmit classical information. Private communication is a specific case of the lesser constrained classical communication through quantum channels. Thus, the classical capacity cannot be lesser than the private capacity.
For proving the other half of the inequality, i.e. $Q(\mathcal{N}) \leq P(\mathcal{N})$, we will consider a pure state that maximizes the quantum capacity. Let $\sigma_{XA'}$ denote the augmented classical-quantum state that correlates with index x.

$$\sigma_{XA'} = \sum_x p_X(x)|x\rangle\langle x| \otimes |\phi_x\rangle\langle\phi_x|_{A'}$$

$$\begin{aligned} Q(\mathcal{N}) &= \max_\rho \left[S(B)_\rho - S(E)_\rho\right] \\ &= S(B)_\sigma - S(E)_\sigma \\ &= S(B)_\sigma - S(B|X)_\sigma - S(E)_\sigma + S(B|X)_\sigma \\ &= I(X;B)_\sigma - S(E)_\sigma + S(E|X)_\sigma \\ &= I(X;B)_\sigma - I(X;E)_\sigma \\ &\leq \max_\rho \left[I(X;B)_\rho - I(X;E)_\rho\right] \\ &= P(\mathcal{N}) \end{aligned}$$

□

The first equality is by definition of quantum capacity. The second equality is by construction of the quantum state. The fourth equality is by the definition of mutual information and by the fact that σ on systems B and E is pure when conditioned on X. The fifth equality is by the definition of mutual information. The last equality is by the definition of private capacity.

6.3 Degradable Quantum Channels

A channel $\mathcal{N} : \mathcal{H}_A \rightarrow \mathcal{H}_B$ can be defined by an isometric embedding $U : \mathcal{H}_A \rightarrow \mathcal{H}_B \otimes \mathcal{H}_E$, followed by a partial trace over the environment system E: $\mathcal{N}(\rho) = \mathrm{Tr}_E U(\rho)$. The complementary channel $\mathcal{N}^c : \mathcal{H}_A \rightarrow \mathcal{H}_E$ is defined by: $\mathcal{N}^c(\rho) = \mathrm{Tr}_B U(\rho)$. We regard a quantum channel as degradable if the transmission from the sender end to the environment is noisier than the transmission to the receiver. This is modeled as a result of an additional application of a “degrading” channel.

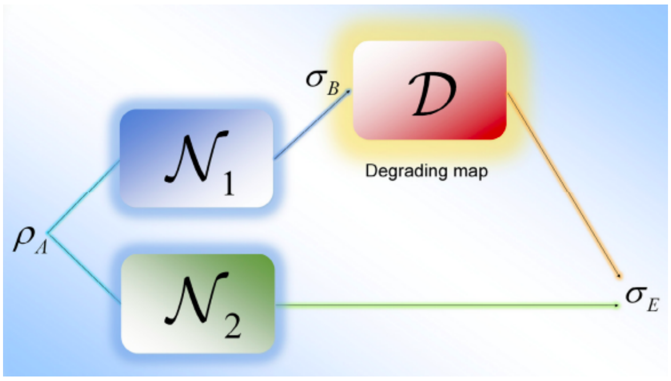


Figure 9: The concept of a degradable quantum channel [4]. The environment state can be simulated by the application of the degrading channel on the quantum state received by Bob.

Definition 6.3 (Degradable Quantum Channel)**.** *A channel $\mathcal{N} : \mathcal{H}_A \to \mathcal{H}_B$ is called degradable when it may be degraded to its complementary channel $\mathcal{N}^c : \mathcal{H}_A \to \mathcal{H}_E$, i.e. when there exists a CPTP map $T : \mathcal{H}_B \to \mathcal{H}_E$ such that*

$$\mathcal{N}^c = T \circ \mathcal{N}$$

Theorem 10 (Additivity of Quantum Capacity)**.** *Let $\mathcal{N}$ and $\mathcal{M}$ represent two degradable quantum channels. The quantum capacity of the combined quantum channel $\mathcal{N} \otimes \mathcal{M}$ is equal to the sum of the individual quantum capacities of $\mathcal{N}$ and $\mathcal{M}$.*

$$Q(\mathcal{N} \otimes \mathcal{M}) = Q(\mathcal{N}) + Q(\mathcal{M})$$

Proof. The reason why $Q(\mathcal{N} \otimes \mathcal{M}) \geq Q(\mathcal{N}) + Q(\mathcal{M})$ holds is easy to see. We can think of a separable input state $\rho_{AB} = \rho_A \otimes \rho_B$ that is maximized indepedently for the two channels. Since the entropy is additive for such states, the joint capacity cannot be lower than the sum of invididual capacities.
For the reverse direction, consider a pure state $\phi_{AA'_1A'_2}$ as the input to the two channels. Let the following denote the result of the channels, where $\rho_{AB_1E_1B_2E_2}$ is a state that maximizes $Q(\mathcal{N} \otimes \mathcal{M})$.

$$\begin{aligned}
\sigma_{AB_1E_1A_2} &= U^N \phi (U^N)^\dagger \\
\theta_{AA_1B_1E_2} &= U^M \phi (U^M)^\dagger \\
\rho_{AB_1E_1B_2E_2} &= (U^N \otimes U^M) \phi ((U^N)^\dagger \otimes (U^M)^\dagger)
\end{aligned}$$

We can write the quantum capacity of the joint channel as follows to show the reverse direction.

$$\begin{aligned}
Q(\mathcal{N} \otimes \mathcal{M}) &= I(A; B_1B_2)_\rho \\
&= S(B_1B_2)_\rho - S(E_1E_2)_\rho \\
&= S(B_1)_\rho - S(E_1)_\rho + S(B_2)_\rho - S(E_2)_\rho - [I(B_1;B_2)_\rho - I(E_1;E_2)_\rho] \\
&\leq S(B_1)_\rho - S(E_1)_\rho + S(B_2)_\rho - S(E_2)_\rho \\
&= S(B_1)_\sigma - S(AA'_2B_1)_\sigma + S(B_2)_\sigma - S(AA'_1B_2)_\sigma \\
&= I(AA'_2; B_1)_\sigma + I(AA'_1; B_2)_\sigma \\
&\leq Q(\mathcal{N}) + Q(\mathcal{M})
\end{aligned}$$

The fourth inequality holds because there is a degrading channel from both B_1 to E_1 and B_2 to E_2. This gives us $I(B_1;B_2)_\rho \geq I(E_1;E_2)_\rho$. The fifth inequality holds because the state σ on systems $AA'_2B_1E_1$ is pure and the state θ on systems $AA'_1B_2E_2$ is pure.
Combining the two inequalities establishes the additivity of quantum capacity of degradable channels. □

Theorem 11 (Additivity of Private Capacity)**.** *Let $\mathcal{N}$ and $\mathcal{M}$ represent two degradable quantum channels. The private capacity of the combined quantum channel $\mathcal{N} \otimes \mathcal{M}$ is equal to the sum of the individual private capacities of $\mathcal{N}$ and $\mathcal{M}$.*

$$P(\mathcal{N} \otimes \mathcal{M}) = P(\mathcal{N}) + P(\mathcal{M})$$

Proof. To prove $P(\mathcal{N} \otimes \mathcal{M}) \geq P(\mathcal{N}) + P(\mathcal{M})$, consider the states ρ and σ maximizing the private capacity of channels $\mathcal{N}$ and $\mathcal{M}$ respectively. Let $\theta = \rho \otimes \sigma$ be the tensor product of the two states. Then using the additivity of mutual information on tensor product states, we can write the following.

$$P(\mathcal{N}) + P(\mathcal{M}) = I(X_1; B_1)_\rho - I(X_1; E_1)_\rho + I(X_2; B_2)_\sigma - I(X_2; E_2)_\sigma$$

$$
\begin{aligned}
&= I(X_1X_2;B_1B_2)_\theta - I(X_1X_2;E_1E_2)_\theta \\
&\leq \max_\alpha \left[I(X_1X_2;B_1B_2)_\alpha - I(X_1X_2;E_1E_2)_\alpha\right] \\
&= P(\mathscr{N} \otimes \mathscr{M})
\end{aligned}
$$

To prove the other direction, consider a state σ that maximizes $P(\mathscr{N} \otimes \mathscr{M})$ and let ρ be the state that arises from sending σ through the joint channel. We can evaluate the quantum capacity of the joint channel as follows.

$$
\begin{aligned}
P(\mathscr{N} \otimes \mathscr{M}) &= I(X;B_1B_2)_\sigma - I(X;E_1E_2)_\sigma \\
&= I(XY;B_1B_2)_\sigma - I(XY;E_1E_2)_\sigma - [I(Y;B_1B_2|X)_\sigma - I(Y;E_1E_2|X)_\sigma] \\
&\leq I(XY;B_1B_2)_\sigma - I(XY;E_1E_2)_\sigma \\
&= S(B_1B_2)_\sigma - S(B_1B_2|XY)_\sigma - S(E_1E_2)_\sigma + S(E_1E_2|XY)_\sigma \\
&= S(B_1B_2)_\sigma - S(B_1B_2|XY)_\sigma - S(E_1E_2)_\sigma + S(B_1B_2|XY)_\sigma \\
&= S(B_1B_2)_\sigma - S(E_1E_2)_\sigma \\
&= S(B_1)_\sigma - S(E_1)_\sigma + S(B_2)_\sigma - S(E_2)_\sigma - [I(B_1;B_2)_\sigma - I(E_1;E_2)_\sigma] \\
&\leq S(B_1)_\sigma - S(E_1)_\sigma + S(B_2)_\sigma - S(E_2)_\sigma \\
&\leq \max_\alpha [S(B_1)_\alpha - S(E_1)_\alpha] + \max_\beta \left[S(B_2)_\beta - S(E_2)_\beta\right] \\
&= Q(\mathscr{N}) + Q(\mathscr{M}) \\
&= P(\mathscr{N}) + P(\mathscr{M})
\end{aligned}
$$

The second equality holds because of the chain rule for mutual information. The third inequality holds because $I(Y;B_1B_2|X)_\sigma \geq I(Y;E_1E_2|X)_\sigma$ for degradable channels. The fourth inequality uses the definition of mutual information. The fifth inequality uses the fact that state σ on systems $B_1B_2E_1E_2$ is pure when conditioning on classical systems X and Y. The seventh inequality uses the definition of mutual information. The second last inequality uses the definition of quantum capacity. The last inequality uses the equivalence of private and quantum capacity for degradable channels (proved later).
Combining the two inequalities establishes the additivity of private capacity of degradable channels. □

Theorem 12 (Equivalence of Private and Quantum Capacity)**.** *For a degradable quantum channel $\mathscr{N}$, the private capacity is equal to the quantum capacity.*

$$P(\mathscr{N}) = Q(\mathscr{N})$$

Proof. We will prove $P(\mathscr{N}) \leq Q(\mathscr{N})$ for any degradable channel $\mathscr{N}$. Consider a state σ that maximizes the private capacity of the channel $\mathscr{N}$. We can write its private capacity as follows.

$$
\begin{aligned}
P(\mathscr{N}) &= I(X;B)_\sigma - I(X;E)_\sigma \\
&= I(XY;B)_\sigma - I(Y;B|X)_\sigma - [I(XY;E)_\sigma - I(Y;E|X)_\sigma] \\
&= I(XY;B)_\sigma - I(XY;E)_\sigma - [I(Y;B|X)_\sigma - I(Y;E|X)_\sigma] \\
&\leq I(XY;B)_\sigma - I(XY;E)_\sigma \\
&= S(B)_\sigma - S(B|XY)_\sigma - S(E)_\sigma + S(E|XY)_\sigma \\
&= S(B)_\sigma - S(B|XY)_\sigma - S(E)_\sigma + S(B|XY)_\sigma \\
&= S(B)_\sigma - S(E)_\sigma
\end{aligned}
$$

$$\begin{aligned}
&\leq \max_{\phi} \left[S(B)_{\phi]} - S(E)_{\phi}\right] \\
&= Q(\mathscr{N})
\end{aligned}$$

The second inequality follows from the chain rule of mutual information. The fourth inequality holds because of the property of the degrading channel. The fifth inequality follows from the definition of mutual information. The sixth inequality holds because the state σ on systems B and E is pure when conditioned on classical systems X and Y. The last inequality follows from the definiton of quantum capacity of a channel.

Combining this inequality with the inequality proved earlier for all quantum channels, i.e. $Q(\mathscr{N}) \leq P(\mathscr{N}) \leq C(\mathscr{N})$, we can conclude that $P(\mathscr{N}) = Q(\mathscr{N})$ for degradable channels. □

7 Super Additivity of Quantum Channels

Superadditivity in quantum information theory refers to the phenomenon where certain types of quantum capacities increase when multiple channels (or multiple uses of the same channel) are considered, rather than treating each channel (or use) independently. This property arises because quantum entanglement and correlations can enhance the effectiveness of the channel when used in a collective or joint manner. Superadditivity is a fundamental property that distinguishes quantum information theory from classical information theory.

This Superadditivity is the result of entanglement and other exotic correlation phenomena that arises in quantum systems. Superadditivity enhances communication capabilities of quantum channels and allows more communication than what their classical counterparts could (sometimes even through zero capacity channels).

7.1 Classical Capacity

One manifestation of superadditivity appears in the classical capacity of quantum channels, particularly through the Holevo information. As we have seen before, the Holevo Information of a quantum channel is an upper bound on the classical information that can be transmitted through it. Also as we have seen before, Classical capacity has been defined in terms of the Holevo information. As a result, it can exhibit superadditivity when multiple channel uses are considered together. For example, when encoding information across multiple uses of the channel in an entangled way or by using complex encoding schemes, the total classical capacity often exceeds what would be achievable by simply using each channel separately. Thus, superadditivity allows us to exploit collective encoding and decoding strategies to maximize the amount of classical information extracted from quantum channels.

Although it is worth noting that we don't yet know the precise mathematics behind why Holevo information is superadditive, we just know from counter examples that additivity is not the general case and super additivity can exist.

Mathematically, we write the Holevo information of two channels acting in parallel as:

$$\chi(\mathscr{N}\otimes\mathscr{M}) \geq \chi(\mathscr{N})+\chi(\mathscr{M})$$

Now, since classical capacity is:

$$C(\mathscr{N}) = \lim_{n\to\infty}\frac{1}{n}\chi(\mathscr{N}^{\otimes n})$$

For channel $\mathscr{N}\otimes\mathscr{M}$ we have:

$$\begin{aligned} C(\mathscr{N}\otimes\mathscr{M}) &= \lim_{n\to\infty}\frac{1}{n}\chi[(\mathscr{N}\otimes\mathscr{M})^{\otimes n}] \\ &\geq \lim_{n\to\infty}\frac{1}{n}[\chi(\mathscr{N}^{\otimes n})+\chi(\mathscr{M}^{\otimes n})] \\ &= \lim_{n\to\infty}\frac{1}{n}\chi(\mathscr{N}^{\otimes n})+\lim_{n\to\infty}\frac{1}{n}\chi(\mathscr{M}^{\otimes n}) \\ &= C(\mathscr{N}+\mathscr{M}) \end{aligned}$$

As the math clearly demonstrates, the Classical capacity is superadditive as a consequence of the superadditivity of the Holevo information.

7.2 Quantum Capacity

The concept of superadditivity is particularly important when discussing the quantum capacity of a channel, which measures the maximum rate at which quantum information can be transmitted reliably. Similar to classical capacity, quantum capacity does not exhibit simple additivity; instead, two or more uses of a channel can yield an overall capacity greater than the sum of the individual capacities. This phenomenon is linked to the coherent information of the channel, similar to how the classical capacity is linked to the Holevo Information of a channel. Since coherent information itself can be superadditive, the overall quantum capacity of a channel may similarly display superadditivity.

Mathematically, we may write this as:

$$Q(\mathscr{N} \otimes \mathscr{M}) \geq Q(\mathscr{N}) + Q(\mathscr{M})$$

One classical analogy would be to consider a system where combining two noisy communication lines produces a more reliable transmission than would be expected from their individual performance, but in quantum mechanics, this is achieved by leveraging entanglement and quantum correlations.

7.3 Private Capacity

It is interesting to see that for quantum channels, not only the quantum and classical capabilities but also the private capacity can be superadditive! As demonstrated by [13], there is possibility for two channels with limited private capacity to combine as a tensor product channel to have private capacity greater than the sum of the two. We do not know the exact formulation of this relation as of now, so we write:

$$P^{?}(\mathscr{N} \otimes \mathscr{M}) > P^{?}(\mathscr{N}) + P^{?}(\mathscr{M})$$

The superadditivity of quantum channels is still a pretty unexplored topic, and there are many open questions still unanswered.

8 Examples of Quantum Channels

8.1 Capacity

We provide two examples of quantum channels here, and calculate various metrics of their capacity:

Erasure Quantum Channel The erasure quantum channel $\mathscr{N}_p$ ”erases” the input state ρ with probability p or transmits the state unchanged with probability $(1-p)$

$$\mathscr{N}_p(\rho) \to (1-p)\rho + (p|e\rangle\langle e|)$$

where $|e\rangle$ is the ”erasure state”. The classical capacity of this channel is given by:

$$C(\mathscr{N}_p) = (1-p)\log(d)$$

Here d is the dimension of the input system. This demonstrates that the channel can transmit some classical information only for $0 \leq p < 1$, otherwise the classical capacity is 0.

On the other hand, its quantum capacity is:

$$Q(\mathscr{N}_p) = (1-2p)\log(d)$$

This is similar to classical capacity, except that information can only be transferred if $0 \leq p < 1/2$.

Phase Erasure channel The Phase Erasure channel ”erases” (as the name indicates) the phase of its inputs with probability p but does not affect the amplitude. The map is expressed as:

$$\mathscr{N}(\rho) \to (1-p)\rho \otimes |0\rangle\langle 0| + p\frac{\rho + Z\rho Z^{\dagger}}{2} \otimes |1\rangle\langle 1|$$

The classical capacity of this channel is:

$$C(\mathscr{N}) = 1$$

The quantum capacity of this channel is:

$$Q(\mathscr{N}) = (1-p)\log(d)$$

where the variables hold the same meaning as the last example.

8.2 Superadditivity

A fascinating example of superadditivity in quantum channels involves the concept of **superactivation**. Superactivation refers to the phenomenon where two quantum channels, each with zero individual quantum capacity, can achieve a positive quantum capacity when used together. This behavior is counterintuitive, and quite the opposite of what you would expect if we were dealing with classical channels instead. We provide an example here:

> In a 2008 study by Graeme Smith and Jon Yard [12], the authors demonstrated the existence of two zero-capacity quantum channels, each individually incapable of transmitting quantum information, which, when used jointly, could support the reliable transmission of quantum information. The study highlights that certain quantum channels, despite having no individual capacity to send quantum data, can ”activate” each other under joint use, allowing

> information to pass through the combined channel setup. This unique behavior is not observed in classical communication channels, where combining two zero-capacity channels will always result in zero capacity for the combined channel.
>
> Smith and Yard's example involved two specific types of quantum channels known as **Private Horodecki channels** and **Symmetric channels**. Private Horodecki channels are designed to have zero quantum capacity while still exhibiting some level of classical and private capacities. Symmetric channels, on the other hand, have zero capacity for both classical and quantum transmission due to symmetry constraints that prevent any meaningful transmission of information (transmission through the symmetric channel would violate the No-cloning theorem). When these two types of channels are used together, however, they demonstrate a form of superadditivity: the combined quantum channel has a positive capacity, enabling it to transmit quantum information reliably.
>
> The mechanics behind superactivation stem from quantum interactions between the two channel types. The private Horodecki channel alone cannot maintain quantum coherence due to limitations related to its structure, while the symmetric channel alone is constrained by its symmetries. However, when these channels are combined in a particular way, each channel's limitations effectively "cancel out" the limitations of the other, allowing a positive quantum capacity to emerge.

This example ilustrates an astounding result in Quantum Information Theory. Not only is it possible for quantum channels to have superadditive capacities, but it is also possible for channels with Zero capacities to combine to form channels with Non-Zero capacities!

This bizzare result is something worth pondering upon, and in our opinion, the perfect food for thought to end this Term paper on.

9 References

[1] L. Gyongyosi, S. Imre, and H. V. Nguyen, “A survey on quantum channel capacities,” *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 1149–1205, 2018.

[2] Anonymous, “Shannon’s noisy channel-coding theorem.”

[3] B. Schumacher, “Quantum coding,” *Phys. Rev. A*, vol. 51, pp. 2738–2747, Apr 1995.

[4] L. Gyongyosi and S. Imre, “Properties of the quantum channel,” *ArXiv*, vol. abs/1208.1270, 2012.

[5] C. E. Shannon, “A mathematical theory of communication,” *The Bell System Technical Journal*, vol. 27, no. 3, pp. 379–423, 1948.

[6] M. M. Wilde, *Quantum Information Theory*. Cambridge University Press, Nov. 2016.

[7] C. H. Bennett, P. W. Shor, J. A. Smolin, and A. V. Thapliyal, “Entanglement-assisted classical capacity of noisy quantum channels,” *Phys. Rev. Lett.*, vol. 83, pp. 3081–3084, Oct 1999.

[8] C. H. Bennett, D. P. DiVincenzo, and J. A. Smolin, “Capacities of quantum erasure channels,” *Phys. Rev. Lett.*, vol. 78, pp. 3217–3220, Apr 1997.

[9] M. Horodecki, P. Horodecki, and R. Horodecki, “Separability of mixed states: necessary and sufficient conditions,” *Physics Letters A*, vol. 223, pp. 1–8, Nov. 1996.

[10] P. Horodecki, “Separability criterion and inseparable mixed states with positive partial transposition,” *Physics Letters A*, vol. 232, pp. 333–339, Aug. 1997.

[11] G. Smith, J. A. Smolin, and A. Winter, “The quantum capacity with symmetric side channels,” *IEEE Transactions on Information Theory*, vol. 54, pp. 4208–4217, Sept. 2008.

[12] G. Smith and J. Yard, “Quantum communication with zero-capacity channels,” *Science*, vol. 321, p. 1812–1815, Sept. 2008.

[13] K. Li, A. Winter, X. Zou, and G. Guo, “Private capacity of quantum channels is not additive,” *Physical Review Letters*, vol. 103, Sept. 2009.