Lecture 5: Quantum Entropy and Information

Javier R. Fonollosa

Universitat Politècnica de Catalunya

javier.fonollosa@upc.edu

October 6, 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

- In this chapter we will introduce the notion of *quantum information* as a way to describe the **information** present in quantum systems and the **mutual correlations** among them.
- The von Newmann entropy or **quantum entropy**, measured in **qubits**, will be defined generalizing the concept of **Shannon entropy**, measured in **bits**.
- As in the classical context, the concept of **quantum entropy** will be generalized to **joint entropy**, **conditional entropy**, **relative entropy**, **mutual information** and **conditional mutual information**.
- We will see that **many** of the properties observed in the classical context are preserved in the quantum world, but **not all**.
- The most prominent example of **discrepancy** between the classic and quantum interpretation associated to entropy is the **conditional quantum entropy**, which can be **negative**.

Separable and Entangled states

Classical composite states

- Consider two discrete random variables X, Y characterized by a joint pmf $p_{XY}(x, y)$. Their joint density matrix is:

$$\begin{aligned}\rho_{XY} &= \sum_{x=0}^{d_X-1}\sum_{y=0}^{d_Y-1} p_{XY}(x,y)(|x\rangle \otimes |y\rangle)(\langle x| \otimes \langle y|) \\ &= \sum_{x=0}^{d_X-1}\sum_{y=0}^{d_Y-1} p_{XY}(x,y)(|x\rangle\langle x| \otimes |y\rangle\langle y|) \\ &= \mathrm{diag}(p_{XY}(0,0), p_{XY}(0,1), \ldots, p_{XY}(d_X-1, d_Y-1))\end{aligned}$$

- If the random variables are independent, i.e. $p_{XY}(x, y) = p_X(x)p_Y(y)$ then

$$\begin{aligned}\rho_{XY} &= \sum_{x=0}^{d_X-1} p_X(x)|x\rangle\langle x| \otimes \sum_{y=0}^{d_Y-1} p_Y(y)|y\rangle\langle y| = \rho_X \otimes \rho_Y \\ &= \mathrm{diag}(p_X(0), \ldots, p_X(d_X-1)) \otimes \mathrm{diag}(p_Y(0), \ldots, p_Y(d_Y-1))\end{aligned}$$

Product, separable and entangled composite states

- **Independent** quantum states (have never interacted before) are characterized by a density matrix formed by the **Kronecker product** of **individual density matrices**, e.g., the joint state between Alice and Bob would be modeled as:

$$\rho_{AB} = \sigma_A \otimes \tau_B.$$

  This is denoted as a **product** state.

- If the states are prepared independently but **conditioned** on the outcome of a **shared random variable** X then they are described as:

$$\rho_{AB} = \sum_{x \in \mathcal{X}} p_X(x) \sigma_A^x \otimes \tau_B^x$$

  Composite states that admit this representation are called **separable**. All other are **entangled**.

- Note that **product** states are a particular case of **separable** states.

Separable states

- For **separable states**, making use of the **spectral decomposition** of the individual density matrices

$$\sigma_A^x = \sum_{y\in\mathcal{Y}} p_Y(y|x)|\phi_y^x\rangle\langle\phi_y^x|_A, \qquad \tau_B^x = \sum_{k\in\mathcal{K}} p_K(k|x)|\psi_k^x\rangle\langle\psi_k^x|_B,$$

$$\begin{aligned}\rho_{AB} &= \sum_{x\in\mathcal{X}}\sum_{y\in\mathcal{Y}}\sum_{k\in\mathcal{K}} p_X(x)p_Y(y|x)p_K(k|x)|\phi_y^x\rangle\langle\phi_y^x|_A \otimes |\psi_k^x\rangle\langle\psi_k^x|_B \\ &= \sum_{z\in\mathcal{Z}} p_Z(z)|\phi_z\rangle\langle\phi_z|_A \otimes |\psi_z\rangle\langle\psi_z|_B.\end{aligned}$$

  where $\mathcal{Z} = \mathcal{X}\times\mathcal{Y}\times\mathcal{K}$ and $|\phi^z\rangle_A$ and $|\psi^z\rangle_B$ are unit vectors.
- **Separable states** can be expressed as a **convex combination** of **pure product states**.
- Note that **pure separable states** are **product** since their rank is one and thus $|\mathcal{Z}| = 1$, i.e. $|\varphi\rangle\langle\varphi|_{AB} = |\phi\rangle\langle\phi|_A \otimes |\psi\rangle\langle\psi|_B$.

Maximally Entangled state: ebit (I)

- Sometimes composite quantum states are **jointly prepared** but then **physically separated** so that **one share** of the quantum system is in the possession of A and the **other share** is with B.
- A prominent example is the **Maximally Entangled state** $\equiv$ **ebit**, between A and B. Its **state vector** and **density matrix** are:

$$|\Phi\rangle_{AB} \equiv \frac{1}{\sqrt{2}}(|00\rangle_{AB} + |11\rangle_{AB}) = \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B),$$

$$\Phi_{AB} \equiv |\Phi\rangle\langle\Phi|_{AB} = \frac{1}{2}\begin{bmatrix}1\\0\\0\\1\end{bmatrix}\begin{bmatrix}1 & 0 & 0 & 1\end{bmatrix} = \frac{1}{2}\begin{bmatrix}1 & 0 & 0 & 1\\0 & 0 & 0 & 0\\0 & 0 & 0 & 0\\1 & 0 & 0 & 1\end{bmatrix}.$$

It can be proved that this **pure** state is also **entangled** since:

$$|\langle\Phi|_{AB}(|\phi\rangle_A \otimes |\psi\rangle_B)|^2 = \frac{1}{2}|\langle\phi|^*|\psi\rangle|^2 \leq \frac{1}{2},$$

for all $|\phi\rangle_A$,$|\psi\rangle_B$ and thus $|\Phi\rangle_{AB}$ **cannot be product**.

Maximally Entangled state: ebit (and II)

- The proof makes use of the Cauchy-Schwarz inequality:

$$\begin{aligned}\langle\Phi|_{AB}(|\phi\rangle_A \otimes |\psi\rangle_B) &= \frac{1}{\sqrt{2}}\big(\langle 0|_A \otimes \langle 0|_B + \langle 1|_A \otimes \langle 1|_B\big)(|\phi\rangle_A \otimes |\psi\rangle_B) \\ &= \frac{1}{\sqrt{2}}\langle\phi|^*_A|\psi\rangle_B\end{aligned}$$

  which by Cauchy-Schwarz implies:

$$|\langle\Phi|_{AB}(|\phi\rangle_A \otimes |\psi\rangle_B)|^2 = \frac{1}{2}|\langle\phi|^*|\psi\rangle|^2 \leq \frac{1}{2}\langle\phi|\phi\rangle\langle\psi|\psi\rangle = \frac{1}{2},$$

  for all $|\phi\rangle_A$,$|\psi\rangle_B$.

- This inequality indicates that there is no way that $|\Phi\rangle_{AB}$ can be expressed as a product vector state $|\phi\rangle_A \otimes |\psi\rangle_B$ and thus $|\Phi\rangle\langle\Phi|_{AB}$ cannot be product either. Since the state is **pure** it must be **entangled**. ■

Classification of composite states

- According to the previous definitions **pure composite states** can be classified as:

Pure composite	$\lvert\psi\rangle_{AB}$	$\lvert\psi\rangle\langle\psi\rvert_{AB}$
Product	$\lvert\phi\rangle_A \otimes \lvert\psi\rangle_B$	$\lvert\phi\rangle\langle\phi\rvert_A \otimes \lvert\psi\rangle\langle\psi\rvert_B$
Entangled	$\neq \lvert\phi\rangle_A \otimes \lvert\psi\rangle_B$	$\neq \lvert\phi\rangle\langle\phi\rvert_A \otimes \lvert\psi\rangle\langle\psi\rvert_B$

- States that are not pure are **mixed**, and the previous definitions yield the following classification of **mixed composite states**:

Mixed composite	ρ_{AB}
Product	$\sigma_A \otimes \tau_B$
Separable	$\sum_{x\in\mathcal{X}} p_X(x)\sigma_A^x \otimes \tau_B^x$
Entangled	$\neq \sum_{x\in\mathcal{X}} p_X(x)\sigma_A^x \otimes \tau_B^x$

- Note that **entangled** states are not constructively defined. The **Schmidt decomposition** provides a better **characterization** of **pure composite** states.

Von Neumann entropy

Definition of quantum entropy

Von Neumann entropy or quantum entropy

For a given quantum state defined in system A by a density matrix $\rho_A \in \mathcal{D}(\mathcal{H}_A)$, the **entropy** of the quantum state is defined as:

$$H(\rho_A) \equiv H(A)_\rho \equiv -\mathrm{tr}\{\rho_A \log \rho_A\}.$$

Note that for ρ_A with **spectral decomposition** $\rho_A = \sum_{i,p_i \neq 0} p_i |\phi_i\rangle\langle\phi_i|_A$, $\log \rho_A \equiv \sum_{i,p_i \neq 0} \log p_i |\phi_i\rangle\langle\phi_i|_A$.

In this case[1],

$$\begin{aligned} H(A)_\rho &= -\mathrm{tr}\{\sum_i p_i |\phi_i\rangle\langle\phi_i|_A \sum_j \log p_j |\phi_j\rangle\langle\phi_j|_A\} \\ &= -\mathrm{tr}\{\sum_i p_i \log p_i |\phi_i\rangle\langle\phi_i|_A\} = -\sum_i p_i \log p_i. \end{aligned}$$

[1] As in the classical definition $x\log(x)$ is assumed 0 for $x = 0$.

Properties of the quantum entropy

- Note that if the density matrix is defined as the **ensemble** of a set of **non-orthogonal** pure states, $\rho_A = \sum_j p_j |\psi_j\rangle\langle\psi_j|_A$, then $H(A)_\rho \neq -\sum_j p_j \log p_j$.
- Inheriting many of the mathematical properties of the classical entropy we can easily prove that **the entropy is non-negative**, it is **zero for pure states** and **maximum**, $\log \dim(\mathcal{H}_A) = \log d_A$, for the **maximally mixed state** $\pi_A = \frac{1}{d_A} I_A$.
- The entropy is isometric invariant

$$H(\rho_A) = H(U \rho_A U^\dagger),$$

  since eigenvalues are **isometric invariant**.
- The entropy is **concave**, i.e.,

$$H(\rho_A) \geq \sum_x p_X(x) H(\rho_A^x), \text{ where } \rho_A = \sum_x p_X(x) \rho_A^x.$$

Joint quantum entropy

Definition of joint quantum entropy

Joint quantum entropy

For a given joint quantum state defined in systems A and B by a density matrix $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the **entropy** $H(AB)_\rho$ of the **joint quantum state** is defined as:

$$H(\rho_{AB}) \equiv H(AB)_\rho \equiv -\mathrm{tr}\{\rho_{AB} \log \rho_{AB}\}$$

Note that $H(A)_\rho = H(\rho_A)$ where $\rho_A = \mathrm{tr}_B\{\rho_{AB}\}$ and $H(B)_\rho$ is similarly defined.

- In the **classical** context we always have,

$$H(X,Y) = H(X) + H(Y|X) \geq H(X)$$
$$H(X,Y) = H(Y) + H(X|Y) \geq H(Y)$$

Is this true in quantum? (No)

Marginal entropies of a pure composite state (I)

Marginal entropies of a pure composite state (MEPCS)

For a given **pure joint quantum state** defined in systems A and B with (rank one) **density matrix** $\phi_{AB} = |\phi\rangle\langle\phi|_{AB}$, the **entropy** $H(AB)_\phi$ is

$$H(AB)_\phi = 0,$$

since the state is **pure**, and,

$$H(A)_\phi = H(B)_\phi.$$

Moreover, the entropies $H(A)_\phi$ and $H(B)_\phi$ are **strictly positive**, i.e.,

$$H(A)_\phi = H(B)_\phi > H(AB)_\phi = 0$$

iff the **Schmidt rank** is greater than one, i.e., the **pure quantum state** ϕ_{AB} is **entangled**.

Marginal entropies of a pure composite state (II)

The proof is **straightforward** using the **Schmidt decomposition** since for

$$|\phi\rangle_{AB} = \sum_{i=0}^{d-1} \sqrt{p_i} |\psi_i\rangle_A \otimes |\varphi_i\rangle_B,$$

$$\phi_{AB} \equiv |\phi\rangle\langle\phi|_{AB} = \sum_{i=0}^{d-1} \sum_{j=0}^{d-1} \sqrt{p_i p_j} |\psi_i\rangle\langle\psi_j|_A \otimes |\varphi_i\rangle\langle\varphi_j|_B,$$

and the density matrix on systems A and B can be expressed as

$$\rho_A = \mathrm{tr}_B\{\phi_{AB}\} = \sum_{i=0}^{d-1} p_i |\psi_i\rangle\langle\psi_i|_A, \quad \rho_B = \mathrm{tr}_A\{\phi_{AB}\} = \sum_{i=0}^{d-1} p_i |\varphi_i\rangle\langle\varphi_i|_B,$$

which implies $H(A)_\phi = H(B)_\phi$ since the entropy depends only on the **eigenvalues**. Note that **iff** $d \geq 2$ the entropies are **strictly positive**. ∎

Generalization for multiple systems

- The previous result can be extended for **pure** states defined on **more than two systems** $|\phi\rangle_{ABCD}$, by considering any arbitrary **cut**:

$$
\begin{aligned}
H(A)_\phi &= H(BCD)_\phi \\
H(AB)_\phi &= H(CD)_\phi \\
H(ABC)_\phi &= H(D)_\phi \\
H(B)_\phi &= H(ACD)_\phi \\
H(BC)_\phi &= H(AD)_\phi \\
H(BD)_\phi &= H(AC)_\phi \\
H(C)_\phi &= H(ABD)_\phi
\end{aligned}
$$

Properties of joint entropy

- Note that for **pure composite states**,

$$H(A)_\phi = H(B)_\phi \text{ and } H(AB)_\phi = 0,$$

whereas for **maximally correlated random variables** X and Y, their joint pmf is $p_{X,Y}(x,y) = p_X(x)\mathbb{1}\{x=y\}$ so we have

$$H(X) = H(Y) \text{ and } H(X,Y) = H(X) = H(Y) \geq 0.$$

- The **joint entropy is additive** for (Kronecker) **product states**,

$$H(\rho_A \otimes \sigma_B) = H(\rho_A) + H(\sigma_B)$$

consequence of the **additivity** of the **Shannon entropy** of their corresponding **eigenvalues**.

Joint entropy of the classical-quantum state (I)

Joint entropy of the classical-quantum state

For the **classical-quantum** state defined as

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

we have

$$H(XB)\rho = H(X) + \sum_x p_X(x)H(\rho_B^x)$$

where $H(X)$ is the **entropy** of **random variable** X with **pmf** $p_X(x)$.

Joint entropy of the classical-quantum state (II)

For the **proof** consider

$$\begin{aligned}\log \rho_{XB} &= \log(\sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_B^x) = \log(\sum_x |x\rangle\langle x|_X \otimes p_X(x)\rho_B^x) \\ &= \sum_x |x\rangle\langle x|_X \otimes \log(p_X(x)\rho_B^x),\end{aligned}$$

then, by **noting** that $\mathrm{tr}\{\rho_A \otimes \rho_B\} = \mathrm{tr}\{\rho_A\}\mathrm{tr}\{\rho_B\}$,

$$\begin{aligned}H(XB)_\rho &= -\mathrm{tr}\{\rho_{XB} \log \rho_{XB}\} \\ &= -\mathrm{tr}\{(\sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_B^x)(\sum_{x'} |x'\rangle\langle x'|_X \otimes \log(p_X(x')\rho_B^{x'}))\} \\ &= -\mathrm{tr}\{\sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_B^x \log(p_X(x)\rho_B^x)\} \\ &= -\sum_x p_X(x)\mathrm{tr}\{\rho_B^x \log(p_X(x)\rho_B^x)\}.\end{aligned}$$

Joint entropy of the classical-quantum state (III)

But

$$\log(p_X(x)\rho_B^x) = \log p_X(x) I + \log \rho_B^x$$

and thus

$$\begin{aligned} H(XB)_\rho &= -\sum_x p_X(x)(\mathrm{tr}\{\log p_X(x)\rho_B^x\} + \mathrm{tr}\{\rho_B^x \log \rho_B^x\}) \\ &= -\sum_x p_X(x) \log p_X(x) - \sum_x p_X(x) \mathrm{tr}\{\rho_B^x \log \rho_B^x\} \\ &= H(X) + \sum_x p_X(x) H(\rho_B^x). \qquad \blacksquare \end{aligned}$$

Conditional entropy and coherent information

Definition of the conditional quantum entropy

Conditional quantum entropy

For a given joint quantum state defined in systems A and B by a density matrix $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the **conditional quantum entropy** $H(A|B)_\rho$ is defined as:

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

Note that the conditional quantum entropy **can be negative** and is in fact **strictly negative** for **entangled pure states**.

Conditioning can not increase entropy

Nevertheless, **conditioning can not increase entropy**. This property, **well known** in **classical**, is **preserved** even if the conditioning system is **quantum**,

$$H(A|B)_\rho \leq H(A)_\rho.$$

Examples (I)

Conditional quantum entropy of the maximally entangled state

The **conditional entropy** of the **maximally entangled** state

$$\Phi_{AB} \equiv |\Phi\rangle\langle\Phi|_{AB} = \frac{1}{d}\sum_{i=0}^{d-1}\sum_{j=0}^{d-1} |i\rangle\langle j|_A \otimes |i\rangle\langle j|_B,$$

with $d = \dim(\mathcal{H}_A) = \dim(\mathcal{H}_B)$ is

$$H(A|B)_\Phi = H(AB)_\Phi - H(B)_\Phi = -H(B)_\Phi = -\log d = H(B|A)_\Phi,$$

since $\rho_B = \mathrm{tr}_A\{\Phi_{AB}\} = \frac{1}{d} I_B = \pi_B$ and similarly $\rho_A = \frac{1}{d} I_A = \pi_A$.

- Note that Φ_{AB} should **not be confused** with the state of two **fully correlated uniformly distributed random variables** X and $Y = X$, where $\bar{\Phi}_{XY} = \frac{1}{d}\sum_{i=0}^{d-1} |i\rangle\langle i|_X \otimes |i\rangle\langle i|_Y$. In this case $H(X|Y) = 0$ and $H(X) = \log|\mathcal{X}| = H(Y)$.

Examples (and II)

Conditional quantum entropy of the classical-quantum state

The **conditional quantum entropy** of a **classical-quantum** state

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

is given by:

$$\begin{aligned} H(B|X)_\rho &= H(XB)_\rho - H(X)_\rho \\ &= H(X) + \sum_x p_X(x) H(\rho_B^x) - H(X) \\ &= \sum_x p_X(x) H(\rho_B^x). \end{aligned}$$

Observe the similarity with the **classical conditional entropy** where:

$$H(Y|X) = \sum_x p_X(x) H(Y|X = x).$$

Definition of coherent information

Coherent information

For a given **joint quantum state** defined in systems A and B by a density matrix $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the **coherent information** $I(A\rangle B)_\rho$ is defined as:

$$I(A\rangle B)_\rho \equiv H(B)_\rho - H(AB)_\rho = -H(A|B)_\rho.$$

As we know already, in the quantum context, $I(A\rangle B)_\rho$ **can be positive**. In fact it can be associated with the concept of **quantum information** and satisfies the **quantum data-processing inequality**, a well known inequality in the classical setting.

Example

Duality of conditional entropy

For a given **joint quantum state** defined in systems A and B by a density matrix $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, consider a **purification** $|\psi\rangle_{ABE}$ defined on system E. We have $\rho_{AB} = \mathrm{tr}_E\{|\psi\rangle\langle\psi|_{ABE}\}$ and thus,

$$\begin{aligned} -H(A|B)_\rho = I(A\rangle B)_\rho &= H(B)_\rho - H(AB)_\rho \\ &= H(B)_\psi - H(AB)_\psi \\ &= H(B)_\psi - H(E)_\psi \\ &= H(AE)_\psi - H(E)_\psi \\ &= -I(A\rangle E)_\psi = H(A|E)_\psi. \end{aligned}$$

Bounds of conditional entropy

Let $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, we have

$$-\log\dim(\mathcal{H}_A) \leq H(A|B)_\rho \leq \log\dim(\mathcal{H}_A)$$

Quantum mutual information

Definition of quantum mutual information

Quantum mutual information

For a given **joint quantum state** defined in systems A and B by a density matrix $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, the **quantum mutual information** $I(A;B)_\rho$ is defined as:

$$
\begin{aligned}
I(A;B)_\rho &\equiv H(A)_\rho + H(B)_\rho - H(AB)_\rho \\
&= H(A)_\rho - H(A|B)_\rho = H(B)_\rho - H(B|A)_\rho .
\end{aligned}
$$

- As in the **classical** context, $I(A;B)_\rho \geq 0$.
- For the **maximally entangled state** Φ_{AB}, that satisfies $\dim(\mathcal{H}_A) = \dim(\mathcal{H}_B) = d$,

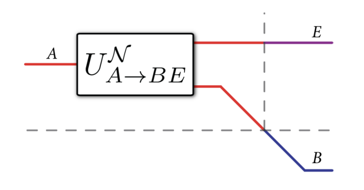
$$I(A;B)_\Phi = 2 \log d.$$

Examples (I)

Pure states and mutual information

Consider a **pure state** $|\psi\rangle_{RA} \in \mathcal{H}_R \otimes \mathcal{H}_A$, a **quantum channel** $\mathcal{N}_{A\to B}$ and an **Isometric Extension** $\mathcal{U}^{\mathcal{N}}_{A\to BE}$ of this channel acting on A share to produce **pure state** $|\phi\rangle_{RBE} \in \mathcal{H}_R \otimes \mathcal{H}_B \otimes \mathcal{H}_E$. We have

$$\begin{aligned}
I(R;A)_\psi &= H(R)_\psi + H(A)_\psi = 2H(R)_\psi = 2H(R)_\phi \\
&= H(R)_\phi + H(B)_\phi - H(RB)_\phi + H(R)_\phi - H(B)_\phi + H(RB)_\phi \\
&= H(R)_\phi + H(B)_\phi - H(RB)_\phi + H(R)_\phi - H(RE)_\phi + H(E)_\phi \\
&= I(R;B)_\phi + I(R;E)_\phi .
\end{aligned}$$



Exercise

Pure states and mutual information

Similarly, consider a **pure state** $|\psi\rangle_{SRA} \in \mathcal{H}_S \otimes \mathcal{H}_R \otimes \mathcal{H}_A$, a **quantum channel** $\mathcal{N}_{A\to B}$ and an **Isometric Extension** $\mathcal{U}^{\mathcal{N}}_{A\to BE}$ of this channel acting on A share to produce $|\phi\rangle_{SRBE} \in \mathcal{H}_S \otimes \mathcal{H}_R \otimes \mathcal{H}_B \otimes \mathcal{H}_E$. Prove

$$I(R;A)_\psi + I(R;S)_\psi = I(R;B)_\phi + I(R;SE)_\phi.$$

Examples (II)

Coherent information and private information

Consider a **pure state** $|\phi\rangle_{ABE}$ in ABE. By the **Schmidt decomposition**, it can be expressed as $|\phi\rangle_{ABE} = \sum_x \sqrt{p_X(x)}|\psi_x\rangle_A \otimes |\varphi_x\rangle_{BE}$. We now apply at A share a **measurement channel** $\{|\psi_x\rangle\langle\psi_x|\}$,

$$\begin{aligned}\bar{\phi}_{XBE} = \mathcal{M}_{A\to X}(\phi_{ABE}) &= \sum_x |x\rangle\langle x|_X \otimes \mathrm{tr}_A\{(|\psi_x\rangle\langle\psi_x|_A \otimes I_{BE})\phi_{ABE}\} \\ &= \sum_x p_X(x)|x\rangle\langle x|_X \otimes |\varphi_x\rangle\langle\varphi_x|_{BE},\end{aligned}$$

$$\begin{aligned}I(A\rangle B)_\phi &= H(B)_\phi - H(AB)_\phi = H(B)_\phi - H(E)_\phi = H(B)_{\bar{\phi}} - H(E)_{\bar{\phi}} \\ &= H(B)_{\bar{\phi}} - H(B|X)_{\bar{\phi}} - H(E)_{\bar{\phi}} + H(B|X)_{\bar{\phi}} \\ &\overset{(a)}{=} H(B)_{\bar{\phi}} - H(B|X)_{\bar{\phi}} - H(E)_{\bar{\phi}} + H(E|X)_{\bar{\phi}} \\ &\overset{(b)}{=} I(X;B)_{\bar{\phi}} - I(X;E)_{\bar{\phi}},\end{aligned}$$

Examples (and III)

- (a) Note that $H(B|X)_{\bar{\phi}} = H(E|X)_{\bar{\phi}}$ since, **conditioned on** X, the **density matrix** at BE is $|\varphi_x\rangle\langle\varphi_x|_{BE}$ and thus **pure**, so we can apply the MEPCS property.
- (b) We have obtained:

$$I(A\rangle B)_\phi = I(X;B)_{\bar{\phi}} - I(X;E)_{\bar{\phi}},$$

which indicates that the **coherent information** between Alice and Bob is related to the **private information capacity**, i.e., the capacity of the **degraded wiretap channel** between **Alice** and **Bob** where E, the **environment** in the quantum setting, plays the role of **Eve**, the **eavesdropper** in the classical setting.

Mutual information of classical-quantum states

Mutual information of classical-quantum states

Consider the following classical-quantum state:

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

The mutual information is

$$I(X;A)_\sigma = H(A)_\sigma - H(A|X)_\sigma = H(\rho_A) - \sum_x p_X(x)H(\rho_A^x)$$

where $\rho_A = \sum_x p_X(x)\rho_A^x = \mathrm{E}_X(\rho_A^x)$ is defined based on the ensemble $\mathcal{E} \equiv \{p_X(x), \rho_A^x\}$.

Note that the fact that the **mutual information is not negative** suffices to prove the **concavity of the quantum entropy**.

Definition of conditional quantum mutual information

Conditional quantum mutual information

For a given **joint quantum state** defined in systems ABC by a density matrix $\rho_{ABC} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B \otimes \mathcal{H}_C)$, the **conditional quantum mutual information** (CQMI), $I(A;B|C)_\rho$ is defined as:

$$I(A;B|C)_\rho \equiv H(A|C)_\rho + H(B|C)_\rho - H(AB|C)_\rho \geq 0.$$

The CQMI is **non-negative**, and this is a **fundamental** result in **quantum information theory**.

Chain rule for CQMI

Also as in the classical case we define the **chain rule** for the CQMI:

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

CQMI of a classical-quantum state

Non-negativity of the **conditional quantum mutual information**, if the conditioning system is classical, follows from the **non-negativity** of the **mutual information**.

CQMI of a classical-quantum state

Consider the classical-quantum state $\sigma_{XAB} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \sigma^x_{AB}$, we have

$$\begin{aligned} I(A;B|X)_\sigma &= H(A|X)_\sigma + H(B|X)_\sigma - H(AB|X)_\sigma \\ &= \sum_x p_X(x)\left(H(\sigma^x_A) + H(\sigma^x_B) - H(\sigma^x_{AB})\right) \\ &= \sum_x p_X(x) I(A;B)_{\sigma^x}. \end{aligned}$$

Quantum relative entropy

Quantum relative entropy preliminaries

Kernel and support

The *kernel* or *null space* of an operator $M \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ is

$$\ker(M) \equiv \{|\psi\rangle \in \mathcal{H}_A : M|\psi\rangle = 0\}.$$

The *support* of M is the subspace in $\mathcal{H}_A$ orthogonal to the kernel:

$$\mathrm{supp}(M) \equiv \{|\psi\rangle \in \mathcal{H}_A : M|\psi\rangle \neq 0\}.$$

If M is Hermitian with spectral decomposition $M = \sum_{i,a_i\neq 0} a_i|\psi_i\rangle\langle\psi_i|$, then

$$\mathrm{supp}(M) = \mathrm{span}\{|\psi_i\rangle : a_i \neq 0\},$$

and the associated projector onto $\mathrm{supp}(M)$,

$$\Pi_M = \sum_{i,a_i\neq 0} |\psi_i\rangle\langle\psi_i|.$$

Definition

Quantum relative entropy

The quantum relative entropy $D(\rho\|\sigma)$ between density matrix $\rho \in \mathcal{D}(\mathcal{H})$ and positive semi-definite operator $\sigma \in \mathcal{L}(\mathcal{H})$ is defined as

$$D(\rho\|\sigma) \equiv \mathrm{tr}\{\rho(\log \rho - \log \sigma)\},$$

if $\mathrm{supp}(\rho) \subseteq \mathrm{supp}(\sigma)$ and $+\infty$ otherwise.

Note that this definition is consistent with the classical definition which, for pmf p and non-negative real function q defined on $\mathcal{X}$,

$$D(p\|q) \equiv \sum_x p(x)(\log p(x) - \log q(x)),$$

If $p(x) > 0$ and $q(x) = 0$ for any $x \in \mathcal{X}$, then $D(p||q) = \infty$.

Relation to other entropic measures

Relative entropy equivalences

For density matrix $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$,

$$\begin{aligned} H(A)_\rho &= -D(\rho_A \| \pi_A) + \log d_A = -D(\rho_A \| I_A), \\ H(A|B)_\rho &= -D(\rho_{AB} \| \pi_A \otimes \rho_B) + \log d_A = -D(\rho_{AB} \| I_A \otimes \rho_B), \\ I(A;B)_\rho &= D(\rho_{AB} \| \rho_A \otimes \rho_B), \\ I(A\rangle B)_\rho &= D(\rho_{AB} \| \pi_A \otimes \rho_B) - \log d_A = D(\rho_{AB} \| I_A \otimes \rho_B). \end{aligned}$$

Positivity of quantum relative entropy

For density matrices $\rho \in \mathcal{D}(\mathcal{H})$ and $\sigma \in \mathcal{D}(\mathcal{H})$,

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

This suffices to prove $I(A;B)_\rho = H(A)_\rho - H(A|B)_\rho \geq 0$.

Classical data processing

We recall first the **classical data processing inequality**,

Data processing inequality (classical)

If random variables X, Y, Z form a Markov chain, $X \leftrightarrow Y \leftrightarrow Z$, i.e., $p(z|x,y) = p(z|y)$ or equivalently $p(x,z|y) = p(x|y)p(z|y)$, then

$$I(X;Y) \geq I(X;Z).$$

- This result can be interpreted as **processing classical data reduces classical correlations**.
- A **similar** result holds in the **quantum** case for both the **coherent information** and the **mutual information**.

Quantum data processing (I)

Monotonicity of quantum relative entropy

For $\rho \in \mathcal{D}(\mathcal{H}_A)$ and positive semi-definite $\sigma \in \mathcal{L}(\mathcal{H}_A)$ and quantum channel $\mathcal{N}_{A\to B}$, then

$$D(\rho\|\sigma) \geq D(\mathcal{N}(\rho)\|\mathcal{N}(\sigma))$$

The **proof** of this inequality is **not trivial** [Wilde, 2017] but easily **yields** the **data processing inequalities** for **mutual** and **coherent** information.

Data processing for mutual information

For density matrix $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and quantum channel $\mathcal{N}_{A\to A'}$ and $\mathcal{M}_{B\to B'}$, let $\sigma_{A'B'} \equiv (\mathcal{N}_{A\to A'} \otimes \mathcal{M}_{B\to B'})(\rho_{AB})$, then

$$I(A;B)_\rho \geq I(A';B')_\sigma.$$

Quantum data processing (II)

We prove the previous inequality by identifying:

$$I(A;B)_\rho = D(\rho_{AB} \| \rho_A \otimes \rho_B),$$

and

$$\begin{aligned}
I(A';B')_\sigma &= D(\sigma_{A'B'} \| \sigma_{A'} \otimes \sigma_{B'}) \\
&= D((\mathcal{N}_{A\to A'} \otimes \mathcal{M}_{B\to B'})(\rho_{AB}) \| \mathcal{N}_{A\to A'}(\rho_A) \otimes \mathcal{M}_{B\to B'}(\rho_B)) \\
&= D((\mathcal{N}_{A\to A'} \otimes \mathcal{M}_{B\to B'})(\rho_{AB}) \| (\mathcal{N}_{A\to A'} \otimes \mathcal{M}_{B\to B'})(\rho_A \otimes \rho_B)).
\end{aligned}$$

■

A similar procedure can be used to prove the **data processing inequality** for coherent information in next slide.

Quantum data processing (and III)

Data processing for coherent information

For density matrix $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$ and **quantum** channel $\mathcal{M}_{B\to B'}$, let $\sigma_{AB'} \equiv (\mathrm{id}_A \otimes \mathcal{M}_{B\to B'})(\rho_{AB})$, then

$$I(A\rangle B)_\rho \geq I(A\rangle B')_\sigma.$$

This result can be generalized to **unital channels** acting on A, i.e., channels for which $\mathcal{N}_{A\to A'}(I_A) = I_{A'}$.

Data processing for coherent information with unital channels on A

For density matrix $\rho_{AB} \in \mathcal{D}(\mathcal{H}_A \otimes \mathcal{H}_B)$, **unital** quantum channel $\mathcal{N}_{A\to A'}$ and **quantum** channel $\mathcal{M}_{B\to B'}$, let $\sigma_{A'B'} \equiv (\mathcal{N}_{A\to A'} \otimes \mathcal{M}_{B\to B'})(\rho_{AB})$, then

$$I(A\rangle B)_\rho \geq I(A'\rangle B')_\sigma.$$

Conditional entropy of classical states $H(X|A)$

As an **example** of **application** of the **data processing inequality** we can show that the **conditional entropy** of a classical state is non-negative.
Consider a **preparation** channel $\mathcal{N}_{Y\to A}$ which uses pure states $|\phi_x\rangle\langle\phi_x|_A$ applied to the Y share of a shared randomness state $\bar{\Phi}_{XY}$:

$$\bar{\Phi}_{XY} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes |x\rangle\langle x|_Y,$$

$$\rho_{XA} = (\mathrm{id}_X \otimes \mathcal{N}_{Y\to A})(\bar{\Phi}_{XY}) = \sum_x p_X(x)|x\rangle\langle x|_X \otimes |\phi_x\rangle\langle\phi_x|_A.$$

By the data processing inequality of mutual information we have:

$$H(X) = I(X;Y)_{\bar{\Phi}} \geq I(X;A)_\rho = H(A)_\rho - H(A|X)_\rho = H(A)_\rho$$

Note that equality is attained for orthogonal states $|\phi_x\rangle\langle\phi_x|_A$. Also,

$$H(X|A)_\rho = H(X) - H(A)_\rho + H(A|X)_\rho = H(X) - H(A)_\rho \geq 0.$$

Trace distance

Quantum Pinsker inequality (I)

We recall first the **classical** Pinsker inequality,

Pinsker inequality (classical)

Let p be a pmf on $\mathcal{X}$ and $q : \mathcal{X} \to [0,1]$ such that $\sum_x q(x) \leq 1$, then

$$D(p\|q) \geq \frac{1}{2\ln 2}\|p-q\|_1^2$$

where $\|p-q\|_1 = \sum_x |p(x)-q(x)|$.

- This inequality relates two measures of similarities between pmfs.
- Using the monotonicity of quantum relative entropy and defining the trace distance, this result can be extended to the quantum setting.

Quantum Pinsker inequality (and II)

Quantum Pinsker inequality

Let $\rho \in \mathcal{D}(\mathcal{H})$ and positive semi-definite $\sigma \in \mathcal{L}(\mathcal{H})$ such that $\mathrm{tr}\{\sigma\} \leq 1$,

$$D(\rho\|\sigma) \geq \frac{1}{2\ln 2}\|\rho - \sigma\|_1^2$$

where for any $M \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$, the trace norm is defined as $\|M\|_1 = \mathrm{tr}\{\sqrt{MM^\dagger}\}$.

Trace norm (I)

Trace norm definition

The **trace** or **Schatten 1** or **nuclear norm** of operator $M \in \mathcal{L}(\mathcal{H}_A, \mathcal{H}_B)$ is defined as:

$$\|M\|_1 \equiv \mathrm{tr}\{|M|\} = \mathrm{tr}\{\sqrt{M^\dagger M}\}$$

Note that for $M = U\Sigma V^\dagger$ the SVD of M, U and V contain the **orthonormal singular vectors** and Σ contains the **non-negative singular values** along the diagonal, thus $M = \sum_{i=0}^{d-1} \lambda_i |u_i\rangle\langle v_i|$.

$$M^\dagger M = (\sum_{i=0}^{d-1} \lambda_i |v_i\rangle\langle u_i|)(\sum_{j=0}^{d-1} \lambda_j |u_j\rangle\langle v_j|) = \sum_{i=0}^{d-1} \lambda_i^2 |v_i\rangle\langle v_i|,$$

and $\|M\|_1 = \mathrm{tr}\{\sqrt{M^\dagger M}\} = \sum_{i=0}^{d-1} \lambda_i$.

Trace norm (II)

Properties of the trace norm

The **trace norm** satisfies the three required properties of any **norm**:

- Non-negative definiteness: $\|M\|_1 \geq 0$.
- Homogeneity: For any constant $c \in \mathbb{C}$, $\|cM\|_1 = |c|\|M\|_1$.
- Triangle inequality: $\|M + N\|_1 \leq \|M\|_1 + \|N\|_1$.

The first property follows considering the SVD of matrix M, the second follows directly from the definition and the third, for square operators M, can be proved using the SVD and the Cauchy-Schwarz inequality.

Trace distance (I)

Trace distance between two density operators

The **trace norm induces** a useful **distance** between **density matrices**. Let $\rho, \sigma \in \mathcal{D}(\mathcal{H}_A)$, then making use of the non-negative definiteness and triangle inequality properties,

$$0 \leq \|\rho - \sigma\|_1 \leq \|\rho\|_1 + \|\sigma\|_1 = 2$$

or equivalently,

$$0 \leq \frac{1}{2}\|\rho - \sigma\|_1 \leq 1$$

Trace distance (and II)

Trace distance as probability difference

The normalized trace distance $\frac{1}{2}\|\rho - \sigma\|_1$ between quantum states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$ is equal to the largest probability difference that the two states could give to the same measurement outcome

$$\frac{1}{2}\|\rho - \sigma\|_1 = \max_{0 \leq \Lambda \leq I,} \mathrm{tr}\{\Lambda(\rho - \sigma)\}$$

Exercise: Use the previous result to find a measurement that achieves the trace distance

$$\|\rho - \sigma\|_1 = \max_{\{\Lambda_x\}} \sum_x |\mathrm{tr}\{\Lambda_x \rho\} - \mathrm{tr}\{\Lambda_x \sigma\}|$$

Trace distance and error probability

- The trace distance is used to evaluate the **performance** of **classical communication** over quantum channels **extending** the concept of **error probability**.
- As such, it becomes important to **relate** it to **conditional entropy** terms as in **Fano's inequality** in the **classical** setting.
- The **AFW inequality** next **does this**.

Alicki-Fannes-Winter inequality (I)

Fano's inequality (classical)

Let $(X, Y) \sim p(x, y)$ and suppose for $\epsilon \in [0, 1]$, $\mathsf{P}\{X \neq Y\} \leq \epsilon$. Then:

$$H(X|Y) \leq H(\epsilon) + \epsilon \log(|\mathcal{X}| - 1) \leq H(\epsilon) + \epsilon \log |\mathcal{X}|.$$

Alicki-Fannes-Winter inequality (proof in [Wilde, 2017])

Let $\rho_{AB}, \sigma_{AB} \in \mathcal{D}(\mathcal{H}_A \times \mathcal{H}_B)$ and suppose for $\epsilon \in [0, 1]$,

$$\frac{1}{2}\|\rho_{AB} - \sigma_{AB}\|_1 \leq \epsilon.$$

Then,

$$|H(A|B)_\rho - H(A|B)_\sigma| \leq 2\epsilon \log d_A + (1 + \epsilon) H(\epsilon/(1 + \epsilon)).$$

Where $H(\epsilon) = -\epsilon \log \epsilon - (1 - \epsilon) \log(1 - \epsilon)$ is the **binary entropy** function.

Alicki-Fannes-Winter inequality (and II)

The AFW becomes **tighter** (the "2" becomes "1") for **classical-quantum states**.

AFW inequality for classical-quantum states

Let ρ_{XB} and σ_{XB} represent classical-quantum states,

$$\rho_{XB} = \sum_x p_X(x)|x\rangle\langle x|_X \otimes \rho_B^x \; ; \; \sigma_{XB} = \sum_x q_X(x)|x\rangle\langle x|_X \otimes \sigma_B^x.$$

Suppose for $\epsilon \in [0,1]$, $\frac{1}{2}\|\rho_{XB} - \sigma_{XB}\|_1 \leq \epsilon$. This implies,

$$|H(X|B)_\rho - H(X|B)_\sigma| \leq \epsilon \log|\mathcal{X}| + (1+\epsilon)H(\epsilon/(1+\epsilon)),$$
$$|H(B|X)_\rho - H(B|X)_\sigma| \leq \epsilon \log d_B + (1+\epsilon)H(\epsilon/(1+\epsilon)).$$

Entropic Uncertainty Principle

Entropic Uncertainty Principle (I)

- The uncertainty principle implies: there is an **unavoidable** uncertainty in the **measurement outcomes** of **incompatible** (non-commuting) **observables**.
- But when measuring Φ_{AB}, if Alice measures the Z **observable**, then Bob can guess the outcome of her **measurement** with **certainty**.
- Also, if Alice were to measure the X **observable**, then Bob would also be able to guess the outcome of her **measurement** with **certainty**.
- This happens **in spite of the fact** that Z and X are **incompatible** observables since $[X, Z] = XZ - ZX = -2iY \neq 0$
- A revision of the **classical definition of the uncertainty principle** is necessary when A and B **share a quantum memory**.

Entropic Uncertainty Principle (II)

- Assume Alice and Bob share state ρ_{AB}. Then Alice applies a POVM defined as $\{\Lambda_A^x\}$ to yield:

$$\sigma_{XB} = \sum_x |x\rangle\langle x|_X \otimes \mathrm{tr}_A\{(\Lambda_A^x \otimes I_B)\rho_{AB}\}$$

- Now assume that Alice instead chooses to use a different POVM $\{\Gamma_A^z\}$. The post measurement state would be:

$$\tau_{ZB} = \sum_z |z\rangle\langle z|_Z \otimes \mathrm{tr}_A\{(\Gamma_A^z \otimes I_B)\rho_{AB}\}$$

- It makes sense to wonder regarding the uncertainty at Bob regarding these measurements to be reflected in the sum:

$$H(X|B)_\sigma + H(Z|B)_\tau$$

- Can we lower bound this? Note that for $\rho_{AB} = \Phi_{AB}$ we have $H(X|B)_\sigma = H(Z|B)_\tau = 0$.

Entropic Uncertainty Principle (and III)

- One way to quantify the incompatibility of POVMs $\{\Lambda_A^x\}$ and $\{\Gamma_A^z\}$ is computing:

$$c \equiv \max_{x,z} \| \sqrt{\Lambda_A^x} \sqrt{\Gamma_A^z} \|_\infty^2$$

where $\|A\|_\infty$ is the maximal eigenvalue of $|A|$.

- If the POVM have one common element then $c = 1$, whereas for the X and Z observables $c = 1/2$.

Uncertainty Principle with Quantum Memory

$$H(X|B)_\sigma + H(Z|B)_\tau \geq \log(1/c) + H(A|B)_\rho$$

Note that for $\rho_{AB} = \Phi_{AB}$ and the X and Z observables the lower bound is zero since $H(A|B)_\rho = -1$. Also note this implies:

$$H(X) + H(Z) \geq \log(1/c) + H(A)_\rho$$

Proof of Uncertainty Principle with Quantum Memory

Uncertainty Principle with Quantum Memory

- Statement of the Uncertainty Principle (I)

$$H(X|B)_\sigma + H(Z|B)_\tau \geq \log(1/c) + H(A|B)_\rho$$

- Part 1 of the proof [Wilde, 2017].

$$H(X|B)_\sigma + H(Z|E)_\omega \geq \log(1/c)$$

where

$$\omega_{ZE} = \sum_z |z\rangle\langle z| \otimes \mathrm{tr}_{AB}\{(\Gamma_A^z \otimes I_{BE})\phi_{ABE}\}$$

and ϕ_{ABE} is a purification of ρ_{AB}.

Proof of Uncertainty Principle with Quantum Memory

Uncertainty Principle with Quantum Memory

- Part 2 of the proof. For

$$\omega_{ZE} = \sum_z |z\rangle\langle z| \otimes \mathrm{tr}_{AB}\{(\Gamma_A^z \otimes I_{BE})\phi_{ABE}\}$$

and ϕ_{ABE} is a purification of ρ_{AB}. show that

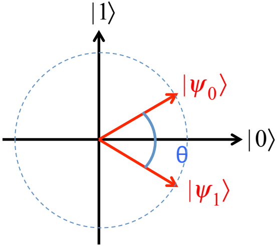
$$H(Z|E)_\omega = H(Z|B)_\tau - H(A|B)_\rho$$

Binary State Discrimination

Binary State Discrimination (from [Pagès-Zamora, 2020])

Problem setup:

- Alice sends one of two states $\{|\psi_0\rangle, |\psi_1\rangle\}$ corresponding to hypothesis $\{H_0, H_1\}$ respectively, with equal probability $p(H_0) = p(H_1) = \frac{1}{2}$.

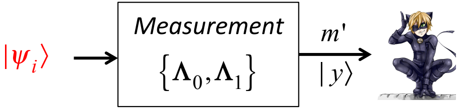


$$|\psi_0\rangle = \begin{bmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{bmatrix} ; |\psi_1\rangle = \begin{bmatrix} \cos\frac{\theta}{2} \\ -\sin\frac{\theta}{2} \end{bmatrix}$$

- Bob decides H_0 or H_1 with an average probability of error

$$P_e = P(H_1|H_0)P(H_0) + P(H_0|H_1)P(H_1)$$

Binary State Discrimination



- What is the measurement Bob needs to attain the minimum P_e?
- Two cases:
  (a) $\{|\psi_0\rangle, |\psi_1\rangle\}$ are orthonormal $(\theta = \pi/2)$ and known
  (b) $\{|\psi_0\rangle, |\psi_1\rangle\}$ are unknown, not necessarily orthonormal.

Binary State Discrimination. Case a)

Case a: States known and $\theta = \frac{\pi}{2}$

- Bob's outcome is $m' \in \{0, 1\}$ with probability

$$m' = \begin{cases} 0 & \text{with } \Pr\{m' = 0\} = \langle \psi_i | \Lambda_0 | \psi_i \rangle \\ 1 & \text{with } \Pr\{m' = 1\} = \langle \psi_i | \Lambda_1 | \psi_i \rangle \end{cases}$$

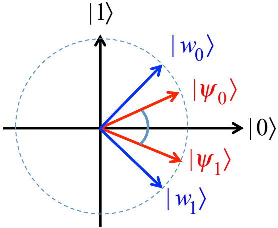
- Bob attains $P_e = 0$ if

$$\Lambda_0 = |\psi_0\rangle\langle\psi_0| \quad ; \quad \Lambda_1 = |\psi_1\rangle\langle\psi_1|$$

which is a valid POVM since $\Lambda_0 + \Lambda_1 = I$.

Case b: States unknown and not necessarily orthonormal.

- Bob uses an orthonormal basis $\{|w_0\rangle, |w_1\rangle\}$ to build his POVM



$$\Lambda_0 = |w_0\rangle\langle w_0| \quad ; \quad \Lambda_1 = |w_1\rangle\langle w_1|$$

- The average probability of error P_e is equal to

$$P(H_1|H_0)P(H_0) + P(H_0|H_1)P(H_1) = \frac{1}{2}\left(\mathrm{tr}\{\Lambda_1|\psi_0\rangle\langle\psi_0|\} + \mathrm{tr}\{\Lambda_0|\psi_1\rangle\langle\psi_1|\}\right)$$

$$= \frac{1}{2} + \frac{1}{2}\langle w_1|\big(|\psi_0\rangle\langle\psi_0|) - |\psi_1\rangle\langle\psi_1|\big)|w_1\rangle$$

where in the last equality we use $\Lambda_0 = I - \Lambda_1$; $\Lambda_1 = |w_1\rangle\langle w_1|$; and the property that $\mathrm{tr}\{\cdot\}$ commutes.

Binary State Discrimination. Case b) (II)

- The value of $|w_1\rangle$ that minimizes P_e is the eigenvector of matrix $|\psi_0\rangle\langle\psi_0|) - |\psi_1\rangle\langle\psi_1|$ associated to the minimum eigenvalue.
- It is not difficult to find that

$$|\psi_0\rangle\langle\psi_0| - |\psi_1\rangle\langle\psi_1| = \begin{bmatrix} 0 & \sin\theta \\ \sin\theta & 0 \end{bmatrix}$$

  with eigenvalues $\{\pm\sin\theta\}$ and associated eigenvectors $\{|\pm\rangle\}$.
- Therefore, $|w_1\rangle = |-\rangle$ and $|w_0\rangle = |+\rangle$, and the POVM are

$$\Lambda_0 = |w_0\rangle\langle w_0| = \frac{1}{2}\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

$$\Lambda_1 = |w_1\rangle\langle w_1| = \frac{1}{2}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$

- The minimum probability of error is equal to $P_{e,min} = \frac{1}{2}(1 - \sin\theta)$

Generalization to arbitrary states

- The optimum POVM $\{\Lambda_0, \Lambda_1\}$ to discriminate among two arbitrary states $\rho, \sigma \in \mathcal{D}(\mathcal{H})$, is the one that maximizes the probability of success which, for equally likely states is

$$\begin{aligned} p_s(\Lambda) &= \mathrm{tr}\{\Lambda_0 \rho\}\frac{1}{2} + \mathrm{tr}\{\Lambda_1 \sigma\}\frac{1}{2} \\ &= \frac{1}{2}(1 + \mathrm{tr}\{\Lambda_0(\rho - \sigma)\} \\ &= \frac{1}{2}(1 + \frac{1}{2}\|\rho - \sigma\|_1) \end{aligned}$$

- Therefore the minimum error probability is given by

$$\begin{aligned} p_e(\Lambda) &= 1 - p_s(\Lambda) \\ &= \frac{1}{2}(1 - \frac{1}{2}\|\rho - \sigma\|_1) \end{aligned}$$

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.

Alba Pagès-Zamora.
Lecture 2. Quantum protocols and channels *lecture notes*.
Quantum communication and computation (230381), ETSETB, UPC.