Lecture 6: Classical Communication over quantum channels

Javier R. Fonollosa

Universitat Politècnica de Catalunya

javier.fonollosa@upc.edu

October 19, 2023

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Departament de Teoria del Senyal
i Comunicacions

UNIVERSITAT POLITÈCNICA DE CATALUNYA
BARCELONATECH
Escola Tècnica Superior d'Enginyeria
de Telecomunicació de Barcelona

Overview

Introduction

Introduction

- We begin by introducing a **Discrete Memoryless Channel** (DMC) model for the transmission of **classical information** making use of one **preparation** channel, an arbitrary **quantum channel** and a **measurement** channel.
- The **capacity** of the resulting DMC, the **accessible information**, is then formulated as a (highly complex) optimization problem.
- The **quantum data processing inequality** is applied to derive an upper bound, the **Holevo information** of the quantum channel which is the solution to a **simpler** optimization.
- We then analyze the **additivity** of the **Holevo information** of a channel to find out that the **additivity** that characterizes the mutual information and private information of **classical** channels is not preserved (in general) in **quantum**.
- We finish presenting the **capacity** for the transmission of **classical** information over quantum channels or HSW Theorem, and some **examples**.

Accesible information

DMC formulation for classical communication

- The transmission of Classical information over quantum channels can be modeled as a DMC by a **preparation** stage in which a quantum state is generated according to a given ensemble $\mathcal{E} \equiv \{p_X(x), \rho_A^x\}$:

$$\rho_{XA} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_A^x, \;\; \rho_A = \sum_x p_X(x)\rho_A^x.$$

- The modulated quantum state is **transmitted** through the channel $\mathcal{N}_{A\to B}$:

$$\rho_{XB} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \mathcal{N}_{A\to B}(\rho_A^x), \;\; \rho_B = \sum_x p_X(x)\mathcal{N}_{A\to B}(\rho_A^x).$$

- A **measurement POVM** is applied to the state at B to yield Y:

$$\begin{aligned}\rho_{XY} &= \sum_{x,y} p_X(x)|x\rangle\langle x|_X \otimes \mathrm{tr}\{\Lambda_y \mathcal{N}_{A\to B}(\rho_A^x)\}|y\rangle\langle y|_Y \\ &= \sum_{x,y} p_X(x)p_{Y|X}(y|x)|x\rangle\langle x|_X \otimes |y\rangle\langle y|_Y.\end{aligned}$$

Accesible information

- Note that this procedure makes **one use** of channel $\mathcal{N}_{A\to B}$ for each DMC use thus does **not exploit any potential entanglement** between the transmitted states at A or any **joint measurement** of the received states at B.
- Incorporating the effect of $\mathcal{N}_{A\to B}$ in the conditional pmf $p_{Y|X}(y|x)$, the communication limits are **well known** (Shannon Capacity), and the DMC channel capacity will be given by the **maximization** of $I(X;Y)$ with respect to the **ensemble** $\mathcal{E} = \{p_X(x), \rho_A^x\}$ and the **measurement** POVM $\{\Lambda_y\}$.

Accesible information $I_{\mathrm{acc}}(\mathcal{E})$ and I^*_{acc}

The accesible information is:

$$I_{\mathrm{acc}}(\mathcal{E}) = \max_{\{\Lambda_y\}} I(X;Y),$$

$$I^*_{\mathrm{acc}} = \max_{\mathcal{E}} I_{\mathrm{acc}}(\mathcal{E}) = \max_{\mathcal{E},\{\Lambda_y\}} I(X;Y).$$

Accesible information upper bound

- **Computation** of the **accessible information** $I_{\text{acc}}(\mathcal{E})$ or I^*_{acc} is **difficult** and would still be subject to **additional optimization** regarding the general strategy (possibly more that one quantum channel use per DMC use - discussion on this issue follows later).
- The quantum data processing inequality $I(A;B)_\rho \geq I(\mathcal{N}(A);\mathcal{M}(B))_\sigma$ provides an upper bound to the accessible information noting that Y is obtained applying a measurement channel to the state at B.

Upper bound to the accesible information

The accessible information of $\mathcal{E} \equiv \{p_X(x), \rho_A^x\}$ when transmitted through $\mathcal{N}_{A\to B}$ is upper bounded by the quantum mutual information $I(X;B)_\rho$:

$$I_{\text{acc}}(\mathcal{E}) = \max_{\{\Lambda_y\}} I(X;Y) \leq I(X;B)_\rho = H(\rho_B) - \sum_x p_X(x) H(\rho_B^x),$$

where $\rho_B = \sum_x p_X(x)\rho_B^x = \sum_x p_X(x)\mathcal{N}_{A\to B}(\rho_A^x)$.

The information of quantum channels

Holevo information of a quantum channel

Holevo information of a quantum channel

The Holevo information $\chi(\mathcal{N})$ of channel $\mathcal{N}_{A\to B}$ is:

$$\chi(\mathcal{N}) \equiv \max_{\rho_{XA}} I(X;B)_\rho \geq \max_{\mathcal{E}} I_{\text{acc}}(\mathcal{E}) = I^*_{\text{acc}}$$

where $\mathcal{E} = \{p_X(x), \rho_A^x\}$ and

$$\rho_{XA} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_A^x,$$
$$\rho_{XB} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \mathcal{N}_{A\to B}(\rho_A^x).$$

The **Holevo Information of a quantum channel** provides an upper bound to the **accessible information**.

Computation of the Holevo information (I)

Pure states are sufficient

The Holevo information of channel $\mathcal{N}_{A\to B}$ can be obtained restricting the maximization to pure states

$$\chi(\mathcal{N}) = \max_{\rho_{XA}} I(X;B)_\rho = \max_{\tau_{XA}} I(X;B)_\tau,$$

where ρ_{XA} and ρ_{XB} are defined as in previous slide and

$$\tau_{XA} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes |\phi_x\rangle\langle\phi_x|_A,$$

and

$$\tau_{XB} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \mathcal{N}_{A\to B}(|\phi_x\rangle\langle\phi_x|_A).$$

Computation of the Holevo information (and II)

The proof follows considering the spectral decomposition,

$$\rho_A^x = \sum_z p_{Z|X}(z|x)|\psi_{x,z}\rangle\langle\psi_{x,z}|_A,$$

defining

$$\sigma_{XZA} = \sum_{x,z} p_X(x)p_{Z|X}(z|x)|x\rangle\langle x|_X \otimes |z\rangle\langle z|_Z \otimes |\psi_{x,z}\rangle\langle\psi_{x,z}|_A,$$

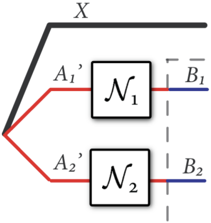
where $\rho_{XA} = \mathrm{tr}_Z\{\sigma_{XZA}\}$. Note that σ_{XZA} is a state of the form τ_{XA} with joint random variable $\mathcal{X} \times \mathcal{Z}$ as the classical part. Let σ_{XZB} be the state that results from sending A through $\mathcal{N}_{A\to B}$:

$$I(X;B)_\rho = I(X;B)_\sigma \leq I(XZ;B)_\sigma.$$

The equality follows from $\rho_{XB} = \mathrm{tr}_Z\{\sigma_{XZB}\}$ and the inequality from the quantum data processing inequality. ■

Superadditivity of the Holevo information (I)

- A more general strategy for the transmission of X through quantum channel $\mathcal{N}_{A\to B}$ is by using more than one quantum channel uses for each x, for instance for **two** quantum channel uses.



$$\rho_{XA_1A_2} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho^x_{A_1A_2},$$

the joint state between X and B is

$$\rho_{XB_1B_2} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes (\mathcal{N}_{A_1\to B_1} \otimes \mathcal{N}_{A_2\to B_2})(\rho^x_{A_1A_2}).$$

Superadditivity of the Holevo information (II)

- Is the Holevo information of the product channel $\mathcal{N}_{A_1\to B_1} \otimes \mathcal{N}_{A_2\to B_2}$ equal to twice the Holevo information of the channel $\mathcal{N}_{A\to B}$?

$$\chi(\mathcal{N}_{A_1\to B_1} \otimes \mathcal{N}_{A_2\to B_2}) \stackrel{?}{=} \chi(\mathcal{N}_{A_1\to B_1}) + \chi(\mathcal{N}_{A_2\to B_2}) = 2\chi(\mathcal{N}_{A\to B}),$$

or

$$\max_{\rho_{XA_1A_2}} I(X;B_1B_2)_\rho \stackrel{?}{=} 2\max_{\rho_{XA}} I(X;B)_\rho,$$

or, in general, for arbitrary channels $\mathcal{N}$ and $\mathcal{M}$,

$$\chi(\mathcal{N}_{A_1\to B_1} \otimes \mathcal{M}_{A_2\to B_2}) \stackrel{?}{=} \chi(\mathcal{N}_{A_1\to B_1}) + \chi(\mathcal{M}_{A_2\to B_2}).$$

Superadditivity of the Holevo information (III)

- Since independent coding is a particular case of joint coding the Holevo information $\chi(\mathcal{N} \otimes \mathcal{M})$ is at least superadditive:

$$\chi(\mathcal{N}_{A_1 \to B_1} \otimes \mathcal{M}_{A_2 \to B_2}) \geq \chi(\mathcal{N}_{A_1 \to B_1}) + \chi(\mathcal{M}_{A_2 \to B_2}).$$

- It was believed to be additive for many years and it is indeed additive for many channels. However, a counterexample demonstrated **strict** superadditivity [Hastings, 2009].

$$\chi(\mathcal{N} \otimes \mathcal{M}) > \chi(\mathcal{N}) + \chi(\mathcal{M}),$$

which yields the **regularized** definition (possibly with strict inequality),

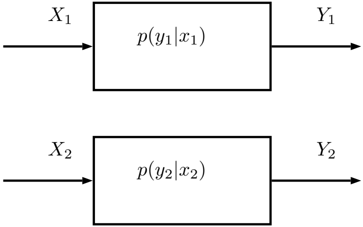
$$\chi_{\mathrm{reg}}(\mathcal{N}) \equiv \lim_{n \to \infty} \frac{1}{n} \chi(\mathcal{N}^{\otimes n}) \geq \chi(\mathcal{N}).$$

- This property has no counterpart in classical information theory where the capacity of the product DMC is additive.

Additivity in classical channels

For two **parallel** DMCs, $p(y_1, y_2|x_1, x_2) = p(y_1|x_1)p(y_2|x_2)$ each of which is characterized by a given **capacity** C_1 and C_2. The **total capacity** C is **equal** to $C_1 + C_2$:

$$C = \max_{p(x_1,x_2)} I(X_1, X_2; Y_1, Y_2) = \max_{p(x_1)} I(X_1; Y_1) + \max_{p(x_2)} I(X_2; Y_2) = C_1 + C_2$$



The HSW theorem

The Holevo-Schumacher-Westmoreland theorem (I)

The Holevo-Schumacher-Westmoreland theorem

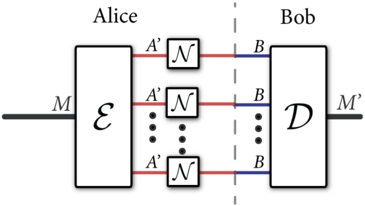
The classical capacity of a quantum channel is equal to the regularization of the Holevo information of the channel:

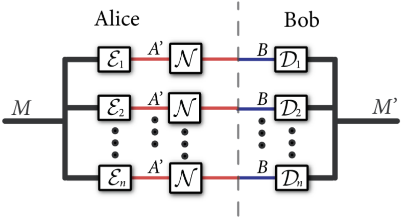
$$\mathrm{C}(\mathcal{N}) = \lim_{k\to\infty} \frac{1}{k}\chi(\mathcal{N}^{\otimes k}) \equiv \chi_{\mathrm{reg}}(\mathcal{N}),$$

where

$$\mathcal{N}^{\otimes k}_{A_1\cdots A_k \to B_1\cdots B_k} = \mathcal{N}_{A_1\to B_1} \otimes \cdots \otimes \mathcal{N}_{A_k\to B_k}.$$

- Regularization is equivalent to **multiletter** characterization in classical information theory $\frac{1}{k}I(X^k;Y^k)$ which, in contrast to **single** letter characterizations, $I(X;Y)$, does not provide much insight into practical coding techniques and is difficult to compute.
- Note that if the Holevo information is **additive** the regularization limit is not necessary and thus, $\mathrm{C}(\mathcal{N}) = \chi(\mathcal{N}) = \max_{\rho_{XA}} I(X;B)_\rho$.





- The **most general** strategy for classical information communication is illustrated **left**. One can make use of entanglement at Alice side and joint decoding at Bob's. On the contrary, the **suboptimal** strategy illustrated **right** uses the block DMC defined by a single use of a quantum channel only. Figures from [Wilde, 2017].
- The HSW theorem requires regularization, and thus investigating the expression $\lim_{k\to\infty} \dfrac{1}{k}\chi(\mathcal{N}^{\otimes k})$, since both strategies are in general not equivalent.

Channel code for classical communication

Channel code for communication over quantum channels

A $(2^{nR}, n, \epsilon)$ code $\mathcal{C}$ for quantum channel $\mathcal{N}_{A\to B}$ is:

- a message set $\mathcal{M} \equiv [1:2^{nR}]$.
- an encoding function $[1:2^{nR}] \to \rho_{A^n}$: that assigns a codeword $\rho^m_{A^n}$ to each message $m \in [1:2^{nR}]$. This state is transmitted over n independent uses on the quantum channel $\mathcal{N}_{A\to B}$ so that the state at Bob end is $\mathcal{N}^{\otimes n}(\rho^m_{A^n})$.
- a decoding POVM $\{\Lambda_m\}$ that assigns an estimate $\hat{m} \in [1:2^{nR}]$ to each received state $\mathcal{N}^{\otimes n}(\rho^m_{A^n})$. The conditional probability of error assuming message m was transmitted is:

$$P_e(m) \equiv \Pr\{\hat{M} \neq m | M = m\} = \mathrm{tr}\{(I - \Lambda_m)\mathcal{N}^{\otimes n}(\rho^m_{A^n})\}$$

The code has ϵ error if $\max_m\{P_e(m)\} \equiv P_e^* \leq \epsilon$.

Random codebook generation

Let $\mathcal{E} = \{p_X(x), \phi_A^x\}$ be the ensemble of pure states $\phi_A^x \equiv |\phi_x\rangle\langle\phi_x|_A$ that attains the Holevo information of $\mathcal{N}$. Generate $x^n(m)$ for $m \in [1 : 2^{nR}]$ according to $p_{X^n}(x^n) = \prod_{i=1}^n p_X(x_i)$. The generated sequences yield the following product density matrix at A:

$$\begin{aligned}
\phi_{A^n} &\equiv \phi_{A_1} \otimes \cdots \otimes \phi_{A_n} \\
&= \sum_{x_1 \in \mathcal{X}} p_X(x_1)\phi_{A_1}^{x_1} \otimes \cdots \otimes \sum_{x_n \in \mathcal{X}} p_X(x_n)\phi_{A_n}^{x_n} \\
&= \sum_{x^n \in \mathcal{X}^n} p_{X^n}(x^n)\phi_{A^n}^{x^n} \equiv \phi_A^{\otimes n}.
\end{aligned}$$

and B, for $\sigma_{B_i}^{x_i} \equiv \mathcal{N}(\phi_{A_i}^{x_i})$:

$$\sigma_{B^n} \equiv \sigma_{B_1} \otimes \cdots \otimes \sigma_{B_n} = \sum_{x^n \in \mathcal{X}^n} p_{X^n}(x^n)\sigma_{B^n}^{x^n} \equiv \sigma_B^{\otimes n}$$

Encoding and decoding

Encoding

To send message $m \in [1:2^{nR}]$ take $\rho_{A^n}^m = \phi_{A^n}^{x^n(m)}$ where:

$$\phi_{A^n}^{x^n(m)} \equiv \phi_{A_1}^{x_1(m)} \otimes \cdots \otimes \phi_{A_n}^{x_n(m)}.$$

Decoding strategy

Let

$$\sigma_{B^n}^{x^n(m)} \equiv \mathcal{N}^{\otimes n}(\phi_{A^n}^{x^n(m)}) \equiv \sigma_{B_1}^{x_1(m)} \otimes \cdots \otimes \sigma_{B_n}^{x_n(m)}$$

be the received quantum state, where $\sigma_{B_i}^{x_i(m)} \equiv \mathcal{N}(\phi_{A_i}^{x_i(m)})$, for $i \in [1:n]$. The receiver declares that $\hat{m} \in [1:2^{nR}]$ was transmitted using a detection POVM $\{\Lambda_m\}$. The quantum typicality properties are used to define the $\{\Lambda_m\}$, where $\Lambda_m \succeq 0$ and $\sum_{m=1}^{2^{nR}} \Lambda_m = I$.

Decoding projectors

- The Shannon typicality decoder would check typicality between the received sequence and the codewords corresponding to each of the m possible messages.
- We will follow a somehow similar procedure in the quantum setting by considering the observed density matrix $\sigma_B^{\otimes n}$, the ones corresponding to the m possibly transmitted messages $\sigma_{B^n}^{x^n(m)}$, and then defining their respective projectors.

Total subspace $\Pi_{B^n}^{\delta}$ and message subspace $\Pi_{B^n|x^n(m)}^{\delta}$ projectors

- The total subspace projector $\Pi_{B^n}^{\delta}$ is defined as the (weakly) typical projector for $\sigma_B^{\otimes n}$.
- The message subspace $\Pi_{B^n|x^n(m)}^{\delta}$ projector is defined as the (weakly) conditional typical projector for $\sigma_{B^n}^{x^n(m)}$.

Quantum Packing Lemma

Quantum Packing Lemma

Assuming strong typicality, we can now apply the properties of quantum typical and conditionally typical sequences:

$$
\begin{aligned}
\mathrm{tr}\{\Pi^{\delta}_{B^n}\rho^{x^n}_{B^n}\} &\geq 1-\epsilon, \\
\mathrm{tr}\{\Pi^{\delta}_{B^n|x^n}\rho^{x^n}_{B^n}\} &\geq 1-\epsilon, \\
\mathrm{tr}\{\Pi^{\delta}_{B^n|x^n}\} &\leq 2^{n(H(B|X)+c\delta)}, \\
\Pi^{\delta}_{B^n}\sigma_B^{\otimes n}\Pi^{\delta}_{B^n} &\preceq 2^{-n(H(B)-c'\delta)}\Pi^{\delta}_{B^n},
\end{aligned}
$$

and then the (derandomized) quantum packing lemma, (see Chapter 16 of [Wilde, 2017] for details) which states that the maximum P_e is:

$$
\begin{aligned}
P_e^* &= \max_m \mathrm{tr}\{(I-\Lambda_m)\sigma^{x^n(m)}_{B^n}\} \\
&\leq 4(\epsilon+2\sqrt{\epsilon}) + 16(1-\epsilon)^{-1}2^{-n(H(B)-H(B|X)-(c+c')\delta)}|\mathcal{M}|.
\end{aligned}
$$

Probability of error and design of the POVM $\{\Lambda_m\}$

Probability of error

The quantum packing lemma shows that P_e^* can be made as small as desired as long as the cardinality of the message set $|\mathcal{M}|$ is small enough,

$$P_e^* \leq 4(\epsilon + 2\sqrt{\epsilon}) + 16(1-\epsilon)^{-1} 2^{-n(I(X;B)-(c+c')\delta)} 2^{nR}$$

and $\lim_{n\to\infty} P_e^* = 0$ if $R < I(X;B) - (c+c')\delta$.

Design of the POVM $\{\Lambda_m\}$

Projectors $\Pi^{\delta}_{B^n|x^n}$ can not be used directly as POVM since there is no guarantee that $\sum_{m=1}^{2^{nR}} \Pi^{\delta}_{B^n|x^n(m)} \equiv \bar{\Pi}^{\delta}_{B^n|x^n} = I$. In order to satisfy this condition one possible solution is:

$$\Lambda_m = (\bar{\Pi}^{\delta}_{B^n|x^n})^{-\frac{1}{2}} \Pi^{\delta}_{B^n|x^n(m)} (\bar{\Pi}^{\delta}_{B^n|x^n})^{-\frac{1}{2}}$$

Remarks on the achievability result

- Note the proof shows achievability for a rate equal to the Holevo information of the quantum channel $\chi(\mathcal{N})$ and not to the possibly higher rate $\chi_{\mathrm{reg}}(\mathcal{N}) = \lim_{k\to\infty} \dfrac{1}{k}\chi(\mathcal{N}^{\otimes k})$.
- But nothing prevents applying the same encoding and decoding procedure to the product channel $\mathcal{N}^{\otimes k}$ showing that for arbitrary k the rate $\dfrac{1}{k}\chi(\mathcal{N}^{\otimes k})$ is achievable.

Randomness distribution over quantum channels

For uniformly distributed messages the ideal **randomness distribution** joint state between Alice and Bob is:

$$\bar{\phi}_{M\hat{M}} \equiv \frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}} |m\rangle\langle m|_M \otimes |m\rangle\langle m|_{\hat{M}},$$

whereas the actual state after encoding and decoding is:

$$\omega_{M\hat{M}} = \frac{1}{|\mathcal{M}|}\sum_{m,\hat{m}\in\mathcal{M}} \mathrm{tr}\{\Lambda_{\hat{m}}\mathcal{N}^{\otimes n}(\rho^m_{A^n})\}|m\rangle\langle m|_M \otimes |\hat{m}\rangle\langle \hat{m}|_{\hat{M}}.$$

Code for randomness distribution over channel $\mathcal{N}_{A\to B}$

A $(2^{nR}, n, \epsilon)$ code for randomness distribution is achievable if there is an encoding-decoding procedure as indicated before for the channel code and

$$\frac{1}{2}\|\bar{\phi}_{M\hat{M}} - \omega_{M\hat{M}}\|_1 \leq \epsilon.$$

- The capacity for randomness distribution cannot be smaller than the one for communication since an error-free classical channel can be used to obtain shared randomness.
- This implies that randomness distribution can be used in converse proofs of classical communication.
- Invoking the AFW inequality:

$$\frac{1}{2}\|\bar{\phi}_{M\hat{M}} - \omega_{M\hat{M}}\|_1 \leq \epsilon,$$

implies

$$\begin{aligned} |H(M|\hat{M})_{\bar{\phi}} - H(M|\hat{M})_{\omega}| &\leq \epsilon \log|\mathcal{M}| + (1+\epsilon)H(\epsilon/(1+\epsilon)) \\ &\equiv f(|\mathcal{M}|, \epsilon). \end{aligned}$$

Converse proof (and II)

- Now applying a procedure similar in spirit to Fano's inequality,

$$\begin{aligned} nR = \log|\mathcal{M}| &= H(M) \\ &= H(M) - H(M|\hat{M})_{\bar{\phi}} \\ &\leq H(M) - H(M|\hat{M})_{\omega} + f(|\mathcal{M}|, \epsilon) \\ &= I(M;\hat{M})_{\omega} + f(|\mathcal{M}|, \epsilon) \\ &\leq I(M;B^n)_{\omega} + f(|\mathcal{M}|, \epsilon) \\ &\leq \chi(\mathcal{N}^{\otimes n}) + f(|\mathcal{M}|, \epsilon). \end{aligned}$$

- Recovering the expression for $f(|\mathcal{M}|, \epsilon)$ and rearranging terms,

$$R(1-\epsilon) \leq \frac{1}{n}\chi(\mathcal{N}^{\otimes n}) + \frac{1}{n}(1+\epsilon)H(\epsilon/(1+\epsilon))$$

meaning that for $n \to \infty$, $R \leq \chi_{\mathrm{reg}}(\mathcal{N})$. ■

Why is regularization needed?

What is different in this converse proof with respect to the classical channel coding converse?

- In classical the mutual information is additive for the block DMC, and in the classical converse proof:

$$nR \leq I(X^n; Y^n) + n\epsilon_n \leq nC + n\epsilon_n$$

- But the Holevo information is not additive!

$$I(M; B^n)_\omega \nleq nI(M; B)_\omega \leq n\chi(\mathcal{N}).$$

Capacity examples

Entanglement-breaking channels

Entanglement-breaking channels

An entanglement breaking channel $\mathcal{N}^{\mathsf{EB}}_{A\to B}$ takes any arbitrary composite state $\rho_{RA} \in \mathcal{D}(\mathcal{H}_R \otimes \mathcal{H}_A)$ to a separable state, i.e.

$$(\mathrm{id}_R \otimes \mathcal{N}^{\mathsf{EB}}_{A\to B})(\rho_{RA}) = \sum_z p_Z(z)\rho^z_R \otimes \tau^z_B$$

One can proof that it suffices to check this condition for the maximally entangled state Φ_{RA} and that a channel is entanglement-breaking iff its Kraus operators are unit rank.

In fact, this proof leads to the corollary that any entanglement-breaking channel is a serial concatenation of a *qc* channel followed by a *cq* channel, i.e., a measurement channel followed by a preparation channel:

$$\mathcal{N}^{\mathsf{EB}}_{A\to B}(\rho_A) = (\mathcal{P}_{Z\to B} \circ \mathcal{M}_{A\to Z})(\rho_A)$$

Capacity of entanglement-breaking channels (I)

Capacity of entanglement-breaking channels

Let $\mathcal{N}^{EB}$ be an entanglement-breaking channel. The capacity of $\mathcal{N}^{EB}$ equals the Holevo information of the channel, i.e.,

$$\mathrm{C}(\mathcal{N}^{EB}) = \chi(\mathcal{N}^{EB}).$$

The proof follows directly from the additivity of the Holevo information of the entanglement-breaking channels.

Additivity of the Holevo information for EB channels

Let $\mathcal{N}^{EB}$ be an entanglement-breaking channel and $\mathcal{M}$ an arbitrary channel. Then,

$$\chi(\mathcal{N}^{EB} \otimes \mathcal{M}) = \chi(\mathcal{N}^{EB}) + \chi(\mathcal{M}).$$

Capacity of entanglement-breaking channels (II)

- The proof requires to show that $\chi(\mathcal{N}^{EB} \otimes \mathcal{M}) \leq \chi(\mathcal{N}^{EB}) + \chi(\mathcal{M})$ since $\chi(\mathcal{N}^{EB} \otimes \mathcal{M}) \geq \chi(\mathcal{N}^{EB}) + \chi(\mathcal{M})$ by definition.
- Let $\rho_{XB_1B_2}$ the state in $\chi(\mathcal{N}^{EB} \otimes \mathcal{M}) = \max_{\rho_{XA_1A_2}} I(X;B_1B_2)_\rho$,

$$\rho_{XB_1B_2} = (\mathrm{id}_X \otimes \mathcal{N}^{EB}_{A_1\to B_1} \otimes \mathcal{M}_{A_2\to B_2})(\rho_{XA_1A_2}),$$
$$\rho_{XA_1A_2} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho^x_{A_1A_2}.$$

After applying $\mathcal{N}^{EB}_{A_1\to B_1}$ to $\rho_{XA_1A_2}$,

$$\begin{aligned}\rho_{XB_1A_2} &= \sum_x p_X(x)|x\rangle\langle x|_X \otimes \mathcal{N}^{EB}_{A_1\to B_1}(\rho^x_{A_1A_2}) \\ &= \sum_x p_X(x)|x\rangle\langle x|_X \otimes \sum_z p_{Z|X}(z|x)\sigma^{x,z}_{B_1} \otimes \theta^{x,z}_{A_2} \\ &= \sum_{x,z} p_X(x)p_{Z|X}(z|x)|x\rangle\langle x|_X \otimes \sigma^{x,z}_{B_1} \otimes \theta^{x,z}_{A_2}.\end{aligned}$$

Capacity of entanglement-breaking channels (III)

- Then, after applying $\mathcal{M}_{A_2 \to B_2}$ to $\rho_{XB_1A_2}$,

$$\rho_{XB_1B_2} = \sum_{x,z} p_{Z|X}(z|x) p_X(x) |x\rangle\langle x|_X \otimes \sigma_{B_1}^{x,z} \otimes \mathcal{M}(\theta_{A_2}^{x,z}).$$

  State for which we can define the extension,

$$\omega_{XZB_1B_2} = \sum_{x,z} p_{Z|X}(z|x) p_X(x) |x\rangle\langle x|_X \otimes |z\rangle\langle z|_Z \otimes \sigma_{B_1}^{x,z} \otimes \mathcal{M}(\theta_{A_2}^{x,z}),$$

  satisfying $\rho_{XB_1B_2} = \mathrm{tr}_Z\{\omega_{XZB_1B_2}\}$.

- Now, by the definition of $\rho_{XB_1B_2}$, the chain rule for mutual information and the definition of $\chi(\mathcal{N}^{EB})$,

$$\begin{aligned} \chi(\mathcal{N}^{EB} \otimes \mathcal{M}) = I(X;B_1B_2)_\rho &= I(X;B_1)_\rho + I(X;B_2|B_1)_\rho \\ &\leq \chi(\mathcal{N}^{EB}) + I(X;B_2|B_1)_\rho. \end{aligned}$$

Capacity of entanglement-breaking channels (IV)

- We concentrate now on the term $I(X;B_2|B_1)_\rho$.

$$\begin{aligned}
I(X;B_2|B_1)_\rho &= I(X;B_2|B_1)_\omega \\
&= I(XB_1;B_2)_\omega - I(B_1;B_2)_\omega \\
&\leq I(XB_1;B_2)_\omega \\
&\leq I(XZB_1;B_2)_\omega \\
&= I(XZ;B_2)_\omega + I(B_1;B_2|XZ)_\omega \\
&= I(XZ;B_2)_\omega \\
&\leq \chi(\mathcal{M}).
\end{aligned}$$

where we have used the chain rule and the positivity for mutual information, the data processing inequality, the chain rule again, the fact that, conditioned on XZ, the $\omega_{XZB_1B_2}$ state is product, and the definition of $\chi(\mathcal{M})$. ■

Capacity of entanglement-breaking channels (and V)

- We showed any $\mathcal{N}^{EB}$ channel can be expressed as a serial concatenation of a measurement channel followed by a preparation channel. Then for $\mathcal{N}$ an arbitrary channel, Alice can simulate an $\mathcal{N}^{EB}$ channel by performing a **measurement** in the $\{|x\rangle\langle x|\}$ basis and **preparing** a state ρ^x conditioned on the outcome so that the resulting density matrix is:

$$\rho_{XB} = \sum_x \langle x|\sigma_A|x\rangle |x\rangle\langle x|_X \otimes \mathcal{N}(\rho^x) = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \mathcal{N}(\rho^x).$$

  the capacity of which is,

$$\max_{p_X(x),\rho^x} I(X;B)_\rho.$$

- This is known as the product-state capacity of $\mathcal{N}$, a lower bound to its true capacity (no entanglement at the encoder).

Capacity of quantum erasure channel (I)

The capacity of the (classical) erasure channel (EC) for $\mathcal{Y} = \mathcal{X} \cup \{e\}$,

$$p_{Y|X}(y|x) = (1-\epsilon)\mathbb{1}\{y = x\} + \epsilon\mathbb{1}\{y = e\}$$

where $\epsilon \in [0, 1]$, is well known, $\mathrm{C}_{\mathrm{EC}} = (1-\epsilon)\log|\mathcal{X}|$. A similar result is obtained in the quantum case.

Capacity of quantum erasure channel

The capacity of the quantum erasure channel defined as

$$\mathcal{N}^{\epsilon}_{A\to B}(\rho_A) = (1-\epsilon)\mathcal{I}_{A\to B}(\rho_A) + \epsilon|e\rangle\langle e|_B$$

where $\epsilon \in [0, 1]$, $d_A = d$, $d_B = d + 1$, and $\{|0\rangle, \cdots, |d-1\rangle\}$ and $\{|0\rangle, \cdots, |d-1\rangle, |e\rangle\}$ form an orthonormal basis in $\mathcal{H}_A$ and $\mathcal{H}_B$ respectively, is

$$\mathrm{C}(\mathcal{N}^{\epsilon}) = (1-\epsilon)\log d.$$

Capacity of quantum erasure channel (II)

Achievability proof (I)

Consider the ensemble $\mathcal{E} \equiv \{1/d; |0\rangle_A, \cdots, |d-1\rangle_A\}$, this yields:

$$\rho_{XB} = \frac{1}{d}\sum_{x=0}^{d-1} |x\rangle\langle x|_X \otimes ((1-\epsilon)|x\rangle\langle x|_B + \epsilon|e\rangle\langle e|_B),$$

$$\begin{aligned}\rho_B = \mathrm{tr}_X\{\rho_{XB}\} &= \frac{1}{d}\sum_{x=0}^{d-1}((1-\epsilon)|x\rangle\langle x|_B + \epsilon|e\rangle\langle e|_B) \\ &= \frac{1-\epsilon}{d}(I_B - |e\rangle\langle e|_B) + \epsilon|e\rangle\langle e|_B.\end{aligned}$$

The eigenvalues of ρ_B are $\frac{1-\epsilon}{d}$ with multiplicity d and ϵ with multiplicity 1,

$$H(B)_\rho = -(1-\epsilon)\log(\frac{1-\epsilon}{d}) - \epsilon\log\epsilon = H(\epsilon) + (1-\epsilon)\log d.$$

Capacity of quantum erasure channel (III)

Achievability proof (II)

$$H(B|X)_\rho = \frac{1}{d}\sum_{x=0}^{d-1} H((1-\epsilon)|x\rangle\langle x|_B + \epsilon|e\rangle\langle e|_B)$$
$$= -(1-\epsilon)\log(1-\epsilon) - \epsilon\log\epsilon = H(\epsilon),$$

since the eigenvalues are $(1-\epsilon)$ and ϵ both with multiplicity 1. Therefore,

$$I(X;B)_\rho = H(B)_\rho - H(B|X)_\rho = (1-\epsilon)\log d.$$

meaning this rate is achievable. ∎

See [Wilde, 2017] for the converse where it is proved that the regularized Holevo information can not exceed $(1-\epsilon)\log d$ an thus it is indeed the capacity.

Private classical and quantum information

Private information of a quantum channel

- We defined the **Holevo information** of a quantum channel $\mathcal{N}_{A\to B}$ as an upper bound to the **accessible information** that can be **transmitted** through the channel:

$$\chi(\mathcal{N}) \equiv \max_{\rho_{XA}} I(X;B)_\rho \geq \max_{\mathcal{E}} I_{\mathrm{acc}}(\mathcal{E}) = I^*_{\mathrm{acc}},$$

  where the **maximization** took place with respect to a classical-quantum state:

$$\rho_{XA} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_A^x,$$

  i.e. the ensemble $\mathcal{E} = \{p_X(x), \rho_A^x\}$.
- Now we want to consider **how much** information can be **transmitted** to Bob **keeping it secret to the rest of the world**, **extending** the concept of the classical **wiretap** channel to the quantum context.

Private information of a quantum channel: definition

Private information of a quantum channel

The private information $P(\mathcal{N})$ of channel $\mathcal{N}_{A\to B}$ is defined as:

$$P(\mathcal{N}_{A\to B}) \equiv \max_{\rho_{XA}} (I(X;B)_\rho - I(X;E)_\rho),$$

for

$$\rho_{XA} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_A^x.$$

The mutual information terms are computed with respect to:

$$\rho_{XBE} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \mathcal{U}^{\mathcal{N}}_{A\to BE}(\rho_A^x),$$

where $\mathcal{U}^{\mathcal{N}}_{A\to BE}$ is an **isometric** extension of channel $\mathcal{N}_{A\to B}$.

Private information of a quantum channel: remarks

- Note that the private information of a quantum channel is **non-negative**, $P(\mathcal{N}) \geq 0$ since for $\rho_{XA} = |0\rangle\langle 0|_X \otimes |\phi\rangle\langle\phi|_A$ we have $P(\mathcal{N}) = 0$.
- Like the Holevo information, the private information is not **additive** in general, but can be additive in **some** cases.
- The **regularized** private information is then defined as

$$P_{\mathrm{reg}}(\mathcal{N}) \equiv \lim_{n\to\infty} \frac{1}{n} P(\mathcal{N}^{\otimes n}).$$

Private information of a classical channel

- The expression for the private information is reminiscent of the capacity of the (classical) **degraded wiretap** channel, i.e. where $X \leftrightarrow Y \leftrightarrow Z$ form a Markov chain,

$$C_S \equiv \max_{p_X(x)} (I(X;Y) - I(X;Z)).$$

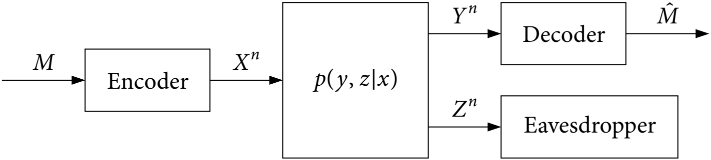


Figure: Wiretap Channel Model

- However, the **private information of a classical channel** is indeed always **additive**.

Coherent information of a quantum channel

Coherent information of a quantum channel

The **coherent information** $Q(\mathcal{N})$ of channel $\mathcal{N}_{A\to B}$ is defined as:

$$Q(\mathcal{N}_{A\to B}) \equiv \max_{\phi_{AA'}} I(A\rangle B)_\rho = -\min_{\phi_{AA'}} H(A|B)_\rho$$

where the maximization takes place with respect to all input **pure** states $\phi_{AA'}$ and

$$\rho_{AB} = \mathcal{N}_{A'\to B}(\phi_{AA'}).$$

- Note that $Q(\mathcal{N})$ is also non-negative, $Q(\mathcal{N}) \geq 0$ since for $\phi_{AA'} = \psi_A \otimes \varphi_{A'}$, we have $\rho_{AB} = \psi_A \otimes \mathcal{N}(\varphi_{A'})$ and,

$$H(A|B) = H(A) = 0.$$

Coherent information and private information

Coherent information and private information

The coherent information $Q(\mathcal{N})$ of any channel $\mathcal{N}$ is never greater than its private information $P(\mathcal{N})$,

$$Q(\mathcal{N}) \leq P(\mathcal{N}).$$

Exercise

Prove that the coherent information of any channel is never greater than its private information, $Q(\mathcal{N}) \leq P(\mathcal{N})$.

- Consider $\phi_{AA'}$ as the **pure state** that maximizes $Q(\mathcal{N})$
- Consider ϕ_{ABE} the state after sending the A' system through $\mathcal{U}^{\mathcal{N}}_{A' \to BE}$, an **isometric** extension of channel $\mathcal{N}_{A' \to B}$.
- Using the spectral decomposition

$$\phi_{A'} = \sum_x |\phi_x\rangle\langle\phi_x|_{A'}$$

create an augmented classical-quantum state of the form:

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

and its extension σ_{XBE} after $\mathcal{U}^{\mathcal{N}}_{A' \to BE}$.

The Private classical capacity

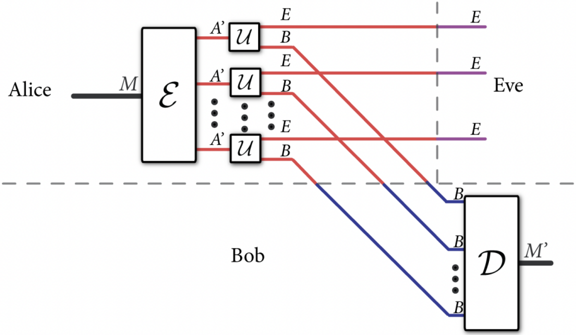


Figure: Private Capacity context

The objective remains as the **maximization** of the classical information transmission **rate** to Bob but without any information **leakage** to Eve.

Reliability and secrecy requirements

- The channel code for private communication over quantum channels must satisfy the same requirements in terms of error probability at Bob:

$$P_e(m) \equiv \Pr\{\hat{M} \neq m | M = m\} = \mathrm{tr}\{(I - \Lambda_m)\mathcal{N}^{\otimes n}(\rho^m_{A^n})\}$$

- In addition there is the secrecy requirement towards Eve. Let $\omega^m_{E^n}$ be the state observed by Eve:

$$\omega^m_{E^n} = \mathrm{tr}_B\{\mathcal{U}^{\mathcal{N}}_{A\to BE}(\rho^m_{A^n})\}$$

- Secrecy is measured by the normalized trace distance between $\omega^m_{E^n}$ and an constant (with respect to m) state σ_{E^n}:

$$\forall m \in [1 : 2^{nR}] : \frac{1}{2}\|\omega^m_{E^n} - \sigma_{E^n}\|_1 \leq \epsilon$$

The Devetak-Cai-Winter-Yeung theorem

The Devetak-Cai-Winter-Yeung theorem

The private classical capacity of a quantum channel is equal to the regularization of the private information of the channel:

$$\mathrm{C}_P(\mathcal{N}) = \lim_{k\to\infty} \frac{1}{k} P(\mathcal{N}^{\otimes k}) \equiv P_{\mathrm{reg}}(\mathcal{N}),$$

with

$$P(\mathcal{N}) \equiv \max_{\rho}(I(X;B)\sigma - I(X;E)_{\sigma}),$$

$$\rho_{XA} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_A^x,$$

$$\sigma_{XBE} = \mathcal{U}^{\mathcal{N}}_{A\to BE}(\rho_{XA}),$$

where $\mathcal{U}^{\mathcal{N}}_{A\to BE}$ is an isometric extension of channel $\mathcal{N}_{A\to B}$.

Sketch of the achievability proof

- Alice constructs a codebook using two-index codewords (m, k), $m \in \mathcal{M}$ and $k \in \mathcal{K}$ that she communicates through n independent uses of the channel $\mathcal{N}_{A\to B}$.
- As long as the condition established by the quantum packing lemma for channel $\mathcal{N}_{A\to B}$ is satisfied, i.e., $|\mathcal{M}||\mathcal{K}| \approx 2^{nI(X;B)}$ Bob can detect both m and k with vanishing maximum error probability. This is known as the **reliability** requirement.
- Also, as long as the condition determined for the quantum covering lemma for the channel $\mathcal{N}^c_{A\to E}$, i.e., $|\mathcal{K}| \approx 2^{nI(X;E)}$ is fulfilled, then the information from the message m leaked to Eve can be made as small as desired. This is the **secrecy** requirement.

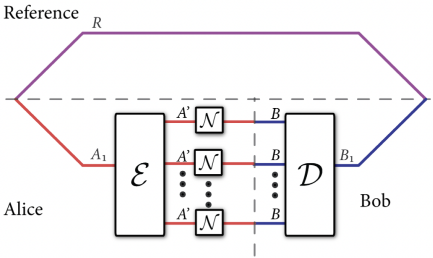


Figure: Information-processing task for entanglement transmission

The objective is that Bob performs a **decoding** of the systems he receives, and the state at the end of the protocol is **close** to the original state **shared** between Alice and the reference.

Quantum Communication requirements

- The state at Bob is:

$$\omega_{RB_1} \equiv \mathcal{D}_{B^n \to B_1}(\mathcal{N}_{A^n \to B^n}(\mathcal{E}_{A_1 \to A^n}(\varphi_{RA_1})))$$

- The encoding and decoding procedure must guarantee:

$$\frac{1}{2}\|\varphi_{RA_1} - \omega_{RB_1}\|_1 \leq \epsilon$$

- The quantum rate Q of this scheme is measured as the number of qbits transmitted per quantum channel use:

$$Q \equiv \frac{1}{n} \log \dim(\mathcal{H}_{A_1})$$

- The quantum capacity $C_Q(\mathcal{N})$ is defined as the supremum of all achievable rates for $\mathcal{N}$.

Quantum capacity theorem

Quantum capacity theorem

The quantum capacity $C_Q(\mathcal{N})$ of a quantum channel is equal to the regularized coherent information of the channel:

$$\mathrm{C}_Q(\mathcal{N}) = \lim_{k\to\infty} \frac{1}{k} Q(\mathcal{N}^{\otimes k}) \equiv Q_{\mathrm{reg}}(\mathcal{N})$$

with

$$Q(\mathcal{N}) \equiv \max_{\phi} I(A\rangle B)_{\sigma}$$

where the maximization is with respect to all pure states $\phi_{AA'}$ and $\sigma_{AB} = \mathcal{N}_{A'\to B}(\phi_{AA'})$

References

References

Mark M. Wilde.
Quantum Information Theory, Second Edition.
Cambridge University Press 2017.

Michael Nielsen and Isaac Chuang.
Quantum Computation and Quantum Information. Tenth Anniversary Edition.
Cambridge University Press 2010.

Hastings M.B.
Superadditivity of communication capacity using entangled inputs
Nature Physics 5, 255-257, 2009.