

LICENCE EN PHYSIQUE
Introduction à l'information quantique

Jordan David
Travail effectué sous la direction du professeur N.Gisin
Université de Genève

February 11, 2002

Contents

1	No cloning theorem & UQCM	5
1.1	Introduction	5
1.2	No cloning theorem	6
1.2.1	Formulation du théorème	6
1.2.2	Démonstration	7
1.3	Essai de copie : Mesure d'intensité <i>gamma</i>	8
1.3.1	Etat du système d'Alice	10
1.3.2	Etat du système de Bob	12
1.3.3	Résumé	14
1.4	Universal Quantum Cloning Machine (UQCM)	14
1.4.1	Introduction	14
1.4.2	Définition de la UQCM	15
1.4.3	No-signaling	16
1.4.4	La PQCM permettrait le <i>signaling</i>	17
1.4.5	UQCM & <i>No-signaling</i>	19
1.4.6	Détermination de <i>êta</i>	22
1.4.7	Cloneuse <i>Hillery-Bužek</i>	23
2	The controlled-NOT gate	33
2.1	Introduction	33
2.1.1	Exemples	34
2.2	Exemple de processeur quantique	35
2.2.1	Exemples d'input	36
2.2.2	Matrices densités partielles	39
2.3	Concentration d'intrication	41
2.3.1	Introduction	41
2.3.2	Concentration sans intrication	42
2.3.3	Principe de purification d'intrication	44

2.4	Correction d'erreur	46
2.4.1	Introduction	46
2.4.2	Procédé de correction d'erreur	46
3	Quantum information theory	51
3.1	Théorie de l'information classique	51
3.1.1	Introduction	51
3.1.2	Entropie de Shannon	52
3.1.3	Exemples	54
3.1.4	Information conditionnelle	56
3.1.5	Information mutuelle	56
3.1.6	Probabilité conjointe	57
3.1.7	Exemples	58
3.1.8	The noisy channel theorem (en bref)	61
3.2	Théorie de l'information quantique	62
3.2.1	Introduction	62
3.2.2	Entropie de Von Neumann	62
3.2.3	Exemple	65
3.2.4	Théorème d'Holevo	67
3.2.5	Relation d'incertitude informationnelle	69

Chapter 1

No cloning theorem & UQCM

1.1 Introduction

On peut commencer par dire que l'information quantique peut être représentée par un état quantique (ψ, ρ) .

Definition 1 *L'état quantique le plus élémentaire (non trivial) est*

$$\psi = c_0 |0\rangle + c_1 |1\rangle \in C^2 \quad (1.1)$$

où $c_0, c_1 \in C$
 $|c_0|^2 + |c_1|^2 = 1$
 $|0\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ et $|1\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$
Cet état est appelé un **qubit**.

On va constater qu'avec l'information quantique un champ de possibilités radicalement nouveau s'ouvre par rapport à celui offert par l'information classique.

On connaît par expérience quelques traitements élémentaires de l'information classique : on peut la copier, la transmettre, la coder...

En revanche, et c'est là un point capital car lourd de conséquences, la théorie de l'information quantique contient un théorème, appelé *No cloning theorem*, qui montre qu'on ne peut pas faire une copie parfaite d'une information de nature quantique.

1.2 No cloning theorem

1.2.1 Formulation du théorème

Soit $\psi = c_0 |0\rangle + c_1 |1\rangle$. On l'énonce trois fois :

1. Une transformation

$$\psi \longrightarrow \psi \otimes \psi \quad \forall \psi \quad (1.2)$$

n'existe pas.

2. Soit H un espace hilbertien t.q. $\dim H \geq 2$ et U un opérateur unitaire défini sur l'espace H . (Rappel: U unitaire $\Rightarrow U^\dagger = U^{-1}$).

Il n'existe pas d'opérateur U t.q.

$$U[\psi \otimes |0\rangle] = \psi \otimes \psi \quad \forall \psi \quad (1.3)$$

Classiquement, $|0\rangle$ représenterait la "la feuille blanche de la photocopieuse".

3. Soit H un espace hilbertien t.q. $\dim H \geq 2$ et U un opérateur unitaire défini sur l'espace H . Soit $H_{auxiliaire}$ un espace hilbertien et $\phi_i, \phi_f \in H_{aux}$.

Il n'existe pas d'opérateur U et d'espace H_{aux} t.q.

$$U[\psi \otimes |0\rangle \otimes \phi_i] = \psi \otimes \psi \otimes \phi_f \quad \forall \psi \quad (1.4)$$

Classiquement, ϕ_i et ϕ_f représenteraient respectivement "la photocopieuse *avant* et *après* avoir effectué la copie".

En bref, ce théorème nous dit que, si $|0\rangle \perp |1\rangle$, alors

$$\begin{array}{lll} |0\rangle & \rightarrow & |00\rangle & \text{est possible} \\ |1\rangle & \rightarrow & |11\rangle & \text{est possible} \\ |0+1\rangle & \rightarrow & |0+1\rangle \otimes |0+1\rangle & \text{est impossible} \end{array} \quad (1.5)$$

1.2.2 Démonstration

La démonstration se base sur le principe de superposition. On commence par démontrer l'énoncé n°2 (point 1) par l'absurde, puis que le n°3 peut s'écrire comme le n°2 (point 2).

1. Hypothèse : il existe un opérateur U . On a donc

$$\begin{aligned} U : |0\rangle \otimes |0\rangle &\rightarrow |0\rangle \otimes |0\rangle \\ |1\rangle \otimes |0\rangle &\rightarrow |1\rangle \otimes |1\rangle \end{aligned} \quad (1.6)$$

Ainsi, si le premier qubit se trouve dans l'état $\psi = c_0 |0\rangle + c_1 |1\rangle$, alors

$$\begin{aligned} U [(c_0 |0\rangle + c_1 |1\rangle) \otimes |0\rangle] &= c_0 U |00\rangle + c_1 U |10\rangle \\ &= c_0 |00\rangle + c_1 |11\rangle \end{aligned} \quad (1.7)$$

D'où

$$U [(c_0 |0\rangle + c_1 |1\rangle) \otimes |0\rangle] \neq (c_0 |0\rangle + c_1 |1\rangle) \otimes (c_0 |0\rangle + c_1 |1\rangle) \quad (1.8)$$

$\Rightarrow$ Contradiction.

L'hypothèse est donc fausse, i.e. il n'existe donc pas un tel opérateur U .

2. Soit ϕ_i et ϕ_f indépendants de ψ et V un opérateur unitaire qui envoie H_{aux} sur H_{aux} et ϕ_f sur ϕ_i . (Comme V unitaire $\Rightarrow V^\dagger$ envoie ϕ_i sur ϕ_f). On introduit encore le projecteur $P = \mathbb{1} \otimes \mathbb{1} \otimes |\phi_i\rangle \langle \phi_i|$.

On calcule

$$P (\mathbb{1} \otimes \mathbb{1} \otimes V) U P (\psi \otimes |0\rangle \otimes \phi_i) = \psi \otimes \psi \otimes \phi_i \quad (1.9)$$

avec

$$U : \psi \otimes |0\rangle \otimes \phi_i \rightarrow \psi \otimes \psi \otimes \phi_f \quad (1.10)$$

On définit $\tilde{U} = P (\mathbb{1} \otimes \mathbb{1} \otimes V) U P$. Ainsi

$$\tilde{U} : H \otimes H \otimes C \rightarrow H \otimes H \otimes \phi_i \quad (1.11)$$

où C est le sous-espace engendré par ϕ_i .

Comme

$$\dim \tilde{U} = \dim (H \otimes H \otimes C) = \dim H * \dim H * \dim C \text{ et } \dim C = 1 \quad (1.12)$$

alors $H \otimes H \otimes C$ est isomorphe à $H \otimes H$, d'où

$$\tilde{U} : H \otimes H \rightarrow H \otimes H \quad (1.13)$$

Ainsi, le *no cloning theorem* nous dit qu'il n'existe aucune "machine unitaire" capable de faire une copie (parfaite) d'états distincts et non-orthogonaux. L'intuition nous pousse à croire qu'on ne peut donc rien faire de cette information.

1.3 Essai de copie : Mesure d'intensité *gamma*

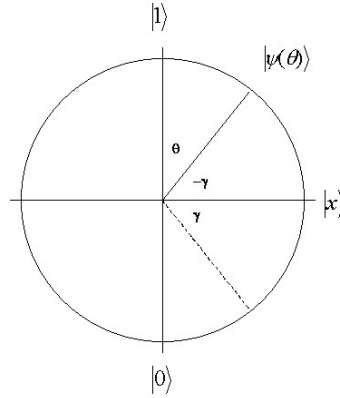
On va néanmoins essayer de produire quelque chose qui va dans le sens d'une copie.

Soit

$$\psi(\theta) = \cos\left(\frac{\theta}{2}\right) |0\rangle + \sin\left(\frac{\theta}{2}\right) |1\rangle \quad (1.14)$$

(rappel : $\psi(\theta) \perp \psi(\theta + \pi)$)

Sur la sphère de Poincaré, les états $\psi(\theta)$ sont représentés par les points du cercle qui passe par les pôles de la sphère symbolisés par $|0\rangle$ et $|1\rangle$. θ est l'angle entre les directions définies par les points représentés par $|1\rangle$ et $\psi(\theta)$.



Soit l'opération U t.q.

$$\begin{aligned}
U : \quad |0\rangle \otimes |x\rangle &\rightarrow |0\rangle \otimes \psi\left(\frac{\pi}{2} + \gamma\right) \\
|1\rangle \otimes |x\rangle &\rightarrow |1\rangle \otimes \psi\left(\frac{\pi}{2} - \gamma\right)
\end{aligned} \tag{1.15}$$

où $|x\rangle = \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle) = \psi\left(\frac{\pi}{2}\right)$ et $\gamma \in \left[0, \frac{\pi}{2}\right]$ paramétrise la quantité d'information copiée.

Remark 1 Les vecteurs $|0\rangle$ et $|1\rangle$ dans (1.15) représentent la donnée (input) qui n'est pas perturbée par l'action de U . Le vecteur $|x\rangle$ correspond à l'essai de copie. On constate en effet que le vecteur $|x\rangle$ sous l'action de U “bouge” dans la direction du qubit donné en entrée (input), ce qui signifie qu'un peu d'information est copié.

Cet “essai de copie” se situe entre les deux cas extrêmes suivants :

1. $\gamma = 0$. C'est le cas où le vecteur $|x\rangle$ ne subit aucune interaction. L'opération (1.15) correspond alors à l'identité.
2. $\gamma = \frac{\pi}{2}$. C'est le cas où la copie est parfaite. A noter que cela ne contredit pas le *no-cloning theorem* (c.f. 1.5)

On va maintenant appliquer l'opération U avec comme donnée d'entrée l'état élémentaire $\psi = c_0 |0\rangle + c_1 |1\rangle$.

$$\begin{aligned}
U(c_0 |0\rangle + c_1 |1\rangle)_A \otimes |x\rangle_B &= c_0 |0\rangle \otimes \psi\left(\frac{\pi}{2} + \gamma\right) + c_1 |1\rangle \otimes \psi\left(\frac{\pi}{2} - \gamma\right) \quad (1.16) \\
&= c_0 \cos\left(\frac{\pi}{4} + \frac{\gamma}{2}\right) |00\rangle + c_0 \sin\left(\frac{\pi}{4} + \frac{\gamma}{2}\right) |01\rangle \\
&\quad + c_1 \cos\left(\frac{\pi}{4} - \frac{\gamma}{2}\right) |10\rangle + c_1 \sin\left(\frac{\pi}{4} - \frac{\gamma}{2}\right) |11\rangle \\
&= \Psi_{out}
\end{aligned}$$

où $U(c_0 |0\rangle + c_1 |1\rangle)_A \otimes |x\rangle_B \in H_{Alice} \otimes H_{Bob}$

On a ainsi calculé l'état de sortie Ψ_{out} de la “machine” U pour un input correspondant à l'état élémentaire. Cet état appartient à l'espace de Hilbert $H_{Alice} \otimes H_{Bob}$. Quel est l'état du système d'Alice? Celui de Bob?

1.3.1 Etat du système d'Alice

L'état du système d'Alice est donné par la matrice densité réduite définie par

$$\rho_A = Tr_B (P_{\Psi_{out}}) \quad (1.17)$$

On pose $c_{\pm} = \cos\left(\frac{\pi}{4} \pm \frac{\gamma}{2}\right)$, $s_{\pm} = \sin\left(\frac{\pi}{4} \pm \frac{\gamma}{2}\right)$ et $\varphi_c = c_0 c_+ |0\rangle + c_1 c_- |1\rangle$, $\varphi_s = c_0 s_+ |0\rangle + c_1 s_- |1\rangle$.

$$\begin{aligned} \rho_A &= Tr_B (|\varphi_c \otimes |0\rangle_B + \varphi_s \otimes |1\rangle_B\rangle \langle \varphi_c \otimes |0\rangle_B + \varphi_s \otimes |1\rangle_B|) \\ &= Tr_B (P_{\varphi_c} \otimes P_{|0\rangle_B} + |\varphi_c\rangle \langle \varphi_s| \otimes |0\rangle \langle 1|_B) \\ &\quad + Tr_B (|\varphi_s\rangle \langle \varphi_c| \otimes |1\rangle \langle 0|_B + P_{\varphi_s} \otimes P_{|1\rangle_B}) \\ &= P_{\varphi_c} Tr (P_{|0\rangle_B}) + |\varphi_c\rangle \langle \varphi_s| \cdot 0 + |\varphi_s\rangle \langle \varphi_c| \cdot 0 + P_{\varphi_s} Tr (P_{|1\rangle_B}) \\ &= (c_0 c_+ |0\rangle + c_1 c_- |1\rangle) (\dots)^\dagger + (c_0 s_+ |0\rangle + c_1 s_- |1\rangle) (\dots)^\dagger \\ &= \begin{pmatrix} |c_0 c_+|^2 + |c_0 s_+|^2 & c_0 c_1^* c_+ c_- + c_0 c_1^* s_+ s_- \\ c_1 c_0^* c_+ c_- + c_1 c_0^* s_+ s_- & |c_1 c_-|^2 + |c_1 s_-|^2 \end{pmatrix} \\ &= \begin{pmatrix} |c_0|^2 & c_0 c_1^* \cos \gamma \\ c_1 c_0^* \cos \gamma & |c_1|^2 \end{pmatrix} \end{aligned} \quad (1.18)$$

(A noter que ρ_A est normalisée, i.e. $Tr(\rho_A) = 1$)

Si on se place aux deux valeurs limites de l'angle γ , la matrice densité réduite ρ_A devient :

1. Si $\gamma = 0 \Rightarrow$

$$\begin{aligned} \rho_A &= \begin{pmatrix} |c_0|^2 & c_0 c_1^* \\ c_0^* c_1 & |c_1|^2 \end{pmatrix} \\ &= P_{c_0|0\rangle + c_1|1\rangle} \end{aligned} \quad (1.19)$$

donc toute l'information de l'input est dans le système d'Alice ($\Rightarrow U = \mathbb{1}$).

2. Si $\gamma = \frac{\pi}{2} \Rightarrow$

$$\begin{aligned} \rho_A &= \begin{pmatrix} |c_0|^2 & 0 \\ 0 & |c_1|^2 \end{pmatrix} \\ &= |c_0|^2 P_{|0\rangle} + |c_1|^2 P_{|1\rangle} \end{aligned} \quad (1.20)$$

c'est-à-dire qu'il ne reste qu'un état de mélange statistique. L'état en sortie a en quelque sorte perdu de sa pureté ($\neq$ projecteur).

La discussion des cas limites en considérant les vecteurs de Bloch devrait apporter une compréhension plus intuitive du phénomène. Un petit commentaire de ces résultats est apporté dans la partie consacrée à Bob.

Vecteur de Bloch

Il est possible de se représenter l'état du système d'Alice par un rayon dans la sphère de Poincaré en calculant le vecteur de Bloch $\vec{M}_{Alice}$. On a par définition

$$\vec{M} = Tr(\vec{\sigma} \rho) \quad (1.21)$$

Dans notre cas, $\vec{M}_{Alice}$ se détermine donc comme suit

$$\begin{aligned} \vec{M}_{Alice} &= Tr(\vec{\sigma} \rho_A) \\ &= Tr(\sigma_x \rho_A, \sigma_y \rho_A, \sigma_z \rho_A) \\ &= Tr\left(\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \rho_A, \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \rho_A, \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \rho_A\right) \\ &= ((c_1 c_0^* + c_0 c_1^*) \cos \gamma, i(c_1 c_0^* - c_0 c_1^*) \cos \gamma, |c_0|^2 - |c_1|^2) \\ &= (m_x \cos \gamma, i m_y \cos \gamma, m_z) \end{aligned} \quad (1.22)$$

où le vecteur

$$\vec{m} = (c_1 c_0^* + c_0 c_1^*, c_1 c_0^* - c_0 c_1^*, |c_0|^2 - |c_1|^2) \quad (1.23)$$

correspond à l'information du système total.

Le calcul du vecteur $\vec{m}$ est expliqué dans la remarque ci-dessous.

Remark 2 *Soit le qubit*

$$\psi = c_0 |0\rangle + c_1 |1\rangle \quad (1.24)$$

$$\begin{aligned} \text{où } c_0, c_1 &\in \mathbb{C} \\ |c_0|^2 + |c_1|^2 &= 1 \end{aligned}$$

Le vecteur $\vec{m}$ est calculé en projetant ψ dans la base des matrices de Pauli.

$$\vec{m} = \langle \vec{\sigma} \rangle_\psi = (\langle \psi | \sigma_x | \psi \rangle, \langle \psi | \sigma_y | \psi \rangle, \langle \psi | \sigma_z | \psi \rangle) \quad (1.25)$$

On explicite en détail le calcul de la composante m_x . Les deux autres se

déterminent de manière similaire.

$$\begin{aligned}
m_x &= \langle \psi | \sigma_x | \psi \rangle \\
&= \langle c_0 | 0 \rangle + c_1 | 1 \rangle \left(\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right) | c_0 | 0 \rangle + c_1 | 1 \rangle \\
&= [c_0^* (10) + c_1^* (01)] \left(\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right) \left[c_0 \begin{pmatrix} 1 \\ 0 \end{pmatrix} + c_1 \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right] \\
&= [c_0^* (10) + c_1^* (01)] \left[c_0 \begin{pmatrix} 0 \\ 1 \end{pmatrix} + c_1 \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right] \\
&= |c_0|^2 \cdot 0 + c_0^* c_1 + c_1^* c_0 + |c_1|^2 \cdot 0 \\
&= c_0^* c_1 + c_1^* c_0
\end{aligned} \tag{1.26}$$

Finalement, on obtient

$$\vec{m} = (c_0^* c_1 + c_1^* c_0, c_1 c_0^* - c_0 c_1^*, |c_0|^2 - |c_1|^2) \tag{1.27}$$

1.3.2 Etat du système de Bob

On garde la notation $c_{\pm}$ et $s_{\pm}$ déjà définie et on pose $\varphi_{c_0} = c_0 c_+ |0\rangle + c_0 s_+ |1\rangle$ et $\varphi_{c_1} = c_1 c_- |0\rangle + c_1 s_- |1\rangle$. Avec des calculs similaires à ceux effectués chez Alice, on trouve dans le système de Bob

$$\begin{aligned}
\rho_B &= Tr_A (P_{\Psi_{out}}) \\
&= Tr_B \left(\left(|0\rangle_A \otimes \varphi_{c_0} + |1\rangle_A \otimes \varphi_{c_1} \right) \left(|0\rangle_A \otimes \varphi_{c_0} + |1\rangle_A \otimes \varphi_{c_1} \right)^\dagger \right) \\
&= (c_0 c_+ |0\rangle + c_0 s_+ |1\rangle) (\dots)^\dagger + (c_1 c_- |0\rangle + c_1 s_- |1\rangle) (\dots)^\dagger \\
&= |c_0|^2 P_{\psi(\frac{\pi}{2}+\gamma)} + |c_1|^2 P_{\psi(\frac{\pi}{2}-\gamma)}
\end{aligned} \tag{1.28}$$

Les deux cas limites donnent chez Bob :

1. Si $\gamma = 0 \Rightarrow$

$$\begin{aligned}
\rho_B &= |c_0|^2 P_{\psi(\frac{\pi}{2})} + |c_1|^2 P_{\psi(\frac{\pi}{2})} \\
&= |c_0|^2 P_{\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)} + |c_1|^2 P_{\frac{1}{\sqrt{2}}(|0\rangle+|1\rangle)} \\
&= (|c_0|^2 + |c_1|^2) \left(\frac{1}{2} |0\rangle + |1\rangle \right) \left(\frac{1}{2} \langle 0| + \langle 1| \right) \\
&= \left(\frac{1}{2} |0\rangle \langle 0| + \frac{1}{2} |0\rangle \langle 1| + \frac{1}{2} |1\rangle \langle 0| + \frac{1}{2} |1\rangle \langle 1| \right) \\
&= \frac{1}{2} \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} \\
&= P_{|x\rangle}
\end{aligned} \tag{1.29}$$

c'est-à-dire qu'aucune information n'est disponible dans le système de Bob ($\Rightarrow U = \mathbb{1}$).

2. Si $\gamma = \frac{\pi}{2} \Rightarrow$

$$\begin{aligned}\rho_B &= |c_0|^2 P_{\psi(0)} + |c_1|^2 P_{\psi(\pi)} \\ &= |c_0|^2 P_{|0\rangle} + |c_1|^2 P_{|1\rangle} \\ &= \begin{pmatrix} |c_0|^2 & 0 \\ 0 & |c_1|^2 \end{pmatrix} \\ &= \rho_A\end{aligned}\tag{1.30}$$

Bob a fait une mesure idéale σ_z sur le système d'Alice.

Dans le cas général, où γ est quelconque, on dit que Bob a fait *une mesure d'intensité* γ sur Alice.

Vecteur de Bloch

On calcule le vecteur de Bloch pour Bob

$$\vec{M}_{Bob} = Tr(\vec{\sigma} \rho_B)\tag{1.31}$$

avec

$$\rho_B = |c_0|^2 P_{\psi(\frac{\pi}{2}+\gamma)} + |c_1|^2 P_{\psi(\frac{\pi}{2}-\gamma)}\tag{1.32}$$

où

$$\begin{aligned}P_{\psi(\frac{\pi}{2}\pm\gamma)} &= \left| \psi\left(\frac{\pi}{2}\pm\gamma\right) \right\rangle \left\langle \psi\left(\frac{\pi}{2}\pm\gamma\right) \right| \\ &= |c_{\pm}|^2 |0\rangle\langle 0| + c_{\pm}s_{\pm} |0\rangle\langle 1| + c_{\pm}s_{\pm} |1\rangle\langle 0| + s_{\pm}^2 |1\rangle\langle 1| \\ &= \begin{pmatrix} c_{\pm}^2 & c_{\pm}s_{\pm} \\ c_{\pm}s_{\pm} & s_{\pm}^2 \end{pmatrix}\end{aligned}\tag{1.33}$$

avec encore une fois $c_{\pm} = \cos\left(\frac{\pi}{4} \pm \frac{\gamma}{2}\right)$ et $s_{\pm} = \sin\left(\frac{\pi}{4} \pm \frac{\gamma}{2}\right)$. Ainsi

$$\begin{aligned}
\vec{M}_{Bob} &= Tr \left[\vec{\sigma} \left(|c_0|^2 P_{\psi(\frac{\pi}{2}+\gamma)} + |c_1|^2 P_{\psi(\frac{\pi}{2}-\gamma)} \right) \right] \\
&= |c_0|^2 Tr \left(\vec{\sigma} P_{\psi(\frac{\pi}{2}+\gamma)} \right) + |c_1|^2 Tr \left(\vec{\sigma} P_{\psi(\frac{\pi}{2}-\gamma)} \right) \\
&= |c_0|^2 Tr \left(\begin{pmatrix} c_+ s_+ & \dots \\ \dots & c_+ s_+ \end{pmatrix}, \begin{pmatrix} -ic_+ s_+ & \dots \\ \dots & ic_+ s_+ \end{pmatrix}, \begin{pmatrix} c_+^2 & \dots \\ \dots & -s_+^2 \end{pmatrix} \right) \\
&\quad + |c_1|^2 Tr(\dots) \\
&= |c_0|^2 (2c_+ s_+, 0, c_+^2 - s_+^2) + |c_1|^2 (2c_- s_-, 0, c_-^2 - s_-^2) \\
&= |c_0|^2 \left(\sin\left(\frac{\pi}{2} + \gamma\right), 0, \cos\left(\frac{\pi}{2} + \gamma\right) \right) \\
&\quad + |c_1|^2 \left(\sin\left(\frac{\pi}{2} - \gamma\right), 0, \cos\left(\frac{\pi}{2} - \gamma\right) \right) \\
&= |c_0|^2 (\cos\gamma, 0, \sin\gamma) + |c_1|^2 (\cos\gamma, 0, -\sin\gamma) \\
&= \left((|c_0|^2 + |c_1|^2) \cos\gamma, 0, (|c_0|^2 - |c_1|^2) \sin\gamma \right) \\
&= (\cos\gamma, 0, m_z \sin\gamma)
\end{aligned} \tag{1.34}$$

Finalement, on peut constater que Bob n'a la possibilité de lire de l'information que dans la direction $\hat{z}$, et plus cette information va chez lui plus l'information d'Alice est perturbée sur les directions $\hat{x}$ et $\hat{y}$ (c.f. équation 1.22). Les vecteurs $\vec{M}_{Alice}$ et $\vec{M}_{Bob}$ sont intriqués.

1.3.3 Résumé

On a essayé de faire *une mesure d'intensité* γ où $\gamma \in [0, \frac{\pi}{2}]$. On a construit un passage continu du cas extrême $\gamma = 0$, qui correspond à zéro perturbation et zéro information copiée, à l'autre cas extrême $\gamma = \frac{\pi}{2}$, qui correspond environ à une mesure idéale avec 100% d'information copiée sur σ_z (où l'information sur les directions $\hat{x}$ et $\hat{y}$ est perdue).

1.4 Universal Quantum Cloning Machine (UQCM)

1.4.1 Introduction

On trace très brièvement en guise de fils conducteur les grandes lignes de cette partie. Jusqu'à présent, on a vu avec le *no cloning theorem* une des propriétés fondamentales de l'information quantique, à savoir qu'un état quantique inconnu ne peut pas être cloné. Supposer qu'un état pur quantique ρ peut

être cloné conduit à une violation de l'unitarité de la mécanique quantique. Compte tenu de cette situation, on a tout de même essayé de construire un procédé qui va dans le sens d'une copie (*mesure d'intensité* γ). Le résultat obtenu est celui d'une copie imparfaite, intriquée avec l'original.

On voudrait maintenant connaître le clonage le moins imparfait que l'on peut espérer faire. Dans ce but, on commence par définir une *cloneuse quantique universelle* (UQCM). A partir d'un état initial $\psi_{initial} \in C^2$, cette "machine" fournit en sortie deux états intriqués de matrice densité ρ_{AB} dont la qualité est définie par un facteur de contraction η . Ainsi, connaître le clonage le moins imparfait possible revient alors à répondre à la question : quelle valeur maximale peut prendre le facteur η ? En se basant sur l'hypothèse de la théorie quantique selon laquelle il n'y a pas de "transmission instantanée de l'information" (appelée *signaling*) et sur celles attachées à la UQCM, on arrive à une expression pour ρ_{AB} qu'il suffit ensuite de diagonaliser pour tomber sur la condition qui détermine justement la valeur maximale que peut prendre η . Finalement, on va étudier une cloneuse qui permet d'atteindre cette valeur (cloneuse *Hillery-Bužek*).

1.4.2 Définition de la UQCM

Soit

$$P_{\psi_{initial}} = \frac{\mathbb{1} + \vec{m} \cdot \vec{\sigma}}{2} \rightarrow \boxed{UQCM} \rightarrow \rho_{AB} \Rightarrow \left\{ \begin{array}{l} \rho_A = \frac{\mathbb{1} + \eta \vec{m} \cdot \vec{\sigma}}{2} \\ \rho_B = \frac{\mathbb{1} + \eta \vec{m} \cdot \vec{\sigma}}{2} \end{array} \right\}$$

avec $\eta \in [0, 1]$

La UQCM affecte tout état initial $\psi_{initial} = |\vec{m}\rangle \in C^2$ de manière identique (d'où la qualification *Universal*), à savoir que la direction dans laquelle pointe l'input n'a pas de répercussion sur son clonage. En effet, on constate que le facteur η est un facteur de contraction indépendant de l'input. Il change la longueur du vecteur initial, mais conserve sa direction.

Ce facteur est lié au fait que ρ_A et ρ_B ne peuvent pas être deux copies d'une qualité parfaite (c.f *no cloning theorem*), ce qui signifie que ρ_{AB} n'est pas un état pur (i.e. un projecteur), mais un état de mélange.

On remarque que $\eta = 0$ donne deux probabilités de $\frac{1}{2}$ (par exple. un qubit dépolarisé) et $\eta = 1$ est impossible, comme déjà mentionné, car refusé par le *no cloning theorem*.

Base de la matrice densité

On a vu que pour un qubit $\in C^2$ la base d'un opérateur agissant sur ce qubit est donnée par $\{\mathbb{1}_2, \vec{\sigma}\}$. Donc, pour deux qubit $\in C^2 \otimes C^2$, la base d'un opérateur agissant sur ces qubits devient

$$\{\mathbb{1}, \vec{\sigma}\} \otimes \{\mathbb{1}, \vec{\sigma}\} \approx \{\mathbb{1} \otimes \mathbb{1}, \mathbb{1} \otimes \vec{\sigma}, \vec{\sigma} \otimes \mathbb{1}, \vec{\sigma} \otimes \vec{\sigma}\} \quad (1.35)$$

Ainsi, de manière générale,

$$\rho_{AB} = \frac{1}{4} \left(\mathbb{1} \otimes \mathbb{1} + \eta \vec{m} \cdot \vec{\sigma} \otimes \mathbb{1} + \eta \mathbb{1} \otimes \vec{m} \cdot \vec{\sigma} + \sum_{i,j=x,y,z} t_{ij} \sigma_i \otimes \sigma_j \right) \quad (1.36)$$

avec $t_{ij} \in C$
 $|\vec{m}\rangle \in C^2$ vecteur normé

De cette expression, on peut extraire

$$\begin{aligned} \rho_A &= Tr_B(\rho_{AB}) \\ &= \frac{1}{4} (2\mathbb{1} + 2\eta \vec{m} \cdot \vec{\sigma}) \\ &= \frac{1}{2} (\mathbb{1} + \eta \vec{m} \cdot \vec{\sigma}) \end{aligned} \quad (1.37)$$

et

$$\begin{aligned} \rho_B &= Tr_A(\rho_{AB}) \\ &= \frac{1}{2} (\mathbb{1} + \eta \vec{m} \cdot \vec{\sigma}) \end{aligned} \quad (1.38)$$

Remark 3 A noter que s'il n'y a pas d'intrication, i.e.

$$\rho_{AB} = \rho_A \otimes \rho_B \quad (1.39)$$

alors

$$(\eta \vec{m} \cdot \vec{\sigma}) \otimes (\eta \vec{m} \cdot \vec{\sigma}) = \sum_{i,j=x,y,z} t_{ij} \sigma_i \otimes \sigma_j \quad (1.40)$$

1.4.3 No-signaling

On considère la situation suivante

$$\boxed{\begin{smallmatrix} \vec{c} \cdot \vec{\sigma} \\ \vec{\sigma}^C \end{smallmatrix}} \leftarrow \text{Source} \rightarrow \boxed{\begin{smallmatrix} \vec{d} \cdot \vec{\sigma} \\ \vec{\sigma}^D \end{smallmatrix}} \xrightarrow{\text{probabilités}} \left\{ \begin{array}{c} P_D \\ 1 - P_D \end{array} \right\}$$

Dans son système, D constate les probabilités P_D et $1 - P_D$ dans ses détecteurs. Sans perte de généralité, on suppose que D sait que C mesure soit dans la direction σ_x , soit σ_z .

En raisonnant par l'absurde, on suppose que la probabilité P_D est différente selon que C choisit σ_x ou $\sigma_z \Rightarrow$ Alors on aurait "une transmission instantanée de l'information", appelée *signaling* (car supposons que $P_D(\sigma_x^C) = 0$, $P_D(\sigma_z^C) = 1$).

La théorie quantique prédit le *no-signaling*. En effet, elle est protégée du *signaling* par l'argument suivant

- Soit $|\psi\rangle \in H_C \otimes H_D$ ($\dim H_C$ et $\dim H_D$ arbitraire) et soit

$$|\psi'\rangle = (U_C(\alpha) \otimes \mathbb{1}) |\psi\rangle \quad (1.41)$$

i.e. C effectue une rotation de sa mesure et D ne fait rien

$\Rightarrow$ On veut que

$$\rho_D = \text{Tr}_C(|\psi\rangle\langle\psi|) = \text{Tr}_C(|\psi'\rangle\langle\psi'|) \quad (1.42)$$

- Preuve de l'argument (1.42) : *La trace est invariante sur l'action d'une opération unitaire.*

Ainsi, comme la trace est la même quelle que soit la base, quoique C fasse de son côté, D ne s'aperçoit de rien du sien.

1.4.4 La PQCM permettrait le *signaling*

On considère la situation suivante

$$\left\{ \begin{array}{c} P_C \\ 1 - P_C \end{array} \right\} \xleftarrow{\text{probabilités}} \begin{array}{c} \vec{c} \cdot \vec{\sigma} \\ \boxed{\vec{\sigma}^C} \\ C \end{array} \leftarrow \text{Source} \rightarrow \begin{array}{c} P=\text{P erfect} \\ \boxed{PQCM} \\ D \end{array} \left\{ \begin{array}{c} \boxed{\vec{\sigma}^A} \xrightarrow{\text{probabilités}} \dots \\ A \\ \boxed{\vec{\sigma}^B} \xrightarrow{\text{probabilités}} \dots \\ B \end{array} \right\}$$

Par exemple, soit à la source l'état singlet (invariant de rotation), alors

$$\begin{aligned} \frac{1}{\sqrt{2}} (|\vec{n}\rangle |-\vec{n}\rangle - |-\vec{n}\rangle |\vec{n}\rangle) &= \frac{1}{\sqrt{2}} (|z\rangle |-z\rangle - |-z\rangle |z\rangle) \\ &= \frac{1}{\sqrt{2}} (|x\rangle |-x\rangle - |-x\rangle |x\rangle) \end{aligned} \quad (1.43)$$

Si on suppose que C mesure dans la direction σ_x , alors

$$\begin{array}{llll} C : & \sigma_x & D & \\ P_C = \frac{1}{2} & | +x \rangle \rightarrow & | -x \rangle \xrightarrow{P_{QCM}} & | -x \rangle_A | -x \rangle_B \\ 1 - P_C = \frac{1}{2} & | -x \rangle \rightarrow & | +x \rangle \xrightarrow{P_{QCM}} & | +x \rangle_A | +x \rangle_B \end{array} \quad (1.44)$$

Ainsi

$$\begin{aligned} \rho_D^x &= \rho_{AB} \\ &= \frac{1}{2} (| -x \rangle_A | -x \rangle_B \langle -x |_A \langle -x |_B) \\ &\quad + \frac{1}{2} (| +x \rangle_A | +x \rangle_B \langle +x |_A \langle +x |_B) \end{aligned} \quad (1.45)$$

où l'indice x se réfère à la mesure faite par C . Ce qui donne (c.f. remarque ci-dessous)

$$\rho_{AB} = \frac{1}{4} (\mathbb{1}_4 + \sigma_x \otimes \sigma_x) \quad (1.46)$$

De façon similaire, si on suppose que C mesure dans la direction σ_z , alors

$$\begin{aligned} \rho_D^z &= \rho_{AB} \\ &= \frac{1}{4} (\mathbb{1}_4 + \sigma_z \otimes \sigma_z) \end{aligned} \quad (1.47)$$

Finalement, on remarque que

$$\rho_D^x \neq \rho_D^z \quad (1.48)$$

ce qui est en contradiction avec l'argument (1.42) qui montre que la trace partielle est indépendante du système où elle s'applique.

Remark 4 Pour arriver à (1.46), on pose

$$|0\rangle \equiv | +z \rangle \text{ et } |1\rangle \equiv | -z \rangle \quad (1.49)$$

états propres de σ_z .

De plus, on rappelle que

$$| \pm x \rangle = \frac{1}{\sqrt{2}} (|0\rangle \pm |1\rangle) \quad (1.50)$$

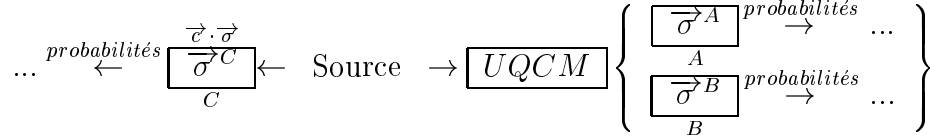
et on signale que

$$|0\rangle \langle 0| = \frac{1}{2} (\mathbb{1} + \sigma_z), \quad |1\rangle \langle 1| = \frac{1}{2} (\mathbb{1} - \sigma_z) \quad (1.51)$$

$$|0\rangle \langle 1| = \frac{1}{2} (\sigma_x + i\sigma_y), \quad |1\rangle \langle 0| = \frac{1}{2} (\sigma_x - i\sigma_y) \quad (1.52)$$

1.4.5 UQCM & *No-signaling*

On considère la situation suivante



On impose la condition de *no-signaling*, à savoir

$$\rho_D^x = \rho_D^z \quad (1.53)$$

où les indices x et z se réfèrent à la mesure faite par C .

On utilise la notation

$$\rho(|\vec{m}\rangle) \quad (1.54)$$

qui signifie qu'on rentre avec $|\vec{m}\rangle$ et qu'on regarde le résultat.

Donc, avec cette notation

$$\rho_{AB}^x = \frac{1}{2}\rho_{out}(|+x\rangle) + \frac{1}{2}\rho_{out}(|-x\rangle) \quad (1.55)$$

et

$$\rho_{AB}^z = \frac{1}{2}\rho_{out}(|+z\rangle) + \frac{1}{2}\rho_{out}(|-z\rangle) \quad (1.56)$$

On utilise les hypothèses attachées à la cloneuse :

- Elle est symétrique

$$\begin{aligned} \rho_A &= \rho_B \\ &= \frac{1}{2}(\mathbb{1} + \eta \vec{m} \cdot \vec{\sigma}) \end{aligned} \quad (1.57)$$

$\Rightarrow$ la forme la plus générale pour ρ_{AB} est

$$\begin{aligned} \rho_{AB}(|\vec{m}\rangle) &= \frac{1}{4}(\mathbb{1}_4 + \eta((\vec{m} \cdot \vec{\sigma}) \otimes \mathbb{1} + \mathbb{1} \otimes (\vec{m} \cdot \vec{\sigma}))) \\ &\quad + \frac{1}{4} \sum_{i,j=x,y,z} t_{ij} \sigma_i \otimes \sigma_j \end{aligned} \quad (1.58)$$

avec $t_{ij} \in \mathbb{C}$.

- Elle est universelle

Remark 5 On rappelle que $\forall U \in SU(2) \exists$ une rotation $R \in SU(3)$ t.q.

$$U \vec{m} \cdot \vec{\sigma} U^\dagger = (R \vec{m}) \cdot \vec{\sigma} \quad (1.59)$$

c'est-à-dire qu'à toute opération unitaire, on peut associer une rotation de $\vec{m}$.

donc

$$\begin{aligned} \rho(U |\vec{m}\rangle) &= \rho(|R\vec{m}\rangle) \\ &= U \otimes U \rho_{out}(|\vec{m}\rangle) U^\dagger \otimes U^\dagger \end{aligned} \quad (1.60)$$

ce qui revient à dire que faire une rotation du vecteur $\vec{m}$ puis une réduction de η équivaut à d'abord faire une réduction de η puis une rotation de $\vec{m}$.

Par conséquent, si on suppose que

$$R \vec{m} = \vec{m} \quad (1.61)$$

à savoir que R est une rotation d'axe $\vec{m}$, alors

$$\begin{aligned} \rho_{out}(|\vec{m}\rangle) &= U \otimes U \rho_{out}(|\vec{m}\rangle) U^\dagger \otimes U^\dagger \\ &= \rho_{AB}(|\vec{m}\rangle) \end{aligned} \quad (1.62)$$

ce qui signifie que si, par exemple, $\vec{m} = \hat{z}$, alors on peut choisir $\hat{x}$ et $\hat{y}$ arbitrairement dans le plan perpendiculaire à $\hat{z}$. Dans l'expression (1.58), ce résultat a pour conséquence de faire tomber les coefficients t_{ij} pour lesquels soit i , soit j , vaut z .

D'où, pour $\vec{m} = \hat{z}$ (pas immédiat)

$$\begin{aligned} \rho_{AB}(|+z\rangle) &= \frac{1}{4} (\mathbb{1}_4 + \eta (\sigma_z \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_z) + t_{zz} \sigma_z \otimes \sigma_z) \\ &\quad + \frac{1}{4} (t_\perp (\sigma_x \otimes \sigma_x + \sigma_y \otimes \sigma_y)) \\ &\quad + \frac{1}{4} (t_{xy} (\sigma_x \otimes \sigma_y - \sigma_y \otimes \sigma_x)) \end{aligned} \quad (1.63)$$

avec $t_\perp = t_{xx} = t_{yy}$.

Remark 6 On peut faire le test $x' = \hat{x} \cos \varphi + \hat{y} \sin \varphi$ et $x' = -\hat{x} \sin \varphi + \hat{y} \cos \varphi$ et constater que ρ_{AB} reste la même.

Avec ce qui précède,

$$\begin{aligned}\rho_{out}(|-z\rangle) &= \sigma_x \otimes \sigma_x \rho_{out}(|+z\rangle) \sigma_x \otimes \sigma_x \\ &= \rho_{AB}(|-z\rangle)\end{aligned}$$

(rappel : $\sigma_x^\dagger = \sigma_x$ et $\sigma_x |+z\rangle = |-z\rangle$).

On a maintenant tous les ingrédients pour écrire la *no-signaling condition* (1.53)

$$\frac{1}{2}\rho_{AB}(|+z\rangle) + \frac{1}{2}\rho_{AB}(|-z\rangle) = \frac{1}{2}\rho_{AB}(|+x\rangle) + \frac{1}{2}\rho_{AB}(|-x\rangle) \quad (1.64)$$

où $\rho_{AB}(|\pm x\rangle)$ s'obtient de $\rho_{AB}(|\pm z\rangle)$ en changeant $z \rightarrow x; x \rightarrow y; y \rightarrow z$.

Il suffit d'égaliser les différentes parties des matrices $\rho_{AB}(\dots)$ de (1.64) pour tirer la condition recherchée :

- $\eta\sigma_z \otimes \mathbb{1} + \eta(-\sigma_z) \otimes \mathbb{1} = \eta\sigma_x \otimes \mathbb{1} + \eta(-\sigma_x) \otimes \mathbb{1} = 0$
- $t_{zz}\sigma_z \otimes \sigma_z + t_{zz}(-\sigma_z) \otimes (-\sigma_z) = t_{xx}\sigma_x \otimes \sigma_x + t_{xx}(-\sigma_x) \otimes (-\sigma_x) \Rightarrow t_{zz} = t_{xx}$
- ...

On trouve ainsi que

$$t_\perp = t_{zz} = t \quad (1.65)$$

Finalement, UQCM et no-signaling donnent

$$\begin{aligned}\rho_{AB}(|+z\rangle) &= \frac{1}{4}(\mathbb{1}_4 + \eta(\sigma_z \otimes \mathbb{1} + \mathbb{1} \otimes \sigma_z) + t\vec{\sigma} \otimes \vec{\sigma}) \\ &\quad + \frac{1}{4}(t_{xy}(\sigma_x \otimes \sigma_y - \sigma_y \otimes \sigma_x))\end{aligned} \quad (1.66)$$

1.4.6 Détermination de $\hat{e}ta$

On commence par écrire $\rho_{AB}(|+z\rangle)$ en explicitant les matrices

$$\begin{aligned}
 4\rho_{AB}(|+z\rangle) = & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} + \eta \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix} + t \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \\
 & + t \begin{pmatrix} 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 \end{pmatrix} + t \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\
 & + t_{xy} \left[\begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & i & 0 \\ 0 & -i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 0 & 0 & -i \\ 0 & 0 & -i & 0 \\ 0 & i & 0 & 0 \\ i & 0 & 0 & 0 \end{pmatrix} \right]
 \end{aligned} \tag{1.67}$$

Ainsi

$$\rho_{AB}(|+z\rangle) = \frac{1}{4} \begin{pmatrix} 1 + 2\eta + t & 0 & 0 & 0 \\ 0 & 1 - t & 2(t + it_{xy}) & 0 \\ 0 & 2(t - it_{xy}) & 1 - t & 0 \\ 0 & 0 & 0 & 1 - 2\eta + t \end{pmatrix} \tag{1.68}$$

Comme $\rho_{AB}(|+z\rangle)$ doit avoir les propriétés d'un état, elle doit être normalisée, self-adjointe, et non négative (rappel: valeurs propres non négatives $\Rightarrow$ probabilité ≥ 0). Les deux premières conditions sont vérifiées et la dernière va nous donner la condition recherchée sur η .

On calcule donc les valeur propre de (1.68). On remarque que $1 + 2\eta + t$ et $1 - 2\eta + t$ correspondent déjà à deux des valeurs propres. Les deux autres s'obtiennent en diagonalisant la matrice

$$\begin{pmatrix} 1 - t & 2(t + it_{xy}) \\ 2(t - it_{xy}) & 1 - t \end{pmatrix} \tag{1.69}$$

On arrive alors aux quatre conditions suivantes

$$1 \pm 2\eta + t \geq 0 \tag{1.70}$$

et

$$1 - t \pm 2\sqrt{t^2 + t_{xy}^2} \geq 0 \quad (1.71)$$

qui doivent être satisfaites par η dont on veut trouver la valeur maximale.

Pour maximiser η , on voit par (1.70) qu'il faut choisir que

$$\eta \leq \frac{1}{2}(1 + t) \quad (1.72)$$

où t est maximum.

Pour estimer la borne supérieure de t , on doit poser dans (1.71) que

$$t_{xy} = 0 \quad (1.73)$$

Ces deux conditions se ramènent alors à

$$1 - 3t \geq 0 \text{ et } 1 + t \geq 0 \Rightarrow t \leq \frac{1}{3} \quad (1.74)$$

La condition qui détermine les valeurs que peut prendre η , est alors obtenue facilement en posant $t = \frac{1}{3}$ dans (1.72). Finalement, on trouve que

$$\eta \leq \frac{2}{3} \quad (1.75)$$

On peut se poser la question suivante : Existe-il une cloneuse qui fournit en sortie des copies avec un facteur de contraction maximum, à savoir avec $\eta_{max} = \frac{2}{3}$? Réponse : oui, la cloneuse *Hillery-Bužek*.

1.4.7 Cloneuse *Hillery-Bužek*

Soit la transformation

$$U : C_A^2 \otimes C_B^2 \otimes C_M^2 \rightarrow C_A^2 \otimes C_B^2 \otimes C_M^2 \quad (1.76)$$

définie comme

$$\begin{aligned} U : |000\rangle_{ABM} &\rightarrow \sqrt{\frac{2}{3}}|000\rangle_{ABM} + \sqrt{\frac{1}{6}}(|01\rangle_{AB} + |10\rangle_{AB}) \otimes |1\rangle_M \\ |100\rangle_{ABM} &\rightarrow \sqrt{\frac{2}{3}}|111\rangle_{ABM} + \sqrt{\frac{1}{6}}(|01\rangle_{AB} + |10\rangle_{AB}) \otimes |0\rangle_M \end{aligned} \quad (1.77)$$

qui décrit la cloneuse *Hillery-Bužek* permettant d'atteindre la qualité maximale d'une copie en mécanique quantique. Les espaces C_A^2 , C_B^2 et C_M^2

représentent respectivement les espaces de l'input, du registre (la copie) et de la machine.

On affirme que

$$U(\psi_A \otimes |00\rangle_{BM}) = \sqrt{\frac{2}{3}}\psi_A \otimes \psi_B \otimes \psi_M^* + \sqrt{\frac{1}{6}}(\psi_A \otimes \psi_B^\perp + \psi_A^\perp \otimes \psi_B) \otimes (\psi_M^\perp)^* \quad (1.78)$$

où

$$\begin{aligned} \psi &= c_0 |0\rangle + c_1 |1\rangle \text{ normé, i.e } |c_0|^2 + |c_1|^2 = 1 \\ \psi^* &= c_0^* |0\rangle + c_1^* |1\rangle \\ \psi^\perp &= c_1^* |0\rangle - c_0^* |1\rangle \Rightarrow \langle \psi | \psi^\perp \rangle = c_0^* c_1^* - c_1^* (-c_0^*) = 0 \\ (\psi^\perp)^* &= c_1 |0\rangle - c_0 |1\rangle = (\psi^*)^\perp \end{aligned} \quad (1.79)$$

Afin de vérifier que cette affirmation est correcte, on contrôle la linéarité de la transformation U , c'est-à-dire que U doit satisfaire

$$U[(\alpha\psi_1 + \beta\psi_2) \otimes |00\rangle] = \alpha U\psi_1 \otimes |00\rangle + \beta U\psi_2 \otimes |00\rangle \quad (1.80)$$

Pour cela, on doit trouver que chacune des composantes de la base vérifie l'égalité de l'affirmation.

- Coefficient $|000\rangle$:

L'égalité suivante doit être vérifiée

$$\begin{aligned} \langle 000 | U(c_0 |000\rangle + c_1 |100\rangle) &= \langle 000 | \left(\sqrt{\frac{2}{3}}\psi \otimes \psi \otimes \psi^* \right) \\ &\quad + \langle 000 | \left(\sqrt{\frac{1}{6}}(\psi \otimes \psi^\perp + \psi^\perp \otimes \psi) \otimes (\psi^\perp)^* \right) \end{aligned} \quad (1.81)$$

En se référant à la définition (1.77), la partie gauche de (1.81) donne

$$\begin{aligned} \langle 000 | U(c_0 |000\rangle + c_1 |100\rangle) &= \langle 000 | c_0 \left(\sqrt{\frac{2}{3}} |000\rangle + \sqrt{\frac{1}{6}}(|011\rangle + |101\rangle) \right) \\ &\quad + \langle 000 | c_1 \left(\sqrt{\frac{2}{3}} |111\rangle + \sqrt{\frac{1}{6}}(|010\rangle + |100\rangle) \right) \\ &= c_0 \sqrt{\frac{2}{3}} + c_0 \cdot 0 + c_1 \cdot 0 + c_1 \cdot 0 \\ &= c_0 \sqrt{\frac{2}{3}} \end{aligned} \quad (1.82)$$

En se référant à la définition (1.77) et à (1.79), on n'explicite dans le calcul de la partie de droite de l'égalité (1.81) que les coefficients $|000\rangle$ des états produit tensoriel formés par les ψ , $\psi^\perp$ et $\psi^{\perp*}$ puisqu'ils sont les seuls à ne pas s'annuler avec le vecteur *bra* $\langle 000|$.

Pour des raisons de lisibilité, on pose

$$I = \langle 000 | \left(\sqrt{\frac{2}{3}} |\psi\psi\psi^*\rangle + \sqrt{\frac{1}{6}} (|\psi\psi^\perp\psi^{\perp*}\rangle + |\psi^\perp\psi\psi^{\perp*}\rangle) \right) \quad (1.83)$$

Ainsi

$$\begin{aligned} I &= \langle 000 | \left[\sqrt{\frac{2}{3}} (c_0 |0\rangle \otimes c_0 |0\rangle \otimes c_0^* |0\rangle) \right] \\ &\quad + \langle 000 | \left[\sqrt{\frac{1}{6}} (c_0 |0\rangle \otimes c_1^* |0\rangle + c_1^* |0\rangle \otimes c_0 |0\rangle) \otimes c_1 |0\rangle \right] \\ &= \sqrt{\frac{2}{3}} c_0 |c_0|^2 + \sqrt{\frac{1}{6}} (c_0 c_1^* + c_1^* c_0) c_1 \\ &= \sqrt{\frac{2}{3}} c_0 |c_0|^2 + \sqrt{\frac{1}{6}} (2 |c_1|^2 c_0) \\ &= \sqrt{\frac{2}{3}} c_0 (|c_0|^2 + |c_1|^2) \\ &= \sqrt{\frac{2}{3}} c_0 \end{aligned} \quad (1.84)$$

L'égalité (1.81) est donc vérifiée.

- Coefficient $|001\rangle$:

Egalité à vérifier

$$\begin{aligned} \langle 001 | U (c_0 |000\rangle + c_1 |100\rangle) &= \langle 001 | \left(\sqrt{\frac{2}{3}} \psi \otimes \psi \otimes \psi^* \right) \\ &\quad + \langle 001 | \left(\sqrt{\frac{1}{6}} (\psi \otimes \psi^\perp + \psi^\perp \otimes \psi) \otimes (\psi^\perp)^* \right) \end{aligned} \quad (1.85)$$

Partie de gauche de (1.85)

$$\begin{aligned} \langle 001 | U (c_0 |000\rangle + c_1 |100\rangle) &= \langle 001 | c_0 \left(\sqrt{\frac{2}{3}} |000\rangle + \sqrt{\frac{1}{6}} (|011\rangle + |101\rangle) \right) \\ &\quad + \langle 001 | c_1 \left(\sqrt{\frac{2}{3}} |111\rangle + \sqrt{\frac{1}{6}} (|010\rangle + |100\rangle) \right) \\ &= 0 \end{aligned} \quad (1.86)$$

Partie de droite de (1.85). Pour des raisons de lisibilité, on pose

$$I' = \langle 001 | \left(\sqrt{\frac{2}{3}} |\psi\psi\psi^*\rangle + \sqrt{\frac{1}{6}} (|\psi\psi^\perp\psi^{\perp*}\rangle + |\psi^\perp\psi\psi^{\perp*}\rangle) \right) \quad (1.87)$$

Ainsi

$$\begin{aligned}
I' &= \langle 001 | \left[\sqrt{\frac{2}{3}} (c_0 |0\rangle \otimes c_0 |0\rangle \otimes c_1^* |1\rangle) \right] \\
&+ \langle 001 | \left[\sqrt{\frac{1}{6}} (c_0 |0\rangle \otimes c_1^* |0\rangle + c_1^* |0\rangle \otimes c_0 |0\rangle) \otimes (-c_0) |1\rangle \right] \\
&= \sqrt{\frac{2}{3}} |c_0|^2 c_1^* + \sqrt{\frac{1}{6}} (c_0 c_1^* + c_1^* c_0) (-c_0) \\
&= \sqrt{\frac{2}{3}} |c_0|^2 c_1^* - \sqrt{\frac{1}{6}} (2 |c_0|^2 c_1^*) \\
&= 0
\end{aligned} \tag{1.88}$$

L'égalité (1.85) est vérifiée.

- Idem pour les autres coefficients

On a donc montré que la transformation U est linéaire.

Pour finir, on va calculer la matrice densité partielle ρ_{AB} , dont on extrait ensuite ρ_A et ρ_B , pour constater que la cloneuse *Hillery-Bužek* décrite par U offre la meilleure qualité de copie possible en mécanique quantique.

Le calcul de ρ_{AB} s'effectue en faisant la trace de ρ_{ABM} sur l'espace de la machine. Comme dans cet espace on trouve les états orthogonaux $|\psi^*\rangle_M$ et $|\psi^{*\perp}\rangle_M$, la trace en question vaut (c.f exercice)

$$\begin{aligned}
\rho_{AB} &= Tr_M (U |\psi 00\rangle \langle \psi 00| U^\dagger) \\
&= \frac{2}{3} P_{\psi \otimes \psi} + \frac{1}{3} P_{\frac{1}{\sqrt{2}} (\psi \otimes \psi^\perp + \psi^\perp \otimes \psi)} \\
&= \frac{2}{3} P_{\psi \otimes \psi} + \frac{1}{6} P_{\psi \otimes \psi^\perp + \psi^\perp \otimes \psi}
\end{aligned} \tag{1.89}$$

Ainsi

$$\begin{aligned}
\rho_A &= Tr_B (\rho_{AB}) \\
&= \frac{2}{3} Tr_B (P_{\psi \otimes \psi}) + \frac{1}{3} Tr_B \left(P_{\frac{1}{\sqrt{2}} (\psi \otimes \psi^\perp + \psi^\perp \otimes \psi)} \right) \\
&= \frac{2}{3} P_\psi + \frac{1}{3} \frac{1}{2} (P_\psi + P_{\psi^\perp}) \\
&= \frac{2}{3} P_\psi + \frac{1}{6} \mathbb{1} \\
&= \frac{2}{3} \frac{\mathbb{1} + \vec{m} \cdot \vec{\sigma}}{2} + \frac{1}{6} \mathbb{1} \\
&= \frac{1}{2} \left(\mathbb{1} + \frac{2}{3} \vec{m} \cdot \vec{\sigma} \right) \\
&= \rho_B
\end{aligned} \tag{1.90}$$

où on utilise également le fait que dans l'espace de Bob les états $|\psi\rangle_B$ et $|\psi^\perp\rangle_B$ sont orthogonaux. Dans (1.90), on constate que le facteur de contraction vaut bien $\eta_{max} = \frac{2}{3}$.

On ne peut pas faire de meilleure copie en mécanique quantique que celle que l'on vient d'étudier sinon on est en contradiction avec le principe de superposition, le fait qu'il n'est pas possible de communiquer de l'information plus vite que la lumière et les inégalités de Heisenberg.

D'ailleurs, on va montrer qu'effectivement la cloneuse *Hillery-Bužek* satisfait ces inégalités.

Cloneuse *Hillery-Bužek* en accord avec les inégalités d'Heisenberg

On considère la situation suivante

$$\psi_{input} \rightarrow \boxed{U} \left\{ \begin{array}{l} \boxed{\sigma_z} \xrightarrow{\text{probabilités}} \left\{ \begin{array}{l} +1 \\ -1 \end{array} \right\} \\ \boxed{\sigma_x} \xrightarrow{\text{probabilités}} \left\{ \begin{array}{l} +1 \\ -1 \end{array} \right\} \end{array} \right\}$$

Dans ce cas, on voit qu'il est possible de mesurer en même temps deux grandeurs incompatibles, i.e. $[\sigma_x, \sigma_z] \neq 0$. Pour comprendre cela, il faut voir que cette mesure n'est pas idéale, mais partielle. En effet, elle n'est effectuée que sur une partie de l'information totale.

- Quelle est la probabilité de trouver $\pm z$?

$$\begin{aligned} \text{Prob}(\pm z) &= \text{Tr} \left(P_{|\pm z\rangle} \cdot \rho_A \right) \\ &= \text{Tr} \left[\frac{1}{2} (\mathbb{1} \pm \sigma_z) \cdot \frac{1}{2} \left(\mathbb{1} + \frac{2}{3} \vec{m} \cdot \vec{\sigma} \right) \right] \\ &= \frac{1}{4} \text{Tr} \left[(\mathbb{1} \pm \sigma_z) \cdot \left(\mathbb{1} + \frac{2}{3} \vec{m} \cdot \vec{\sigma} \right) \right] \\ &= \frac{1}{4} \text{Tr} \left[\mathbb{1} \mathbb{1} + \frac{2}{3} \mathbb{1} \vec{m} \cdot \vec{\sigma} \pm \sigma_z \mathbb{1} \pm \frac{2}{3} \sigma_z \vec{m} \cdot \vec{\sigma} \right] \\ &= \frac{1}{4} \text{Tr} \left(\mathbb{1} + \frac{2}{3} \vec{m} \cdot \vec{\sigma} \pm \sigma_z \pm \frac{2}{3} m_z \mathbb{1} \right) \\ &= \frac{1}{4} \left(2 + 0 \pm 0 \pm \frac{2}{3} m_z 2 \right) \\ &= \frac{1}{2} \left(1 \pm \frac{2}{3} m_z \right) \end{aligned} \tag{1.91}$$

D'où, par exemple,

$$\text{Si } m_z = +1 \Rightarrow \left\{ \begin{array}{l} \text{Prob}(+z) = \frac{5}{6} \text{ (la fidélité)} \\ \text{Prob}(-z) = \frac{1}{6} \end{array} \right\} \tag{1.92}$$

ainsi, même si $m_z = +1$ (i.e. $\psi_{input} = |1\rangle$), la probabilité de trouver le résultat -1 est non nulle, elle vaut $\frac{1}{6}$.

- Quelle est la probabilité de trouver $\pm x$?

$$\begin{aligned}
 \text{Prob}(\pm x) &= \text{Tr} \left(P_{|\pm x\rangle} \cdot \rho_B \right) \\
 &= \text{Tr} \left[\frac{1}{2} (\mathbb{1} \pm \sigma_x) \cdot \left(\frac{2}{3} P_\psi + \frac{1}{6} \mathbb{1} \right) \right] \\
 &= \frac{1}{2} \text{Tr} \left(\mathbb{1} \frac{2}{3} P_\psi \right) + \frac{1}{2} \text{Tr} \left(\mathbb{1} \frac{1}{6} \mathbb{1} \right) \\
 &\quad + \frac{1}{2} \text{Tr} \left(\pm \sigma_x \frac{2}{3} P_\psi \right) + \frac{1}{2} \text{Tr} \left(\pm \sigma_x \frac{1}{6} \mathbb{1} \right) \\
 &= \frac{1}{2} \left[\frac{2}{3} \langle \mathbb{1} \rangle_\psi + \frac{1}{3} + \frac{2}{3} \langle \pm \sigma_x \rangle_\psi + 0 \right] \\
 &= \frac{1}{2} \left(1 \pm \frac{2}{3} m_x \right)
 \end{aligned} \tag{1.93}$$

On fait cette expérience un grand nombre de fois et on regarde la moyenne des résultats mesurés en z que l'on note $\text{Mean}(z)$. On peut ensuite avec cette moyenne déterminer l'écart quadratique Δ_z^2 de la manière suivante

$$\begin{aligned}
 \Delta_z^2 &= \text{Mean}((\pm z)^2) - (\text{Mean}(\pm z))^2 \\
 &= (+1)^2 \text{Prob}(+z) + (-1)^2 \text{Prob}(-z) \\
 &\quad - ((+1) \text{Prob}(+z) + (-1) \text{Prob}(-z))^2 \\
 &= 1 - \left(\frac{2}{3} m_z \right)^2
 \end{aligned} \tag{1.94}$$

et par symétrie

$$\Delta_x^2 = 1 - \left(\frac{2}{3} m_x \right)^2 \tag{1.95}$$

Finalement, on arrive à l'inégalité

$$\begin{aligned}
 \Delta_z^2 \cdot \Delta_x^2 &= 1 - \frac{4}{9} (m_z^2 + m_x^2) + \frac{16}{81} m_z^2 m_x^2 \\
 &\geq 1 - \frac{4}{9} (1 - m_y^2) \\
 &= \frac{1}{9} (5 + 4m_y^2) \\
 &\geq m_y^2
 \end{aligned} \tag{1.96}$$

On rappelle que les inégalités de Heisenberg s'écrivent

$$\Delta\sigma_z \cdot \Delta\sigma_x \geq m_y \tag{1.97}$$

avec $\Delta\sigma_x = \langle \sigma_x^2 \rangle_\psi - \langle \sigma_x \rangle_\psi^2$
 $m_y = \frac{1}{2} \left| \langle [\sigma_z, \sigma_x] \rangle_\psi \right|$

Il faut noter que les inégalités (1.96) et (1.97) ont des significations différentes :

- (1.96) illustre qu'il n'existe pas de mesure conjointe de σ_x et σ_z t.q. la statistique des résultats ne soit pas en accord avec (1.97).
- (1.97) montre d'autre part qu'il n'existe pas d'état ψ dans l'espace de Hilbert t.q.

$$\Delta\sigma_z \cdot \Delta\sigma_x \geq \langle \sigma_y \rangle_\psi \quad (1.98)$$

En guise de conclusion, on fait la remarque suivante. Dans la théorie, une mesure est égale à un projecteur. En pratique, on arrive presque jamais à atteindre ce résultat. Ce qu'on vient de voir ici va dans ce sens.

Cloneuse *Hillery-Bužek* et résultats conjoints

Dans ce qui constitue la dernière partie de ce chapitre, on se penche encore sur la probabilité des résultats conjoints $Prob(+z \& +x)$ et ainsi on montre qu'il y a intrication entre ρ_A et ρ_B , c'est-à-dire que

$$\rho_{AB} \neq \rho_A \otimes \rho_B \quad (1.99)$$

La probabilité recherchée s'obtient en calculant

$$\begin{aligned} Prob(+z \& +x) &= Tr \left(P_{|+z\rangle} \otimes P_{|+x\rangle} \cdot \rho_{AB} \right) \\ &= Tr \left[P_{|+z\rangle} \otimes P_{|+x\rangle} \cdot \left(\frac{2}{3} P_{\psi \otimes \psi} + \frac{1}{3} P_{\frac{1}{\sqrt{2}}(\psi \otimes \psi^\perp + \psi^\perp \otimes \psi)} \right) \right] \\ &= Tr \left[\frac{1}{2} (\mathbb{1} + \sigma_z) \otimes \frac{1}{2} (\mathbb{1} + \sigma_x) \cdot \frac{2}{3} P_{\psi \otimes \psi} \right] \\ &\quad + Tr \left[\frac{1}{2} (\mathbb{1} + \sigma_z) \otimes \frac{1}{2} (\mathbb{1} + \sigma_x) \cdot \frac{1}{3} P_{\frac{1}{\sqrt{2}}(\psi \otimes \psi^\perp + \psi^\perp \otimes \psi)} \right] \\ &= \frac{2}{3} \left[\frac{1}{2} (1 + \langle \sigma_z \rangle_\psi) \frac{1}{2} (1 + \langle \sigma_x \rangle_\psi) \right] \\ &\quad + \frac{1}{3} \left\langle \frac{1}{4} (\mathbb{1}_4 + \mathbb{1}_2 \otimes \sigma_x + \sigma_z \otimes \mathbb{1}_2 + \sigma_z \otimes \sigma_x) \right\rangle_{\frac{1}{\sqrt{2}}(\psi \otimes \psi^\perp + \psi^\perp \otimes \psi)} \\ &= \frac{2}{3} \frac{1}{2} (1 + m_z) \frac{1}{2} (1 + m_x) \\ &\quad + \frac{1}{3} \left\langle \frac{1}{4} (\mathbb{1}_4 + \mathbb{1}_2 \otimes \sigma_x + \sigma_z \otimes \mathbb{1}_2 + \sigma_z \otimes \sigma_x) \right\rangle_{\frac{1}{\sqrt{2}}(\psi \otimes \psi^\perp + \psi^\perp \otimes \psi)} \end{aligned} \quad (1.100)$$

avec $\psi = c_0 |0\rangle + c_1 |1\rangle$ normé, i.e $|c_0|^2 + |c_1|^2 = 1$

$$\psi^\perp = c_1^* |0\rangle - c_0^* |1\rangle$$

$$\vec{m} = (c_0^* c_1 + c_0 c_1^*, c_0^* c_1 - c_0 c_1^*, |c_0|^2 - |c_1|^2) \quad (\text{c.f. 1.27})$$

On explicite séparément le calcul des termes compris dans $\langle \dots \rangle_{\frac{1}{\sqrt{2}}(\psi \otimes \psi^\perp + \psi^\perp \otimes \psi)}$ de (1.100) :

- Terme $\langle \sigma_z \otimes \mathbb{1}_2 \rangle_{\frac{1}{\sqrt{2}}(\psi \otimes \psi^\perp + \psi^\perp \otimes \psi)}$:

On pose pour des questions de lisibilité

$$I = \frac{1}{2} \langle \psi \psi^\perp + \psi^\perp \psi | \sigma_z \otimes \mathbb{1}_2 | \psi \psi^\perp + \psi^\perp \psi \rangle \quad (1.101)$$

Alors

$$\begin{aligned} I &= \frac{1}{2} \langle \psi^\perp \psi | \sigma_z \otimes \mathbb{1}_2 | \psi^\perp \psi \rangle + \frac{1}{2} \langle \psi^\perp \psi | \sigma_z \otimes \mathbb{1}_2 | \psi \psi^\perp \rangle \\ &\quad + \frac{1}{2} \langle \psi \psi^\perp | \sigma_z \otimes \mathbb{1}_2 | \psi^\perp \psi \rangle + \frac{1}{2} \langle \psi \psi^\perp | \sigma_z \otimes \mathbb{1}_2 | \psi \psi^\perp \rangle \\ &= -\frac{1}{2} m_z + 0 + 0 + \frac{1}{2} m_z \\ &= 0 \end{aligned} \quad (1.102)$$

(c.f. la remarque ci-après pour plus de détails de calculs)

- Terme $I' = \langle \sigma_z \otimes \sigma_x \rangle_{\frac{1}{\sqrt{2}}(\psi \otimes \psi^\perp + \psi^\perp \otimes \psi)}$:

$$\begin{aligned} I' &= \frac{1}{2} \langle \psi^\perp \psi | \sigma_z \otimes \sigma_x | \psi^\perp \psi \rangle + \frac{1}{2} \langle \psi^\perp \psi | \sigma_z \otimes \sigma_x | \psi \psi^\perp \rangle \\ &\quad + \frac{1}{2} \langle \psi \psi^\perp | \sigma_z \otimes \sigma_x | \psi^\perp \psi \rangle + \frac{1}{2} \langle \psi \psi^\perp | \sigma_z \otimes \sigma_x | \psi \psi^\perp \rangle \\ &= \frac{1}{2} (-m_z) m_x + \frac{1}{2} (2c_0 c_1) (c_1^* c_1^* - c_0^* c_0^*) \\ &= \frac{1}{2} (2c_0^* c_1^*) (c_1 c_1 - c_0 c_0) + \frac{1}{2} m_z (-m_x) \\ &= -m_z m_x + (|c_1|^2 c_0 c_1^* - |c_0|^2 c_0^* c_1) + (|c_1|^2 c_0^* c_1 - |c_0|^2 c_0 c_1^*) \\ &= -m_z m_x + c_0 c_1^* (|c_1|^2 - |c_0|^2) + c_0^* c_1 (|c_1|^2 - |c_0|^2) \\ &= -m_z m_x + (c_0 c_1^* + c_0^* c_1) (-m_z) \\ &= -2m_z m_x \end{aligned} \quad (1.103)$$

(c.f. la remarque ci-dessous)

Remark 7 Dans le calcul des expressions (1.102) et (1.103), on a utilisé pour σ_z les résultats suivants

$$\begin{aligned} \langle \psi | \sigma_z | \psi \rangle &= (c_0^*, c_1^*) \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} c_0 \\ c_1 \end{pmatrix} \\ &= |c_0|^2 - |c_1|^2 \\ &= m_z \end{aligned} \quad (1.104)$$

$$\langle \psi^\perp | \sigma_z | \psi^\perp \rangle = -m_z \quad (1.105)$$

$$\begin{aligned}\langle \psi | \sigma_z | \psi^\perp \rangle &= (c_0^*, c_1^*) \begin{pmatrix} c_1^* \\ c_0^* \end{pmatrix} \\ &= 2c_0^* c_1^*\end{aligned}\quad (1.106)$$

$$\begin{aligned}\langle \psi^\perp | \sigma_z | \psi \rangle &= (c_1, -c_0) \begin{pmatrix} c_0 \\ -c_1 \end{pmatrix} \\ &= 2c_0 c_1\end{aligned}\quad (1.107)$$

et pour σ_x

$$\begin{aligned}\langle \psi | \sigma_x | \psi \rangle &= (c_0^*, c_1^*) \begin{pmatrix} c_1 \\ c_0 \end{pmatrix} \\ &= m_x\end{aligned}\quad (1.108)$$

$$\langle \psi^\perp | \sigma_x | \psi^\perp \rangle = -m_x \quad (1.109)$$

$$\begin{aligned}\langle \psi | \sigma_x | \psi^\perp \rangle &= (c_0^*, c_1^*) \begin{pmatrix} -c_0^* \\ c_1^* \end{pmatrix} \\ &= c_1^* c_1^* - c_0^* c_0^*\end{aligned}\quad (1.110)$$

$$\begin{aligned}\langle \psi^\perp | \sigma_x | \psi \rangle &= (c_1, -c_0) \begin{pmatrix} c_1 \\ c_0 \end{pmatrix} \\ &= c_1 c_1 - c_0 c_0\end{aligned}\quad (1.111)$$

Finalement

$$\begin{aligned}\text{Prob}(+z \& +x) &= \frac{2}{3} \left[\frac{1}{2} (1 + m_z) \frac{1}{2} (1 + m_x) \right] + \frac{1}{12} (1 - 2m_x m_z) \\ &= \frac{1}{12} (3 + 2m_x + 2m_z) \\ &= \frac{1}{4} + \frac{1}{6} (m_x + m_z)\end{aligned}\quad (1.112)$$

Il reste à comparer le résultat (1.112) avec

$$\begin{aligned}\text{Prob}(+z) \text{Prob}(+x) &= \frac{1}{2} \left(1 + \frac{2}{3} m_z \right) \frac{1}{2} \left(1 + \frac{2}{3} m_x \right) \\ &= \frac{1}{4} + \frac{1}{6} (m_x + m_z) + \frac{1}{9} (m_x m_z) \\ &\neq \text{Prob}(+z \& +x)\end{aligned}\quad (1.113)$$

Donc les probabilités sur A et B ne sont pas indépendantes, il y a donc intrication, i.e. (1.99) est vérifiée.

Chapter 2

The controlled-NOT gate

2.1 Introduction

Le *controlled-NOT gate* ou *CNOT* (ou encore *XOR*) est la transformation à la base des circuits quantiques. Son utilité est en quelque sorte comparable à celle du transistor.

Definition 2 Soit l'opérateur unitaire U_χ représentant la transformation à deux qubit *CNOT*

$$U_\chi = |11\rangle\langle 10| + |10\rangle\langle 11| + |00\rangle\langle 00| + |01\rangle\langle 01| \quad (2.1)$$

Ainsi, on trouve que

$$U_\chi : \begin{array}{l} |00\rangle \rightarrow |00\rangle \\ |01\rangle \rightarrow |01\rangle \\ |10\rangle \rightarrow |11\rangle \\ |11\rangle \rightarrow |10\rangle \end{array} \quad (2.2)$$

ce qui revient à

$$U_\chi : |\sigma, \omega\rangle \rightarrow |\sigma, \sigma \oplus \omega\rangle \quad (2.3)$$

avec $\sigma, \omega \in \{1, 0\}$
 $\oplus$ l'additivité modulo 2

Dans la définition, le premier qubit est appelé le contrôle (source) et le second la cible (target). L'opération du *CNOT* consiste comme on le voit, à inverser la cible si le contrôle est dans l'état $|1\rangle$ et laisser le contrôle inchangé.

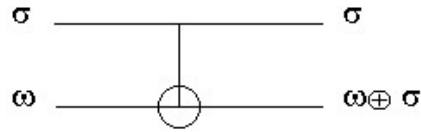
Un théorème montre l'importance de cette opération.

Theorem 3 Soit la transformation unitaire

$$U : (C^2)^{\otimes n} \rightarrow (C^2)^{\otimes n} \quad (2.4)$$

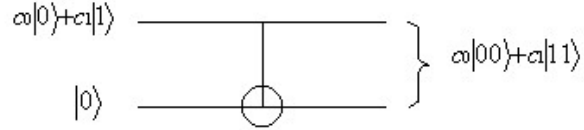
Alors toute transformation U peut être réalisée à l'aide de (beaucoup) transformations U_χ et de rotations à 1 qubit.

Cette opération est symbolisée par le schéma suivant



2.1.1 Exemples

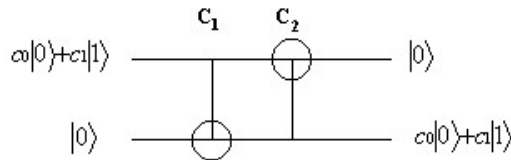
1. On place au contrôle le qubit élémentaire $\psi = c_0 |0\rangle + c_1 |1\rangle$ et à la source $|0\rangle$. Alors



$$(c_0 |0\rangle + c_1 |1\rangle) \otimes |0\rangle \xrightarrow{U_\chi} c_0 |00\rangle + c_1 |11\rangle \quad (2.5)$$

On part ainsi de l'état non intriqué en input $c_0 |00\rangle + c_1 |10\rangle$ et on trouve en sortie l'état $c_0 |00\rangle + c_1 |11\rangle$, qui lui, est intriqué .

2. Soit la situation où on note les transformations U_χ par C_i avec $i = 1, 2$.

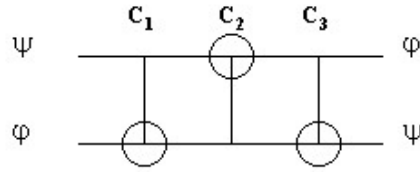


On fait le calcul

$$(c_0 |0\rangle + c_1 |1\rangle) \otimes |0\rangle \Rightarrow \left\{ \begin{array}{l} |00\rangle \xrightarrow{C_1} |00\rangle \xrightarrow{C_2} |00\rangle \\ |10\rangle \xrightarrow{C_1} |11\rangle \xrightarrow{C_2} |01\rangle \end{array} \right\} \Rightarrow |0\rangle \otimes (c_0 |0\rangle + c_1 |1\rangle) \quad (2.6)$$

On remarque qu'il y a simplement permutation entre le contrôle et la cible.

3. Soit le *swap* où $\psi = c_0 |0\rangle + c_1 |1\rangle$ et $\varphi = a_0 |0\rangle + a_1 |1\rangle$ sont deux états



quelconques. En entrée, on a

$$\psi \otimes \varphi = c_0 a_0 |00\rangle + c_0 a_1 |01\rangle + c_1 a_0 |10\rangle + c_1 a_1 |11\rangle \quad (2.7)$$

Pour chaque $|\sigma, \omega\rangle$ ($\sigma, \omega \in \{0, 1\}$) de (2.7), le passage dans le *swap* donne

$$\begin{aligned} |\sigma, \omega\rangle &\xrightarrow{C_1} |\sigma, \omega \oplus \sigma\rangle \\ &\xrightarrow{C_2} |\sigma \oplus (\omega \oplus \sigma), \omega \oplus \sigma\rangle \\ &\xrightarrow{C_3} |\sigma \oplus (\omega \oplus \sigma), \omega \oplus \sigma \oplus (\sigma \oplus (\omega \oplus \sigma))\rangle \\ &= |\omega, \sigma\rangle \end{aligned} \quad (2.8)$$

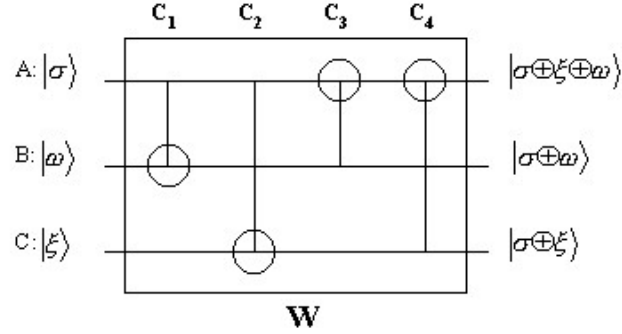
ainsi en sortie on trouve

$$a_0 c_0 |00\rangle + a_1 c_0 |10\rangle + a_0 c_1 |01\rangle + a_1 c_1 |11\rangle = \varphi \otimes \psi \quad (2.9)$$

A la suite de ces exemples simples de combinaisons de *CNOT*, on va en étudier de plus compliquées qui offrent des propriétés intéressantes.

2.2 Exemple de processeur quantique

Soit la transformation W avec $\sigma, \omega, \xi \in \{0, 1\}$.



Le calcul donne

$$\begin{aligned}
 |\sigma, \omega, \xi\rangle &\xrightarrow{C_1} |\sigma, \omega \oplus \sigma, \xi\rangle \\
 &\xrightarrow{C_2} |\sigma, \omega \oplus \sigma, \xi \oplus \sigma\rangle \\
 &\xrightarrow{C_3} |\sigma \oplus \omega \oplus \sigma, \omega \oplus \sigma, \xi \oplus \sigma\rangle = |\omega, \omega \oplus \sigma, \xi \oplus \sigma\rangle \\
 &\xrightarrow{C_4} |\omega \oplus \xi \oplus \sigma, \omega \oplus \sigma, \xi \oplus \sigma\rangle
 \end{aligned} \tag{2.10}$$

d'où on vérifié que

$$W |\sigma, \omega, \xi\rangle = |\omega \oplus \xi \oplus \sigma, \omega \oplus \sigma, \xi \oplus \sigma\rangle \tag{2.11}$$

Le but va être de montrer que l'on peut considérer le canal A du processeur W comme l'input et les canaux B et C comme le "programme". Ce faisant, on va voir qu'en sortie on obtient un output sur deux canaux et un canal que l'on peut interpréter comme du déchet.

2.2.1 Exemples d'input

1. Soit l'input

$$A : |\psi\rangle = c_0 |0\rangle + c_1 |1\rangle \tag{2.12}$$

et

$$BC : |\phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \tag{2.13}$$

où l'état $|\phi^+\rangle$ est un état de Bell.

On calcule

$$\begin{aligned}
 W \left(|\psi\rangle_A \otimes |\phi^+\rangle_{BC} \right) &= \frac{1}{\sqrt{2}} (c_0 W |000\rangle + c_0 W |011\rangle) \\
 &+ \frac{1}{\sqrt{2}} (c_1 W |100\rangle + c_1 W |111\rangle) \\
 &= \frac{1}{\sqrt{2}} (c_0 |000\rangle + c_0 |011\rangle + c_1 |111\rangle + c_1 |100\rangle) \\
 &= |\psi\rangle_A \otimes |\phi^+\rangle_{BC}
 \end{aligned} \tag{2.14}$$

ce qui signifie que l'état $|\psi\rangle_A \otimes |\phi^+\rangle_{BC}$ est un état propre de W . Ainsi, avec cet état comme input, W n'est rien d'autre que l'identité.

2. Soit l'input

$$A : |\psi\rangle = c_0 |0\rangle + c_1 |1\rangle \tag{2.15}$$

et

$$B : |0\rangle \tag{2.16}$$

$$C : |x\rangle = \frac{1}{\sqrt{2}} (|0\rangle + |1\rangle) \tag{2.17}$$

On calcule

$$\begin{aligned}
 W (|\psi, 0, x\rangle_{ABC}) &= \frac{1}{\sqrt{2}} (c_0 W |000\rangle + c_0 W |001\rangle) \\
 &+ \frac{1}{\sqrt{2}} (c_1 W |100\rangle + c_1 W |101\rangle) \\
 &= \frac{1}{\sqrt{2}} (c_0 |000\rangle + c_0 |101\rangle) \\
 &+ \frac{1}{\sqrt{2}} (c_1 |111\rangle + c_1 W |010\rangle) \\
 &= \frac{1}{\sqrt{2}} |0\rangle_A (c_0 |0\rangle_B + c_1 |1\rangle_B) |0\rangle_C \\
 &+ \frac{1}{\sqrt{2}} |1\rangle_A (c_0 |0\rangle_B + c_1 |1\rangle_B) |1\rangle_C \\
 &= |\psi\rangle_B \left[\frac{1}{\sqrt{2}} (|0\rangle_A |0\rangle_C + |1\rangle_A |1\rangle_C) \right] \\
 &= |\psi\rangle_B \otimes |\phi^+\rangle_{AC}
 \end{aligned} \tag{2.18}$$

Pour résumer, on a donc en considérant (2.14) et (2.18) que

$$|\psi\rangle_A \otimes |\phi^+\rangle_{BC} \xrightarrow{W} |\psi\rangle_A \otimes |\phi^+\rangle_{BC} \tag{2.19}$$

et

$$|\psi\rangle_A \otimes |0\rangle_B \otimes |x\rangle_C \xrightarrow{W} |\psi\rangle_B \otimes |\phi^+\rangle_{AC} \tag{2.20}$$

Alors, on obtient

$$|\psi\rangle_A (\alpha |\phi^+\rangle + \beta |0, x\rangle)_{BC} \xrightarrow{W} \alpha |\psi\rangle_A |\phi^+\rangle_{BC} + \beta |\psi\rangle_B |\phi^+\rangle_{AC} \tag{2.21}$$

avec $\alpha, \beta \in \mathbb{R}$ où l'output

$$\begin{aligned}\Psi_{out} &= \alpha |\psi\rangle_A |\phi^+\rangle_{BC} + \beta |\psi\rangle_B |\phi^+\rangle_{AC} \\ &= \frac{1}{\sqrt{2}} \alpha [c_0 (|000\rangle + |011\rangle) + c_1 (|100\rangle + |111\rangle)] \\ &\quad + \frac{1}{\sqrt{2}} \beta [c_0 (|000\rangle + |101\rangle) + c_1 (|010\rangle + |111\rangle)]\end{aligned}\tag{2.22}$$

est normalisé si l'input Ψ_{in} l'était étant donné que la transformation W est unitaire. On trouve la condition de normalisation de l'état

$$\begin{aligned}\Psi_{in} &= \alpha |\phi^+\rangle + \beta |0, x\rangle \\ &= \frac{1}{\sqrt{2}} \alpha (|00\rangle + |11\rangle) + \frac{1}{\sqrt{2}} \beta (|00\rangle + |01\rangle) \\ &= \frac{1}{\sqrt{2}} [(\alpha + \beta) |00\rangle + \alpha |11\rangle + \beta |01\rangle]\end{aligned}\tag{2.23}$$

en calculant le produit scalaire

$$\begin{aligned}\langle \Psi_{in} | \Psi_{in} \rangle &= \frac{1}{2} [(\alpha + \beta)^2 + \alpha^2 + \beta^2] \\ &= \alpha^2 + \beta^2 + \alpha\beta\end{aligned}\tag{2.24}$$

et en imposant que

$$\alpha^2 + \beta^2 + \alpha\beta = 1 \text{ où } \alpha, \beta \in \mathbb{R}\tag{2.25}$$

On va voir avec le calcul des matrices densités partielles ρ_A et ρ_B que l'on a “fabriqué” avec ce processeur quantique W une cloneuse (asymétrique), dont l'input est un état ψ quelconque et le “programme” une superposition de l'état de Bell $|\phi^+\rangle$ et de l'état $|0, x\rangle$.

2.2.2 Matrices densités partielles

Dans les espaces BC de l'état Ψ_{out} on trouve les vecteurs $|00\rangle_{BC} \perp |10\rangle_{BC} \perp |01\rangle_{BC} \perp |11\rangle_{BC}$. On obtient alors pour la matrice densité partielle ρ_A que

$$\begin{aligned}
\rho_A &= Tr_{BC}(P_{\Psi_{out}}) \\
&= \left[\left(\frac{1}{\sqrt{2}}\alpha c_0 + \frac{1}{\sqrt{2}}\beta c_0 \right) |0\rangle + \frac{1}{\sqrt{2}}\alpha c_1 |1\rangle \right] \cdot adjoint \\
&+ \left[\frac{1}{\sqrt{2}}\alpha c_0 |0\rangle + \left(\frac{1}{\sqrt{2}}\alpha c_1 + \frac{1}{\sqrt{2}}\beta c_1 \right) |1\rangle \right] \cdot adjoint \\
&+ \left[\frac{1}{\sqrt{2}}\beta c_0 |1\rangle \right] \left[\langle 1| \frac{1}{\sqrt{2}}\beta c_0^* \right] \\
&+ \left[\frac{1}{\sqrt{2}}\beta c_1 |0\rangle \right] \left[\langle 0| \frac{1}{\sqrt{2}}\beta c_1^* \right] \\
&= \left[\frac{1}{\sqrt{2}}\alpha\psi + \frac{1}{\sqrt{2}}\beta c_0 |0\rangle \right] \cdot adjoint \\
&+ \left[\frac{1}{\sqrt{2}}\alpha\psi + \frac{1}{\sqrt{2}}\beta c_1 |1\rangle \right] \cdot adjoint \\
&+ \frac{1}{2}\beta^2 |c_0|^2 P_{|1\rangle} + \frac{1}{2}\beta^2 |c_1|^2 P_{|0\rangle} \\
&= \frac{1}{2}\alpha^2 P_\psi + \left(\frac{1}{2}\alpha\beta c_0^* |\psi\rangle \langle 0| + \frac{1}{2}\alpha\beta c_0 |0\rangle \langle \psi| \right) + \frac{1}{2}\beta^2 |c_0|^2 P_{|0\rangle} \\
&+ \frac{1}{2}\alpha^2 P_\psi + \left(\frac{1}{2}\alpha\beta c_1^* |\psi\rangle \langle 1| + \frac{1}{2}\alpha\beta c_1 |1\rangle \langle \psi| \right) + \frac{1}{2}\beta^2 |c_1|^2 P_{|1\rangle} \\
&+ \frac{1}{2}\beta^2 |c_0|^2 P_{|1\rangle} + \frac{1}{2}\beta^2 |c_1|^2 P_{|0\rangle} \\
&= \alpha^2 P_\psi + \frac{1}{2}\beta^2 (P_{|0\rangle} + P_{|1\rangle}) \\
&+ \frac{1}{2}\alpha\beta (|\psi\rangle \langle c_0 | 0\rangle + c_1 |1\rangle + |c_0 | 0\rangle + c_1 |1\rangle) \langle \psi| \\
&= \alpha^2 P_\psi + \frac{1}{2}\beta^2 \mathbb{1} + \alpha\beta P_\psi
\end{aligned} \tag{2.26}$$

Finalement

$$\begin{aligned}
\rho_A &= (\alpha^2 + \alpha\beta) P_\psi + \frac{1}{2}\beta^2 \mathbb{1} \\
\rho_B &= (\beta^2 + \alpha\beta) P_\psi + \frac{1}{2}\alpha^2 \mathbb{1}
\end{aligned} \tag{2.27}$$

Avec la condition de normalisation (2.25), on peut analyser quelques cas intéressants :

- Si

$$\alpha = 1 \Rightarrow \beta = 0 \tag{2.28}$$

alors

$$\rho_A = P_\psi \text{ et } \rho_B = \frac{1}{2}\mathbb{1} \tag{2.29}$$

Dans ce cas le processeur ne fait rien, c'est-à-dire que l'input est égal à l'output.

(A noter que $\mathbb{1}_2$ représente le centre sur la sphère de Poincaré, ce qui correspond par exemple à un photon complètement dépolarisé.)

- Si

$$\alpha = 0 \Rightarrow \beta = 1 \quad (2.30)$$

alors

$$\rho_A = \frac{1}{2}\mathbb{1} \text{ et } \rho_B = P_\psi \quad (2.31)$$

Dans ce cas le processeur correspond à un opérateur de permutation.

- Si

$$\alpha = \beta \quad (2.32)$$

alors

$$\begin{aligned} \rho_A &= \rho_B \\ &= \frac{2}{3}P_\psi + \frac{1}{3}\frac{1}{2}\mathbb{1} \\ &= \frac{2}{3}\frac{\mathbb{1} + \vec{m} \cdot \vec{\sigma}}{2} + \frac{1}{3}\frac{1}{2}\mathbb{1} \\ &= \frac{1}{2}\left(\frac{2}{3}\mathbb{1} + \frac{1}{3}\mathbb{1} + \frac{2}{3}\vec{m} \cdot \vec{\sigma}\right) \\ &= \frac{1}{2}\left(\mathbb{1} + \frac{2}{3}\vec{m} \cdot \vec{\sigma}\right) \end{aligned} \quad (2.33)$$

Dans ce cas le processeur correspond à la *UQCM* optimale (c.f. cloneuse symétrique *Hillery-Bužek* équation (1.90)).

Remark 8 Ces résultats se retrouvent en raisonnant sur un concept appelé la fidélité (qualité), noté F , que l'on définit dans l'introduction de la partie suivante concentration d'intrication (éq. 2.37). En effet, par définition

$$\begin{aligned} F_A &= \langle \rho_A \rangle_\psi \\ &= \text{Tr}(\rho_A \cdot P_\psi) \\ &= (\alpha^2 + \alpha\beta) \text{Tr}(P_\psi P_\psi) + \frac{1}{2}\beta^2 \text{Tr}(\mathbb{1} P_\psi) \\ &= (\alpha^2 + \alpha\beta) \text{Tr}(P_\psi) + \frac{1}{2}\beta^2 \text{Tr}(P_\psi) \end{aligned} \quad (2.34)$$

ce qui donne

$$\begin{aligned} F_A &= (\alpha^2 + \alpha\beta) + \frac{1}{2}\beta^2 \\ F_B &= (\beta^2 + \alpha\beta) + \frac{1}{2}\alpha^2 \end{aligned} \quad (2.35)$$

Ainsi, par exemple, pour avoir une qualité égale chez Alice et chez Bob, on arrive (en omettant le cas trivial) à

$$F_A = F_B \Rightarrow \alpha = \beta \quad (2.36)$$

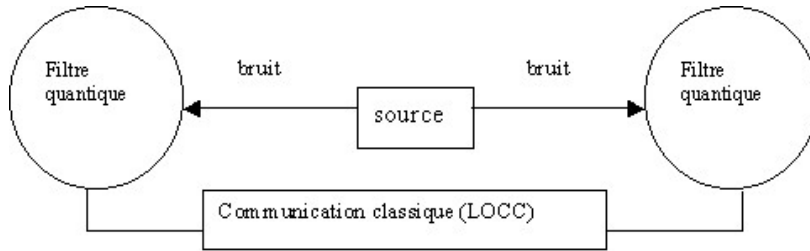
Pour conclure, on remarque que d'une manière générale les facteurs α et β déterminent la "programmation" du processeur.

2.3 Concentration d'intrication

2.3.1 Introduction

Les états intriqués sont essentiels dans les processus comme le traitement d'information quantique, la téléportation quantique, la cryptographie quantique... En général, ces applications ont un fonctionnement optimum avec des états quantiques purs, intriqués d'une façon maximale. Cependant en pratique, à cause des phénomènes de dissipation et de décohérence ("bruit"), les états disponibles n'ont pas l'intrication maximale recherchée, ou sont partiellement mélangés (i.e. pas purs), voire les deux à la fois. Afin de remédier à ce problème, plusieurs procédés de distillation d'intrication, de purification et concentration d'états existent. Ici, on va se cantonner à étudier le principe de la purification d'intrication découvert par C.H. Bennett *et al.* en 1995.

Pour nos calculs, on considère de petits systèmes, à savoir des systèmes à deux qubits. Soit $\rho_{source} \in C^2 \otimes C^2$.



Le but est donc de concevoir un "filtre quantique" qui va réduire l'effet du "bruit" de sorte qu'après son passage, on trouve une plus forte concentration (i.e. *fidélité* plus élevée) d'états intriqués qu'à son entrée. On va s'apercevoir que ce processus de purification engendre une perte d'au moins 50% des particules émises par la source.

Definition 4 On introduit la notion de *fidélité (qualité)* de la source

$$F_{source} = \langle \rho_{source} \rangle_{\psi^{(-)}} \quad (2.37)$$

où $\psi^{(-)}$ est l'état intriqué de façon maximale auquel on veut tendre à la sortie des filtres. Comme on le voit, la *fidélité* représente le chevauchement le plus important entre l'état en sortie et l'état à la source.

Soit F_{source} la fidélité de la source, F_{ligne} la fidélité de la source après avoir passé dans la ligne bruitée et $F_{concentré}$ la fidélité de la source en sortie des filtres. Ainsi,

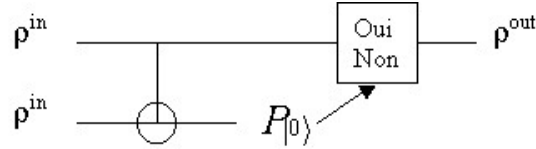
$$F_{source} \geq F_{ligne} \leq F_{concentré} \approx 1 \quad (2.38)$$

où on peut montrer qu'il est possible de s'approcher arbitrairement proche de 1.

2.3.2 Concentration sans intrication

Avant d'en venir au principe de purification d'intrication proprement dit, on passe par une approche simplifiée du problème (i.e. sans intrication) pour se familiariser avec le circuit dont on va se servir.

Soit un état non intriqué dans C^2 avec la base $\{|0\rangle, |1\rangle\}$. Donc pour un état dans $C^2 \otimes C^2$, on a la base $\{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$. Soit $\rho^{in}, \rho^{out} \in C^2 \otimes C^2$. On voit qu'un test est effectué dans le fil du bas et si sa réponse est



positive, c'est-à-dire que l'état zéro est mesuré, on laisse passer l'état du fil du haut en fermant l'interrupteur *oui-non*. Dans le cas contraire, on ouvre cet interrupteur, bloquant ainsi le passage. Et on répète l'opération. Il y a donc au mieux un qubit qui ressort lorsque deux entrent, ce qui correspond à 50 % de perte mais une plus grande concentration d'un certain type d'état.

Le processus *concentration sans intrication* se formule

$$\mathbb{1}_2 \otimes P_{|0\rangle} U_\chi (\rho^{in} \otimes \rho^{in}) U_\chi^\dagger (\mathbb{1}_2 \otimes P_{|0\rangle})^\dagger = 1 \cdot \rho^{out} \otimes P_{|0\rangle} + 0 \cdot \rho^{out} \otimes P_{|1\rangle} \quad (2.39)$$

où le terme $0 \cdot \rho^{out} \otimes P_{|1\rangle}$ signifie que rien ne sort.

On va expliciter les matrices de (2.39) qui entrent en jeu. En entrée, on a

$$\rho^{in} \otimes \rho^{in} = \begin{pmatrix} \rho_{00}^{in} \rho^{in} & \rho_{01}^{in} \rho^{in} \\ \rho_{10}^{in} \rho^{in} & \rho_{11}^{in} \rho^{in} \end{pmatrix} \quad (2.40)$$

On pose

$$A = U_\chi (\rho^{in} \otimes \rho^{in}) U_\chi^\dagger \quad (2.41)$$

alors on trouve pour la matrice A

$$A = \begin{pmatrix} \rho_{00}^{in} \rho^{in} & \rho_{01}^{in} \rho^{in} \sigma_x \\ \rho_{10}^{in} \sigma_x \rho^{in} & \rho_{11}^{in} \sigma_x \rho^{in} \sigma_x \end{pmatrix} \quad (2.42)$$

Il faut noter que les opérateurs

$$U_\chi = \begin{pmatrix} \mathbb{1}_2 & 0 \\ 0 & \sigma_x \end{pmatrix} \quad (2.43)$$

et

$$\mathbb{1}_2 \otimes P_{|0\rangle} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad (2.44)$$

sont self-adjoints. On peut encore préciser que

$$\begin{aligned} \rho^{out} \otimes P_{|0\rangle} &= \begin{pmatrix} (\rho_{00}^{in})^2 & (\rho_{01}^{in})^2 \\ (\rho_{10}^{in})^2 & (\rho_{11}^{in})^2 \end{pmatrix} \otimes \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \\ &= \begin{pmatrix} (\rho_{00}^{in})^2 & 0 & (\rho_{01}^{in})^2 & 0 \\ 0 & 0 & 0 & 0 \\ (\rho_{10}^{in})^2 & 0 & (\rho_{11}^{in})^2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned} \quad (2.45)$$

Finalement, en comparant (2.40) et (2.45), on constate que

$$\begin{aligned} \rho_{00}^{out} &= (\rho_{00}^{in})^2 \\ \rho_{01}^{out} &= (\rho_{01}^{in})^2 \\ \rho_{10}^{out} &= (\rho_{10}^{in})^2 \\ \rho_{11}^{out} &= (\rho_{11}^{in})^2 \end{aligned} \quad (2.46)$$

Ainsi, la matrice ρ^{out} correspond à la matrice ρ^{in} dont chaque élément a été pris au carré.

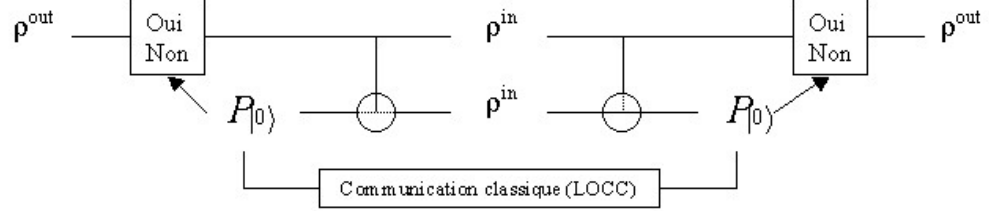
Avec ce qui précède, on peut calculer la probabilité de succès, à savoir la probabilité d'avoir un qubit en sortie :

$$Prob(succès) = Tr(\rho^{out}) = (\rho_{00}^{in})^2 + (\rho_{11}^{in})^2 \leq 1 \quad (2.47)$$

On voit que l'état de sortie ρ^{out} n'est en général pas normalisé. En effet, si ρ^{in} est normalisé (i.e. $Tr(\rho^{in}) = 1$), alors on trouve forcément (2.47). L'égalité est obtenue s'il y a un état pur en entrée, par exemple $\rho^{in} = P_{|0\rangle}$. Dans ce processus, de l'information a donc été perdue.

2.3.3 Principe de purification d'intrication

La situation à étudier se schématise de la façon suivante



Une paire d'intrication plus grande est créée en partant de deux paires moins intriquées dont un membre de chacune de ces paires a été envoyé l'un à Alice et l'autre à Bob. Aussi bien Alice que Bob font subir dans leur système une opération CNOT aux deux particules. Puis, ils effectuent une mesure sur les particules de la cible (fil du bas) et comparent les résultats mesurés via une communication classique (Local Operation Classical Communication). Si le résultat mesuré est zéro de part et d'autre, ils gardent la paire du contrôle (fil du haut) qui a alors un degré d'intrication plus élevé que les deux paires initiales. En revanche, s'ils obtiennent d'autres résultats, ils la laissent tomber. En répétant la même procédure, il est possible de distiller des paires d'une qualité d'intrication arbitrairement haute. A noter que tout cela fonctionne efficacement pourvu que la qualité de la paire initiale soit suffisamment élevée, c'est-à-dire qu'elle pointe déjà en partie dans la direction de l'état que l'on cherche à concentrer. (Extrait tiré de l'article *J-W. Pan, C. Simon, C. Bruckner & A. Zeilinger, Entanglement purification for quantum communication, Lett. to Nature, vol. 410, p. 1068 (2001).*)

Ce processus se formule

$$\mathbb{1}_4 \otimes P_{|00\rangle} U_\chi U_\chi (\rho^{in} \otimes \rho^{in}) U_\chi^\dagger U_\chi^\dagger (\mathbb{1}_4 \otimes P_{|00\rangle})^\dagger = \rho^{out} \otimes P_{|00\rangle} \quad (2.48)$$

Après calculs, on retrouve

$$\rho_{ij}^{out} = (\rho_{ij}^{in})^2 \quad (2.49)$$

avec $i, j \in \{|00\rangle, |01\rangle, |10\rangle, |11\rangle\}$.

Pour l'explication du principe de purification d'intrication, on suppose maintenant qu'Alice et Bob voudraient partager la paire spécifique de photons dans l'état d'intrication maximum

$$|\phi^+\rangle = \frac{1}{\sqrt{2}} (|00\rangle + |11\rangle) \quad (2.50)$$

Avant la purification, les paires de photons qu'ils se partagent se trouvent dans l'état de mélange

$$\rho^{in} = \lambda P_{\phi^+} + (1 - \lambda) \frac{1}{4} \mathbb{1}_4 \quad (2.51)$$

avec $\lambda \in [0, 1]$ où $(1 - \lambda) \frac{1}{4} \mathbb{1}_4$ représente le “bruit”. On précise que :

- Si $\lambda = 1 \Rightarrow \rho^{in}$ est un état pur intriqué de façon maximale.
- Si $\lambda = 0 \Rightarrow \rho^{in}$ est un état de mélange maximum.
- On peut montrer que ρ^{in} est intriqué $\Leftrightarrow \lambda > \frac{1}{3}$.

A noter, compte tenu de (2.49) et (2.51), que

$$\rho^{in} = \frac{1}{4} \begin{pmatrix} 1 + \lambda & 0 & 0 & 2\lambda \\ 0 & 1 - \lambda & 0 & 0 \\ 0 & 0 & 1 - \lambda & 0 \\ 2\lambda & 0 & 0 & 1 + \lambda \end{pmatrix} \quad (2.52)$$

et

$$\rho^{out} = \frac{1}{16} \begin{pmatrix} (1 + \lambda)^2 & 0 & 0 & (2\lambda)^2 \\ 0 & (1 - \lambda)^2 & 0 & 0 \\ 0 & 0 & (1 - \lambda)^2 & 0 \\ (2\lambda)^2 & 0 & 0 & (1 + \lambda)^2 \end{pmatrix} \quad (2.53)$$

où encore une fois ρ^{out} n'est pas normalisée. En effet, comme $\lambda \in [0, 1]$

$$\begin{aligned} Tr(\rho^{out}) &= \left(\frac{1}{4}\right)^2 [2(1 + \lambda)^2 + 2(1 - \lambda)^2] \\ &= \left(\frac{1}{4}\right)^2 (4 + 4\lambda^2) \\ &= \frac{1}{4} (1 + \lambda^2) \\ &< 1 \end{aligned} \quad (2.54)$$

On va comparer la fidélité en entrée F^{in} avec celle en sortie F^{out} , et constater que la purification de la source a bien eu lieu :

- Fidélité en entrée

$$\begin{aligned} F^{in} &= \langle \rho^{in} \rangle_{\phi^+} \\ &= \lambda + \frac{1}{4}(1 - \lambda) \\ &= \frac{1}{4}(1 + 3\lambda) \end{aligned} \quad (2.55)$$

- Fidélité en sortie (à normaliser)

$$\begin{aligned}
 F^{out} &= \frac{1}{Tr(\rho^{out})} \langle \rho^{out} \rangle_{\phi^+} \\
 &= \frac{1}{\frac{1}{4}(1+\lambda^2)} \left(\frac{1}{4}\right)^2 (1 + 2\lambda + 5\lambda^2) \\
 &= \frac{1}{(1+\lambda^2)} \frac{1}{4} (1 + 2\lambda + 5\lambda^2)
 \end{aligned} \tag{2.56}$$

Enfin, on peut voir que

$$\begin{aligned}
 F^{out} > F^{in} &\Leftrightarrow F^{in} > \frac{1}{2} \\
 &\Leftrightarrow \lambda^{in} > \frac{1}{3}
 \end{aligned} \tag{2.57}$$

2.4 Correction d'erreur

2.4.1 Introduction

La plupart des systèmes analogiques sont très sensibles aux petites perturbations. Dans les ordinateurs “classiques”, le système de correction d'erreur se trouve dans les processus numériques. Un exemple simple de correction classique d'erreur est un code répété : on remplace un bit par trois copies de ce bit,

$$\begin{aligned}
 0 &\rightarrow (000) \\
 1 &\rightarrow (111)
 \end{aligned}$$

Il se peut maintenant qu'une erreur se produise causant l'inversion d'un des trois bits (*bit flip*); par exemple, l'inversion du premier

$$\begin{aligned}
 (000) &\rightarrow (100) \\
 (111) &\rightarrow (011)
 \end{aligned}$$

Il s'en suit malgré cette erreur, qu'il est toujours possible de décoder correctement le bit par un vote majoritaire : deux zéros sur trois est lu comme un zéro.

En mécanique quantique, comment corriger les erreurs? On va étudier l'analogie quantique du *bit flip*, introduit par Peter Shor.

2.4.2 Procédé de correction d'erreur

La transmission d'information peut être faite dans le temps (mémoire) ou dans l'espace (d'un point A à un point B). On va considérer le cas d'une

transmission d'information dans l'espace au travers d'un canal quantique bruité. Le bruit correspond à introduire des erreurs dans l'input du canal. On choisit l'erreur particulière suivante

$$\psi = c_0 |0\rangle + c_1 |1\rangle \xrightarrow{\text{bruit}} \begin{cases} c_0 |0\rangle + c_1 |1\rangle & \text{avec probabilité } 1 - P \\ c_0 |1\rangle + c_1 |0\rangle & \text{avec probabilité } P \end{cases} \quad (2.58)$$

qui n'est en fait que l'application de la transformation $U = \sigma_x$ sur l'état ψ avec une probabilité P .

Remark 9 *Au lieu d'un bit flip complet, on peut imaginer une erreur plus faible. Soit une transformation unitaire*

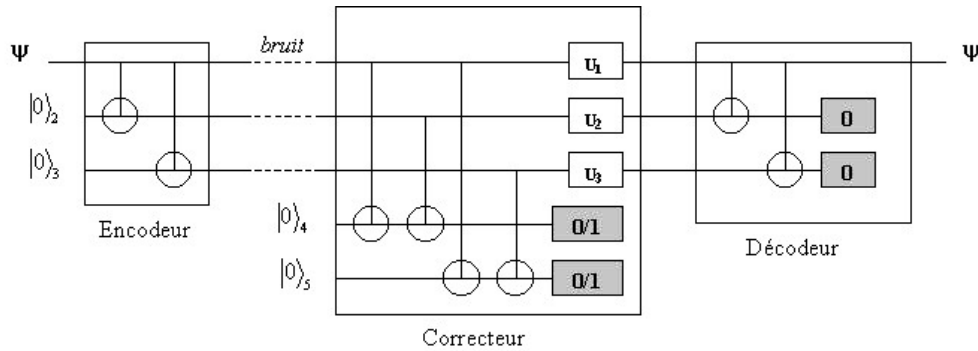
$$U_{\text{bruit}} : \begin{aligned} |0\rangle &\rightarrow \alpha |0\rangle + \beta |1\rangle \\ |1\rangle &\rightarrow \alpha |1\rangle - \beta |0\rangle \end{aligned} \quad (2.59)$$

avec $\alpha^2 + \beta^2 = 1$ et $\alpha, \beta \in \mathbb{R}$. Si le bruit est très faible $\Rightarrow \alpha^2 \gg \beta^2$.

On va montrer qu'à l'aide des *CNOT*, il est possible de construire une suite de procédés (encodeur, correcteur, décodeur) capable de corriger l'erreur particulière que l'on s'est donnée.

Ce circuit se compose de cinq canaux. Les résultats obtenus par les détecteurs des canaux 4 et 5 du correcteur déterminent les opérations unitaires U_i ($i = 1, 2, 3$) à effectuer. Les calculs montrent également que les détecteurs du décodeur donnent toujours le même résultat.

Soit le circuit



Soit l'état

$$\psi = c_0 |0\rangle + c_1 |1\rangle \quad (2.60)$$

Encodeur

A la sortie de l'encodeur, l'état est

$$\psi_1 \otimes |00\rangle_{23} = c_0 |000\rangle_{123} + c_1 |100\rangle_{123} \xrightarrow{\text{Encodeur}} c_0 |000\rangle_{123} + c_1 |111\rangle_{123} \quad (2.61)$$

où les indices 1, 2, 3 se réfèrent aux différents canaux.

Bruit

Après le passage dans le canal bruité où l'on a supposé qu'on a au plus une erreur, on peut trouver les états suivants

$$\begin{aligned} c_0 |000\rangle_{123} + c_1 |111\rangle_{123} &\xrightarrow{\text{bruit}} c_0 |100\rangle_{123} + c_1 |011\rangle_{123} \\ \text{ou} & c_0 |010\rangle_{123} + c_1 |101\rangle_{123} \\ \text{ou} & c_0 |001\rangle_{123} + c_1 |110\rangle_{123} \\ \text{ou} & c_0 |000\rangle_{123} + c_1 |111\rangle_{123} \end{aligned} \quad (2.62)$$

Correcteur & décodeur

On va prendre les différents cas de (2.62) et suivre leur passage au travers du correcteur et du décodeur

1. On commence par le cas sans erreur et on vérifie ainsi que le procédé ne rajoute pas d'erreur

$$c_0 |00000\rangle_{12345} + c_1 |11100\rangle_{12345} \xrightarrow{\text{Correcteur}} (c_0 |00000\rangle + c_1 |11100\rangle)_{12345} \quad (2.63)$$

La paire de résultats des détecteurs du correcteur est (0,0). Il faut donc que

$$U_i = \mathbb{1}_2 \quad i = 1, 2, 3 \quad (2.64)$$

D'où, après que les canaux 4 et 5 aient été mesurés et les transformations unitaires (2.64) effectuées

$$(c_0 |000\rangle + c_1 |111\rangle) \otimes |00\rangle \rightarrow c_0 |000\rangle + c_1 |111\rangle \quad (2.65)$$

Et enfin,

$$c_0 |000\rangle + c_1 |111\rangle \xrightarrow{\text{Décodeur}} c_0 |000\rangle + c_1 |100\rangle = \psi_1 \otimes |00\rangle_{23} \quad (2.66)$$

Les détecteurs du décodeurs mesurent (0,0) et en sortie on retrouve l'état initial ψ .

2. Erreur sur le premier qubit

$$(c_0 |10000\rangle + c_1 |01100\rangle) \xrightarrow{\text{Correcteur}} (c_0 |100\rangle + c_1 |011\rangle)_{123} \otimes |11\rangle_{45} \quad (2.67)$$

La paire de résultats des détecteurs du correcteur est (1,1). Le premier qubit doit être permuté. Il faut donc que

$$U_1 = \sigma_x \text{ et } U_2 = U_3 = \mathbb{1}_2 \quad (2.68)$$

D'où, après que les canaux 4 et 5 aient été mesurés et les transformations unitaires (2.68) effectuées

$$(c_0 |100\rangle + c_1 |011\rangle) \otimes |11\rangle \rightarrow c_0 |000\rangle + c_1 |111\rangle \quad (2.69)$$

Et enfin,

$$c_0 |000\rangle + c_1 |111\rangle \xrightarrow{\text{Décodeur}} c_0 |000\rangle + c_1 |100\rangle = \psi_1 \otimes |00\rangle_{23} \quad (2.70)$$

Les détecteurs du décodeurs mesurent (0,0) et en sortie on retrouve l'état initial ψ .

3. Erreur sur le deuxième qubit

$$(c_0 |01000\rangle + c_1 |10100\rangle) \xrightarrow{\text{Correcteur}} (c_0 |010\rangle + c_1 |101\rangle)_{123} \otimes |10\rangle_{45} \quad (2.71)$$

La paire de résultats des détecteurs du correcteur est (1,0). Il faut donc que

$$U_2 = \sigma_x \text{ et } U_1 = U_3 = \mathbb{1}_2 \quad (2.72)$$

D'où

$$(c_0 |010\rangle + c_1 |101\rangle) \otimes |10\rangle \rightarrow c_0 |000\rangle + c_1 |111\rangle \quad (2.73)$$

Et enfin,

$$c_0 |000\rangle + c_1 |111\rangle \xrightarrow{\text{Décodeur}} \psi_1 \otimes |00\rangle_{23} \quad (2.74)$$

Les détecteurs du décodeurs mesurent (0,0) et en sortie on retrouve l'état initial ψ .

4. Erreur sur le troisième qubit : similaire au point 3.

On a donc montré qu'il est possible de corriger un certain type d'erreur. On localise l'erreur, on la corrige et cela sans acquérir la moindre information sur l'état ψ . En effet, le processus ne modifie pas les paramètres c_0 et c_1 .

Pour conclure, on peut se demander dans quelle mesure tout cela est efficace. Cette question est légitime car en fin de compte ce procédé ne sert à remplacer qu'une simple ligne. On peut montrer que la correction offerte par ce circuit devient valable si l'erreur (le bruit) générée par les *CNOT*, les transformations U_i et les détecteurs est inférieure à 1%.

Chapter 3

Quantum information theory

3.1 Théorie de l'information classique

3.1.1 Introduction

Avant d'exposer quelques points importants de la théorie de l'information quantique, on passe par une introduction des concepts classiques principaux. Claude Shannon a établi en 1948 les deux résultats qui sont au coeur de la théorie de l'information classique. Les deux problèmes qu'il est parvenu à résoudre sont :

1. De quel ordre peut être comprimé un message ? Ce qui revient à : quelle est la redondance d'une information.
2. A quel taux peut-on communiquer de manière sûre au travers d'un canal bruité? Ce qui revient à : quelle redondance doit être incorporée à un message pour le protéger contre les erreurs?

Les réponses à ces questions se formulent en deux théorèmes : *the noiseless coding theorem* (dont on montre le résultat d'une façon intuitive) pour la première et *the noisy channel coding theorem* (que l'on explique brièvement en guise d'information) pour la seconde.

On le voit, les deux questions soulèvent la notion de la redondance, c'est-à-dire le degré d'imprévisibilité (en moyenne) de la prochaine lettre d'un message. Donc, pour commencer, le problème se résume à déterminer le concept à utiliser pour quantifier la redondance. La clef a été trouvée par Shannon. Il a vu que l'*entropie* fournissait une solution convenable.

3.1.2 Entropie de Shannon

Un message est une suite de lettres choisies dans un alphabet composé de k lettres

$$\{x_1, x_2, \dots, x_k\} \quad (3.1)$$

où l'on suppose que les lettres du message sont statistiquement indépendantes et que la probabilité d'avoir dans le message la lettre x_i est donnée par $p(x_i)$ avec $\sum_{i=1}^k p(x_i) = 1$.

Prenons par exemple le cas simple d'un alphabet binaire

$$\{0, 1\} \quad (3.2)$$

où 0 a lieu avec la probabilité $1 - p$ et 1 avec p ($0 \leq p \leq 1$). Soit un long message composé de n lettres, $n \gg 1$. Est-il alors possible de compresser le message en une chaîne de lettres plus courte transmettant essentiellement la même information?

Pour n très grand, la loi des grands nombres nous dit que les chaînes typiques contiendront (pour le cas binaire) environ $n(1 - p)$ 0 et np 1. Le nombre de chaînes distinctes de cette forme est donné par le coefficient binomial $\binom{n}{np}$, et en utilisant l'approximation de Stirling $\log n! = n \log n - n$, on trouve

$$\begin{aligned} \log \binom{n}{np} &= \log \left(\frac{n!}{(np)![n(1-p)]!} \right) \\ &\cong n \log n - n - (np \log np - np) \\ &\quad - [n(1-p) \log n(1-p) - n(1-p)] \\ &= nH(p) \end{aligned} \quad (3.3)$$

où

$$H(p) = -p \log p - (1 - p) \log(1 - p) \quad (3.4)$$

est l'entropie. On précise que les logarithmes sont en base deux (sauf spécifié). Ainsi, le nombre des chaînes typiques de lettres est de l'ordre de $2^{nH(p)}$ (Rappel : $\log_a u = x \Leftrightarrow u = a^x$).

Pour transmettre essentiellement toute l'information contenue dans une chaîne de n bits, il suffit de choisir un code (*block code*) qui assigne un index (un entier positif, une lettre,...) à chacune des chaînes typiques. Ce code a donc environ $2^{nH(p)}$ index (équibrobables), ce qui signifie que chacun d'eux (i.e. chacune des chaînes typiques) peut être spécifiés en utilisant une chaîne binaire (logarithme en base 2) d'une longueur de

$$\log_2 2^{nH(p)} = nH(p) \quad (3.5)$$

Comme $0 \leq H(p) \leq 1$ pour $0 \leq p \leq 1$, et $H(p) = 1$ seulement pour $p = \frac{1}{2}$, le code raccourcit le message pour tous $p \neq \frac{1}{2}$ (i.e. chaque fois que 1 et 0 ne sont pas équiprobables). C'est le résultat de Shannon. Le point important est que l'on n'a pas besoin d'un code pour chaque séquences de lettres, seules les séquences typiques comptent. La probabilité d'avoir un message atypique devient négligeable dans la limite $n \rightarrow \infty$.

Ce raisonnement se généralise facilement au cas d'un alphabet de k lettres, où la lettre x_i apparaît avec une probabilité $p(x_i)$. Dans une chaîne de n lettres, x_i apparaît typiquement $np(x_i)$ fois, et le nombre de chaînes typiques est de l'ordre de

$$\frac{n!}{\prod_{i=1}^k (np(x_i))!} \cong 2^{nH(X)} \quad (3.6)$$

où encore une fois on a utilisé l'approximation de Stirling et

$$H(X) = - \sum_{i=1}^k p(x_i) \log p(x_i) \quad (3.7)$$

est l'entropie de Shannon de l'ensemble $X = \{x_i, p(x_i)\}$ avec $i = 1, \dots, k$. A noter que $0 \leq H(X) \leq \log_2 k$ bits.

De même, en adoptant un code qui assigne une chaîne binaire aux séquences typiques, l'information d'une chaîne de n lettres peut être compressée d'un facteur $H(X)$ bits (logarithme base 2). Dans ce sens, une lettres x_i choisie dans l'ensemble porte en moyenne $H(X)$ bits d'information. Il est également possible d'exprimer la compression $H(X)$ en nats (logarithme base e), c'est-à-dire en utilisant non plus une chaîne binaire pour spécifier les chaînes typiques mais un alphabet composé de n lettres.

Avec ce qui précède, on formule le *noiseless coding theorem* de Shannon de la façon suivante : Il existe de nombreuses façon de coder un message. Le code optimal est celui qui compresse en moyenne chaque lettre d'un message de n lettres d'un facteur $H(X)$ bits dans la limite $n \rightarrow \infty$.

Pour une bonne compréhension de la littérature, il faut se rendre compte que l'entropie $H(X)$ renseigne sur le *gain d'information* acquis une fois le message X reçu. Il est important de ne pas confondre la notion du gain d'information avec celle de la *signification* de l'information. On voit par exemple qu'un message complètement aléatoire (i.e. $p(x_i) = \frac{1}{k}$) a une entropie maximale $H(X) = \log_2 k$ bits. En d'autres termes, on peut dire que le gain d'information à sa réception est maximum, mais son sens n'est pas du tout pris en considération dans cette interprétation de l'entropie.

Remark 10 Une illustration très schématique (i.e. à ne pas prendre au pied de la lettre) devrait aider à comprendre cette notion du gain d'information. Par exemple, on affirme que le message $X = \text{“exzwcty”}$ a une entropie plus grande que le message $X' = \text{“beaucoup”}$. En effet, en français “beaucoup” est le seul mot commençant par les lettres “beauc”. Dans ce cas, après avoir reçu les cinq premières lettres du mot, on connaît à coup sûr les trois dernières. En revanche, la suite des lettres composant X est complètement imprévisible; il renferme donc un gain d'information plus grand pour le même nombre de lettres que X' (même si en français X n'a aucun sens).

3.1.3 Exemples

Exemple 1

Soit à la source l'alphabet de quatre lettres $X = \{a, b, c, d\}$ avec les probabilités $p_a = \frac{1}{2}$, $p_b = \frac{1}{4}$, $p_c = p_d = \frac{1}{8}$. On suppose que l'on veut envoyer un message $A = \{y_1, \dots, y_n\}$, $y_i \in X$ de n lettres. Naïvement, on décide pour coder les 4 lettres de l'alphabet d'utiliser

$$\log_2 4 = 2 \text{ bits/lettre} \quad (3.8)$$

ce qui donne

$$\begin{aligned} a &\rightarrow 00 \\ b &\rightarrow 01 \\ c &\rightarrow 10 \\ d &\rightarrow 11 \end{aligned} \quad (3.9)$$

Ainsi, par exemple, le message A

$$aabaabcbabd\dots$$

devient

$$0000010000000110000111\dots$$

Il faut donc logiquement pour coder ce message

$$n \left(\frac{1}{2} \cdot 2 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 2 + \frac{1}{8} \cdot 2 \right) = 2n \text{ bits} \quad (3.10)$$

D'autre part, on sait que le facteur de compression optimal que peut offrir un code est donné par l'entropie de Shannon

$$\begin{aligned}
 H(X) &= \left(\sum_{x=1}^4 -p_x \log p_x \right) \\
 &= \left(-\frac{1}{2} \log \frac{1}{2} - \frac{1}{4} \log \frac{1}{4} - \frac{1}{8} \log \frac{1}{8} - \frac{1}{8} \log \frac{1}{8} \right) \\
 &= \left(\frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 3 \right) \\
 &= \frac{7}{4}
 \end{aligned} \tag{3.11}$$

On remarque qu'en choisissant un code qui tient compte des probabilités p_i

$$\begin{aligned}
 a &\rightarrow 0 \\
 b &\rightarrow 10 \\
 c &\rightarrow 110 \\
 d &\rightarrow 111
 \end{aligned} \tag{3.12}$$

et qui demande pour coder le message A

$$\begin{aligned}
 n \left(\frac{1}{2} \cdot 1 + \frac{1}{4} \cdot 2 + \frac{1}{8} \cdot 3 + \frac{1}{8} \cdot 3 \right) &= \frac{7}{4}n \text{ bits} \\
 &= H(X) n
 \end{aligned} \tag{3.13}$$

on parvient à atteindre la compression maximale. Ce résultat inopiné est dû au fait que les probabilités p_i sont des puissances de 2.

Exemple 2

Soit un message X composé de n lettres équiprobables, i.e. $p(x_i) = \frac{1}{n}$. Alors une fois le message reçu

$$\begin{aligned}
 H(X) &= - \sum_{i=1}^n \frac{1}{n} \log_2 \frac{1}{n} \\
 &= - \log_2 \frac{1}{n} \\
 &= \log_2 n \text{ (bits)}
 \end{aligned} \tag{3.14}$$

ou

$$\begin{aligned}
 H(X) &= - \sum_{i=1}^n \frac{1}{n} \log_n \frac{1}{n} \\
 &= \sum_{i=1}^n \frac{1}{n} \\
 &= 1 \text{ (nat)}
 \end{aligned} \tag{3.15}$$

le gain d'information (entropie) est maximum. Le message ne peut donc pas être compressé.

Exemple 3

Soit un message X composé d'une seule lettre, i.e. $p(x_1) = 1$ et $p(x_i) = 0$ où $i \neq 1$. Alors

$$\begin{aligned} H(X) &= -1 \log 1 \\ &= 0 \end{aligned} \quad (3.16)$$

le gain d'information est nul.

3.1.4 Information conditionnelle

L'entropie $H(X)$ quantifie en moyenne l'information transmise par une lettre tirée d'un ensemble X , pour laquelle elle nous dit combien de bits sont nécessaires (dans la limite $n \rightarrow \infty$, où n est le nombre de lettres tirées de l'ensemble) pour codifier cette information.

De ce qui précède, on introduit la notion d'*information conditionnelle* $H(A | B)$. Elle représente le gain d'information obtenu à la réception du message A , compte tenu du fait que B est connu. Donc, $H(A | B)$ correspond au gain d'information contenu dans A qui ne se trouve pas déjà dans B .

Soit le message A composé des lettres x_i et soit le message B composé des lettres y_j associées aux poids (=probabilités de présence) $p(y_j) = q_j$. Alors

$$\begin{aligned} H(A | B) &= \sum_j q_j \left\{ - \sum_i p(x_i | y_j) \log [p(x_i | y_j)] \right\} \\ &= - \sum_j q_j \sum_i p(x_i | y_j) \log [p(x_i | y_j)] \end{aligned} \quad (3.17)$$

3.1.5 Information mutuelle

L'*information mutuelle* $I(A; B)$ quantifie la corrélation de deux messages (A et B). En d'autres termes, elle représente la quantité d'information qu'ont en commun les messages A et B .

Alors

$$I(A; B) = H(A) - H(A | B) \quad (3.18)$$

avec $I(A; B) = I(B; A) \in [0, 1]$.

On remarque que :

- Si $A = B$ alors

$$p(x_i | y_j) = \delta_{ij} \Rightarrow H(A | y_j) = 0 \quad \forall j \quad (3.19)$$

d'où

$$\begin{aligned} I(A; B) &= H(A) - H(A | B) \\ &= H(A) - 0 \\ &= H(A) \end{aligned} \quad (3.20)$$

En effet, si les deux messages sont identiques, toute l'information leur est commune.

- Si A et B n'ont rien en commun, alors

$$p(x_i | y_j) = p(x_i) \quad \forall j \quad (3.21)$$

d'où

$$\begin{aligned} I(A; B) &= H(A) - H(A | B) \\ &= H(A) - H(A) \\ &= 0 \end{aligned} \quad (3.22)$$

3.1.6 Probabilité conjointe

La *probabilité conjointe* $p(x_i \& y_j)$ représente la probabilité d'avoir la lettre x_i et la lettre y_j .

Soit $p(x_i)$ la probabilité d'avoir la lettre x_i et $p(y_j | x_i)$ la probabilité conditionnelle d'avoir y_j si x_i a eu lieu. Alors

$$\begin{aligned} p(x_i \& y_j) &= p(x_i) p(y_j | x_i) \\ &= p(y_j) p(x_i | y_j) \\ &= p(y_j \& x_i) \end{aligned} \quad (3.23)$$

On peut alors énoncer le *théorème de Bayes*

$$p(x_i | y_j) = \frac{p(x_i \& y_j)}{p(y_j)} \quad (3.24)$$

et trouver une autre expression pour l'information conditionnelle

$$\begin{aligned} H(A | B) &= - \sum_{ij} q_j p(x_i | y_j) \log[p(x_i | y_j)] \\ &= - \sum_{ij} p(y_j) \frac{p(x_i \& y_j)}{p(y_j)} \log \left[\frac{p(x_i \& y_j)}{p(y_j)} \right] \\ &= - \sum_{ij} p(x_i \& y_j) [\log p(x_i \& y_j) - \log p(y_j)] \\ &= - \sum_{ij} p(x_i \& y_j) \log p(x_i \& y_j) + \sum_{ij} p(x_i \& y_j) \log p(y_j) \\ &= H(A \& B) + \sum_j p(y_j) \log p(y_j) \\ &= H(A \& B) - H(B) \end{aligned} \quad (3.25)$$

ainsi que pour l'information mutuelle

$$I(A; B) = H(A) + H(B) - H(A \& B) \quad (3.26)$$

où l'on peut constater plus aisément que dans l'expression (3.18) la symétrie de l'information mutuelle.

On retrouve que :

- Si $A = B$ alors

$$\begin{aligned} I(A; B) &= H(A) + H(B) - H(A \& B) \\ &= H(A) + H(A) - H(A) \\ &= H(A) \end{aligned} \quad (3.27)$$

- Si A et B n'ont rien en commun, alors

$$\begin{aligned} I(A; B) &= H(A) + H(B) - H(A \& B) \\ &= H(A) + H(B) - [H(A) + H(B)] \\ &= 0 \end{aligned} \quad (3.28)$$

3.1.7 Exemples

Exemple 1

Soit $A = \{(00); (01); (10); (11)\}$ un ensemble de quatre lettres équiprobables (i.e. probabilité de $\frac{1}{4}$ pour chaque lettre) et soit $B = \{0, 1\}$ un ensemble de deux lettres.

On fait l'hypothèse

$$B = X_{or} A \quad (3.29)$$

c'est-à-dire que B est une fonction de A . Les probabilités conditionnelles sont

$$\begin{aligned} p(B = 0 \mid A = (00)) &= p(B = 0 \mid A = (11)) = 1 \\ p(B = 0 \mid A = (01)) &= p(B = 0 \mid A = (10)) = 0 \end{aligned} \quad (3.30)$$

ce qui signifie que

$$\begin{aligned} p(B = 0) &= \sum_{x \in A} p(B = 0 \mid x) p(x) \\ &= \left(\frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot 0 \right) \\ &= \frac{1}{2} \end{aligned} \quad (3.31)$$

et donc $p(B = 1) = \frac{1}{2}$ également. Ainsi

$$\begin{aligned} H(A) &= - \sum_{i=1}^4 \frac{1}{4} \log 2^{-2} \\ &= \sum_{i=1}^4 \frac{1}{4} 2 \\ &= 2 \end{aligned} \quad (3.32)$$

et

$$\begin{aligned} H(B) &= - \sum_{i=1}^2 \frac{1}{2} (-1) \log 2 \\ &= 1 \end{aligned} \quad (3.33)$$

De plus, comme par hypothèse B est entièrement déterminé par l'information de A

$$\begin{aligned} H(A \& B) &= 2 \\ &= H(A) \end{aligned} \quad (3.34)$$

(toute l'information de B est connue dès qu'on connaît A).

D'où

$$\begin{aligned} H(A | B) &= H(A \& B) - H(B) \\ &= 1 \end{aligned} \quad (3.35)$$

ou encore

$$\begin{aligned} H(A | B) &= p(B = 0) H(A | B = 0) + p(B = 1) H(A | B = 1) \\ &= \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot 1 \\ &= 1 \end{aligned} \quad (3.36)$$

Ainsi, l'information mutuelle vaut

$$\begin{aligned} I(A; B) &= H(A) - H(A | B) \\ &= 2 - 1 \\ &= 1 \end{aligned} \quad (3.37)$$

Par le théorème de Bayes, on peut encore calculer

$$\begin{aligned} p(A = (00) | B = 0) &= \frac{p(A=(00) \& B=0)}{p(B=0)} \\ &= \frac{p(B=0 | A=(00)) \cdot p(A=(00))}{p(B=0)} \\ &= 2 \cdot 1 \cdot \frac{1}{4} \\ &= \frac{1}{2} \end{aligned} \quad (3.38)$$

de même pour

$$p(A = (11) \mid B = 0) = \frac{1}{2} \quad (3.39)$$

et

$$\begin{aligned} p(A = (10) \mid B = 0) &= p(A = (01) \mid B = 0) \\ &= 0 \end{aligned} \quad (3.40)$$

Exemple 2

Soit l'ensemble $A = \{x_0, x_1, \dots, x_{n-1}\}$ et l'ensemble de distributions de probabilités $B = \{\rho_0, \rho_1, \dots, \rho_{k-1}\}$ avec la probabilité $p(\rho_j) = q_j$.

De plus, on pose que $\rho_j(x_i) = p(x_i \mid \rho_j) \in [0, 1]$ où $i = 0, \dots, n-1$ et $j = 0, \dots, k-1$ avec $\sum_i \rho_j(x_i) = 1$ et $\sum_j q_j \rho_j(x_i) = p(x_i)$.

En particulier, on choisit $A = \{0, 1\}$ et $B = \{\rho_0, \rho_1\}$ avec $\rho_0(0) = \rho_0(1) = \frac{1}{2}$, $\rho_1(0) = \frac{1}{4}$ et $\rho_1(1) = \frac{3}{4}$, $q_j = \frac{1}{2}$ avec $j = 0, 1$.

De ce qui précède, on trouve

$$\begin{aligned} p(A = 0) &= \sum_{j=0}^1 q_j \rho_j(A = 0) \\ &= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{1}{4} \\ &= \frac{3}{8} \end{aligned} \quad (3.41)$$

et

$$\begin{aligned} p(A = 1) &= \sum_{j=0}^1 q_j \rho_j(A = 1) \\ &= \frac{1}{2} \cdot \frac{1}{2} + \frac{1}{2} \cdot \frac{3}{4} \\ &= \frac{5}{8} \end{aligned} \quad (3.42)$$

Ainsi, l'entropie du message A vaut

$$\begin{aligned} H(A) &= - \sum_{i=0}^1 p(x_i) \log p(x_i) \\ &= -\frac{3}{8} \log \frac{3}{8} - \frac{5}{8} \log \frac{5}{8} \\ &\simeq 0.9544 \end{aligned} \quad (3.43)$$

Pour calculer l'information conditionnelle entre les deux messages, on commence par déterminer

$$\begin{aligned} H(A \mid \rho_0) &= - \sum_{i=0}^1 p(x_i \mid \rho_0) \log p(x_i \mid \rho_0) \\ &= -\frac{1}{2} \log \frac{1}{2} - \frac{1}{2} \log \frac{1}{2} \\ &= -\frac{1}{2} (-1) - \frac{1}{2} (-1) \\ &= 1 \end{aligned} \quad (3.44)$$

et

$$\begin{aligned} H(A | \rho_1) &= - \sum_{i=0}^1 p(x_i | \rho_1) \log p(x_i | \rho_1) \\ &= -\frac{1}{4} \log \frac{1}{4} - \frac{3}{4} \log \frac{3}{4} \\ &\simeq 0.813 \end{aligned} \quad (3.45)$$

d'où

$$\begin{aligned} H(A | B) &= \sum_{j=0}^1 q_j H(A | \rho_j) \\ &= \frac{1}{2} H(A | \rho_0) + \frac{1}{2} H(A | \rho_1) \\ &\simeq 0.906 \\ &< H(A) \end{aligned} \quad (3.46)$$

Donc, l'information mutuelle est forcément supérieure à zéro

$$\begin{aligned} I(A; B) &= H(A) - H(A | B) \\ &= 0.954 - 0.906 \\ &= 0.048 \\ &> 0 \end{aligned} \quad (3.47)$$

Compte tenu de nos hypothèses, on remarque logiquement que le gain d'information $H(A)$ est supérieur à celui obtenu pour $H(A | B)$. En effet, dans le cas où B est déjà connu, le gain d'information est plus faible une fois le message A reçu que lorsqu'on ne connaît rien.

3.1.8 The noisy channel theorem (en bref)

Si l'on veut communiquer au travers d'un canal bruité, il est évident que l'on peut améliorer la fiabilité de la transmission en utilisant la redondance. Mais est-il toujours possible dans un canal bruité donné de trouver un code qui puisse assurer une fiabilité arbitrairement grande (quand $n \rightarrow \infty$) ? Et que peut-on dire sur le taux d'un tel code, c'est-à-dire sur le nombre de bits requis pour coder chaque lettre du message?

En fait, Shannon a montré que tout canal peut être utilisé pour une communication d'une qualité arbitrairement fiable à un taux fini (différent de zéro), tant qu'il existe une corrélation entre l'input et l'output. Il a aussi pu déterminer une expression pour le taux optimal pouvant être atteint, appelé la capacité du canal. Ces résultats résument le contenu du théorème.

3.2 Théorie de l'information quantique

3.2.1 Introduction

On voudrait généraliser les considérations de la partie classique à l'information quantique. Les lettres sont représentées par des états quantiques et la source consiste en un mélange d'états quantiques ρ_x , associé à une probabilité spécifique p_x , que l'on mesure avec l'observable A . A la source, on a donc

$$\rho = \sum_x p_x \rho_x \quad (3.48)$$

De plus, on sait que pour toute matrice densité, il existe une représentation diagonale telle que

$$\rho = \sum_b \lambda_b |b\rangle \langle b| \quad (3.49)$$

avec $\langle b | b' \rangle = \delta_{bb'}$.

La nouveauté de l'information quantique, c'est le choix de la base (i.e. de l'alphabet).

3.2.2 Entropie de Von Neumann

Soit une source quantique ρ . Supposons que l'on choisisse pour *lettres* les états d'une décomposition orthogonale

$$\rho_x \equiv |b\rangle \langle b| \quad (3.50)$$

Alors on peut choisir une observable

$$A = \sum_b a_b |b\rangle \langle b| \quad (3.51)$$

avec $a_b \neq a_{b'}$ pour $b \neq b'$. Dans ce cas, si l'on obtient le résultat a_b , on sait avec certitude que la lettre était $|b\rangle \langle b|$.

Puisque la lettre $|b\rangle$ apparaît avec probabilité λ_b , on a l'entropie de cette source

$$H(\{|b\rangle, \lambda_b\}) \equiv H(B) = - \sum_b \lambda_b \log \lambda_b \quad (3.52)$$

Conclusion : en choisissant comme lettres les états d'une décomposition orthogonale de ρ , la source quantique se comporte comme une source classique définie par $\{|b\rangle, \lambda_b\}$.

On remarque que

$$-\sum_b \lambda_b \log \lambda_b = -Tr(\rho \log \rho) \quad (3.53)$$

En effet

$$\rho \log \rho = -\sum_{bb'} (\lambda_b P_b) (\log \lambda_{b'} P_{b'}) \text{ et } Tr(P_{bb'}) = \delta_{bb'} \quad (3.54)$$

Pour cette (ou n'importe quelle) matrice densité, on définit l'*entropie de Von Neumann* comme

$$S(\rho) \equiv -Tr(\rho \log \rho) \quad (3.55)$$

Encore une fois, quand l'alphabet de la source est constitué d'états quantiques purs mutuellement orthogonaux (i.e. parfaitement discernables), la source quantique se réduit à une source classique. La source quantique devient plus intéressante lorsque les états de l'ensemble ne commutent pas mutuellement.

Rappel : décomposition d'une matrice densité

On choisit par exemple une source qui produit soit $|+z\rangle$, soit $|+x\rangle$, des états non-orthogonaux de spin $\frac{1}{2}$, chacun avec une probabilité $p = \frac{1}{2}$. Donc la source peut s'écrire

$$\rho = \frac{1}{2}P_{|+z\rangle} + \frac{1}{2}P_{|+x\rangle} \quad (3.56)$$

Celle-ci n'est manifestement pas la décomposition orthogonale. On peut montrer (c.f. exercice) que la décomposition orthogonale est

$$\begin{aligned} \rho &= \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right) P_{|+m\rangle} + \left(\frac{1}{2} - \frac{1}{2\sqrt{2}}\right) P_{|-m\rangle} \\ &= p_+ P_{|+m\rangle} + p_- P_{|-m\rangle} \end{aligned} \quad (3.57)$$

avec $\vec{m} = \frac{1}{\sqrt{2}}(\hat{x} + \hat{z})$.

Cela signifie que ρ étant donné

- il est impossible de savoir si ρ a été préparé en mélangeant $|+z\rangle$ et $|+x\rangle$ en qualité égale, ou $|+m\rangle$ et $|-m\rangle$ avec les poids p_+ et p_- .
- ou, en terme d'information : on peut "lire" dans ρ des états $|+z\rangle$ et $|+x\rangle$, ou des états $|+m\rangle$ et $|-m\rangle$. (et beaucoup d'autres...).

Or, si on veut lire $|+z\rangle$ et $|+x\rangle$, c'est-à-dire si on choisit ces états comme lettres, on sort du cadre de l'information classique car ces deux lettres ne peuvent pas être discernées avec certitude, par aucune mesure.

Pourquoi s'intéresser à de tels choix de lettres? Deux raisons :

1. compréhension : on va pouvoir reformuler les particularités du monde quantique en termes d'information, à savoir l'information encodée dans l'état, celle qu'il est possible d'extraire, etc.
2. application : des protocoles comme la cryptographie quantique montrent que les sources quantiques peuvent être *utiles*, en particulier pour "cacher" l'information à toute personne non-autorisée.

Bornes de l'entropie

Reprenons l'exemple précédent :

$$\begin{aligned}\rho &= \frac{1}{2}P_{|+z\rangle} + \frac{1}{2}P_{|+x\rangle} \\ &= \left(\frac{1}{2} + \frac{1}{2\sqrt{2}}\right)P_{|+m\rangle} + \left(\frac{1}{2} - \frac{1}{2\sqrt{2}}\right)P_{|-m\rangle}\end{aligned}\quad (3.58)$$

Si l'alphabet est $|+m\rangle, |-m\rangle$, on sait calculer l'entropie de cette source, c'est l'entropie de Shannon $H\left(\frac{1}{2} \pm \frac{1}{2\sqrt{2}}\right)$. Mais si l'alphabet est $|+z\rangle, |+x\rangle$?

Entropie de préparation L'entropie de préparation est ici

$$H\left(\frac{1}{2}, \frac{1}{2}\right) = \log 2 \quad (3.59)$$

le maximum.

Soit un ensemble $X = \{|k\rangle, p_k\}$ et soit une décomposition

$$\rho = \sum_k p_k |k\rangle \langle k| \quad (3.60)$$

pour la matrice densité, avec $\langle k | k' \rangle \neq 0$ en général.

Alors

$$H(X) = H(\{p_k\}) \leq S(\rho) \quad (3.61)$$

où l'égalité est obtenue si les états $|k\rangle$ sont mutuellement orthogonaux, i.e. dès que $\langle k | k' \rangle = \delta_{kk'}$. Cela signifie que, ρ étant donné, la décomposition

orthogonale est celle qui minimise l'entropie de préparation. Ainsi, cela indique que l'on perd en discernabilité lorsque des états purs non orthogonaux sont mélangés. Il n'est pas possible de complètement retrouver l'information sur l'état qui a été préparé car le gain d'information atteint par une mesure est bornée inférieurement par $S(\rho)$.

Entropie de mesure On suppose que l'on mesure l'observable

$$A = \sum_y a_y |y\rangle \langle y| \quad (3.62)$$

avec, si ρ est la source, la probabilité d'obtenir le résultat a_y

$$p(y) = \langle y | \rho | y \rangle \quad (3.63)$$

On a alors l'alphabet $Y \equiv \{a_y, p(y)\}$. On peut montrer que l'entropie de Shannon de l'ensemble des résultats de mesures satisfait

$$H(Y) \geq S(\rho) \quad (3.64)$$

où l'égalité est obtenue lorsque A commute avec ρ , c'est-à-dire si et seulement si la base des $|y\rangle$ forment la base qui diagonalise ρ .

Mathématiquement, cela signifie que $S(\rho)$ augmente si l'on remplace tous les éléments non diagonaux de la matrice ρ par zéro ($\simeq$ cas classique), pour n'importe quelle base.

Physiquement, cela établit que le hasard sur le résultat de la mesure est minimisé si l'on choisit de mesurer une observable qui commute avec la matrice densité (i.e. si A et ρ ont des vecteurs propres communs). Ainsi, ρ étant donné, tout choix d'alphabet autre que la décomposition orthogonale rend le message plus "aléatoire".

3.2.3 Exemple

Soit une source produisant des états "up" ou "horizontaux" avec une probabilité $\frac{1}{2}$, i.e. $\xi = \{|+z\rangle, |x\rangle\}$ où $p_1 = p_2 = \frac{1}{2}$ et soit l'observable $A = \sigma_z$.

Ainsi

$$\begin{aligned} H(A | |+z\rangle) &= -p_{+z}^{+z} \log p_{+z}^{+z} - p_{-z}^{+z} \log p_{-z}^{+z} \\ &= 0 \end{aligned} \quad (3.65)$$

où

$$p_{\pm z}^{+z} = |\langle \pm z | +z \rangle|^2 \quad (3.66)$$

donne la probabilité de mesurer $\pm z$ si l'état en entrée est $+z$. En effet, si l'on est informé que l'état sur lequel on mesure A est "up", alors on sait déjà que le résultat sera $+1$, ce qui signifie que le gain d'information après la mesure est nul.

En revanche,

$$H(A \mid | +x \rangle) = 1 \quad (3.67)$$

car connaissant l'orientation horizontale de l'état à mesurer, on ne dispose d'aucune information sur le résultat de la mesure A , i.e. le gain d'information après la mesure est maximal.

D'où

$$\begin{aligned} H(A \mid \xi) &= \frac{1}{2}H(A \mid | +z \rangle) + \frac{1}{2}H(A \mid | +x \rangle) \\ &= \frac{1}{2} \end{aligned} \quad (3.68)$$

Soit l'état de mélange $\rho_\xi = \frac{1}{2}P_{|+z\rangle} + \frac{1}{2}P_{|+x\rangle}$ et $A = \sigma_z$. On introduit

$$\begin{aligned} p_+ &= p_1 p(A = 1 \mid | +z \rangle) + p_2 p(A = 1 \mid | +x \rangle) \\ &= \frac{1}{2} \cdot 1 + \frac{1}{2} \cdot \frac{1}{2} \\ &= \frac{3}{4} \end{aligned} \quad (3.69)$$

la probabilité d'obtenir le résultat $+1$. Par conséquent, la probabilité de mesurer -1 vaut

$$\begin{aligned} p_- &= 1 - p_+ \\ &= \frac{1}{4} \end{aligned} \quad (3.70)$$

Ainsi

$$\begin{aligned} H(A \mid \rho_\xi) &= -p_+ \log p_+ - p_- \log p_- \\ &\simeq 0.81 \\ &> H(A \mid \xi) \end{aligned} \quad (3.71)$$

ce qui signifie que le gain d'information sur le résultat de la mesure A est plus important lorsque l'on connaît ρ_ξ plutôt que ξ . Le fait de mélanger les états crée une perte irréversible d'information, c'est-à-dire qu'à partir d'un état de mélange, on ne peut acquérir aucune information sur la préparation. Ce qui peut être confirmé par le calcul de l'information mutuelle

$$I(A; \xi) = H(A \mid \rho_\xi) - H(A \mid \xi) \quad (3.72)$$

où on voit alors que

$$\begin{aligned} I(A; \xi) &= H(A \mid \rho_\xi) - H(A \mid \xi) \\ &\simeq 0.81 - \frac{1}{2} \\ &\simeq 0.31 \\ &> 0 \end{aligned} \quad (3.73)$$

et que d'autre part

$$\begin{aligned} I(A; \rho_\xi) &= H(A | \rho_\xi) - H(A | \rho_\xi) \\ &= 0 \end{aligned} \quad (3.74)$$

3.2.4 Théorème d'Holevo

Soit l'ensemble $\xi = \{\rho_x, p_x\}$ et soit

$$\rho = \sum_x p_x \rho_x \quad (3.75)$$

la matrice densité tirée de ξ . On définit l'*information d'Holevo*, l'analogue quantique en quelque sorte de l'information mutuelle classique, comme

$$\chi(\xi) \equiv S(\rho) - \sum_x p_x S(\rho_x) \quad (3.76)$$

qui détermine de combien, en moyenne, l'entropie de Von Neumann d'un ensemble est réduite lorsque la préparation choisie de l'état quantique tiré de ξ est connue. On remarque d'ailleurs que l'information d'Holevo se réduit à l'entropie de Von Neumann $S(\rho)$ pour un ensemble d'états purs, car dans ce cas $S(\rho_x) = 0$.

Compte tenu des notations utilisées précédemment, on peut formuler le *théorème d'Holevo* de la façon suivante

$$I(A; \xi) \leq \chi(\xi) \quad \forall A \quad (3.77)$$

c'est-à-dire qu'il n'est pas possible d'obtenir une information mutuelle entre une observable A et la préparation de l'ensemble ξ qui soit supérieure à la valeur définie par $\chi(\xi)$.

Illustration du théorème

En particulier, appliquons ce théorème à une situation connue : soit l'ensemble $\xi = \{|+z\rangle, |+x\rangle\}$ où $p_1 = p_2 = \frac{1}{2}$ et $\rho_\xi = \sum_x p_x \rho_x = \frac{1}{2}P_{|+z\rangle} + \frac{1}{2}P_{|+x\rangle}$ avec $S(P_{|+z\rangle}) = S(P_{|+x\rangle}) = 0$ (états purs).

Ainsi, on trouve $\forall A$

$$\begin{aligned} I(A; \xi) &\leq S(\rho_\xi) - \frac{1}{2}S(P_{|+z\rangle}) - \frac{1}{2}S(P_{|+x\rangle}) \\ &= S(\rho_\xi) \\ &\simeq 0.6 \end{aligned} \quad (3.78)$$

Choix d'observables On peut essayer de trouver l'observable qui offre l'information mutuelle la plus proche de 0.6, la borne supérieure fixée par le théorème d'Holevo. Pour ce faire, on pose

$$\begin{aligned} A(\theta, \psi) &= \hat{n} \cdot \vec{\sigma} \\ &= \sin\theta \cos\psi \cdot \sigma_x + \sin\theta \sin\psi \cdot \sigma_y + \cos\theta \cdot \sigma_z \end{aligned}$$

et on choisit de rester dans le plan xy , c'est-à-dire

$$A(\theta, \psi = 0) = \sin\theta \cdot \sigma_x + \cos\theta \cdot \sigma_z$$

ainsi

$$|+n\rangle = \begin{pmatrix} \cos\frac{\theta}{2} \\ \sin\frac{\theta}{2} \end{pmatrix} \text{ et } |-n\rangle = \begin{pmatrix} -\sin\frac{\theta}{2} \\ \cos\frac{\theta}{2} \end{pmatrix}$$

Pour déterminer $I(A(\theta); \xi)$, on doit calculer

$$\begin{aligned} H(A(\theta) | \xi) &= \frac{1}{2} H(A(\theta) | | +z \rangle) + \frac{1}{2} H(A(\theta) | | +x \rangle) \\ &= \frac{1}{2} \left(-p_{+n}^{+z} \log p_{+n}^{+z} - p_{-n}^{+z} \log p_{-n}^{+z} \right) \\ &\quad + \frac{1}{2} \left(-p_{+n}^{+x} \log p_{+n}^{+x} - p_{-n}^{+x} \log p_{-n}^{+x} \right) \end{aligned}$$

où

$$p_{\pm n}^{+z} = |\langle \pm n | +z \rangle|^2 \text{ et } p_{\pm n}^{+x} = |\langle \pm n | +x \rangle|^2$$

correspondent à la probabilité de mesurer $\pm n$ si l'état est $|+z\rangle$, respectivement $|+x\rangle$, et

$$H(A(\theta) | \rho_\xi) = -p_{+n}^{\rho_\xi} \log p_{+n}^{\rho_\xi} - p_{-n}^{\rho_\xi} \log p_{-n}^{\rho_\xi}$$

où

$$\begin{aligned} p_{\pm n}^{\rho_\xi} &= \text{Tr}(\rho_\xi \cdot P_{|\pm n\rangle}) \\ &= \langle \pm n | \rho_\xi | \pm n \rangle \end{aligned}$$

correspond à la probabilité de mesurer $\pm n$ si à la source on a ρ_ξ .

Enfin, en faisant varier l'angle θ , on cherche alors la valeur maximale de l'information mutuelle

$$I(A(\theta); \xi) = H(A(\theta) | \rho_\xi) - H(A(\theta) | \xi)$$

et on trouve

$$I_{\max} \left(A \left(\theta \simeq \frac{14}{15} \pi \right); \xi \right) \simeq 0.36$$

ce qui est à peine supérieur au résultat obtenu avec l'observable $A ((\theta = 2\pi)) = \sigma_z$ (équation 3.73) pour laquelle

$$I(A(\theta = 2\pi); \xi) \simeq 0.31$$

Dans ce cas particulier, la borne supérieure de 0.6 fixée par le théorème d'Holevo n'est donc pas atteinte.

3.2.5 Relation d'incertitude informationnelle

Cette partie est un peu l'équivalent des relations d'incertitude de Heisenberg dans le domaine de l'information quantique. En fait, on peut même dire que cela représente une manière moderne de les formuler.

On énonce le théorème sans en faire la démonstration (car pas évidente du tout).

Soit l'espace H de dimension N et soit A_1 et A_2 deux observables avec $A^j |\alpha_i^j\rangle = \alpha_i^j |\alpha_i^j\rangle$ et $\langle \alpha_i^j | \alpha_i^j \rangle = 1$ où $i = 1, \dots, N$ et $j = 1, 2$. Alors $\forall \xi$

$$I(A_1; \xi) + I(A_2; \xi) \leq 2 \log_2 (N \cdot C) \quad (3.79)$$

où $C = \max_{i,k} |\langle \alpha_i^1 | \alpha_k^2 \rangle|$.

Illustration de la relation d'incertitude

Exemple 1 Si $[A_1, A_2] = 0 \Rightarrow A_1$ et A_2 ont des vecteurs propres en commun donc $C = 1$. Alors

$$I(A_1; \xi) + I(A_2; \xi) \leq 2 \log_2 N \quad (3.80)$$

où $I(A_1; \xi) \leq \log_2 N$.

Exemple 2 Soit A_1 et A_2 conjuguées de façon maximales, i.e. leurs vecteurs propres sont toujours le plus loin possible les uns des autres (exple σ_x et σ_z).

Ainsi, toutes les paires d'états propres ont le même recouvrement, i.e. $C = \sqrt{\frac{1}{N}}$. Alors

$$\begin{aligned} I(A_1; \xi) + I(A_2; \xi) &\leq 2 \log_2 \left(N^{\frac{1}{2}} \right) \\ &= 2 \left(\frac{1}{2} \right) \log_2 N \\ &= \log_2 N \end{aligned} \quad (3.81)$$

donc $I(A_i; \xi) \in [0, \log_2 N]$.

Exemple 3 Soit l'ensemble $\xi = \{|+z\rangle, |-z\rangle, |+x\rangle, |-x\rangle\}$ avec $p_i = \frac{1}{4}$ où $i = 1, \dots, 4$ et soit $A_1 = \sigma_z$ et $A_2 = \sigma_x$, donc $I(A_1; \xi) = I(A_2; \xi)$. D'où (c.f. exemples 3.2.3)

$$\begin{aligned} H(A | \xi) &= \frac{1}{4}H(A_1 | |+z\rangle) + \frac{1}{4}H(A_1 | |-z\rangle) \\ &\quad + \frac{1}{4}H(A_1 | |+x\rangle) + \frac{1}{4}H(A_1 | |-x\rangle) \\ &= 0 + 0 + \frac{1}{4} + \frac{1}{4} \\ &= \frac{1}{2} \end{aligned} \quad (3.82)$$

et

$$\begin{aligned} H(A | \rho_\xi) &= -p_+ \log p_+ - p_- \log p_- \\ &= -\frac{1}{2} \log \frac{1}{2} - \frac{1}{2} \log \frac{1}{2} \\ &= 1 \end{aligned} \quad (3.83)$$

avec

$$\begin{aligned} p_+ &= p_1 p(A = 1 | |+z\rangle) + p_2 p(A = 1 | |-z\rangle) \\ &\quad + p_3 p(A = 1 | |+x\rangle) + p_4 p(A = 1 | |-x\rangle) \\ &= \frac{1}{4} \cdot 1 + \frac{1}{4} \cdot 0 + \frac{1}{4} \cdot \frac{1}{2} + \frac{1}{4} \cdot \frac{1}{2} \\ &= \frac{1}{2} \\ &= p_- \end{aligned} \quad (3.84)$$

la probabilité d'obtenir le résultat +1 (respectivement -1).

Ainsi

$$\begin{aligned} I(A_1; \xi) &= H(A | \rho_\xi) - H(A | \xi) \\ &= \frac{1}{2} \end{aligned} \quad (3.85)$$

d'où

$$I(A_1; \xi) + I(A_2; \xi) = 1 \quad (3.86)$$

Finalement,

$$\begin{aligned} 1 = I(A_1; \xi) + I(A_2; \xi) &\leq 2 \log_2 \left(2 \cdot \frac{1}{\sqrt{2}} \right) \\ &= 1 \end{aligned} \quad (3.87)$$

car pour la valeur de C , il faut se souvenir que les vecteurs propres de A_1 sont $|\pm z\rangle$ et de A_2 $|\pm x\rangle = \frac{1}{\sqrt{2}}(|+z\rangle \pm |-z\rangle)$.

Cas particulier du théorème

On énonce un corollaire moins général que (3.80) pour s'apercevoir de la façon dont une expression de ce type arrive à être démontrée.

Soit deux observables A et B auxquelles on associe les probabilités p_A et p_B et soit l'ensemble $\xi = \{\rho\}$.

Lemme

$$H(A | \rho) + H(B | \rho) \geq -2 \log_2 \left(\frac{1+C}{2} \right) \quad (3.88)$$

Démonstration

$$\begin{aligned} H(A | \rho) + H(B | \rho) &= - \sum_a p_A \log p_A - \sum_b p_B \log p_B \\ &= - \sum_{ab} p_A p_B \log(p_A p_B) \end{aligned} \quad (3.89)$$

avec

$$\begin{aligned} p_A p_B &= |\langle a | \varphi \rangle|^2 |\langle b | \varphi \rangle|^2 \\ &= \cos^2 \left(\frac{\alpha - \varphi}{2} \right) \cos^2 \left(\frac{\beta - \varphi}{2} \right) \\ &= \frac{1}{4} \left[\cos \left(\frac{\alpha - \beta}{2} \right) + \cos \left(\frac{\alpha + \beta - 2\varphi}{2} \right) \right]^2 \end{aligned} \quad (3.90)$$

donc

$$\begin{aligned} p_A p_B &\leq \frac{1}{4} \left[1 + \cos \left(\frac{\alpha - \beta}{2} \right) \right]^2 \\ &= \left[\frac{1}{2} (1 + |\langle a | b \rangle|) \right]^2 \end{aligned} \quad (3.91)$$

Finalement

$$\begin{aligned} H(A | \rho) + H(B | \rho) &\geq - \sum_{ab} p_A p_B \log_2 \left(\frac{1+C}{2} \right)^2 \\ &= -2 \log_2 \left(\frac{1+C}{2} \right) \end{aligned} \quad (3.92)$$

avec $C = \max_{a,b} |\langle a | b \rangle|$.

Corollaire Soit A et B non-dégénérées, donc autant de valeurs propres que la dimension de H , notée N . En imaginant qu'on ne connaît rien sur la préparation de notre système, on a

$$\begin{aligned} H \left(A \mid \frac{1}{N} \mathbb{1} \right) &= H \left(B \mid \frac{1}{N} \mathbb{1} \right) \\ &= -\frac{1}{N} \log_2 \frac{1}{N} \\ &= \log_2 N \end{aligned}$$

d'où

$$\begin{aligned}
 I(A; \rho) + I(B; \rho) &= H\left(A \mid \frac{1}{N} \mathbb{1}\right) - H(A \mid \rho) \\
 &\quad + \left(H\left(B \mid \frac{1}{N} \mathbb{1}\right) - H(B \mid \rho)\right) \\
 &= H\left(A \mid \frac{1}{N} \mathbb{1}\right) + H\left(B \mid \frac{1}{N} \mathbb{1}\right) \\
 &\quad - (H(A \mid \rho) + H(B \mid \rho)) \\
 &\leq 2 \log_2 N + 2 \log_2 \left(\frac{1+C}{2}\right) \\
 &= 2 \log_2 \left(\frac{1+C}{2} \cdot N\right)
 \end{aligned} \tag{3.93}$$

On remarque que ce corollaire est en effet moins fort que le théorème (3.80) car la borne qu'il impose à la somme des informations mutuelles est supérieure à $2 \log_2 (N \cdot C)$.