Université de Montréal

Analyse et comparaison de protocoles de purification de l'intrication quantique

par

Antoine Schoeb
D.I.R.O.
Faculté des arts et des sciences

Mémoire pésenté à la Faculté des études supérieures
en vue de l'obtention du grade de
Maître ès sciences (M.Sc.)
en informatique

août, 1999

Université de Montréal
Faculté des études supérieures

Ce mémoire intitulé :

Analyse et comparaison de protocoles de purification
de l'intrication quantique

présenté par :

Antoine Schoeb

a été évalué par un jury composé des personnes suivantes :

Mémoire accepté le :

Sommaire

Ce mémoire se consacre à l'étude des protocoles de purification. Ces protocoles ont été développé afin de corriger des erreurs pouvant survenir dans le monde quantique. Leur intérêt provient aussi du fait qu'ils révèlent une facette importante de la notion d'intrication quantique. Dans ce mémoire, nous allons analyser en profondeur cette facette en comparant différents protocoles de purification dont quelques uns originaux.

Le premier chapitre a pour but de définir de façon globale l'intrication quantique. Pour ce faire, nous devons rappeler les notions de base sous-jacentes à celle-ci. On y définit entre autres les notions d'états purs, de mesures orthogonales, d'opérations unitaires et de mélanges statistiques.

Le second chapitre commence par introduire les protocoles de purification dans un contexte général. Les notions rattachées aux protocoles de purification, tel le rendement, y sont aussi définies. Le premier protocole discuté est la méthode de récurrence. Une analyse originale de ce protocole aboutit sur une condition forte afin d'optimiser son rendement. Ensuite, on présente les méthodes de purification directe et de filtrage. Nous avons trouvé des variantes originales pour ces deux protocoles. La comparaison de ces méthodes est faite à l'aide de graphiques. À la fin de ce chapitre, nous présentons la méthode de hashing ainsi qu'une généralisation moins connue de celle-ci.

Le troisième chapitre discute aussi de protocoles de purification mais adaptés au modèle d'erreur des codes correcteurs quantiques. Dans ce chapitre, nous allons limiter notre étude aux protocoles de purification unidirectionnels car ceux-ci sont directements associables à des codes correcteurs quantiques. Par une méthode originale, on a réussi

à générer plusieurs codes correcteurs. En autre, on obtient un circuit à dix portes pour le cas le plus connu ($t = 1, m = 1, n = 5$). Ce circuit possède un nombre de portes plus petit que ceux retrouvés dans la littérature.

D'autres résultats ont été trouvés pour les cas ($t = 1, m = 2, n = 8$), ($t = 1, m = 3, n = 9$) et ($t = 2, m = 1, n = 15$). Pour ces derniers cas, aucun résultat n'a été trouvé dans la littérature.

Table des matières

Liste des tableaux

Liste des figures

Remerciements

Je voudrais remercier mon directeur de maîtrise Gilles Brassard pour ses nombreuses suggestions ainsi que ses commentaires utiles qui m'ont aidé à réaliser ce mémoire. De plus, je voudrais souligner la qualité de son enseignement, lequel a suscité mon intérêt pour le domaine de l'informatique quantique.

J'aimerais aussi remercier mon co-directeur de maîtrise Claude Crépeau pour ses bons conseils.

Chapitre 1

Notions de base en informatique quantique

1.1 Lois de la mécanique quantique

La théorie de la mécanique quantique est un modèle mathématique visant à expliquer le monde réel. La validité de cette théorie est généralement admise dans la communauté scientifique. Dans cette section, nous allons exposer les grandes lignes de ce modèle.

1.1.1 États et espace de Hilbert

D'abord, il faut spécifier comment le modèle représente les états. Un état est une description complète d'un système physique. En mécanique quantique, un état est un *rayon* dans un espace de Hilbert[1]. Soit $\mathcal{H}$ un espace de Hilbert de dimension n et soit

$$\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix}$$

un vecteur de $\mathcal{H}$ exprimé en termes de ses composantes selon une base orthonormée arbitraire. Suivant la notation de Dirac, ce vecteur est dénoté par un *ket* $|\psi\rangle$. Le *bra* $\langle\psi|$ dénote le vecteur conjugué transposé de $|\psi\rangle$:

$$\langle\psi| = (\alpha_1^*, \alpha_2^*, \ldots, \alpha_n^*),$$

où * représente l'opération de conjugaison complexe. La juxtaposition d'un bra et d'un ket est le produit scalaire des deux vecteurs. Ce produit envoie une paire ordonnée de vecteurs dans $\mathbb{C}$. Par exemple, soit :

$$|\phi\rangle = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix}$$

alors

$$\begin{aligned} \langle\psi|\phi\rangle &= (\alpha_1^*, \alpha_2^*, \ldots, \alpha_n^*) \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} \\ &= \sum_{i=1}^{n} \alpha_i^* \beta_i. \end{aligned}$$

Le produit scalaire possède les propriétés suivantes :

1. positivité : $\langle\psi|\psi\rangle$ est un nombre réel non négatif,

[1] Un espace vectoriel sur les nombres complexes $\mathbb{C}$.

2. linéarité : $\langle\phi|\Big(a|\psi_1\rangle + b|\psi_2\rangle\Big) = a\langle\phi|\psi_1\rangle + b\langle\phi|\psi_2\rangle \qquad (a, b \in \mathbb{C})$,
3. symétrie : $\langle\phi|\psi\rangle = \langle\psi|\phi\rangle^*$.

La norme d'un vecteur $|\psi\rangle$ est donnée par $\|\,|\psi\rangle\,\| = \langle\psi|\psi\rangle^{1/2}$.

Dans un espace de Hilbert, un rayon est une classe de vecteurs qui diffèrent d'une multiplication par une constante complexe. On dénote un rayon par un représentant de sa classe qui possède une norme égale à l'unité :

$$\langle\psi|\psi\rangle = 1.$$

On note que pour tout θ réel, $e^{i\theta}|\psi\rangle$ est aussi un représentant valide de la classe. On obtient alors que $|\psi\rangle$ et $e^{i\theta}|\psi\rangle$ décrivent le même état physique. En réalité, cette dernière affirmation a servi de justification à la définition d'état quantique ; elle ne devrait pas être vue comme une conséquence de celle-ci.

1.1.2 Système bipartite

Une autre opération importante agissant dans l'espace de Hilbert est le produit tensoriel. Ce produit intervient lorsque l'on considère un système bipartite[2].

Supposons donc que deux systèmes quantiques A et B interagissent. Soit $\mathcal{H}_A$ et $\mathcal{H}_B$ les deux espaces de Hilbert qui servent à représenter les états de A et B respectivement. On note :

$$\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$$

l'espace de Hilbert du système composé des sous-systèmes $\mathcal{H}_A$ et $\mathcal{H}_B$. L'opération $\otimes$ désigne le produit tensoriel entre deux espaces. Si $\{|\mu_i\rangle\}$ et $\{|\nu_i\rangle\}$ sont des bases orthonormées pour $\mathcal{H}_A$ et $\mathcal{H}_B$ respectivement, alors

$$\{|\mu_i\rangle \otimes |\nu_j\rangle\}$$

est une base orthonormée pour $\mathcal{H}$. En général, si $u \in \mathcal{H}_A$ et $v \in \mathcal{H}_B$, alors $u \otimes v \in \mathcal{H}$. Le produit tensoriel est analogue au produit cartésien entre deux ensembles. Par

[2]Système composé de deux sous-systèmes.

exemple, le produit tensoriel de deux kets $|\psi\rangle$ et $|\phi\rangle$ de $\mathcal{H}_2$ est un ket de $\mathcal{H}_{2\times 2}$ dont les composantes se calculent comme suit :

$$|\psi\rangle \otimes |\phi\rangle = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \otimes \begin{pmatrix} \beta_1 \\ \beta_2 \end{pmatrix} = \begin{pmatrix} \alpha_1\beta_1 \\ \alpha_1\beta_2 \\ \alpha_2\beta_1 \\ \alpha_2\beta_2 \end{pmatrix}.$$

Dans l'espace $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, le produit scalaire de $|v\rangle = |v_A\rangle \otimes |v_B\rangle$ et $|u\rangle = |u_A\rangle \otimes |u_B\rangle$ revient à :

$$\langle v|u\rangle = \langle v_A|u_A\rangle\langle v_B|u_B\rangle.$$

1.1.3 Opérateurs et mesures orthogonales

Il y a plusieurs opérations que l'on peut appliquer à un état quantique. On définit un opérateur comme une application linéaire de vecteurs à vecteurs. L'application de l'opérateur A sur le ket $|\psi\rangle$ est dénotée par $A|\psi\rangle$. Puisqu'il s'agit d'une application linéaire, on a que :

$$A(a|\phi\rangle + b|\varphi\rangle) = aA|\phi\rangle + bA|\varphi\rangle.$$

Dans la notation matricielle, ces opérateurs sont simplement des matrices de nombres complexes. Par exemple, le produit d'un ket et d'un bra :

$$\begin{aligned} |\psi\rangle\langle\phi| &= \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{pmatrix} (\beta_1^*, \beta_2^*, \ldots, \beta_n^*) \\ &= \begin{pmatrix} \alpha_1\beta_1^* & \alpha_1\beta_2^* & \ldots & \alpha_1\beta_n^* \\ \alpha_2\beta_1^* & \alpha_2\beta_2^* & \ldots & \alpha_2\beta_n^* \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_n\beta_1^* & \alpha_n\beta_2^* & \ldots & \alpha_n\beta_n^* \end{pmatrix} \end{aligned}$$

est un opérateur qui envoie un vecteur $|\xi\rangle$ dans $|\psi\rangle\langle\phi|\xi\rangle = \langle\phi|\xi\rangle|\psi\rangle$, i.e. un vecteur de longueur $\langle\phi|\xi\rangle$ dans la direction $|\psi\rangle$. Un cas particulier de cette classe d'opérateurs survient lorsque $|\psi\rangle = |\phi\rangle$ et que $\| \, |\phi\rangle \, \| = 1$; on parle alors d'*opérateurs de projection unidimensionnelle*.

Les opérateurs hermitiens constituent une classe importante qui englobe les opérateurs de projection unidimensionnelle. L'opérateur A est hermitien si $A = A^\dagger$, où $\dagger$ désigne l'opération composée de conjugaison et de transposition. Chaque opérateur hermitien représente un *observable*, i.e. une propriété d'un système physique qui peut être mesurée. Un opérateur hermitien A agissant dans un espace de Hilbert $\mathcal{H}$ possède une représentation spectrale ; ses vecteurs propres forment une base orthonormée de $\mathcal{H}$. On peut exprimer A comme :

$$A = \sum_n a_n P_n,$$

où les a_n sont les valeurs propres réelles de A et les P_n sont les projections orhogonales dans les sous-espaces propres associés aux a_n. Soit $\{|n_i\rangle\}$ une base orthonormée de l'espace propre associé à a_n, alors on peut écrire que $P_n = \sum_i |n_i\rangle\langle n_i|$. (Si a_n est non dégénéré, P_n se réduit à un opérateur de projection unidimensionnelle.) Les P_n sont des projections orthogonales puisqu'ils satisfont :

$$\begin{aligned} P_n P_m &= \delta_{n,m} P_n, \\ P_n^\dagger &= P_n, \\ \sum_n P_n &= I, \end{aligned}$$

où I désigne la matrice identité de l'espace dans lequel la mesure est effectuée.

En mécanique quantique, le résultat numérique de la mesure d'un observable A est l'une de ses valeurs propres. Si l'état quantique juste avant la mesure est $|\psi\rangle$, le résultat a_n est obtenu avec probabilité :

$$\mathrm{prob}(a_n) = \|P_n|\psi\rangle\|^2 = \langle\psi|P_n|\psi\rangle.$$

On note que la somme des probabilités est bien égale à 1 :

$$\begin{aligned}\sum_n \mathrm{prob}(a_n) &= \sum_n \langle\psi|P_n|\psi\rangle \\ &= \langle\psi|(\sum_n P_n)|\psi\rangle \\ &= \langle\psi|I|\psi\rangle \\ &= \langle\psi|\psi\rangle \\ &= 1.\end{aligned}$$

À la suite de l'obtention du résultat a_n, l'état quantique devient :

$$\frac{P_n|\psi\rangle}{(\langle\psi|P_n|\psi\rangle)^{1/2}}.$$

(On note que $P_n|\psi\rangle$ n'est pas normalisé, d'où le besoin d'introduire une constante de normalisation.)

En vue d'introduire les matrices de densité, il est intéressant d'exprimer la probabilité d'obtenir le résultat a_n sous une autre forme. Pour en arriver là, nous avons besoin d'une nouvelle définition : la *trace* d'un opérateur est la somme des éléments diagonaux de la matrice qui représente cet opérateur ; cette valeur est indépendante de la base dans laquelle la matrice est exprimée. De plus, elle est invariante sous une permutation cyclique, i.e. $\mathrm{Tr}(ABC) = \mathrm{Tr}(CBA)$. Soit une base orthonormée $\{|\mu_i\rangle\}$ et un opérateur B quelconque. On a que :

$$\begin{aligned}\mathrm{Tr}(|\psi\rangle\langle\psi|B) &= \sum_i \langle\mu_i|\psi\rangle\langle\psi|B|\mu_i\rangle \\ &= \sum_i \langle\psi|B|\mu_i\rangle\langle\mu_i|\psi\rangle \\ &= \langle\psi|B\Big(\sum_i |\mu_i\rangle\langle\mu_i|\Big)|\psi\rangle \\ &= \langle\psi|B|\psi\rangle \qquad \text{car} \qquad \sum_i |\mu_i\rangle\langle\mu_i| = I.\end{aligned}$$

En utilisant cette dernière équation, on obtient que :

$$\begin{aligned}\mathrm{prob}(a_n) &= \langle\psi|P_n|\psi\rangle \\ &= \mathrm{Tr}(\rho P_n),\end{aligned}$$

où $\rho = |\psi\rangle\langle\psi|$ est l'opérateur de projection sur l'état $|\psi\rangle$. On constate donc que ρ peut servir de représentation pour l'état $|\psi\rangle$ puisqu'il permet de rendre compte des mêmes probabilités suite à une mesure. Nous allons revenir sur ce point dans la prochaine sous-section.

Une autre classe d'opérateurs importante est celle des opérateurs unitaires. Un opérateur U est unitaire si $U^\dagger = U^{-1}$. Un système quantique isolé sur lequel on n'effectue aucune mesure évolue, après un temps t, de façon unitaire :

$$|\psi(t)\rangle = U|\psi(0)\rangle.$$

On remarque que $|\psi(t)\rangle$ est normalisé car :

$$\langle\psi(t)|\psi(t)\rangle = \langle\psi(0)|U^\dagger U|\psi(0)\rangle = \langle\psi(0)|U^{-1}U|\psi(0)\rangle = \langle\psi(0)|\psi(0)\rangle = 1.$$

La nature de l'opérateur unitaire U dépend bien sûr du système quantique que l'on considère. La relation qui relie les deux, bien que fondamentale en théorie, n'est cependant pas importante dans ce que nous allons faire. Puisque toute opération unitaire peut être produite par un système physique particulier, nous avons le droit en principe d'utiliser n'importe laquelle de ces opérations unitaires dans nos protocoles quantiques. Nous n'allons pas nous soucier de l'implantation physique de chacune de ces opérations.

1.1.4 Mélanges statistiques et matrice de densité

Jusqu'à maintenant, nous avons donné une description tout à fait correcte et générale de la théorie quantique. Cependant, dans certaines circonstances, il semble que cette description soit inappropriée ou même incorrecte. Le problème est que le formalisme de la mécanique quantique est défini pour des systèmes parfaitement isolés tel l'univers entier. La plupart du temps, le système que l'on considère n'est pas isolé puisque celui-ci fait partie d'un système plus grand avec lequel il interagit. Dans ce cas, l'état du système sous considération n'est pas un rayon de l'espace de Hilbert et, plus important encore, l'évolution de cet état n'est pas unitaire. Notre définition d'état quantique doit donc être revisée ; à l'avenir, un rayon dans un espace de Hilbert sera appelé un *état pur*. Nous allons maintenant voir un autre type d'états : les *mélanges statistiques.*

Pour introduire les mélanges statistiques, nous avons besoin de revenir sur les systèmes bipartites. Soit donc, tel que vu dans la sous-section 1.1.2, un espace de Hilbert $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$ avec $\{|\mu_i\rangle\}$ et $\{|\nu_i\rangle\}$ comme bases orthonormées pour $\mathcal{H}_A$ et $\mathcal{H}_B$ respectivement. Un état pur arbitraire de $\mathcal{H}$ peut s'écrire sous la forme :

$$|\psi\rangle = \sum_{i,j} a_{ij}|\mu_i\rangle \otimes |\nu_j\rangle, \tag{1.1}$$

où $\sum_{ij}|a_{ij}|^2 = 1$. Soit A et B deux observables agissant sur $\mathcal{H}_A$ et $\mathcal{H}_B$ respectivement. Ces deux observables possèdent les représentations spectrales suivantes :

$$\begin{aligned} A &= \sum_n a_n P_n \\ B &= \sum_m b_m Q_m. \end{aligned}$$

Considérons aussi C, un observable agissant sur le système global, dont l'effet est la mesure conjointe des observables A et B. On peut exprimer C comme :

$$C = \sum_{n,m} c_{n,m} P_n \otimes Q_m,$$

où $c_{n,m}$ représente la valeur propre reliée à l'obtention des résultats a_n et b_m. Comme on a vu, si on applique l'observable C sur l'état $|\psi\rangle$, on a que :

$$\text{prob}(c_{n,m}) = \langle\psi|P_n \otimes Q_m|\psi\rangle.$$

Prenons maintenant la position d'un observateur du sytème de $\mathcal{H}_A$ qui ne possède aucune connaissance du système de $\mathcal{H}_B$. Suite à la mesure de l'observable C, on observera que :

$$\begin{aligned} \text{prob}(a_n) &= \sum_m \text{prob}(c_{n,m}) \\ &= \sum_m \langle\psi|P_n \otimes Q_m|\psi\rangle \\ &= \langle\psi|P_n \otimes (\sum_m Q_m)|\psi\rangle \\ &= \langle\psi|P_n \otimes I|\psi\rangle. \end{aligned}$$

Il est rassurant de noter que cette dernière expression pour la probabilité d'obtenir le résultat a_n ne dépend pas de B. Cependant, elle fait encore intervenir $|\psi\rangle$, l'état du système global. Afin de se débarrasser de cette dépendance au système global, nous devons introduire les matrices de densité réduites. Le but de ces matrices est de renfermer l'information reliée à un seul des sous-systèmes. Ainsi, pour obtenir la matrice de densité ρ_A reliée au sous-système $\mathcal{H}_A$, on fait la *trace partielle* des degrés de liberté du sous-système $\mathcal{H}_B$ sur le système global. Cette opération se traduit comme suit :

$$\begin{aligned}
\rho_A &= \mathrm{Tr}_B(|\psi\rangle\langle\psi|) \\
&\equiv \sum_m \langle\nu_m|\Big(|\psi\rangle\langle\psi|\Big)|\nu_m\rangle.
\end{aligned}$$

Afin de montrer que cette définition a du sens, considérons le cas où $|\psi\rangle$ est le simple produit tensoriel de deux états purs $|\mu\rangle$ et $|\nu\rangle$. On obtiendrait alors :

$$\begin{aligned}
\rho_A &= \mathrm{Tr}_B(|\psi\rangle\langle\psi|) \\
&= \sum_m \langle\nu_m|\Big(|\psi\rangle\langle\psi|\Big)|\nu_m\rangle \\
&= \sum_m \langle\nu_m|\Big(|\mu\rangle\otimes|\nu\rangle\langle\mu|\otimes\langle\nu|\Big)|\nu_m\rangle \\
&= |\mu\rangle\Big(\sum_m \langle\nu_m|\nu\rangle\langle\nu|\nu_m\rangle\Big)\langle\mu| \\
&= |\mu\rangle\Big(\sum_m \langle\nu|\nu_m\rangle\langle\nu_m|\nu\rangle\Big)\langle\mu| \\
&= |\mu\rangle\langle\nu|\Big(\sum_m |\nu_m\rangle\langle\nu_m|\Big)|\nu\rangle\langle\mu| \\
&= |\mu\rangle\langle\nu|\nu\rangle\langle\mu| \qquad \text{car} \qquad \sum_m |\nu_m\rangle\langle\nu_m| = I \\
&= |\mu\rangle\langle\mu|,
\end{aligned}$$

tel qu'attendu. Dans le cas général, on obtient :

$$\begin{aligned}
\rho_A &= \mathrm{Tr}_B(|\psi\rangle\langle\psi|) \\
&= \sum_m \langle\nu_m|\Big(|\psi\rangle\langle\psi|\Big)|\nu_m\rangle
\end{aligned}$$

$$
\begin{aligned}
&= \sum_m \langle \nu_m| \Big(\sum_{i,j} a_{ij} |\mu_i\rangle \otimes |\nu_j\rangle \Big) \Big(\sum_{k,l} a^*_{kl} \langle \mu_k| \otimes \langle \nu_l| \Big) |\nu_m\rangle \\
&= \sum_m \sum_{i,j} a_{ij} |\mu_i\rangle \delta_{jm} \sum_{k,l} a^*_{kl} \langle \mu_k| \delta_{lm} \\
&= \sum_{i,j,k,l} a_{ij} a^*_{kl} |\mu_i\rangle\langle \mu_k| \sum_m \delta_{jm}\delta_{lm} \\
&= \sum_{i,j,k,l} a_{ij} a^*_{kl} |\mu_i\rangle\langle \mu_k| \delta_{jl} \\
&= \sum_{i,j,k} a_{ij} a^*_{kj} |\mu_i\rangle\langle \mu_k|.
\end{aligned}
$$

Maintenant, on peut vérifier facilement que $\mathrm{prob}(a_n) = \mathrm{Tr}(\rho_A P_n)$. Il suffit de noter que :

$$
\begin{aligned}
\mathrm{Tr}(\rho_A P_n) &= \mathrm{Tr}\Big(\sum_{i,j,k} a_{ij} a^*_{kj} |\mu_i\rangle\langle \mu_k| P_n\Big) \\
&= \sum_l \langle \mu_l| \sum_{i,j,k} a_{ij} a^*_{kj} |\mu_i\rangle\langle \mu_k| P_n |\mu_l\rangle \\
&= \sum_{i,j,k,l} a_{ij} a^*_{kj} \langle \mu_l|\mu_i\rangle\langle \mu_k| P_n |\mu_l\rangle \\
&= \sum_{i,j,k,l} a_{ij} a^*_{kj} \delta_{il} \langle \mu_k| P_n |\mu_l\rangle \\
&= \sum_{i,j,k} a_{ij} a^*_{kj} \langle \mu_k| P_n |\mu_i\rangle,
\end{aligned}
$$

et que :

$$
\begin{aligned}
\mathrm{prob}(a_n) &= \langle \psi|(P_n \otimes I_B)|\psi\rangle \\
&= \Big(\sum_{k,l} a^*_{kl} \langle \mu_k| \otimes \langle \nu_l| \Big) (P_n \otimes I_B) \Big(\sum_{i,j} a_{ij} |\mu_i\rangle \otimes |\nu_j\rangle \Big) \\
&= \sum_{i,j,k,l} a^*_{kl} a_{ij} \langle \mu_k| P_n |\mu_i\rangle \langle \nu_l|\nu_j\rangle \\
&= \sum_{i,j,k} a^*_{kj} a_{ij} \langle \mu_k| P_n |\mu_i\rangle \qquad \text{car } \langle \nu_l|\nu_j\rangle = \delta_{lj} \\
&= \mathrm{Tr}(\rho_A P_n).
\end{aligned}
$$

Cette expression est une généralisation de celle obtenue à la section précédente pour les états purs. Dans le présent cas, ρ_A ne s'exprime plus nécessairement sous la forme

$|\phi\rangle\langle\phi|$ comme dans le cas des états purs. En fait, ρ_A appartient à une classe d'opérateurs plus vaste (qui englobe les opérateurs de projection unidimensionnelle) que l'on appelle les *opérateurs de densité*. Ces opérateur possèdent les propriétés suivantes :

1. ρ_A est hermitien : $\rho_A = \rho_A^\dagger$.

2. ρ_A est positif :

$$\begin{aligned} \langle\phi|\rho_A|\phi\rangle &= \langle\phi|(\sum_{i,j,k} a_{ij}a_{kj}^*|\mu_i\rangle\langle\mu_k|)|\phi\rangle \\ &= \sum_j(\sum_i a_{ij}\langle\phi|\mu_i\rangle)(\sum_k a_{kj}^*\langle\mu_k|\phi\rangle) \\ &= \sum_j |\sum_i a_{ij}\langle\phi|\mu_i\rangle|^2 > 0 \qquad \text{pour tout } |\phi\rangle. \end{aligned}$$

3. $\mathrm{Tr}(\rho_A) = 1$:

$$\begin{aligned} \mathrm{Tr}(\rho_A) &= \sum_l \langle\mu_l|\rho_A|\mu_l\rangle \\ &= \sum_{i,j,k,l} a_{ij}a_{kj}^*\langle\mu_l|\mu_i\rangle\langle\mu_k|\mu_l\rangle \\ &= \sum_{i,j,k,l} a_{ij}a_{kj}^*\delta_{li}\delta_{kl} \\ &= \sum_{j,l} a_{lj}a_{lj}^* \\ &= \sum_{j,l} |a_{lj}|^2 \\ &= 1. \end{aligned}$$

Grâce à ces propriétés, ρ_A peut être diagonalisé et possède des valeurs propres réelles qui somment à un. Un exemple particulier de matrice de densité est la matrice $\frac{1}{n}I_n$ qui représente un état quantique de $\mathcal{H}_n$ parfaitement aléatoire.

En résumé, lorsque l'on considère un sous-système A d'un système plus large décrit par un état pur, l'état du sous-système n'est pas représenté par un rayon, mais plutôt par un opérateur de densité ρ_A. En général, une matrice de densité, exprimée dans la base qui la rend diagonale, s'écrit sous la forme :

$$\rho_A = \sum_a p_a |\psi_a\rangle\langle\psi_a|,$$

où $0 < p_a \leq 1$ et $\sum_a p_a = 1$. Si l'état n'est pas pur, il y a plus d'un terme dans cette somme et on parle alors de mélange statistique. À partir de cela, on peut montrer qu'une matrice de densité ρ représente un état pur si et seulement si $\rho^2 = \rho$.

On voit que l'on peut donc interpréter ρ_A comme une collection d'états purs dans laquelle l'état $|\psi_a\rangle$ intervient avec probabilité p_a. Dans cette collection "spectrale", les $|\psi_a\rangle$ sont orthonormés. D'autre part, il existe d'autres collections d'états purs, ceux-là pas nécessairement orthogonaux, qui peuvent aussi bien servir à préparer ρ_A. Chacune de ces collections représente une recette différente pour préparer ρ_A.

Cependant, un postulat central en mécanique quantique dit qu'il est impossible, par quelque mesure que ce soit, de savoir comment un mélange statistique a été préparé. Les matrices de densité constituent donc un bon formalisme pour représenter les mélanges statistiques ; elles permettent d'encapsuler toute l'information pertinente sur un état sans laisser d'indice sur la façon dont celui-ci a été préparé.

1.1.5 Evolution de l'opérateur de densité

Lorsqu'on considère un système bipartite AB dans lequel il n'y a pas de couplage entre les deux sous-systèmes A et B, l'opérateur d'évolution du système combiné :

$$U_{AB} = U_A \otimes U_B$$

est décomposable en deux opérateurs d'évolution indépendants agissant sur chaque sous-système. Après l'application de cet opérateur d'évolution, l'état (1.1) devient :

$$|\psi'\rangle = \sum_{i,j} a_{ij} |\mu_i'\rangle \otimes |\nu_j'\rangle,$$

où

$$|\mu_i'\rangle = U_A|\mu_i\rangle \qquad \forall i,$$
$$|\nu_j'\rangle = U_B|\nu_j\rangle \qquad \forall j,$$

définit une nouvelle base orthonormée pour $\mathcal{H}_A$ et $\mathcal{H}_B$. En faisant la trace partielle, on trouve que ρ_A subit une évolution unitaire :

$$\begin{aligned} \rho'_A &= \sum_{i,j,k} a_{ij} a^*_{kj} |\mu'_i\rangle\langle\mu'_k| \\ &= \sum_{i,j,k} a_{ij} a^*_{kj} U_A |\mu_i\rangle\langle\mu_k| U_A^\dagger \\ &= U_A \rho_A U_A^\dagger. \end{aligned}$$

En général, l'opérateur d'évolution U_{AB} n'est pas décomposable en deux parties ; il est alors impossible de l'écrire sous forme d'un produit tensoriel. Dans ce cas, l'évolution de ρ_A est beaucoup plus complexe et, en particulier, non unitaire. Heureusement, dans ce mémoire nous allons presque toujours considérer des opérateurs unitaires décomposables.

1.2 Introduction à l'informatique quantique

Nous avons maintenant une assez bonne connaissance de la théorie quantique pour présenter les bases de l'informatique quantique nécessaires à notre développement.

En informatique classique, le bit constitue l'unité de base de l'information ; il peut prendre les deux valeurs possibles $\{0, 1\}$. L'unité correspondante d'information quantique est appelé le bit quantique ou *qubit*.

Le plus petit espace de Hilbert non trivial est de dimension deux. Il est convenable de dénoter $\{|0\rangle, |1\rangle\}$ une base orthonormée pour un tel espace. Dans cette base, appelée souvent *base standard*, l'état quantique le plus général peut être exprimé comme :

$$a|0\rangle + b|1\rangle,$$

où a et b sont des nombres complexes tels que $|a|^2 + |b|^2 = 1$. Un qubit est n'importe quel état de cette forme. Comme on a vu, on peut faire une mesure qui projette le qubit dans la base standard. Nous allons alors obtenir le résultat $|0\rangle$ avec probabilité $|a|^2$, et le résultat $|1\rangle$ avec probabilité $|b|^2$. De plus, excepté dans les cas où $a = 0$ ou

$b = 0$, la mesure change de façon irrévocable l'état initial. Si la valeur du qubit est inconnue au départ, aucune mesure ne permettra de déterminer les valeurs de a et b. Cependant, après la mesure, le qubit a été préparé dans un état connu : soit $|0\rangle$ ou $|1\rangle$.

Des systèmes quantiques de plus grande dimension peuvent être mis en jeu si l'on considère plusieurs qubits. Par exemple, pour deux qubits la base standard devient le produit tensoriel des bases standards des deux qubits séparés :

$$\begin{gathered} |0\rangle \otimes |0\rangle, \\ |0\rangle \otimes |1\rangle, \\ |1\rangle \otimes |0\rangle, \\ |1\rangle \otimes |1\rangle. \end{gathered}$$

Habituellement, on écrit simplement $|00\rangle, |01\rangle, |10\rangle$ et $|11\rangle$. Une autre base de $\mathcal{H}_4$ souvent utilisée est la *base de Bell*. Dans la base standard, cette base s'exprime de la façon suivante :

$$\begin{aligned} |\Phi^+\rangle &= \tfrac{1}{\sqrt{2}}(|00\rangle + |11\rangle), \\ |\Phi^-\rangle &= \tfrac{1}{\sqrt{2}}(|00\rangle - |11\rangle), \\ |\Psi^+\rangle &= \tfrac{1}{\sqrt{2}}(|01\rangle + |10\rangle), \\ |\Psi^-\rangle &= \tfrac{1}{\sqrt{2}}(|01\rangle - |10\rangle). \end{aligned}$$

L'état $|\Psi^-\rangle$ est appelé l'état *singlet*. Parfois, on utilise $\uparrow$ au lieu de 0 et $\downarrow$ au lieu de 1. La raison est que le système physique le plus connu pour représenter un qubit est le spin de l'électron. Lorsque mesuré selon l'axe z, la direction du spin peut être soit parallèle à cet axe ($\uparrow$) ou soit anti-parallèle à celui-ci ($\downarrow$).

1.2.1 Portes quantiques élémentaires

En pratique, les opérateurs unitaires sont représentés par des portes quantiques. Ces portes servent à construire toutes sortes de circuits quantiques.

Les portes les plus simples sont les portes qui agissent sur un seul qubit. Parmi celles-ci, les plus souvent utilisées sont l'inversion de phase et la négation. L'effet sur un qubit de ces opérations s'exprime comme :

$$a|0\rangle + b|1\rangle \longrightarrow b|0\rangle + a|1\rangle,$$

pour la négation, et comme :

$$a|0\rangle + b|1\rangle \longrightarrow a|0\rangle - b|1\rangle,$$

pour l'inversion de phase. Sous forme matricielle, ces opérations peuvent être représentées à l'aide des *matrices de Pauli* :

$$\begin{aligned} \sigma_x &= \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \\ \sigma_y &= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \\ \sigma_z &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

Sauf si mentionné autrement, on utilise la base standard pour représenter les opérateurs sous forme matricielle. La négation correspond à σ_x, tandis que l'inversion de phase correspond à σ_z. La matrice σ_y n'a pas de signification physique particulière ; elle découle simplement de la composition d'une inversion de phase et d'une négation (à un facteur de phase global près).

Si l'on dispose seulement de portes à un qubit, il est impossible de simuler une opération générale sur plusieurs qubits. En effet, ces seules opérations ne permettent pas de faire interagir deux qubits. Pour cette raison, il est essentiel d'introduire une porte quantique à deux qubits, en l'occurrence la porte XOR :

$$\mathrm{XOR} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Cette opération est appelé le *ou excusif* puisqu'elle s'exprime comme suit dans la base standard :

$$|x\rangle|y\rangle \longrightarrow |x\rangle|x \oplus y\rangle.$$

Comme les portes précédentes à un qubit, la porte XOR est auto-inverse, i.e. la double application d'un XOR revient à appliquer l'identité. De plus, il se trouve que n'importe quelle opération unitaire peut être construite à l'aide de portes XOR et de portes à un qubit. Ce résultat est d'une grande importance technologique puisqu'il nous permet, comme dans le cas classique, de décomposer une opération complexe en un circuit constitué uniquement de portes élémentaires. Cependant, ainsi que dans le cas classique, la taille du circuit peut être exponentielle dans le nombre de qubits.

1.3 Introduction à l'intrication quantique

Dans cette section, nous allons aborder l'intrication quantique. Cette notion est centrale au développement de l'informatique quantique. En particulier, le présent travail y est intimement lié.

L'intrication quantique est une notion complexe. Cela est dû au fait qu'il semble y avoir plusieurs types d'intrication dans la nature. Nous allons donc l'introduire prudemment, en considérant d'abord le cas des états purs, et ensuite celui des mélanges statistiques. Mais avant, nous avons besoin d'un dernier outil.

1.3.1 Décomposition de Schmidt

Un état pur bipartite peut être exprimé sous une forme souvent très utile grâce au procédé de *décomposition de Schmidt*.

Pour arriver à cette forme, notons que $|\psi\rangle$ de l'expression (1.1) peut se réécrire comme :

$$|\psi\rangle = \sum_i |\mu_i\rangle \otimes |\nu_i'\rangle,$$

où on a défini :

$$|\nu_i'\rangle \equiv \sum_j a_{ij}|\nu_j\rangle.$$

Notons aussi que les $|\nu_i'\rangle$ ne sont pas nécessairement orthogonaux ou orthonormés. Supposons maintenant que la base $\{|\mu_i\rangle\}$ soit choisie comme étant la base dans laquelle ρ_A est diagonal :

$$\rho_A = \sum_i p_i|\mu_i\rangle\langle\mu_i|. \tag{1.2}$$

Nous pouvons aussi calculer ρ_A en faisant la trace partielle :

$$\begin{aligned}
\rho_A &= \mathrm{Tr}_B(|\psi\rangle\langle\psi|) \\
&= \mathrm{Tr}_B(\sum_{i,j} |\mu_i\rangle\langle\mu_j| \otimes |\nu_i'\rangle\langle\nu_j'|) \\
&= \sum_{i,j}\langle\nu_j'|\nu_i'\rangle(|\mu_i\rangle\langle\mu_j|).
\end{aligned} \tag{1.3}$$

On obtient la dernière égalité en notant que

$$\begin{aligned}
\mathrm{Tr}(|\nu_i'\rangle\langle\nu_j'|) &= \sum_k\langle\nu_k|\nu_i'\rangle\langle\nu_j'|\nu_k\rangle \\
&= \sum_k\langle\nu_j'|\nu_k\rangle\langle\nu_k|\nu_i'\rangle \\
&= \langle\nu_j'|\Big(\sum_k |\nu_k\rangle\langle\nu_k|\Big)|\nu_i'\rangle \\
&= \langle\nu_j'|\nu_i'\rangle \qquad \text{car} \qquad \sum_k |\nu_k\rangle\langle\nu_k| = I.
\end{aligned}$$

En comparant les éléments de matrice $\langle\mu_i|\rho_A|\mu_j\rangle$ calculés à partir des deux équations (1.2) et (1.3), on obtient que :

$$\langle\nu_i'|\nu_j'\rangle = p_i\delta_{ij}.$$

Ceci montre que les $\{|\nu_i'\rangle\}$ sont orthogonaux après tout ! En normalisant ces vecteurs, on obtient des vecteurs orthonormés :

$$|\tilde{\nu_i}\rangle = p_i^{-1/2}|\nu_i'\rangle,$$

ce qui permet d'écrire $|\psi\rangle$ sous la forme :

$$|\psi\rangle = \sum_i \sqrt{p_i}|\mu_i\rangle|\tilde{\nu_i}\rangle. \tag{1.4}$$

Cette équation est la décomposition de Schmidt de l'état pur bipartite $|\psi\rangle$. Tout état pur bipartite peut être exprimé sous cette forme. Évidemment, les bases orthonormées $\{|\mu_i\rangle\}$ et $\{|\tilde{\nu_j}\rangle\}$ dépendent de l'état pur considéré.

À partir de l'équation (1.4), si on fait la trace partielle sur les degrés de liberté de A, on obtient :

$$\begin{aligned}\rho_B &= \mathrm{Tr}_A(|\psi\rangle\langle\psi|)\\ &= \sum_i p_i|\tilde{\nu_i}\rangle\langle\tilde{\nu_i}|.\end{aligned}$$

On constate donc que ρ_A et ρ_B ont les mêmes valeurs propres non nulles (soit les p_i). Cependant, puisque $\mathcal{H_A}$ et $\mathcal{H_B}$ n'ont pas nécessairement la même dimension, le nombre de valeurs propres nulles de ρ_A peut différer de celui de ρ_B.

1.3.2 Intrication des états purs

Nous allons maintenant définir la notion d'intrication quantique pour un état pur bipartite. On dira qu'un état pur bipartite $|\psi\rangle$ est *séparable* (ou non intriqué) s'il peut s'écrire comme un produit tensoriel d'états purs de $\mathcal{H}_A$ et $\mathcal{H}_B$:

$$|\psi\rangle = |\phi\rangle_A \otimes |\varphi\rangle_B,$$

sinon, on dira qu'il est intriqué (ou non séparable). Dans le cas où $|\psi\rangle$ est séparable, les matrices de densités réduites $\rho_A = |\phi\rangle\langle\phi|$ et $\rho_B = |\varphi\rangle\langle\varphi|$ sont pures. Dans le cas contraire, $|\psi\rangle$ ne peut pas être exprimé par un tel produit tensoriel ; ρ_A et ρ_B sont alors des mélanges statistiques impurs.

Quantitativement, l'intrication d'un état pur, dénotée par E, peut être mesurée par l'entropie de von Neumann :

$$E(|\psi\rangle) = S(\rho_A) = S(\rho_B),$$

où $S(\rho)$ est defini de la façon suivante[3] :

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

où les λ_i sont les valeurs propres non nulles de ρ. Si ρ est pur, sa seule valeur propre non nulle est égale à 1 ce qui confirme le fait que E est bien nulle dans ce cas.

En utilisant cette formule, on trouve par exemple que l'intrication de chaque état de Bell est égale à l'unité. En effet, si ρ est une matrice de densité d'un des quatre états de Bell, on note que les matrices de densité réduites, ρ_A et ρ_B, satisfont :

$$\rho_A = \rho_B = \frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

par conséquent $\lambda_1 = \lambda_2 = 1/2$ et $S(\rho_A) = 1$.

La quantité E varie entre 0 pour un état séparable jusqu'à $\lg n$ pour un état maximalement intriqué de $\mathcal{H}_n \otimes \mathcal{H}_n$. On définit un *ebit* comme étant la quantité d'intrication d'un état à deux qubits maximalement intriqué, ou de tout autre état bipartite pour lequel $E = 1$. Voici quelques autres propriétés de E démontrées dans [4] :

- L'intrication de systèmes indépendants est additive. Par exemple, n singlets possèdent n ebits d'intrication.
- L'intrication de formation E est conservée à la suite d'opérations unitaires dites locales, i.e. décomposables sous la forme d'un produit tensoriel $U = U_A \otimes U_B$.

[3]On voit souvent $S(\rho) = -\mathrm{Tr}\rho \lg \rho$.

– L'espérance de E ne peut pas augmenter à la suite d'opérations locales non unitaires. Formellement, si un état pur bipartite $|\Gamma\rangle$ subit une opération locale non unitaire (par exemple une mesure sur $\mathcal{H}_A$) produisant les états purs résiduels $|\Gamma_i\rangle$ avec les probabilités respectives p_i, alors l'espérance de l'intrication de l'état final est inférieure ou égale à celle de l'état initial :

$$\sum_i p_i E(|\Gamma_i\rangle) \leq E(|\Gamma\rangle).$$

Mais notons cependant qu'il se peut $E(|\Gamma_i\rangle) > E(|\Gamma\rangle)$ pour un certain i. Autrement dit, l'intrication peut augmenter si on est prêt à parier !

1.3.3 Intrication des mélanges statistiques

Récemment, beaucoup de travail a été fait dans le but d'étendre la notion d'intrication aux mélanges statistiques. Un mélange statistique intriqué peut survenir lorsqu'un état pur $|\Gamma\rangle$ interagit de façon intentionnelle ou non avec d'autres degrés de liberté quantiques. Cette interaction est décrite par une évolution non unitaire de l'état pur $|\Gamma\rangle$ vers un mélange statistique ρ.

L'un des buts de ce mémoire est de montrer que les mélanges statistiques, ou les canaux bruités qui les produisent, peuvent quand même être utilisés pour transmettre de l'information quantique de façon fiable. Pour ce faire, nous allons présenter différents protocoles de purification qui peuvent être utilisés pour transmettre "parfaitement" des états quantiques via un canal bruité. Ces protocoles de purification sont le sujet du prochain chapitre. Mais avant, nous avons besoin d'établir une définition appropriée pour l'intrication d'un mélange statistique ρ.

Contrairement aux états purs, il n'existe pas un paramètre unique qui caractérise complètement l'intrication des mélanges statistiques. En effet, pour un mélange statistique arbitraire, la relation qui existe entre la quantité d'intrication que l'on peut "distiller" de cet état et de celle qui est nécessaire à sa formation n'est pas facile à déterminer[4]. Cette méconnaisance nous force à définir au moins deux mesures d'intrication pour les mélanges statistiques : l'intrication de distillation (qui sera vastement discutée dans le prochain chapitre) ; et l'intrication de formation pour laquelle nous continuons d'utiliser le symbole E.

L'intrication de formation d'un mélange statistique $E(\rho)$ est définie comme étant le minimum, pris sur toutes les collections d'états purs réalisant ρ, de l'intrication espérée d'une de ces collections. La positivité de $E(\rho)$ nous servira encore de critère de non-séparabilité. Autrement dit, ρ sera considéré non intriqué s'il peut être exprimé par une collection d'états purs séparables.

1.3.4 Définition formelle de l'intrication de formation

De façon formelle, soit $\mathcal{E} = \{p_i, |\Gamma_i\rangle\}$ une collection d'états purs telle que $\rho = \sum_i p_i |\Gamma_i\rangle\langle\Gamma_i|$, l'intrication de formation de $\mathcal{E}$ est simplement donnée par l'intrication moyenne des états purs qui composent cette collection soit $\sum_i p_i E(|\Gamma_i\rangle)$. Il est clair que certaines collections sont plus économiques que d'autres. Par exemple, un mélange statistique parfaitement aléatoire de deux qubits peut être préparé à un coût nul à partir d'un mélange égal de quatre états séparables, ou bien, à un coût unitaire, à partir d'un mélange égal des quatre états de Bell.

L'intrication de formation d'un mélange statistique quelconque ρ est donc donnée par le minimum de $E(\mathcal{E})$ pris sur toutes les collections d'états purs $\mathcal{E} = \{p_i, |\Gamma_i\rangle\}$ réalisant ρ.

Pour s'assurer que cette définition de l'intrication de formation ait du sens, il faut montrer que l'espérance de $E(\rho)$ ne peut pas augmenter suite à des opérations locales conjuguées à de la communication classique. Ceci est fait dans [4].

[4]Pour un état pur, on sait que ces deux quantités sont égales [5].

1.3.5 Borne inférieure sur l'intrication de formation

Dans cette sous-section, nous allons présenter une borne inférieure sur l'intrication de formation d'un mélange statistique ρ. Ensuite, nous allons montrer que cette borne est entre autres atteinte par deux classes d'états : les états purs et les mélanges d'états de Bell. Seulement le premier des deux résultats sera démontré.

En général, il n'est pas évident de calculer l'intrication de formation d'un état ρ. Cela provient du fait qu'il y a plusieurs façons de décomposer un mélange statistique en une collection d'états purs et, surtout, qu'il n'existe pas de moyen simple d'engendrer toutes ces décompositions. Pour cette raison, on se contente souvent d'estimer la valeur de l'intrication de formation. Une façon élégante de faire cela consiste à trouver la *fraction d'intrication maximale* de l'état considéré. En général, l'intrication de formation croît avec cette valeur.

La fraction d'intrication maximale d'un état ρ est défini comme étant le maximum de $\langle e|M|e\rangle$ pris sur tous les états parfaitement intriqués $|e\rangle$. On dénote cette quantité $f(\rho)$. Contrairement à l'intrication de formation, la fraction d'intrication maximale se calcule aisément. Il faut d'abord exprimer ρ dans une base spéciale appellée la base magique. Cette base est identique à la base de Bell à des facteurs de phase près :

$$\begin{aligned}
|e_1\rangle &= |\Phi^+\rangle, \\
|e_2\rangle &= i|\Phi^-\rangle, \\
|e_3\rangle &= i|\Psi^+\rangle, \\
|e_4\rangle &= |\Psi^-\rangle.
\end{aligned}$$

Dans cette base, il se trouve que tous les vecteurs à composantes réelles correspondent à des états parfaitement intriqués. Donc, lorsque ρ est écrite dans cette base, $f(\rho)$ correspond à la valeur maximale de $\langle e|M|e\rangle$ prise sur tous les vecteurs à composantes réelles $|e\rangle$. Cette valeur est simplement donnée par la plus grande valeur propre de la partie réelle de ρ.

On peut exprimer l'intrication de formation d'un état pur à l'aide des composantes de cet état dans la base magique. À partir de cela, il a été démontré toujours dans [4] que l'intrication de formation satisfait :

$$E(\rho) \geq h(f(\rho)),$$

où h est défini comme :

$$h(f) = \begin{cases} H(\frac{1}{2} + \sqrt{f(1-f)}) & \text{pour } f \geq \frac{1}{2} \\ 0 & \text{pour } f < \frac{1}{2}. \end{cases}$$

Ici, $H(p) = -p \lg p - (1-p) \lg(1-p)$ est la fonction d'entropie binaire.

En guise d'exercice, nous allons montrer que cette borne est atteinte pour les états purs. Nous avons vu qu'à l'aide d'opérations locales[5], il est possible de transformer tout état pur sous la forme de Schmidt :

$$|\phi\rangle = \alpha|00\rangle + \beta|11\rangle,$$

où α et β sont des nombres réels positifs qui satisfont $\alpha^2 + \beta^2 = 1$. Si on exprime $|\phi\rangle$ dans la base magique, on obtient :

$$\begin{aligned} |\phi\rangle &= \alpha \left(\frac{|e_1\rangle - i|e_2\rangle}{\sqrt{2}} \right) + \beta \left(\frac{|e_1\rangle + i|e_2\rangle}{\sqrt{2}} \right) \\ &= \tfrac{1}{\sqrt{2}}(\alpha + \beta)|e_1\rangle - \tfrac{i}{\sqrt{2}}(\alpha - \beta)|e_2\rangle, \end{aligned}$$

ou, sous forme de matrice de densité :

$$\rho = |\phi\rangle\langle\phi| = \begin{pmatrix} \frac{1}{2}(\alpha+\beta)^2 & -\frac{i}{2}(\alpha^2 - \beta^2) & 0 & 0 \\ \frac{i}{2}(\alpha^2 - \beta^2) & \frac{1}{2}(\alpha - \beta)^2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

La partie réelle de cette matrice contient seulement deux éléments diagonaux de sorte que la fraction d'intrication maximale de $|\phi\rangle\langle\phi|$, donné par la valeur propre maximale de $\mathrm{Re}\,\rho$, est simplement le maximum de ces deux valeurs, soit :

$$f(\rho) = \frac{1}{2}(\alpha + \beta)^2 = \frac{1}{2} + \alpha\beta \geq \frac{1}{2}.$$

[5] On sait que l'intrication n'est pas affectée par de telles opérations.

On peut maintenant calculer la borne sur l'intrication de formation de ρ :

$$\begin{aligned}
h(f(\rho)) &= H\left(\frac{1}{2}+\sqrt{f(\rho)(1-f(\rho))}\right) \\
&= H\left(\frac{1}{2}+\sqrt{\frac{1}{4}(\alpha+\beta)^2(\alpha-\beta)^2}\right) \\
&= H\left(\frac{1}{2}+\frac{1}{2}(\alpha+\beta)(\alpha-\beta)\right) \\
&= H\left(\frac{1}{2}+\frac{1}{2}(\alpha^2-\beta^2)\right) \\
&= H\left(\frac{1}{2}+\frac{1}{2}(\alpha^2-(1-\alpha^2))\right) \\
&= H\left(\frac{1}{2}+\frac{1}{2}(2\alpha^2-1)\right) \\
&= H(\alpha^2),
\end{aligned}$$

qui n'est autre que l'intrication de formation de ρ. En effet, puisque α et β sont les coefficients de la décomposition de Schmidt, alors α^2 et β^2 constituent les valeurs propres non nulles de $\rho_A = \mathrm{Tr}_B\rho$. On a donc :

$$E(\rho) = S(\rho_A) = H(\alpha^2) = h(f(\rho)).$$

Il est possible de montrer que cette borne est aussi atteinte pour les mélanges d'états de Bell. Nous laissons au lecteur le soin de consulter la preuve dans [4]. Un mélange d'états de Bell W peut toujours s'écrire de la façon suivante :

$$W = \sum_i p_i |e_i\rangle\langle e_i|.$$

On note que nous avons utilisé la base magique au lieu de la base de Bell pour représenter W ; cela ne fait pas de différence puisque ces deux bases diffèrent seulement par des facteurs de phase. La fraction d'intrication maximale de W est simplement le plus grand des p_i. On obtient donc :

$$E(W) = h\left(\max_i(p_i)\right).$$

1.3.6 Condition nécessaire et suffisante à l'inséparabilité d'états de sytèmes de $\mathcal{H}_2 \otimes \mathcal{H}_2$ ou de $\mathcal{H}_2 \otimes \mathcal{H}_3$

Comme on le sait, la définition que nous avons donnée à l'intrication de formation de mélanges statistiques, bien qu'elle soit très claire et parfaitement acceptable en théorie, n'est pas très utile en pratique. Pour cette raison, la découverte par Peres [11] d'une condition simple nécessaire à l'inséparabilité de mélanges statistiques constitue un résultat important. Il revient aux Horodecki [7] le mérite d'avoir démontré la suffisance de cette condition pour les états d'un système à deux qubits ou d'un système à un qubit et un qutrit[6]. Puisque les discussions de ce mémoire sont reliées à de tels systèmes, ce résultat, énoncé dans le prochain théorème, nous sera très utile.

Théorème 1 *Un mélange statistique ρ agissant sur $\mathcal{H}_2 \otimes \mathcal{H}_2$ ou sur $\mathcal{H}_2 \otimes \mathcal{H}_3$ est séparable (ou non intriqué) ssi sa transposée partielle est un opérateur positif.*

Ici, la transpostion partielle est définie comme l'opération de transposée usuelle appliquée seulement sur les degrés de liberté du deuxième sous-système. Si on suppose que $\{|\mu_i\rangle\}$ et $\{|\nu_j\rangle\}$ sont deux bases orthonormées des deux sous-systèmes et que

$$\rho_{ij,kl} = \langle \mu_i \nu_j | \rho | \mu_k \nu_l \rangle,$$

alors on définit la transposée partielle de ρ comme ceci :

$$\rho^{T_2}_{ij,kl} \equiv \rho_{il,kj}.$$

Vu sous l'angle matriciel, cette opération est très simple : elle consiste à faire la transposée des sous-matrices agissant sur le second sous-système. En effet, un état ρ d'un sytème $\mathcal{H}_M \otimes \mathcal{H}_N$ peut s'écrire de la façon suivante :

$$\rho = \begin{bmatrix} A_{11} & \dots & A_{1M} \\ \vdots & \ddots & \vdots \\ A_{M1} & \dots & A_{MM} \end{bmatrix},$$

[6]Un qutrit est un système dont l'espace d'états est $\mathcal{H}_3$.

où les A_{ij} sont des matrices $N \times N$ agissant sur $\mathcal{H}_N$. La transposée partielle de ρ est simplement donnée par :

$$\rho^{T_2} = \begin{bmatrix} A_{11}^T & \dots & A_{1M}^T \\ \vdots & \ddots & \vdots \\ A_{M1}^T & \dots & A_{MM}^T \end{bmatrix}.$$

Pour déterminer la positivité de ρ^{T_2}, on peut d'abord calculer son déterminant. S'il est négatif, alors ρ^{T_2} n'est pas un opérateur positif et par conséquent ρ est inséparable. Par contre, s'il est positif, on ne peut rien conclure ; il vaut alors la peine de calculer les valeurs propres pour tester la positivité de ρ^{T_2}.

Considérons maintenant un exemple concret. Soit la famille d'états suivante :

$$\rho = p|\psi_1\rangle\langle\psi_1| + (1-p)|\psi_2\rangle\langle\psi_2|,$$

où $|\psi_1\rangle = a|00\rangle + b|11\rangle$ et $|\psi_2\rangle = a|01\rangle + b|10\rangle$ avec $a, b > 0$ réels tels que $a^2 + b^2 = 1$. Dans ce cas, on a que :

$$\rho = \begin{bmatrix} pa^2 & 0 & 0 & pab \\ 0 & (1-p)a^2 & (1-p)ab & 0 \\ 0 & (1-p)ab & (1-p)b^2 & 0 \\ pab & 0 & 0 & pb^2 \end{bmatrix},$$

et donc :

$$\rho^{T_2} = \begin{bmatrix} pa^2 & 0 & 0 & (1-p)ab \\ 0 & (1-p)a^2 & pab & 0 \\ 0 & pab & (1-p)b^2 & 0 \\ (1-p)ab & 0 & 0 & pb^2 \end{bmatrix}.$$

On obtient par un calcul simple que le déterminant de ρ^{T_2} est égal à $-a^4b^4(2p-1)^2$. Si $ab \neq 0$ et que $p \neq \frac{1}{2}$, le déterminant est négatif et donc ρ est inséparable. Par contre, si a ou b sont nuls, $|\psi_1\rangle$ et $|\psi_2\rangle$ sont séparables et par conséquent ρ l'est.

Finalement, pour $p = \frac{1}{2}$, on a que $\rho^{T_2} = \rho$ ce qui nous assure de la positivité de ρ^{T_2}. Dans ce cas, la collection d'états purs séparables qui réalisent ρ n'est pas triviale. Elle est constituée de deux états équiprobables :

1. $|\phi_1\rangle = (a|0\rangle + b|1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle + |1\rangle)$,
2. $|\phi_2\rangle = (a|0\rangle - b|1\rangle) \otimes \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)$.

Il est facile de vérifier que $\rho = \frac{1}{2}|\phi_1\rangle\langle\phi_1| + \frac{1}{2}|\phi_2\rangle\langle\phi_2|$.

1.4 Téléportation quantique

Nous allons maintenant discuter du schème de téléportation quantique introduit pour la première fois dans [1]. Ce schème indique comment deux parties A et B peuvent se transmettre un état quantique arbitraire inconnu même s'ils ne disposent pas de canal quantique lors de la transmission. L'intrication quantique joue un rôle clé dans l'accomplissement de cette tâche. Nous allons présenter ici une version légèrement modifiée du scheme original. Le résultat final est tout à fait équivalent.

Il est commun d'associer au système A une personne appelée Alice. De même, on associe au système B une personne appelée Bob. Alice et Bob ont le droit d'appliquer n'importe quelle opération sur leurs systèmes respectifs. De plus, ils disposent d'un canal classique qui leur permet de communiquer autant d'information classique qu'ils le désirent.

Supposons qu'Alice veuille transmettre à Bob un qubit arbitraire $|\psi\rangle = a|0\rangle + b|1\rangle$. Si Alice connait les coefficients a et b, elle peut simplement les communiquer classiquement à Bob qui, par la suite, peut reproduire $|\psi\rangle$ dans son propre système. Cependant, si Alice ne connait pas ces coefficients, il semble impossible d'accomplir correctement la tâche de téléportation. Cela peut être fait si Alice et Bob partagent au préalable un état de deux qubits parfaitement intriqué tel $|\Phi^+\rangle$. Initialement, Alice et Bob disposent donc de l'état global :

$$\begin{aligned}|\psi\rangle_T \otimes |\Phi^+\rangle_{AB} &= (a|0\rangle_T + b|1\rangle_T) \otimes \frac{1}{\sqrt{2}}(|0\rangle_A \otimes |0\rangle_B + |1\rangle_A \otimes |1\rangle_B) \\ &= \frac{1}{\sqrt{2}}(a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle),\end{aligned}$$

où l'indice T identifie la particule à téléporter, tandis que les indices A et B identifient les deux particules formant l'état $|\Phi^+\rangle$. Dans la dernière équation, ces indices ont été omis ; par convention, on conserve l'ordre déjà établi soit T, A, B.

La première étape du scheme consiste pour Alice à faire une mesure orthogonale dans la base de Bell sur ses deux qubits (ceux indicés par T et A). Les applications respectives des quatre projecteurs de Bell sur l'état initial mènent aux quatre expressions suivantes :

1.
$$\begin{aligned}
&(|\Phi^+\rangle\langle\Phi^+|)_{TA}\Big(|\psi\rangle_T \otimes |\Phi^+\rangle_{AB}\Big)\\
&= |\Phi^+\rangle_{TA}\tfrac{1}{\sqrt{2}}(\langle 00| + \langle 11|)_{TA}\tfrac{1}{\sqrt{2}}(a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle)\\
&= \tfrac{1}{2}|\Phi^+\rangle_{TA}\Big(a|0\rangle_B + b|1\rangle_B\Big)\\
&= \tfrac{1}{2}|\Phi^+\rangle_{TA} \otimes |\psi\rangle_B.
\end{aligned}$$

2.
$$\begin{aligned}
&(|\Psi^+\rangle\langle\Psi^+|)_{TA}\Big(|\psi\rangle_T \otimes |\Phi^+\rangle_{AB}\Big)\\
&= |\Psi^+\rangle_{TA}\tfrac{1}{\sqrt{2}}(\langle 01| + \langle 10|)_{TA}\tfrac{1}{\sqrt{2}}(a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle)\\
&= \tfrac{1}{2}|\Psi^+\rangle_{TA}\Big(a|1\rangle_B + b|0\rangle_B\Big)\\
&= \tfrac{1}{2}|\Psi^+\rangle_{TA} \otimes \sigma_x|\psi\rangle_B.
\end{aligned}$$

3.
$$\begin{aligned}
&(|\Phi^-\rangle\langle\Phi^-|)_{TA}\Big(|\psi\rangle_T \otimes |\Phi^+\rangle_{AB}\Big)\\
&= |\Phi^-\rangle_{TA}\tfrac{1}{\sqrt{2}}(\langle 00| - \langle 11|)_{TA}\tfrac{1}{\sqrt{2}}(a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle)\\
&= \tfrac{1}{2}|\Phi^-\rangle_{TA}\Big(a|0\rangle_B - b|1\rangle_B\Big)\\
&= \tfrac{1}{2}|\Phi^-\rangle_{TA} \otimes \sigma_z|\psi\rangle_B.
\end{aligned}$$

4.
$$\begin{aligned}
&(|\Psi^-\rangle\langle\Psi^-|)_{TA}\Big(|\psi\rangle_T \otimes |\Phi^+\rangle_{AB}\Big)\\
&= |\Psi^-\rangle_{TA}\tfrac{1}{\sqrt{2}}(\langle 01| - \langle 10|)_{TA}\tfrac{1}{\sqrt{2}}(a|000\rangle + a|011\rangle + b|100\rangle + b|111\rangle)\\
&= \tfrac{1}{2}|\Psi^-\rangle_{TA}\Big(a|1\rangle_B - b|0\rangle_B\Big)\\
&= \tfrac{-i}{2}|\Psi^-\rangle_{TA} \otimes \sigma_y|\psi\rangle_B.
\end{aligned}$$

Suite à cette mesure, Alice envoie deux bits classiques à Bob pour lui indiquer lequel des quatre résultats possibles elle a observé. Selon le résultat que Bob reçoit, celui-ci décide d'appliquer sur sa particule soit :

1. rien du tout.
2. σ_x.
3. σ_z.
4. σ_y.

Puisque les matrices de Pauli sont auto-inverses, Bob retrouvera donc sans erreur l'état $|\psi\rangle$ qu'Alice a perdu. On note que si Alice n'envoie pas le résultat de la mesure à Bob, l'état de la particule de Bob demeure :

$$\frac{1}{4}\begin{pmatrix} a^2 & ab \\ ab & b^2 \end{pmatrix}+\frac{1}{4}\begin{pmatrix} b^2 & ab \\ ab & a^2 \end{pmatrix}+\frac{1}{4}\begin{pmatrix} a^2 & -ab \\ -ab & b^2 \end{pmatrix}+\frac{1}{4}\begin{pmatrix} b^2 & -ab \\ -ab & a^2 \end{pmatrix}=\frac{1}{2}\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

soit l'état d'un qubit aléatoire. Cela n'est pas surprenant car si Bob pouvait apprendre quelque chose sur $|\psi\rangle$, il aurait obtenu de l'information de façon supraluminique ce qui violerait le principe de causalité.

Chapitre 2

Protocoles de purification et intrication de distillation

Les *protocoles de purification* constituent le sujet principal de ce mémoire. Plusieurs exemples de ces protocoles seront discutés au cours des prochaines sections. Mais avant, voyons dans quel contexte les protocoles de purification apparaissent.

Supposons que deux parties, que l'on continuera d'appeler Alice et Bob, désirent échanger de l'information. Étant éloignés l'un de l'autre, les seuls moyens dont ils disposent pour communiquer sont : un canal quantique bruité[1] et un canal classique. Malgré qu'il leur permette d'échanger autant d'information classique qu'ils le désirent, en l'absence d'intrication ce canal classique est impuissant à transmettre un qubit inconnu arbitraire $|\Psi\rangle$ de façon fidèle.

Supposons maintenant qu'Alice veuille transmettre sans erreur un message de m qubits à Bob. On sait que s'ils partageaient m singlets, ils pourraient les utiliser via le procédé de téléportation pour transmettre le message sans erreur. Une stratégie possible pour arriver à cette fin est la suivante : Alice prépare dans son laboratoire un grand nombre n de singlets et envoie ensuite une des particules de chaque singlet à Bob par le canal quantique. L'effet du canal bruité sera d'incorporer des erreurs dans l'état

[1] Un canal qui transmet des qubits avec fidélité imparfaite.

des particules transmises. À la suite de la transmission, Alice et Bob ne partagent donc pas n singlets, mais plutôt n paires de particules décrites par un mélange statistique ρ. S'ils ont de telles paires en quantité suffisante, ils peuvent arriver, par l'intermédiaire des protocoles de purification, à obtenir $m < n$ singlets asymptotiquement parfaits dans la limite où n tend vers l'infini. Le développement des protocoles de purification est donc intimement lié à la correction d'erreurs suite à une transmission par un canal quantique bruité ; voilà une raison importante qui motive leur étude. Pour commencer, présentons une description générale de leur fonctionnement.

Initialement, Alice et Bob partagent n paires de particules. Chaque paire est décrite par la même matrice de densité ρ. Le produit tensoriel de ces n paires est l'état $\boldsymbol{\rho} = (\rho)^n$. Alice et Bob procèdent en répétant trois étapes :

- Appliquer des transformations unitaires sur des sous-ensembles de leurs particules respectives (ceci exclut la possibilité d'utiliser des opérations quantiques globales sur tout le système).
- Effectuer des mesures sur des sous-ensembles de leurs particules respectives (encore une fois, il s'agit d'opérations locales).
- Echanger par voie de communication classique les résultats des mesures afin de déterminer la prochaine étape à accomplir.

Le but du protocole est d'arriver à *distiller* le plus de singlets possibles de l'état de départ $\boldsymbol{\rho}$, i.e. de sacrifier certaines particules de façon à transformer l'état des autres paires non mesurées en singlets $|\Psi^-\rangle$, ou du moins en une bonne approximation. Si le nombre de singlets obtenus est m $(0 \leq m \leq n)$, le *rendement* du protocole sera donné par le rapport m/n. Bien entendu, pour un état de départ ρ donné, le rendement d'un protocole $R(\rho)$ constituera un critère important dans la comparaison des différents protocoles.

L'échange qui s'effectue à la troisième étape peut être fait de deux façons : soit par une communication *unidirectionnelle*, ou soit par une communication *bidirectionnelle*. La première permet à Alice, mais pas à Bob, d'envoyer de l'information classique à son partenaire ; tandis que la seconde, moins restrictive, alloue cette permission aux deux partenaires. Cette distinction définit en fait deux types de protocoles. Les protocoles de purification bidirectionnels constituent les protocoles les plus généraux et seront dénotés par l'abbréviation PP2. Les protocoles de purification unidirectionnels (PP1), quoique moins puissants, sont d'une grande importance puisqu'ils permettent la création de codes correcteurs quantiques [4].

Nous sommes maintenant prêts à définir deux quantités importantes reliées à l'intrication d'un état ρ. L'*intrication de distillation unidirectionnelle* $D_1(\rho)$ est le rendement maximal que l'on peut obtenir à l'aide d'un PP1 avec ρ comme état de départ, dans la limite où n tend vers l'infini. L'*intrication de distillation bidirectionnelle* $D_2(\rho)$ est définie de façon analogue en remplacant PP1 par PP2 dans la définition précédente. Ces deux notions sont définies au sens asymptotique, i.e. le nombre n de paires qui forment l'état ρ tend vers l'infini. Le rendement est donc donné par $\lim_{n\to\infty} m/n$. De plus, on n'exige pas que les m paires résultantes soient des singlets parfaits ; on demande plutôt que chacune de ces paires, décrite par la matrice de densité T, satisfasse $lim_{n\to\infty}\langle\Psi^-|T|\Psi^-\rangle = 1$, i.e. que la fidélité des paires produites par rapport au singlet tende vers l'unité lorsque le nombre initial de paires tend vers l'infini.

De ces deux définitions découle la double inégalité $D_1(\rho) \leq D_2(\rho) \leq E(\rho)$. La première provient du fait qu'un PP1 est un PP2, tandis que la seconde nous est imposée par une propriété de l'intrication de formation. En effet, si Alice et Bob pouvaient obtenir plus que $E(\rho)$ singlets par paire, ils auraient augmenté l'intrication de formation par des opérations locales, ce qui contredit une propriété fondamentale de l'intrication de formation.

Si ρ est un état pur, les trois quantités $D_1(\rho), D_2(\rho), E(\rho)$ sont égales [5] et aucune distinction n'est à faire entre l'intrication de formation et les intrications de distillation. Cependant, si ρ n'est pas un état pur (c'est le cas qui nous intéresse ici), la relation entre ces trois quantités est beaucoup moins évidente. Il se pourrait par exemple que $D_2 = E$ pour tous les états d'un système 2×2. On sait par contre que D_1 et D_2 sont différents pour certains de ces états.

2.1 Quelques outils de base

Avant d'aborder le premier exemple de protocole de purification, nous allons introduire quelques notions qui nous serviront à développer ces protocoles.

2.1.1 Les états de Werner

Un protocole de purification est conçu en fonction d'un certain état ρ dont on veut extraire le plus d'intrication possible. Ceci pourrait laisser croire qu'il faille développer un nouveau protocole pour chaque état ρ différent. Heureusement, tel n'est pas le cas ; beaucoup de protocoles de purification se révèlent efficaces sur un grand nombre d'états caractéristiques. Pour cette raison, il est utile de considérer une famille d'états particulière : les états de Werner. Ceux-ci sont paramétrisés par F, une valeur réelle comprise entre 0 et 1 que l'on nomme fidélité :

$$W_F = F|\Psi^-\rangle\langle\Psi^-| + \frac{1-F}{3}\Big(|\Psi^+\rangle\langle\Psi^+| + |\Phi^+\rangle\langle\Phi^+| + |\Phi^-\rangle\langle\Phi^-|\Big).$$

Les états de Werner font partie de la classe plus large des mélanges d'états de Bell. Ceux-ci seront habituellement dénotés par W sans indice. Puisque que les mélanges d'états de Bell atteignent la borne sur l'intrication de formation présentée à la section (1.3.5), on a que :

$$E(W_F) = h\left(\max\left(F, \frac{1-F}{3}\right)\right),$$

en simplifiant, on obtient

$$E(W_F) = \begin{cases} H(1/2 + \sqrt{F(1-F)}) & \text{si } F > 1/2 \\ 0 & \text{sinon.} \end{cases}$$

Donc si $F \leq 1/2$, on a que $E(W_F) = 0$ et, par conséquent, $D_1(W_F) = D_2(W_F) = 0$. On ne peut appliquer aucun protocole de purification à un tel état pour obtenir des singlets purs. Par contre si $F > 1/2$, W_F est potentiellement "purifiable" puisque son intrication de formation $E(W_F)$ est non nulle.

Les états de Werner sont importants pour plusieurs raisons. D'abord, ce type d'états peut être produit en mélangeant une quantité quelconque de singlets avec un mélange de deux qubits dans des états parfaitement aléatoires. C'est exactement le résultat que l'on obtient en envoyant un $|\Psi^-\rangle$ dans un canal de dépolarisation. Un canal de dépolarisation-x est un canal par lequel un état est transmis sans erreur avec probabilité $1-x$ et est remplacé par un mélange statistique de qubits aléatoires avec probabilité x. L'état qui décrit le résultat d'une transmission de $|\Psi^-\rangle$ via un canal de dépolarisation-x est donné par :

$$\rho_x = (1-x)|\Psi^-\rangle\langle\Psi^-| + \frac{x}{4}\cdot I_4$$

où $\frac{1}{4}I_4$ est la matrice de densité qui représente un mélange statistique de deux qubits "aléatoires". Il est facile de voir que I_4 peut s'écrire de la façon suivante :

$$I_4 = |\Psi^-\rangle\langle\Psi^-| + |\Psi^+\rangle\langle\Psi^+| + |\Phi^+\rangle\langle\Phi^+| + |\Phi^-\rangle\langle\Phi^-|.$$

On obtient donc que :

$$\rho_x = \Big(1 - \frac{3}{4}x\Big)|\Psi^-\rangle\langle\Psi^-| + \frac{x}{4}\Big(|\Psi^+\rangle\langle\Psi^+| + |\Phi^+\rangle\langle\Phi^+| + |\Phi^-\rangle\langle\Phi^-|\Big).$$

Cet état n'est rien d'autre que $W_{1-\frac{3}{4}x}$. Ceci permet d'établir la relation $F = 1 - \frac{3}{4}x$. Pour le canal de dépolarisation-50%, il résulte de la transmission d'un singlet parfait l'état de Werner le plus connu :

$$W_{5/8} = \frac{5}{8}|\Psi^-\rangle\langle\Psi^-| + \frac{1}{8}\Big(|\Psi^+\rangle\langle\Psi^+| + |\Phi^+\rangle\langle\Phi^+| + |\Phi^-\rangle\langle\Phi^-|\Big).$$

Cet état a la propriété que $D_1(W_{5/8}) = 0 < D_2(W_{5/8}) \leq E(W_{5/8}) = 0.1176$. L'inégalité stricte sera démontrée dans la prochaine section, tandis que la première égalité est déduite de jolie façon dans [4]. Elle découle aussi d'un résultat plus fort de Knill et Laflamme [10], à savoir que $D_1(W_F) = 0$ lorsque $F < 3/4$.

En plus de résulter de la transmission d'un singlet par un canal dépolarisant, l'état de Werner peut être préparé à partir de n'importe quel mélange statistique ρ, et cela en conservant la fraction d'intrication maximale de ρ (sans cette contrainte, l'énoncé serait vrai de façon triviale car on peut toujours préparer l'état $W_{1/4}$). Le procédé, que l'on nomme “twirling”, par lequel on convertit un mélange statistique ρ quelconque, avec $F = f(\rho)$, dans l'état W_F est décrit à l'annexe A. Il est important de noter que ce procédé est irréversible et donc non unitaire. Par conséquent, il en résulte possiblement un gain d'entropie qui s'accompagne, dans la plupart des cas, d'une perte d'intrication de formation. En effet :

$$E(W_F) = h(f(W_F)) = h(F) = h(f(\rho)) \leq E(\rho)$$

La première égalité ainsi que l'inégalité découlent de la section 1.3.5 tandis que $f(W_F) = F$ car $F \geq 1/4$. En général, cette perte est relativement faible par rapport à l'intrication totale de ρ puisque sa fraction d'intrication maximale est conservée.

Dans la plupart des protocoles, on exige seulement que l'état ρ soit mis sous la forme d'un mélange d'états de Bell (dont l'état de Werner est un cas particulier). Une fois que l'état ρ a été rendu sous cette forme, on peut le soumettre aux protocoles conçus pour des mélanges d'états de Bell. Par la suite, on n'a plus besoin de se préoccuper de l'état de départ ρ ou du canal bruité (dépolarisant ou autre) qui l'a généré ; cela s'avère très pratique dans le développement des protocoles de purification.

2.1.2 Transformations unitaires utiles

Vu l'importance des mélanges d'états de Bell, il est important de prendre connaissance des effets des opérations unitaires locales sur ces états. Il se trouve que certaines de ces opérations, lorsqu'appliquées sur les états de Bell, agissent comme de simples permutations sur des sous-ensembles de ces états. Les transformations unitaires présentées ici appartiennent à ce groupe.

On distinguera deux types d'opérations : les opérations unilatérales qui sont effectuées par Alice ou Bob (mais pas les deux), et les opérations bilatérales qui sont effectuées par Alice et Bob à la manière d'un produit tensoriel $U_A \otimes U_B$. Dans la plupart des protocoles de purification, on utilise trois ensembles d'opérations. Le premier comprend les rotations unilatérales de π radians qui correspondent aux trois matrices de Pauli σ_x, σ_y et σ_z. Le second ensemble est constitué des rotations bilatérales de $\pi/2$ radians que l'on dénote B_x, B_y et B_z. Enfin le dernier comprend une seule opération qui est le *ou-exculsif bilatéral* (ou BXOR). Cette opération consiste, pour Alice, à appliquer le XOR sur deux de ses particules, et pour Bob, à faire de même sur ses particules respectives. Le BXOR est la seule opération parmi ce groupe qui s'applique à deux paires distinctes ; les autres s'appliquent seulement sur l'un ou les deux membres d'une paire unique. L'effet de ces opérations est exposé dans le tableau 2.1.

2.1.3 Mesures locales

En plus d'effectuer des opérations unitaires locales sur leurs paires impures, Alice et Bob devront mesurer certaines paires afin d'obtenir l'information qui leur permettra de purifier les paires restantes non mesurées. Or, il est important de savoir ce qu'ils peuvent apprendre d'une mesure locale, i.e. une mesure appliquée non pas sur le système global, mais effectuée séparément dans les laboratoires d'Alice et Bob. Ils peuvent sûrement mesurer dans la base standard pour obtenir un des quatre résultats $|00\rangle$, $|01\rangle$, $|10\rangle$ ou $|11\rangle$. Ils ne peuvent cependant pas effectuer une mesure dans la base intriquée des états de Bell. Cela est évident puisqu'à la suite d'une mesure appliquée sur $|00\rangle$ par exemple, ils disposeraient d'une paire parfaitement intriquée ($|\Phi^+\rangle$ ou $|\Phi^-\rangle$) créant ainsi de l'intrication de façon locale.

Par contre, ils peuvent différencier un état $|\Phi\rangle$ d'un état $|\Psi\rangle$. Pour ce faire, ils peuvent simplement mesurer la paire dans la base standard et échanger ensuite les resultats de leurs mesures. Si les deux résultats sont identiques, il s'agissait d'un $|\Phi\rangle$, sinon il s'agissait d'un $|\Psi\rangle$. De plus, en raison des opérations unitaires décrites plus haut, n'importe quel sous-ensemble de deux états de Bell peut être distingué localement de l'ensemble formé des deux autres. Cependant, après une telle mesure, on perd toute autre information concernant l'état mesuré. Autrement dit, une telle mesure retire le

Rotations unilatérales :

	source			
	Ψ^-	Φ^-	Φ^+	Ψ^+
I	Ψ^-	Φ^-	Φ^+	Ψ^+
σ_x	Φ^-	Ψ^-	Ψ^+	Φ^+
σ_y	Φ^+	Ψ^+	Ψ^-	Φ^-
σ_z	Ψ^+	Φ^+	Φ^-	Ψ^-

Rotations bilatérales :

	source			
	Ψ^-	Φ^-	Φ^+	Ψ^+
I	Ψ^-	Φ^-	Φ^+	Ψ^+
B_x	Ψ^-	Φ^-	Ψ^+	Φ^+
B_y	Ψ^-	Ψ^+	Φ^+	Φ^-
B_z	Ψ^-	Φ^+	Φ^-	Ψ^+

Ou-exclusif bilatéral :

	source				
cible	Ψ^-	Φ^-	Φ^+	Ψ^+	
	Ψ^+	Φ^+	Φ^-	Ψ^-	(source)
Ψ^-	Φ^-	Ψ^-	Ψ^-	Φ^-	(cible)
	Ψ^+	Φ^+	Φ^-	Ψ^-	(source)
Φ^-	Ψ^-	Φ^-	Φ^-	Ψ^-	(cible)
	Ψ^-	Φ^-	Φ^+	Ψ^+	(source)
Φ^+	Ψ^+	Φ^+	Φ^+	Ψ^+	(cible)
	Ψ^-	Φ^-	Φ^+	Ψ^+	(source)
Ψ^+	Φ^+	Ψ^+	Ψ^+	Φ^+	(cible)

TAB. 2.1: Rotations unilatérales et bilatérales qu'Alice et Bob utilisent pour transformer les états de Bell en d'autres états de Bell. Chaque entrée du tableau du BXOR a deux lignes, la première indique l'état résultant de la paire source, la seconde indique l'état résultant de la paire cible.

maximum d'information, i.e un bit sur deux.

Dans ce qui suit, nous emploierons la notation :

$$|\Phi^+\rangle = 00,$$
$$|\Psi^+\rangle = 01,$$
$$|\Phi^-\rangle = 10,$$
$$|\Psi^-\rangle = 11.$$

Le bit de droite identifie la propriété Φ/Ψ des états de Bell. Cette propriété est appelée l'*amplitude*. Le bit de gauche identifie la propriété $+/-$ des états de Bell. Celle-ci est appelée la *phase*. Comme autre convention, nous utiliserons $|\Phi^+\rangle$, au lieu de $|\Psi^-\rangle$, comme étant l'état dominant de ρ (celui affecté de la probabilité F). C'est par rapport à cet état que nous tenterons d'augmenter la fidélité. Ce choix est dû au fait que $|\Phi^+\rangle$ reste inchangé lorsqu'il est utilisé comme source et cible du BXOR. De toutes façons, on pourra toujours ramener l'état sous forme de Werner par après. Nous disposons maintenant de tous les outils qu'il faut pour aborder un premier protocole de purification.

2.2 Méthode de récurrence

Le protocole que nous allons décrire ci-dessous est un PP2 proposé pour la première fois par Bennett, Brassard, Popescu, Schumacher, Smolin et Wooters dans [2]. Ce protocole, appelée la méthode de récurrence, constitue le premier protocole de purification jamais réalisé. Il a été inspiré de deux protocoles développés dans le domaine de la cryptographie [3][6]. Le premier des deux est d'une importance particulière puisqu'il constitue l'analogue classique de la méthode de récurrence.

Nous allons maintenant débuter la description de la méthode de récurrence. Ce protocole est conçue pour travailler avec des paires tirées d'un mélange statistique ρ diagonal dans la base de Bell (pas nécessairement sous la forme de Werner). L'état des paires initiales est donc complètement décrit par le quadruplet $(p_{00}, p_{01}, p_{10}, p_{11})$ qui correspond aux éléments diagonaux de ρ lorsqu'exprimé dans la base de Bell. Tel qu'argumenté plus haut, on pose $p_{00} = F$.

La première étape du protocole consiste, pour Alice et Bob, à grouper leurs paires deux à deux pour ensuite appliquer un BXOR sur la paire source et la paire cible de chacun de ces groupes (voir figure 2.1).

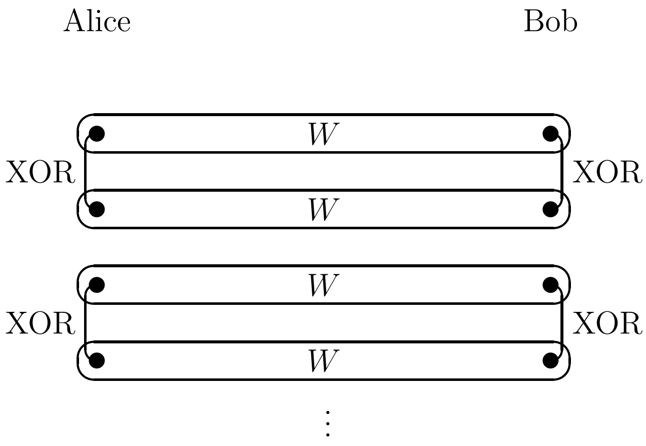

FIG. 2.1: Application des BXORs par Alice et Bob.

Après avoir fait cela, ils mesurent (toujours de façon locale) l'amplitude de chaque paire cible résultante, rendant ainsi aléatoire la phase de chacune d'elles. Selon le résultat de la mesure (0 ou 1), ils obtiennent deux mélanges statistiques différents pour les nouvelles paires sources. Ces deux mélanges statistiques demeurent diagonaux dans la base de Bell (voir tableau 2.2).

Commencons par analyser le mélange statistique qui découle de l'obtention du résultat 1. Celui-ci est décrit par les nouvelles probabilités :

$$p'_{00} = p'_{01} = (p_{00}p_{01} + p_{10}p_{11})/(1 - p_{pass})$$
$$p'_{10} = p'_{11} = (p_{00}p_{11} + p_{01}p_{10})/(1 - p_{pass})$$

avec

$$p_{pass} = (p_{00} + p_{10})^2 + (p_{01} + p_{11})^2,$$

	initial		après BXOR		Résultat
Probabiltité	S	C	S	C	de la mesure
p_{00}^2	00	00	00	00	S
$p_{00}p_{01}$	00	01	00	01	E
$p_{00}p_{10}$	00	10	10	10	S
$p_{00}p_{11}$	00	11	10	11	E
$p_{01}p_{00}$	01	00	01	01	E
p_{01}^2	01	01	01	00	S
$p_{01}p_{10}$	01	10	11	11	E
$p_{01}p_{11}$	01	11	11	10	S
$p_{10}p_{00}$	10	00	10	00	S
$p_{10}p_{01}$	10	01	10	01	E
p_{10}^2	10	10	00	10	S
$p_{10}p_{11}$	10	11	00	11	E
$p_{11}p_{00}$	11	00	11	01	E
$p_{11}p_{01}$	11	01	11	00	S
$p_{11}p_{10}$	11	10	01	11	E
p_{11}^2	11	11	01	10	S

TAB. 2.2: Résultat de la mesure du bit amplitude de la paire cible suite au BXOR. S indique un succès (le bit amplitude de la cible est 0) et E indique un échec (ce même bit est 1).

où p_{pass} représente la probabilité que le résultat de la mesure soit 0. Il est facile de voir que chacun de ces nouveaux éléments diagonaux est inférieur ou égal à 1/2. Soit ρ' la nouvelle matrice de densité du second mélange statistique, alors on a que $E(\rho') = h(f(\rho')) = 0$ car $f(\rho') \leq 1/2$. Puisque l'intrication de formation de cet état est nulle, il est impossible d'en retirer la moindre parcelle d'intrication et, par conséquent, les paires issues de ce mélange statistique sont simplement mises de côté.

Quant au second mélange statistique, découlant de l'obtention du résultat 0, il est décrit par les quatre nouveaux éléments diagonaux :

$$\begin{aligned} p'_{00} &= (p_{00}^2 + p_{10}^2)/p_{pass} & \qquad p'_{01} &= (p_{01}^2 + p_{11}^2)/p_{pass} \\ p'_{10} &= 2p_{00}p_{10}/p_{pass} & \qquad p'_{11} &= 2p_{01}p_{11}/p_{pass} \end{aligned}$$

Ce sont les paires issues de ce mélange qu'Alice et Bob vont garder. Il est important de noter que c'est précisément ici que le protocole requiert une communication bidirectionnelle. Alice et Bob doivent chacun connaître le résultat de l'autre afin d'éliminer les "mauvaises" paires.

Maintenant, analysons ce dernier mélange statistique. On voudrait bien que $p'_{00} > p_{00}$ pour ainsi obtenir un mélange d'états de Bell avec une fidélité supérieure à celle du mélange de départ. Cela nous permettrait de répéter le procédé pour amener la fidélité des paires arbitrairement proche de 1 (ce qui correspond à obtenir des singlets asymptotiquement parfaits). Mais quelles conditions p_{00}, p_{01}, p_{10} et p_{11} doivent-ils satisfaire pour que $p'_{00} > p_{00}$? L'analyse que nous allons présenter ici est entièrement originale. Elle procure une réponse simple et complète à la question. Pour en arriver là, considérons d'abord la fonction :

$$p'_{00} = f(p_{00}, p_{01}, p_{10}, p_{11}) = \frac{(p_{00}^2 + p_{10}^2)}{(p_{00} + p_{10})^2 + (p_{01} + p_{11})^2},$$

sous les contraintes $p_{00} + p_{01} + p_{10} + p_{11} = 1$ et $0 \leq p_i \leq 1$. Rappelons que si la fraction d'intrication maximale est inférieure ou égale à 1/2, l'intrication de formation du mélange statistique est nulle et tout ce procédé ne sert à rien. Sans perte de généralité, on peut donc supposer que $p_{00} > 1/2$. En se servant des contraintes, on peut réécrire

p'_{00} comme une fonction à deux arguments :

$$p'_{00} = g(p_{00}, p_{10}) = \frac{(p_{00}^2 + p_{10}^2)}{(p_{00} + p_{10})^2 + (1 - p_{00} - p_{10})^2},$$

avec $1/2 < p_{00} \leq 1$ et $0 \leq p_{10} \leq 1 - p_{00}$. En calculant la dérivée partielle de g par rapport à p_{10}, on obtient le résultat suivant (voir Annexe B) :

$$\frac{\partial g(p_{00}, p_{10})}{\partial p_{10}} < 0.$$

Cela implique que si l'on fixe p_{00}, la valeur maximale de p'_{00} est atteinte lorsque $p_{10} = 0$ et la valeur minimale est atteinte lorsque $p_{10} = 1 - p_{00}$. Ces deux valeurs sont :

$$\begin{aligned} p'_{00_{max}} &= g(p_{00}, 0) &&= \frac{p_{00}^2}{p_{00}^2 + (1 - p_{00})^2}, \\ p'_{00_{min}} &= g(p_{00}, 1 - p_{00}) &&= p_{00}^2 + (1 - p_{00})^2. \end{aligned}$$

Il est facile de voir que $p'_{00_{min}} < p_{00} < p'_{00_{max}}$ si $p_{00} > 1/2$. Comme démontré à l'Annexe B, la valeur p_{eqv} qui satisfait $g(p_{00}, p_{eqv}) = p_{00}$ est donnée par l'expression :

$$p_{eqv} = \sqrt{p_{00}}(1 - \sqrt{p_{00}}).$$

Ceci nous permet de conclure que $\boldsymbol{p'_{00} > p_{00}}$ **ssi** $\boldsymbol{p_{10} < p_{eqv}}$. Par exemple, si l'état est initialement sous la forme de Werner, alors $p'_{00} > p_{00}$ car $p_{10} = \frac{1-p_{00}}{3} < p_{eqv}$ si $p_{00} > 1/2$. Si l'état n'est pas sous la forme de Werner, c'est encore mieux ! En effet, il existe alors un i tel que $p_i < \frac{1-p_{00}}{3}$. Par conséquent, il suffit de choisir la bonne rotation pour faire en sorte que p_{10} soit égal à p_i, obtenant ainsi une meilleur amélioration de la fidélité que dans le cas où l'état était sous la forme de Werner.

La stratégie à adopter pour augmenter la fidélité au-delà d'un certain seuil dépend donc du mélange statistique dont on dispose au départ. Sans perte de généralité, on suppose que ce mélange est diagonal dans la base de Bell (pas nécessairement sous forme de Werner) avec $F = p_{00} > 1/2$. Si une des trois autres composantes est nulle (on va rencontrer ce genre d'états plus tard), la stratégie est simple : utiliser la bonne rotation pour que p_{10} devienne nul et appliquer le procédé décrit plus haut. On a vu que dans ce cas, la hausse de fidélité est optimale. De plus, puisque $p'_{10} = 0$ lorsque $p_{10} = 0$, le nouveau mélange est prêt à subir une nouvelle amélioration de sa fidélité en réappliquant le procédé, et cela, autant de fois qu'il le faut pour atteindre la fidélité voulue.

Si le mélange initial est sous la forme de Werner, on a vu qu'une application du procédé va augmenter la fidélité. Cependant, le nouveau mélange issu de cette première application ne respecte pas la condition $p_{10} < p_{eqv}$ (ce qui veut dire qu'une seconde application du procédé mènerait à une détérioration de la fidélité). Deux stratégies ont été suggérées pour pallier à ce problème. La première consiste à remettre, par l'opération de "twirling", le nouveau mélange sous la forme de Werner. Ce nouvel état de Werner est ensuite prêt à subir une hausse de sa fidélité par une autre application du procédé. La deuxième stratégie, découverte par Macchiavello, utilise une rotation déterministe $\sigma_z B_x \sigma_z$ qui a pour effet d'échanger p_{10} et p_{11}. À la suite de ce changement, on remarque que $p_{10} < \frac{1-p_{00}}{3}$. Lors de la prochaine itération, la fidélité sera donc haussée davantage que si l'état avait été mis sous la forme de Werner. Machiavello a remarqué qu'après chaque application subséquente du procédé, la composante la plus petite du mélange résultant est toujours p_{11} ; la rotation $\sigma_z B_x \sigma_z$ demeure donc la meilleure à utiliser.

Cependant, même dans le meilleur cas, la méthode de récurrence est plutôt inefficace en ce qui a trait au rendement. À chaque itération, plus de la moitié des paires sont sacrifiées (toutes les paires cibles en plus des paires sources qui tombent dans le mauvais ensemble). Plus précisément, la proportion de paires restantes après une itération est égale à $p_{pass}/2$. Puisqu'une fidélité asymtotiquement parfaite est seulement atteinte si le nombre de d'itérations tend vers l'infini, il est clair que le rendement d'un tel protocole est nul.

Toutefois, un rendement positif peut être obtenu si, après avoir augmenté la fidélité au-delà d'un certain seuil avec la méthode de récurrence, on utilise un autre protocole de purification nous fournissant un rendement positif sur l'état résultant. Un tel protocole est décrit à la section 2.6 ; il s'agit de la *méthode de hashing*, un PP1 qui possède un rendement positif sur les mélanges d'états de Bell d'entropie inférieure à 1. Le protocole général, que l'on pourrait appeler la méthode de récurrence-hashing, consiste à utiliser, l'une à la suite de l'autre, la méthode de récurrence et la méthode de hashing. Le moment de la transition doit être choisi de façon judicieuse afin de maximisier le rendement du protocole global. Dans le but d'alléger la lecture, la méthode de récurrence-hashing sera appelé la méthode de récurrence. De toute façon, la méthode de récurrence sans hashing n'est d'aucune utilité car son rendement est nul.

2.2.1 Rendement de la méthode de récurrence

Dans cette section, nous allons présenter sous forme graphique le rendement de la méthode de récurrence pour différentes familles de mélanges d'états de Bell. Sur un même graphique, nous avons placé des courbes de rendement pour différentes variantes de la méthode de récurrence-hashing. De plus, dans chacun de ces graphiques, nous avons tracé en pointillé la courbe de l'intrication de formation en fonction de la fidélité.

Le premier cas que nous allons aborder est la famille composée des états de Werner. Comme mentionné dans la section précédente, tout mélange d'états de Bell peut être mis sous cette forme (tout en conservant la fraction d'intrication maximale). De plus, ces états découlent de la transmission d'un état parfaitement intriqué par un canal dépolarisant.

Dans la figure 2.2, nous avons tracé trois courbes de rendement versus fidélité pour trois protocoles de purification différents. La courbe en plus foncée représente le rendement de la méthode de hashing. Nous verrons à la section 2.6 que cette méthode possède un rendement positif lorsque $F > 0.8107$ pour un état de Werner. Les deux autres courbes, comme la légende l'indique, sont associées aux protocoles de Machiavello et Bennett *et al.* Ce sont les deux protocoles dont nous avons discuté à la section précédente dans le cas des états de Werner. Rappelons aussi que ces deux protocoles sont conjugués à la méthode de hashing afin d'obtenir un rendement positif. On re-

FIG. 2.2: Comparaison des rendements de la purification de W_F.

marque d'ailleurs que les courbes de rendement des méthodes de récurrence rejoignent celle de la méthode de hashing.

Comme deuxième cas, nous allons considérer une famille plus particulière. Il s'agit de la famille d'états de Bell décrite par le mélange statistique suivant :

$$X_F = F|\Psi^-\rangle\langle\Psi^-| + \frac{1-F}{2}\Big(|\Psi^+\rangle\langle\Psi^+| + |\Phi^+\rangle\langle\Phi^+|\Big)$$

où $1/2 < F < 1$. Par une opération similaire au “twirling”, tout mélange d'états de Bell ayant une composante nulle peut être mis sous cette forme (tout en conservant la fraction d'intrication maximale). Bien sûr, il est toujours possible de ramener cet état sous forme de Werner. Cependant, tel qu'argumenté dans la section précédente, il ne s'agit sûrement pas de la meilleure stratégie. La figure 2.3 démontre cela.

Cette figure contient trois courbes de rendement versus fidélité. Celle en plus foncée représente le rendement du protocole de hashing. On remarque que le rendement est positif à partir d'une valeur inférieure à 0.8107. En effet, puisque $S(X_F) < S(W_F)$, le

FIG. 2.3: Comparaison des rendements de la purification de X_F.

rendement du protocole de hashing est meilleur dans ce cas (voir section 2.6).

La courbe désignée sous le nom "récurrence" représente le rendement optimal de la méthode de récurrence, tandis que la courbe désignée sous le nom "Machiavello" représente le rendement de la méthode de Machiavello si, comme étape préliminaire, l'état est mis sous forme de Werner. On remarque que l'écart entre les deux courbes est assez important.

Comme dernier cas, considérons la famille d'états suivante :

$$Y_F = F|\Psi^-\rangle\langle\Psi^-| + (1-F)|\Psi^+\rangle\langle\Psi^+|$$

où $1/2 < F < 1$. Il se trouve que le protocole de hashing donne un rendement positif sur tous les états de cette famille (puisque l'entropie de ces états est toujours inférieure à 1). Nous avons vérifié que la meilleure stratégie de purification dans ce cas consiste à utiliser la méthode hashing directement sans faire aucune itération de récurrence au préalable ! Le graphique du rendement se trouve à la figure 2.4.

FIG. 2.4: Rendement de la purification de Y_F.

Les cas que nous venons d'analyser ne sont que trois exemples de mélanges d'états de Bell. La recette que nous avons développée à la section précédente nous permet d'utiliser efficacement la méthode de récurrence sur ces états. Dans le cas d'un état de Werner, notre recette se réduit à la méthode de Machiavello, mais dans d'autres cas (par exemple X_F) elle fournit un meilleur rendement.

Un mélange d'états de Bell possède une intrication de formation positive seulement lorsque sa fraction d'intrication maximale est supérieure à $1/2$. La méthode de récurrence obtient un rendement positif dans ce cas. Cependant, l'écart important qu'il y a entre les courbes de rendement et celle de l'intrication de formation porte à croire qu'un meilleur rendement peut être atteint par d'autres protocoles.

La méthode de récurrence nous donne néanmoins une borne inférieure positive sur $D_2(\rho)$ lorsque $f(\rho) > 1/2$ (ρ est quelconque). Mais qu'en est-il si $f(\rho) \leq 1/2$? Est-il possible de retirer de l'intrication d'un tel état ? La section suivante montre qu'on peut le faire dans certains cas.

2.3 Purification directe de mélanges non diagonaux dans la base de Bell

Supposons que l'on veuille purifier un mélange statistique ρ avec $f(\rho) \leq 1/2$. La méthode de récurrence ne s'applique pas directement dans ce cas. En effet, l'opération de “twirling” aurait pour effet de perdre toute l'intrication contenue dans l'état ρ. Il faut donc prendre davantage conscience de la structure de l'état ρ pour espérer en distiller quelque chose.

On présente dans cette section une méthode particulière pour épurer des états non purifiables par la méthode de récurrence. Considérons d'abord la famille d'états suivante :

$$\rho_p = (1-p)|00\rangle\langle 00| + p|\Psi^+\rangle\langle\Psi^+|,$$

où p est une valeur réelle comprise entre 0 et 1. Dans la base standard, ρ_p est représentée par la matrice de densité :

$$\rho_p = \begin{pmatrix} 1-p & 0 & 0 & 0 \\ 0 & p/2 & p/2 & 0 \\ 0 & p/2 & p/2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

La fraction d'intrication maximale est donné par le maximum des valeurs propres de $\mathrm{Re}\,\rho_p$ lorsque ρ_p est exprimé dans la base magique :

$$\rho_p = (1-p)\frac{1}{\sqrt{2}}(|e_1\rangle - i|e_2\rangle)\frac{1}{\sqrt{2}}(\langle e_1| + i\langle e_2|) + p|e_3\rangle\langle e_3|.$$

Dans la notation matricielle, ρ_p s'écrit :

$$\rho_p = \begin{pmatrix} \frac{1}{2}(1-p) & \frac{i}{2}(1-p) & 0 & 0 \\ -\frac{i}{2}(1-p) & \frac{1}{2}(1-p) & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \qquad \text{avec} \qquad \mathrm{Re}\,\rho_p = \begin{pmatrix} \frac{1-p}{2} & 0 & 0 & 0 \\ 0 & \frac{1-p}{2} & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

De cela, on conclut que :

$$f(\rho_p) = \max\left(\frac{1-p}{2}, p\right).$$

On remarque que si $p \leq 1/2$, alors $f(\rho_p) \leq 1/2$ et la méthode de récurrence est inutile à la purification de ρ_p. Maintenant, puisqu'on s'intéresse à l'intrication de ρ_p, écrivons $\rho_p^{T_2}$:

$$\rho_p^{T_2} = \begin{pmatrix} 1-p & 0 & 0 & p/2 \\ 0 & p/2 & 0 & 0 \\ 0 & 0 & p/2 & 0 \\ p/2 & 0 & 0 & 0 \end{pmatrix}.$$

Le déterminant de cette matrice est égal à $-p^4/16$ ce qui montre que $E(\rho_p) > 0$ ssi $p > 0$. Donc, malgré le fait que $f(\rho_p) \leq 1/2$ lorsque $0 < p \leq 1/2$, l'état ρ_p contient de l'intrication que l'on aimerait bien retirer par purification.

Considérons maintenant le protocole proposé par [4] : Alice et Bob appliquent un BXOR sur leurs particules groupées deux à deux et mesurent ensuite les paires cibles dans la base standard. À chacun des quatre résultats possibles de la mesure correspond un nouveau mélange statistique qui décrit l'état de la paire source. La détermination de ces quatre états constitue une étape essentielle à la description du reste du protocole. Pour cette raison, il est nécessaire de la faire de façon complète et précise.

D'abord, la matrice de densité N qui représente l'état global d'une paire source et d'une paire cible groupées ensemble (avant l'application du BXOR) est donnée par le produit tensoriel suivant :

$$N = \rho_p^S \otimes \rho_p^C = \begin{pmatrix} (1-p)\rho_p & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \frac{p}{2}\rho_p & \frac{p}{2}\rho_p & \mathbf{0} \\ \mathbf{0} & \frac{p}{2}\rho_p & \frac{p}{2}\rho_p & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix},$$

où S et C désignent respectivement la paire source et la paire cible. Chaque élément de la matrice ci-dessus est une matrice de dimension 4×4. L'élément $\mathbf{0}$ dénote la matrice nulle de dimension 4×4. Au total, la dimension de N, une matrice décrivant un état d'un système à quatre qubits, donne bien 16×16. La base dans laquelle est exprimée N est le simple produit tensoriel de la base standard du système de la paire source avec celle du système de la paire cible :

$$\begin{pmatrix} |0_A 0_B\rangle \\ |0_A 1_B\rangle \\ |1_A 0_B\rangle \\ |1_A 1_B\rangle \end{pmatrix}_S \otimes \begin{pmatrix} |0_A 0_B\rangle \\ |0_A 1_B\rangle \\ |1_A 0_B\rangle \\ |1_A 1_B\rangle \end{pmatrix}_C ,$$

où A et B identifient les particules appartenant à Alice et Bob. Les opérateurs de projection de la paire cible dans la base standard sont décrits par :

$$\begin{aligned} P_{00} &= I_S \otimes (|0_A 0_B\rangle\langle 0_A 0_B|)_C, \\ P_{01} &= I_S \otimes (|0_A 1_B\rangle\langle 0_A 1_B|)_C, \\ P_{10} &= I_S \otimes (|1_A 0_B\rangle\langle 1_A 0_B|)_C, \\ P_{11} &= I_S \otimes (|1_A 1_B\rangle\langle 1_A 1_B|)_C, \end{aligned}$$

où I_S est l'opérateur identité dans le sous-espace de la paire source.

Appelons U la matrice unitaire qui représente l'opération du ou-excusif bilatéral. U s'exprime aussi comme un produit tensoriel :

$$U = \mathrm{XOR}_A \otimes \mathrm{XOR}_B \qquad \text{avec} \qquad \mathrm{XOR} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

Il est important de noter que ce produit tensoriel ne s'applique pas sur les même sous-systèmes que précédemment (i.e. source versus cible), mais plutôt sur le sous-système d'Alice versus celui de Bob. On peut décrire l'effet d'un XOR de la façon suivante :

$$|\alpha\beta\rangle \xrightarrow{\mathrm{XOR}} |\alpha(\alpha \oplus \beta)\rangle, \tag{2.1}$$

où $|\alpha\beta\rangle$ dénote la base standard en notation binaire. Cela sera pratique dans ce qui suit.

Nous sommes maintenant prêts à calculer la matrice de densité ρ qui représente l'état du système global après la mesure, mais avant la connaissance de son résultat :

$$\rho = \sum_x P_x U N U^\dagger P_x \qquad \text{avec } x \in \{00, 01, 10, 11\}.$$

Puisque U est réelle et symétrique, on a que $U^\dagger = U$. En utilisant la propriété 2.1, on peut donc facilement calculer les $U^\dagger P_x$ égaux aux UP_x :

$$\begin{aligned}
UP_{00} &= U\Big(I_S \otimes (|0_A 0_B\rangle\langle 0_A 0_B|)_C\Big) \\
&= U\Big((|0_A 0_B\rangle\langle 0_A 0_B| + |0_A 1_B\rangle\langle 0_A 1_B| + |1_A 0_B\rangle\langle 1_A 0_B| + |1_A 1_B\rangle\langle 1_A 1_B|)_S \\
&\quad \otimes(|0_A 0_B\rangle\langle 0_A 0_B|)_C\Big) \\
&= U\Big(|0_{AS}0_{BS}0_{AC}0_{BC}\rangle\langle 0_{AS}0_{BS}0_{AC}0_{BC}| + |0_{AS}1_{BS}0_{AC}0_{BC}\rangle\langle 0_{AS}1_{BS}0_{AC}0_{BC}| + \\
&\quad |1_{AS}0_{BS}0_{AC}0_{BC}\rangle\langle 1_{AS}0_{BS}0_{AC}0_{BC}| + |1_{AS}1_{BS}0_{AC}0_{BC}\rangle\langle 1_{AS}1_{BS}0_{AC}0_{BC}|\Big) \\
&= \Big(\mathrm{XOR}_A|0_{AS}0_{AC}\rangle \otimes \mathrm{XOR}_B|0_{BS}0_{BC}\rangle\Big)\langle 0_{AS}0_{BS}0_{AC}0_{BC}| + \\
&\quad \Big(\mathrm{XOR}_A|0_{AS}0_{AC}\rangle \otimes \mathrm{XOR}_B|1_{BS}0_{BC}\rangle\Big)\langle 0_{AS}1_{BS}0_{AC}0_{BC}| + \\
&\quad \Big(\mathrm{XOR}_A|1_{AS}0_{AC}\rangle \otimes \mathrm{XOR}_B|0_{BS}0_{BC}\rangle\Big)\langle 1_{AS}0_{BS}0_{AC}0_{BC}| + \\
&\quad \Big(\mathrm{XOR}_A|1_{AS}0_{AC}\rangle \otimes \mathrm{XOR}_B|1_{BS}0_{BC}\rangle\Big)\langle 1_{AS}1_{BS}0_{AC}0_{BC}| \\
&= |0_{AS}0_{BS}0_{AC}0_{BC}\rangle\langle 0_{AS}0_{BS}0_{AC}0_{BC}| + |0_{AS}1_{BS}0_{AC}1_{BC}\rangle\langle 0_{AS}1_{BS}0_{AC}0_{BC}| + \\
&\quad |1_{AS}0_{BS}1_{AC}0_{BC}\rangle\langle 1_{AS}0_{BS}0_{AC}0_{BC}| + |1_{AS}1_{BS}1_{AC}1_{BC}\rangle\langle 1_{AS}1_{BS}0_{AC}0_{BC}|\,.
\end{aligned}$$

De la même façon, on obtient pour les autres UP_x :

$$\begin{aligned}
UP_{01} &= |0_{AS}0_{BS}0_{AC}1_{BC}\rangle\langle 0_{AS}0_{BS}0_{AC}1_{BC}| + |0_{AS}1_{BS}0_{AC}0_{BC}\rangle\langle 0_{AS}1_{BS}0_{AC}1_{BC}| + \\
&\quad |1_{AS}0_{BS}1_{AC}1_{BC}\rangle\langle 1_{AS}0_{BS}0_{AC}1_{BC}| + |1_{AS}1_{BS}1_{AC}0_{BC}\rangle\langle 1_{AS}1_{BS}0_{AC}1_{BC}|, \\
UP_{10} &= |0_{AS}0_{BS}1_{AC}0_{BC}\rangle\langle 0_{AS}0_{BS}1_{AC}0_{BC}| + |0_{AS}1_{BS}1_{AC}1_{BC}\rangle\langle 0_{AS}1_{BS}1_{AC}0_{BC}| + \\
&\quad |1_{AS}0_{BS}0_{AC}0_{BC}\rangle\langle 1_{AS}0_{BS}1_{AC}0_{BC}| + |1_{AS}1_{BS}0_{AC}1_{BC}\rangle\langle 1_{AS}1_{BS}1_{AC}0_{BC}|, \\
UP_{11} &= |0_{AS}0_{BS}1_{AC}1_{BC}\rangle\langle 0_{AS}0_{BS}1_{AC}1_{BC}| + |0_{AS}1_{BS}1_{AC}0_{BC}\rangle\langle 0_{AS}1_{BS}1_{AC}1_{BC}| + \\
&\quad |1_{AS}0_{BS}0_{AC}1_{BC}\rangle\langle 1_{AS}0_{BS}1_{AC}1_{BC}| + |1_{AS}1_{BS}0_{AC}0_{BC}\rangle\langle 1_{AS}1_{BS}1_{AC}1_{BC}|.
\end{aligned}$$

Finalement, avec un peu d'algèbre matricielle, on obtient :

$$
\begin{aligned}
P_{00}UNU^{\dagger}P_{00} &= \begin{pmatrix} (1-p)^2 & 0 & 0 & 0 \\ 0 & p^2/4 & p^2/4 & 0 \\ 0 & p^2/4 & p^2/4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}_S \otimes (|0_A0_B\rangle\langle 0_A0_B|)_C, \\
P_{01}UNU^{\dagger}P_{01} &= \begin{pmatrix} p(1-p)/2 & 0 & 0 & 0 \\ 0 & p(1-p)/2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}_S \otimes (|0_A1_B\rangle\langle 0_A1_B|)_C, \\
P_{10}UNU^{\dagger}P_{10} &= \begin{pmatrix} p(1-p)/2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & p(1-p)/2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}_S \otimes (|1_A0_B\rangle\langle 1_A0_B|)_C, \\
P_{11}UNU^{\dagger}P_{11} &= \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & p^2/4 & p^2/4 & 0 \\ 0 & p^2/4 & p^2/4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}_S \otimes (|1_A1_B\rangle\langle 1_A1_B|)_C.
\end{aligned}
$$

La dernière équation indique que si lors de la mesure Alice et Bob obtiennent le résultat associé au vecteur propre $|11\rangle$ (cela va se produire avec probabilité $\mathrm{Tr}(P_{11}UNU^{\dagger}P_{11}) = p^2/2$), la paire source se retrouvera alors dans l'état pur $|\Psi^+\rangle\langle\Psi^+|$. En seulement une étape, on obtient un rendement positif égal à $p^2/4$! (On divise par deux la probabilté d'obtenir le résultat $|11\rangle$ car la moitié des paires, soit les paires cibles, sont sacrifiées.) Dans [4], le protocole se termine ici. Pourtant, il n'est pas exclu que l'état des paires sources issues des autres résultats contienne de l'intrication distillable. Analysons donc les autres possibilités.

Si on considère l'état de la paire source après la sortie du résultat associé au vecteur propre $|01\rangle$, on s'apperçoit que cet état n'est plus intriqué. Pour s'en convaincre, il suffit de constater que $\rho_S = \frac{\mathrm{Tr}_C(P_{01}UNU^\dagger P_{01})}{\mathrm{Tr}(P_{01}UNU^\dagger P_{01})}$ est diagonal dans la base standard et donc que $\rho_S^{T_2} = \rho_S$. La même remarque s'applique pour le résultat relié au vecteur propre $|10\rangle$. Cependant, pour ce qui est du premier résultat qui survient avec probabilité $(1-p)^2 + p^2/2$, on a que la matrice de densité qui représente l'état de la paire source est :

$$\rho_S = \frac{\mathrm{Tr}_C(P_{00}UNU^\dagger P_{00})}{\mathrm{Tr}(P_{00}UNU^\dagger P_{00})} = \frac{1}{(1-p)^2 + p^2/2} \begin{pmatrix} (1-p)^2 & 0 & 0 & 0 \\ 0 & p^2/4 & p^2/4 & 0 \\ 0 & p^2/4 & p^2/4 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Si on pose $p' = \frac{p^2/2}{(1-p)^2+p^2/2}$, ρ_S s'écrit :

$$\rho_S = \begin{pmatrix} (1-p') & 0 & 0 & 0 \\ 0 & p'/2 & p'/2 & 0 \\ 0 & p'/2 & p'/2 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

qui n'est rien d'autre que $\rho_{p'}$. Puisque $p' > 0$, ρ_S contient de l'intrication que l'on peut distiller de nouveau en réappliquant le procédé. Le rendement global de ce PP2 sur l'état ρ_p s'exprime récursivement comme suit :

$$R(\rho_p) = \frac{p^2}{4} + \frac{1}{2}\left((1-p)^2 + \frac{p^2}{2}\right) R(\rho_{p'}).$$

Dans le cas où $p > 1/2$, il est intéressant de comparer la méthode de récurrence et la méthode directe présentée dans cette section. D'abord, remarquons que ρ_p s'écrit comme suit dans la base de Bell :

$$\rho_p = \begin{pmatrix} (1-p)/2 & (1-p)/2 & 0 & 0 \\ (1-p)/2 & (1-p)/2 & 0 & 0 \\ 0 & 0 & p & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

FIG. 2.5: Comparaison des rendements de la méthode de récurrence et de la méthode par purification directe.

de sorte qui si on applique la première étape du “twirling”, on peut s’arranger ensuite pour avoir $p_{00} = p$ et $p_{10} = 0$. Pour un tel état, on connaît le rendement de la méthode de récurrence (voir figure 2.3). La comparaison de ces deux courbes de rendement est donnée par la figure 2.5.

On voit qu’au-delà de la valeur $p \approx 0.75$, la méthode de récurrence est préférable à la méthode directe. D’après la figure, on remarque aussi que $D_2(\rho_p) \geq R(\rho_p) > 0$ pour $p > 0$. Mais puisque $E(\rho_p) > 0$ ssi $p > 0$, cela permet de conclure que :

$$E(\rho_p) > 0 \Rightarrow D_2(\rho_p) > 0 \qquad \forall p.$$

Comme on a vu, les mélanges d’états de Bell respectent aussi cette propriété. Il est pertinent de se demander si cette propriété s’étend à un mélange statistique arbitraire. Ce problème a été résolu en partie par les Horodecki [8] qui ont exhibé un état ρ d’un système à deux qutrits (3×3) qui ne respecte pas cette propriété, i.e. ρ satisfait $E(\rho) > 0$ et $D_2(\rho) = 0$. Ce résultat très important implique qu’en général $D_2 \neq E$, et donc qu’il existe bien deux types d'intrication différente, à savoir l'intrication de formation et celle de distillation. Les implications de ce résultat sont profondes et ouvrent la voie à des interprétations qui dépassent le niveau de ce mémoire ; c'est pourquoi nous nous garderons d'en discuter davantage.

D'autre part, une autre partie du problème a été éclaircie lorsque ces mêmes Horodecki [9] ont montré que la propriété restait valide pour les systèmes 2×2 et 2×3. Pour ce faire, ils ont développé un protocole de purification qui possède un rendement positif sur n'importe quel état à deux qubits d'intrication de formation non nulle. Ce protocole, baptisé méthode de filtrage, est le sujet de la prochaine section.

2.4 Méthode de filtrage

On a vu qu'un état ρ d'un système 2×2 est inséparable (i.e. $E(\rho) > 0$) si et seulement si sa transposée partielle ρ^{T_2} n'est pas un opérateur positif. La méthode de filtrage prend avantage de la connaissance d'un tel critère pour augmenter, par l'application d'un "filtre" approprié, la fraction d'intrication maximale de ρ au-delà de $1/2$. Le nouvel état peut ensuite être purifié par la méthode de récurrence obtenant ainsi un rendement positif. Mais avant de poursuivre, nous devons clarifier ce que nous entendons par le terme "filtre". Pour cela, il nous faut introduire la notion de mesure généralisée (POVM).

Dans le chapitre 1, nous avons mentionné que si on se restreint à une vue locale du système global, il est possible qu'un état subisse une évolution non unitaire. L'état du sous-système après l'opération est alors obtenu en faisant une trace partielle sur l'état du système global.

Maintenant, supposons qu'une mesure est effectuée sur le système global. Cette mesure est caractérisée par un ensemble de projecteurs orthogonaux associés aux différents résultats de la mesure. Au niveau du sous-système, l'effet de la mesure n'est pas pareil ; il n'est en général plus possible d'y associer un ensemble de projecteurs. La mesure dite "généralisée" peut cependant être caractérisée par un ensemble d'opérateurs positifs $\{V_i\}$ qui forment une partition de l'identité :

$$\sum_{i=1}^{n} V_i V_i^\dagger = I.$$

(Contrairement au cas des mesures orthogonales, on note que n peut excéder la dimension du système.) Supposons que ρ représente l'état du sous-système avant la mesure. Si une mesure est effectuée alors on obtient, avec probabilité $p_i = \mathrm{Tr}(V_i \rho V_i^\dagger)$, le $i^{ième}$ résultat relié à la mesure généralisée et l'état ρ devient :

$$\rho_i = \frac{V_i \rho V_i^\dagger}{\mathrm{Tr}(V_i \rho V_i^\dagger)}.$$

Malgé le fait que cette caractérisation ne rende pas compte de toutes les mesures généralisées, elle sera suffisamment générale pour ce qui suit.

Habituellement, l'application d'une mesure généralisée se fait en deux étapes : ajouter un système ancillaire au système sous considération et appliquer une mesure orthogonale au nouveau système global. Par la suite, le système ancillaire peut être mis de côté. Le résultat net de ces opérations équivaut à l'application d'une mesure généralisée. Bien entendu, une mesure orthogonale constitue un cas particulier d'une mesure généralisé.

Lorsqu'on parle de filtre, on fait référence à un élément particulier de la partition d'une mesure généralisée. Le filtre est donc associé à un opérateur V_f avec $f \in 1, ..., n$. La plupart du temps, il s'agit d'un opérareur sur une particule. Lorsqu'on applique la mesure et qu'on obtient le résultat associé à V_f, on dit que la particule a passé le filtre. Dans le cas contraire, on dit que la particule a été absorbée par le filtre, auquel cas on jette l'état résultant. Nous sommes maintenant prêts à aborder la méthode de filtrage.

Voyons d'abord, partant d'un état ρ inséparable, comment déterminer les paramètres du filtre. Par le critère d'inséparabilité, on sait qu'il existe un vecteur propre $|\varphi\rangle$ associé à une valeur propre négative de ρ^{T_2}. En utilisant la décomposition de Schmidt, on peut écrire $|\varphi\rangle$ de la façon suivante :

$$|\varphi\rangle = a|\alpha_1\rangle \otimes |\beta_1\rangle + b|\alpha_2\rangle \otimes |\beta_2\rangle,$$

où a et b sont des nombres réels positifs qui satisfont $a^2 + b^2 = 1$. De plus, les $\{|\alpha_i\rangle\}$ et les $\{|\beta_i\rangle\}$ forment deux bases orthonormées de $\mathcal{H}_2$. Si on dénote $\{|b_i\rangle\}$ la base standard dans $\mathcal{H}_2$, alors il existe deux matrices unitaires U_1 et U_2 qui permettent de passer des bases $\{|\alpha_i\rangle\}$ et $\{|\beta_i\rangle\}$ respectivement à la base $\{|b_i\rangle\}$:

$$\{|\alpha_i\rangle\} \xrightarrow{U_1} \{|b_i\rangle\},$$
$$\{|\beta_i\rangle\} \xrightarrow{U_2} \{|b_i\rangle\}.$$

Maintenant, si Alice et Bob appliquent les opérations unitaires U_1 et U_2 sur leurs systèmes respectifs, $|\varphi\rangle$ se transforme comme suit :

$$\begin{aligned} |\varphi'\rangle &= (U_1 \otimes U_2)|\varphi\rangle \\ &= (U_1 \otimes U_2)\Big(a|\alpha_1\rangle \otimes |\beta_1\rangle + b|\alpha_2\rangle \otimes |\beta_2\rangle\Big) \\ &= aU_1|\alpha_1\rangle \otimes U_2|\beta_1\rangle + bU_1|\alpha_2\rangle \otimes U_2|\beta_2\rangle \\ |\varphi'\rangle &= a|b_1\rangle \otimes |b_1\rangle + b|b_2\rangle \otimes |b_2\rangle. \end{aligned}$$

Sans perte de généralité, on peut donc supposer que $|\varphi\rangle$ est de la forme :

$$|\varphi\rangle = a|00\rangle + b|11\rangle.$$

Soit W le filtre local défini par la matrice :

$$W = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}.$$

On note que l'opérateur $I \otimes W$ est bien un opérateur positif car $a, b > 0$ (montré plus loin). Soit $\tilde{\rho}$ l'état qui surgit après l'application par Bob de ce filtre W. En utilisant le fait que $(I \otimes W)^\dagger = I \otimes W$, on a que :

$$\tilde{\rho} = \frac{(I \otimes W)\rho(I \otimes W)}{\mathrm{Tr}((I \otimes W)\rho(I \otimes W))}.$$

L'état des paires qui ont passé le filtre local est donc représenté par $\tilde{\rho}$. Si on pose $p = \mathrm{Tr}((I \otimes W)\rho(I \otimes W))$, alors p est la probabilité pour chaque paire de passer le filtre. Pour montrer que $p > 0$, notons d'abord que :

$$\begin{aligned} p &= a^2\rho_{11} + b^2\rho_{22} + a^2\rho_{33} + b^2\rho_{44} \\ &\geq \min(a^2, b^2)\mathrm{Tr}\rho = \min(a^2, b^2). \end{aligned}$$

De plus, $a > 0$ puisque sinon $\langle\varphi|\rho^{T_2}|\varphi\rangle = \langle 11|\rho^{T_2}|11\rangle = \rho_{44} \geq 0$ contredisant le fait que $|\varphi\rangle$ est un vecteur propre associé à une valeur propre négative de ρ^{T_2}. Un argument similaire permet de conclure que $b > 0$ de sorte que $p > 0$.

Il nous reste à montrer que $f(\tilde{\rho}) > 1/2$ pour obtenir le résultat voulu. Mais avant, nous avons besoin d'un résultat intermédiaire : soit $P = |\Phi^+\rangle\langle\Phi^+|$, on veut prouver que $\mathrm{Tr}(P^{T_2}\tilde{\rho}) < 0$. Notons d'abord que :

$$P = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \qquad \text{et} \qquad P^{T_2} = \frac{1}{2}\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

À l'aide de cela, on obtient :

$$\begin{aligned} \langle 00|P^{T_2} &= \frac{1}{2}\langle 00|, \\ \langle 01|P^{T_2} &= \frac{1}{2}\langle 10|, \\ \langle 10|P^{T_2} &= \frac{1}{2}\langle 01|, \\ \langle 11|P^{T_2} &= \frac{1}{2}\langle 11|, \end{aligned}$$

de sorte que :

$$\begin{aligned} \mathrm{Tr}(P^{T_2}\tilde{\rho}) &= \langle 00|P^{T_2}\tilde{\rho}|00\rangle + \langle 01|P^{T_2}\tilde{\rho}|01\rangle + \langle 10|P^{T_2}\tilde{\rho}|10\rangle + \langle 11|P^{T_2}\tilde{\rho}|11\rangle \\ &= \frac{1}{2}\langle 00|\tilde{\rho}|00\rangle + \langle 10|\tilde{\rho}|01\rangle + \langle 01|\tilde{\rho}|10\rangle + \langle 11|\tilde{\rho}|11\rangle \\ &= \frac{1}{2p}(a^2\langle 00|\rho|00\rangle + ab\langle 10|\rho|01\rangle + ab\langle 01|\rho|10\rangle + b^2\langle 11|\rho|11\rangle). \end{aligned}$$

D'autre part, on a que :

$$\langle\varphi|\rho^{T_2}|\varphi\rangle = (a\langle 00| + b\langle 11|)\rho^{T_2}(a|00\rangle + b|11\rangle)$$

$$
\begin{aligned}
&= a^2\langle 00|\rho^{T_2}|00\rangle + ab\langle 00|\rho^{T_2}|11\rangle + ab\langle 11|\rho^{T_2}|00\rangle + b^2\langle 11|\rho^{T_2}|11\rangle \\
&= a^2\langle 00|\rho|00\rangle + ab\langle 01|\rho|10\rangle + ab\langle 10|\rho|01\rangle + b^2\langle 11|\rho|11\rangle .
\end{aligned}
$$

La dernière équation se dérive à partir de la règle :

$$
\rho^{T_2}_{m\mu,n\nu} = \rho_{m\nu,n\mu} \equiv \langle m\nu|\rho|n\mu\rangle \qquad m,\mu,n,\nu \in \{0,1\}.
$$

On a donc que :

$$
\mathrm{Tr}(P^{T_2}\tilde{\rho}) = \frac{1}{2p}\langle\varphi|\rho^{T_2}|\varphi\rangle < 0, \tag{2.2}
$$

car $\langle\varphi|\rho^{T_2}|\varphi\rangle < 0$ et $p > 0$. Maintenant que nous avons notre résultat intermédiaire en mains, calculons $\langle\Psi^-|\tilde{\rho}|\Psi^-\rangle$:

$$
\begin{aligned}
\langle\Psi^-|\tilde{\rho}|\Psi^-\rangle &= \frac{1}{2}(\langle 01| - \langle 10|)\tilde{\rho}(|01\rangle - |10\rangle) \\
&= \frac{1}{2}(\langle 01|\tilde{\rho}|01\rangle - \langle 01|\tilde{\rho}|10\rangle - \langle 10|\tilde{\rho}|01\rangle + \langle 10|\tilde{\rho}|10\rangle).
\end{aligned}
$$

Puisque $\mathrm{Tr}\tilde{\rho} = 1$, on peut écrire :

$$
\begin{aligned}
\langle\Psi^-|\tilde{\rho}|\Psi^-\rangle &= \frac{1}{2}(1 - \langle 00|\tilde{\rho}|00\rangle - \langle 01|\tilde{\rho}|10\rangle - \langle 10|\tilde{\rho}|01\rangle - \langle 11|\tilde{\rho}|11\rangle) \\
&= \frac{1}{2} - \mathrm{Tr}(P^{T_2}\tilde{\rho}).
\end{aligned}
$$

De l'inégalité 2.2 découle directement le résultat voulu, à savoir que $\langle\Psi^-|\tilde{\rho}|\Psi^-\rangle > 1/2$. Ensuite, $\tilde{\rho}$ peut être épuré par la méthode de récurrence.

Résumons le protocole : dans un premier temps, Alice et Bob doivent déterminer les paramètres du filtre W (ils dérivent ces paramètres à partir de l'état inséparable ρ). Ensuite, Bob applique ce filtre à ses particules et informe Alice desquelles ont été “absorbées” par le filtre. Alice se débarasse alors des particules qui ont perdu leur contrepartie chez Bob. L'état résultant des paires survivantes possède une fraction d'intrication maximale supérieure à 1/2 et peut donc être soumis au protocole de récurrence. Si ϵ est le rendement de ce dernier protocole, le rendement du protocole global (filtrage + récurrence) est donné par $p\epsilon$ où $p = \mathrm{Tr}((I \otimes W)\rho(I \otimes W))$ est la probabilité qu'une particule de Bob passe le filtre. Il est plutôt surprenant de constater que la partie du filtrage ne requiert qu'une communication unidirectionnelle (de Bob vers Alice). Cependant, puisque la méthode de récurrence est un PP2, le protocole globale en demeure un.

Un meilleur rendement peut être parfois obtenu si on remplace le filtre par une mesure généralisée dont un des résultats, se produisant avec probabilité p, a le même effet que l'application du filtre. Par rapport au simple filtrage, l'avantage de la mesure généralisée est clair : elle évite le gaspillage de certaines particules et permet donc, par un procédé récursif, d'obtenir un meilleur rendement. Ce procédé est illustré dans la prochaine section.

2.5 Mesure généralisée avant purification

Pour un état ρ inséparable quelconque, on peut se demander si le choix du filtre W, tel que décrit dans la section précédente, est optimal. Autrement dit, existe-t-il un autre filtre qui permette d'obtenir un rendement supérieur ? Il est difficile de répondre avec justesse à cette question. Considérons par exemple, la famille d'états disctuée dans [9] et introduite dans la section sur l'inséparabilité :

$$\rho_p(c,d) = p|\Phi\rangle\langle\Phi| + (1-p)|\Psi\rangle\langle\Psi|,$$

où

$$|\Phi\rangle = c|00\rangle + d|11\rangle,$$
$$|\Psi\rangle = c|01\rangle + d|10\rangle,$$

avec $c, d > 0$ et $c^2 + d^2 = 1$. On a vu que si $p \neq 1/2$, alors ρ_p est inséparable. Avec un peu d'algèbre, on trouve que la fraction d'intrication maximale de $\rho_p(c,d)$ est donnée par la formule :

$$f(\rho_p(c,d)) = \max(p, 1-p)\left(\frac{1}{2} + cd\right).$$

On s'intéresse surtout aux états de la famille qui satisfont $f(\rho_p(c,d)) \leq 1/2$, i.e. les états $\rho_p(c,d)$ pour lesquels :

$$\max(p, 1-p) \leq \frac{1}{1+2cd}.$$

Ici, au lieu de suivre la méthode de filtrage, utilisons plutôt le filtre proposé dans [9] :

$$W = \begin{pmatrix} d & 0 \\ 0 & c \end{pmatrix}.$$

Si Alice applique ce filtre à ses particules, l'état résultant des paires restantes devient :

$$\tilde{\rho}_p(c,d) = \frac{(W \otimes I)\rho_p(c,d)(W \otimes I)}{\mathrm{Tr}((W \otimes I)\rho_p(c,d)(W \otimes I))}.$$

Après quelques calculs, on à que :

$$(W \otimes I)\rho_p(c,d)(W \otimes I) = \begin{pmatrix} pc^2d^2 & 0 & 0 & pc^2d^2 \\ 0 & (1-p)c^2d^2 & (1-p)c^2d^2 & 0 \\ 0 & (1-p)c^2d^2 & (1-p)c^2d^2 & 0 \\ pc^2d^2 & 0 & 0 & pc^2d^2 \end{pmatrix},$$

avec

$$\mathrm{Tr}((W \otimes I)\rho_p(c,d)(W \otimes I)) = 2pc^2d^2 + 2(1-p)c^2d^2 = 2c^2d^2 > 0,$$

car c et d sont positifs. En combinant les deux derniers résultats, on obtient :

$$\tilde{\rho}_p(c,d) = \tilde{\rho}_p = \frac{1}{2}\begin{pmatrix} p & 0 & 0 & p \\ 0 & (1-p) & (1-p) & 0 \\ 0 & (1-p) & (1-p) & 0 \\ p & 0 & 0 & p \end{pmatrix},$$

ce qui revient à écrire :

$$\tilde{\rho}_p = p|\Phi^+\rangle\langle\Phi^+| + (1-p)|\Psi^+\rangle\langle\Psi^+|.$$

On remarque que $\tilde{\rho}_p$ est diagonal dans la base de Bell et que $f(\tilde{\rho}_p) = \max(p, 1-p) > 1/2$ (car $p \neq 1/2$). De plus, puisque deux des quatre composantes de Bell sont nulles, l'état

$\tilde{\rho}_p$ peut être soumis directement à la méthode de hashing. Nous allons voir dans la prochaine section que le rendement de cet état est égale à $1 - H(p)$.

Au lieu du simple filtre W, nous avons trouvé une mesure généralisée qui permet d'obtenir un meilleur rendement. Ce POVM est décrit par les opérateurs positifs suivants :

$$\begin{aligned} V_1 &= W \otimes I \\ V_2 &= \tilde{W} \otimes I \qquad \text{où} \qquad \tilde{W} = \begin{pmatrix} c & 0 \\ 0 & d \end{pmatrix}. \end{aligned}$$

On a bien que $V_1 V_1^\dagger + V_2 V_2^\dagger = I$ car $c^2 + d^2 = 1$. De plus, on sait déjà que :

$$\rho_1 = \frac{V_1 \rho_p(c,d) V_1^\dagger}{\mathrm{Tr}(V_1 \rho_p(c,d) V_1^\dagger)} = \tilde{\rho}_p.$$

Pour ce qui est de ρ_2, on obtient que :

$$V_2 \rho_p(c,d) V_2^\dagger = \begin{pmatrix} pc^4 & 0 & 0 & pc^2d^2 \\ 0 & (1-p)c^4 & (1-p)c^2d^2 & 0 \\ 0 & (1-p)c^2d^2 & (1-p)d^4 & 0 \\ pc^2d^2 & 0 & 0 & pd^4 \end{pmatrix},$$

avec

$$\mathrm{Tr}(V_2 \rho_p(c,d) V_2^\dagger) = p(c^4 + d^4) + (1-p)(c^4 + d^4) = c^4 + d^4.$$

Si on pose :

$$c' = \frac{c^2}{\sqrt{c^4 + d^4}} \qquad \text{et} \qquad d' = \frac{d^2}{\sqrt{c^4 + d^4}},$$

on peut écrire ρ_2 de la façon suivante :

$$\rho_2 = \begin{pmatrix} pc'^2 & 0 & 0 & pc'd' \\ 0 & (1-p)c'^2 & (1-p)c'd' & 0 \\ 0 & (1-p)c'd' & (1-p)d'^2 & 0 \\ pc'd' & 0 & 0 & pd'^2 \end{pmatrix},$$

de cela, il découle que :

$$\rho_2 = \rho_p(c', d').$$

FIG. 2.6: Comparaison des rendements de la méthode avec POVM et de la méthode de récurrence pour l'état $\rho_p(\sqrt{0.1}, \sqrt{0.9})$.

Puisque c' et d' sont supérieurs à zéro, ρ_2 est inséparable ce qui permet de réappliquer la même mesure généralisée $\{V_1, V_2\}$ sur cet état. Le rendement du processus global est donnée par la formule récursive :

$$R(\rho_p(c, d)) = 2c^2d^2(1 - H(p)) + (c^4 + d^4)R(\rho_p(c', d')),$$

où $1 - H(p)$ est le rendement de la méthode de hashing sur l'état $\tilde{\rho}_p$.

Dans la figure 2.6, on compare ce rendement avec celui de la méthode de récurrence pour un cas particulier ($c^2 = 0.1$ et $d^2 = 0.9$). Les courbes sont tracés en fonction du paramètre p. Notons que lorsque $0.385 \leq p \leq 0.625$ le rendement de la méthode de récurrence est nul car la fraction d'intrication maximale est plus petite ou égale à $1/2$. Même à l'extérieur de cet intervalle, la méthode basée sur la mesure généralisée est avantageuse en ce qui a trait au rendement.

Il aurait été intéressant de comparer le filtre utilisé dans cette section avec celui produit par la méthode originale de filtrage. Le problème est que les paramètres de ce dernier filtre se calculent difficilement. De plus, afin de déterminer le rendement de la méthode de récurrence après filtrage, il aurait fallu calculer l'état résultant du filtrage et ensuite calculer la fraction d'intrication maximale de celui-ci.

Néanmoins, cette section démontre par un exemple explicite qu'il est possible d'utiliser avec succès les POVMs à des fins de distillation. Contrairement à la méthode de filtrage où seulement un des opérateurs de la mesure généralisée est mis à profit, on a montré qu'il est avantageux de considérer tous les états possibles pouvant survenir à la suite d'une mesure généralisée.

2.6 Méthode de hashing

Nous allons maintenant présenter la méthode de hashing, un PP1 qui sert à épurer les mélanges statistiques diagonaux dans la base de Bell et d'entropie inférieure à 1. Soit W un tel état, alors on a que $S(W) < 1$. Ce protocole utilise des idées similaires à la méthode de hashing universelle : en sacrifiant à peu près $nS(W)$ paires de particules, Alice et Bob obtiennent environ $nS(W)$ bits d'information qui leur permettent de déterminer, avec une grande probabilité de réussite, l'état presque pur des $n(1-S(W))$ paires restantes. En appliquant les bonnes rotations, ils peuvent par la suite transformer ces paires restantes en $|\Phi^+\rangle$ obtenant ainsi un rendement égal à $1 - S(W)$.

Tout d'abord, il est convenable de représenter les n paires impures par une chaîne de $2n$ bits suivant la notation introduite à la fin de la section 2.1. Par exemple, la séquence $|\Phi^+\rangle|\Phi^-\rangle|\Psi^+\rangle$ est représentée par la chaîne 001001. La parité d'une chaîne est définie comme étant la somme modulo 2 sur les bits qui forment cette chaîne. Dans notre exemple, la parité de 001001 est 0. La parité d'un sous-ensemble s d'une chaîne de bits x peut être exprimée comme un produit scalaire $s \cdot x$, i.e. la somme modulo 2 du ET bit à bit des chaînes s et x. Par exemple, $111100 \cdot 001001 = 1$ puisque la somme modulo 2 des quatre premiers bits de x est 1.

Voici trois résulats importants dont la méthode de hashing tire profit :

1. Soit $\delta > 0$ et P_{x_0} la distribution des séquences initiales x_0 tirées du mélange statistique $(W)^n$. P_{x_0} est donc le produit de n distributions identiques et indépendantes. Soit $\mathcal{L}$ l'ensemble qui comprend les $2^{n(S(W)+\delta)}$ chaînes les plus probables de P_{x_0}, alors la probabilité qu'une chaîne x ne soit pas comprise dans $\mathcal{L}$ est dans $\mathcal{O}(e^{-\delta^2 n})$.
2. Tel que démontré à l'annexe C, l'utilisation adéquate des opérations locales des tableaux 2.1 et 2.2 suivie d'une mesure locale sur une paire permettent à Alice et Bob d'apprendre la parité d'un sous-ensemble arbitraire s des bits d'une séquence d'états de Bell inconnue x. Après la mesure, la nouvelle séquence qui décrit les paires non mesurées est donnée par $f_s(x)$ où f_s est une fonction connue. On note que $f_s(x)$ contient deux bits de moins que x.
3. Pour deux chaînes distinctes $x \neq y$ de même longueur, la probabilité que $s \cdot x = s \cdot y$ pour un choix aléatoire de s, est exactement $1/2$. Ceci est une conséquence de la loi de distributivité $s \cdot (x \oplus y) = (s \cdot x) \oplus (s \cdot y)$:

$$\begin{aligned} P(s \cdot x = s \cdot y) &= P((s \cdot x) \oplus (s \cdot y) = 0) \\ &= P(s \cdot (x \oplus y) = 0) \\ &= P((x \oplus y) \cdot s = 0) \\ &= 1/2. \end{aligned}$$

   La dernière équation découle du fait que, puisque $x \neq y$, $x \oplus y$ n'est pas le sous-ensemble vide et il est clair que la somme modulo 2 d'un sous-ensemble non vide de la chaîne aléatoire s a autant de chances d'être 0 ou 1.

Soit x_0 la séquence initiale d'états de Bell inconnue. Nous allons voir où et comment ces résultats interviennent dans le protocole de hashing. Alice choisit d'abord s_0, une chaîne aléatoire de bits, et l'envoie à Bob. Ensuite, ils appliquent la bonne suite d'opérations décrite à l'annexe C afin d'apprendre le résultat de $s_0 \cdot x_0$. Ils se retrouvent désormais avec une nouvelle séquence $x_1 = f_s(x_0)$ de deux bits plus courts que x_0. Alice et Bob répètent alors ces étapes un nombre de fois suffisant pour être capable de déterminer de façon unique la séquence qui représente les paires restantes.

A chaque ronde, une paire est mesurée révélant un bit d'information sur la séquence inconnue. Nous allons montrer que le nombre de rondes (correspondant aux nombre de bits d'information) nécessaire et suffisant pour déterminer de façon unique la séquence des paires restantes tend vers $nS(W)$ lorsque n tend vers l'infini.

Considérons les "trajectoires" de deux séquences arbitraires mais distinctes $x_0 \neq y_0$ lors de ce protocole. Soit x_k et y_k les images de x_0 et y_0 respectivement après k rondes. On suppose ici que la même séquence de chaînes aléatoires $s_0, s_1, \ldots, s_{k-1}$ est utilisée pour les deux trajectoires. Dénotons $E(r)$ l'événement qui correspond à obtenir, après r rondes sur les deux trajectoires, deux séquences de résultats de mesure identiques avec des images finales distinctes, i.e. :

$$E(r) = \begin{cases} (x_0 \neq y_0) & \text{si } r = 0. \\ (x_r \neq y_r) \wedge \forall_{k=0}^{r-1} (s_k \cdot x_k = s_k \cdot y_k) & \text{si } r = 1, 2, \ldots, n-1. \end{cases}$$

On veut montrer que $P(E(r)) \leq 2^{-r}$ pour $r = 0, 1, \ldots, n-1$. Cela se démontre facilement par induction sur r. La base ($r = 0$) est vérifiée trivialement car toute probabilité est inférieure ou égale à 1. Avant de passer au pas d'induction, remarquons que :

$$x_{i+1} \neq y_{i+1} \Longrightarrow x_i \neq y_i \qquad \text{pour } i = 0, 1, \ldots, n-2.$$

(Il est plus évident de voir que la contraposée est vraie : si $x_i = y_i$, alors $x_{i+1} = f_{s_i}(x_i) = f_{s_i}(y_i) = y_{i+1}$.) A cause de cela, on peut dire que :

$$E(i+1) = E(i) \wedge (s_i \cdot x_i = s_i \cdot y_i) \wedge (x_{i+1} \neq y_{i+1}) \qquad \text{pour } i = 0, 1, \ldots, n-2.$$

Il faut maintenant montrer que :

$$P(E(i)) \leq 2^{-i} \Longrightarrow P(E(i+1)) \leq 2^{-(i+1)} \qquad \text{pour } i = 0, 1, \ldots, n-2.$$

Pour $i = 0, 1, \ldots, n-2$, on a justement que :

$$\begin{aligned} P(E(i+1)) &= P(E(i) \wedge (s_i \cdot x_i = s_i \cdot y_i) \wedge (x_{i+1} \neq y_{i+1})) \\ &\leq P(E(i) \wedge (s_i \cdot x_i = s_i \cdot y_i)) \end{aligned}$$

$$\begin{aligned}
&\leq P(s_i \cdot x_i = s_i \cdot y_i | E(i))P(E(i)) \\
&\leq \frac{1}{2} \cdot 2^{-i} = 2^{-(i+1)}.
\end{aligned}$$

On a que $P(s_i \cdot x_i = s_i \cdot y_i | E(i)) = 1/2$ car $E(i) \Longrightarrow (x_i \neq y_i)$ et, par le point 3, on obtient le résultat.

Maintenant, après s'être rappelé du point 1, qui dit que l'ensemble $\mathcal{L}$ contient seulement $2^{n(S(W)+\delta)}$ chaînes, mais qu'avec une probabilité plus grande que $1-\mathcal{O}(e^{-\delta^2 n})$ la séquence initiale x_0 est comprise dans $\mathcal{L}$, une stratégie pour Bob serait de trouver parmi $\mathcal{L}$ un candidat qui concorde avec les résultats des mesures. On peut montrer qu'à la suite de m rondes, la probabilité d'erreur, i.e. d'avoir que $x_0 \notin \mathcal{L}$ ou bien que deux candidats dans $\mathcal{L}$ produisent l'événement $E(m)$, est au plus $2^{n(S(W)+\delta)-m} + \mathcal{O}(\exp(-\delta^2 n))$. En effet :

$$\begin{aligned}
P(erreur) &= P(erreur | x_0 \in \mathcal{L})P(x_0 \in \mathcal{L}) + P(erreur | x_0 \notin \mathcal{L})P(x_0 \notin \mathcal{L}) \\
&\leq P(erreur | x_0 \in \mathcal{L}) + P(x_0 \notin \mathcal{L}) \\
&\leq \sum_{\substack{y_0 \in \mathcal{L} \\ y_0 \neq x_0}} P(E(m)) + \mathcal{O}(e^{-\delta^2 n}) \\
&\leq 2^{n(S(W)+\delta)} \cdot 2^{-m} + \mathcal{O}(e^{-\delta^2 n}) \\
&\leq 2^{n(S(W)+\delta)-m} + \mathcal{O}(e^{-\delta^2 n}).
\end{aligned}$$

Si on pose $m = n(S(W) + 2\delta)$, la probabilité devient :

$$P(erreur) \leq 2^{-n\delta} + \mathcal{O}(e^{-\delta^2 n}).$$

On note donc que la probabilité d'erreur converge de façon exponnentielle vers zéro lorsque n tend vers l'infini. Le rendement du protocole est alors égal à $1 - S(W) - 2\delta$.

En posant $\delta \approx n^{-1/4}$, le rendement atteint la valeur de $1 - S(W)$ dans la limite où n tend vers l'infini. Cependant, la probabilité d'erreur converge alors moins vite vers zéro dans ce cas. En effet :

$$P(erreur) \leq 2^{-n\delta} + \mathcal{O}(e^{-\delta^2 n})$$

$$\begin{aligned} &\leq 2^{-n^{3/4}} + \mathcal{O}(e^{-n^{1/2}}) \\ &\leq \mathcal{O}(e^{-n^{1/2}}). \end{aligned}$$

Il est à noter que ce protocole ne requiert qu'une communication unidirectionnelle. Une fois qu'Alice a terminé sa partie du protocole, ayant mesuré m qubits, elle envoie à Bob la séquence de chaînes aléatoires $s_0, s_1, \ldots, s_{m-1}$ ainsi que les résultats de ses mesures. Celui-ci peut ensuite mesurer de façon adéquate m de ses qubits pour déterminer la chaîne x_m qui décrit l'état des $n-m$ paires restantes.

Le rendement de la méthode de hashing est donné par $1 - S(W)$. Pour un état de Werner, on a que :

$$S(W_F) = -F \lg F - (1-F) \lg \frac{1-F}{3},$$

ce qui donne un rendement positif si $F > 0.8107$. Mais qu'en est-il si $F \leq 0.8107$? Est-il possible d'obtenir une borne inférieure positive sur D_1 ? La marge de manoeuvre est assez faible puisqu'on sait d'après Knill et Laflamme [10] que $D_1(W_F) = 0$ si $F \leq 3/4$. La prochaine section présente un autre PP1 qui constitue une généralisation de la méthode de hashing et qui permet d'abaisser la borne jusqu'à 0.8096.

2.7 Méthode de hashing généralisée

On a vu que le rendement de la méthode de hashing est simplement donné par l'expression $1 - S(W)$, où W décrit l'état initial qui est diagonal dans la base de Bell. La méthode de hashing généralisée [12] vise, par l'application d'opérations locales, à tranformer la matrice W en une collection de mélanges statistiques diagonaux dans la base de Bell, $\mathcal{E} = \{p_i, W_i\}$, d'entropie plus petite que W (i.e. $S(\mathcal{E}) < S(W)$), obtenant ainsi une meilleure base pour la méthode de hashing. Bien que les W_i soient diagonaux dans la base de Bell, il n'est pas clair que la méthode de hashing puisse agir sur les collections d'états tel $\mathcal{E}$. Ce point sera éclairci plus loin.

La première étape du protocole consiste pour Alice et Bob à grouper leurs paires de particules en blocs de taille égale à k. Ensuite, pour chaque bloc, ils appliquent $k-1$

fois l'opération BXOR avec toujours la première paire comme paire source et les $k-1$ autres paires tour à tour comme paire cible (voir figure 2.7).

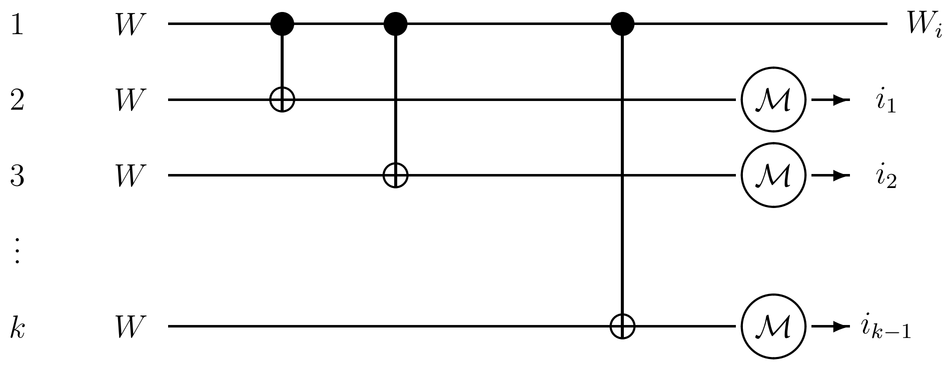

FIG. 2.7: Circuit découvert par Shor et Smolin pour diminuer l'entropie d'un mélange d'états de Bell. La mesure d'un bit amplitude est désignée par $\mathcal{M}$.

Par après, ils peuvent apprendre les bits d'amplitude des $k-1$ paires cibles dans le but d'identifier laquelle des 2^{k-1} matrices de densité (qui correspond aux 2^{k-1} résultats possibles) décrit l'état de la paire source. Comme on va le voir plus loin, toutes ces matrices sont diagonales dans la base de Bell.

Si on dénote $\mathcal{E}$ la collection de ces 2^{k-1} états de la paire source pouvant survenir à la suite des $k-1$ mesures, on obtient que $S(\mathcal{E})$, i.e. l'entropie espérée du nouvel état de la paire source, est donnée par :

$$S(\mathcal{E}) = \sum_{i=0}^{2^{k-1}-1} p_i S(W_i),$$

où p_i est la probabilité d'obtenir la chaîne i écrite en binaire comme résultat des bits amplitudes laissant la paire source dans l'état W_i. Il se trouve que cette même entropie peut s'exprimer de façon plus simple par une fonction récursive :

$\mathcal{S}(n, M) \quad :=$

si $n = 1$ alors retourner $S(M)$,

sinon retourner $P_0(M)\mathcal{S}(n-1, M_0(M)) + P_1(M)\mathcal{S}(n-1, M_1(M))$,

où $S(M) = -\mathrm{Tr}(M \lg M)$. Pour définir les fonction P_0, P_1, M_0 et M_1, considérons la situation suivante : Alice et Bob appliquent un BXOR avec M comme paire source et W comme paire cible et mesurent par la suite le bit d'amplitude de la paire cible. Avec probabilité $P_0(M)$, ils vont obtenir le résultat 0 et le nouvel état de la paire source sera décrit par la matrice de densité $M_0(M)$. On définit $P_1(M)$ et $M_1(M)$ de façon similaire dans le cas où le resultat de la mesure est 1. On remarque qu'un appel de la fonction $\mathcal{S}(k, W)$ retourne la valeur de l'entropie espérée, i.e. $S(\mathcal{E})$.

En utilisant la table du BXOR, on peut trouver des formules simples pour les fonctions P_0, P_1, M_0 et M_1. Si on utilise la notation :

$$p_{ij}(M) = \langle\varphi|M|\varphi\rangle \qquad \text{où } |\varphi\rangle \text{ est l'état de Bell dénoté par } ij,$$

et en supposant que M est diagonal dans la base de Bell, on obtient que :

$$\begin{aligned} P_0(M) &= (p_{00}(M) + p_{10}(M))(p_{00}(W) + p_{10}(W)) + (p_{01}(M) + p_{11}(M))(p_{01}(W) + p_{11}(W)), \\ P_1(M) &= 1 - P_0(M). \end{aligned}$$

Et pour M_0 :

$$\begin{aligned} p_{00}(M_0(M)) &= \frac{p_{00}(M)p_{00}(W) + p_{10}(M)p_{10}(W)}{P_0(M)}, \\ p_{01}(M_0(M)) &= \frac{p_{01}(M)p_{01}(W) + p_{11}(M)p_{11}(W)}{P_0(M)}, \\ p_{10}(M_0(M)) &= \frac{p_{00}(M)p_{10}(W) + p_{10}(M)p_{00}(W)}{P_0(M)}, \\ p_{11}(M_0(M)) &= \frac{p_{01}(M)p_{11}(W) + p_{11}(M)p_{01}(W)}{P_0(M)}. \end{aligned}$$

Et pour M_1 :

$$p_{00}(M_1(M)) = \frac{p_{00}(M)p_{01}(W) + p_{10}(M)p_{11}(W)}{P_1(M)},$$

$$\begin{aligned}
p_{01}(M_1(M)) &= \frac{p_{01}(M)p_{00}(W) + p_{11}(M)p_{10}(W)}{P_1(M)}, \\
p_{10}(M_1(M)) &= \frac{p_{00}(M)p_{11}(W) + p_{10}(M)p_{01}(W)}{P_1(M)}, \\
p_{11}(M_1(M)) &= \frac{p_{01}(M)p_{10}(W) + p_{11}(M)p_{00}(W)}{P_1(M)}.
\end{aligned}$$

Heureusement, on remarque que $M_0(M)$ et $M_1(M)$ sont diagonales dans la base de Bell. Cela est important car lors de l'évaluation de $\mathcal{S}(k, W)$, nous serons amenés à calculer les valeurs suivantes : $M_0(W), M_0(M_0(W)), \ldots, M_0^{k-1}(W)$. Puisque W est diagonal, il en découle que toutes ces matrices de densité le sont aussi ; et donc, bien que les équations ci-dessus soient valides seulement dans le cas où M est diagonal, elles sont suffisantes à notre développement.

Maintenant, si $S(\mathcal{E}) = \mathcal{S}(k, W) < 1$, Alice et Bob peuvent appliquer la méthode de hashing sur les paires sources restantes. Puisqu'il y a une paire source par bloc de taille k, les autres paires cibles étant perdues lors des mesures, le rendement du protocole global sera égal à $\frac{1-\mathcal{S}(k,W)}{k}$. Il faut cependant prendre garde lorsqu'on applique la méthode de hashing aux paires sources restantes puisque celles-ci ne sont plus décrites par une simple matrice de densité W, mais plutôt par une collection d'états $\mathcal{E} = \{p_i, W_i\}$. Or, on a seulement montré que la méthode de hashing fonctionnait sur les mélanges statistiques tels W.

Heureusement, on peut facilement généraliser la méthode de hashing aux collections d'états. Une fois que Bob reçoit les résultats d'Alice et les conjugue aux siens, il peut savoir à laquelle des 2^{k-1} matrices de densité chaque paire source correspond. Puisqu'ils ont plusieurs tels résultats, l'entropie moyenne observée tend vers la valeur de l'entropie espérée. Autrement dit, en choisissant n, le nombre de paires initiales, suffisamment grand, on peut rendre l'entropie moyenne observée arbitrairement proche de l'entropie espérée. On peut donc supposer qu'Alice et Bob, disposent d'un ensemble de paires décrites par différentes matrices de densité W_i, mais dont l'entropie moyenne est égale à $S(\mathcal{E}) = \sum p_i S(W_i)$.

La distribution P_{x_0} des séquences initiales n'est donc plus un produit de n distributions identiques indépendantes, mais plutôt un produit de n distributions indépendantes possiblement différentes. Il demeure cependant, que si on définit $\mathcal{L}$ comme l'ensemble qui comprend les $2^{n(S(\mathcal{E})+\delta)}$ séquences les plus probables de P_{x_0}, la probabilité que x tombe en-dehors de $\mathcal{L}$ est dans $\mathcal{O}(e^{-\delta^2 n})$. On note que seulement Bob connait l'état des paires sources, mais cela ne change rien au protocole car c'est à lui que revient la tâche d'énumérer les séquences les plus probables afin de déterminer le syndrôme d'erreur. Alice ne fait que choisir les chaînes aléatoires s_k et appliquer sa part d'opérations unitaires pour apprendre la parité $s_k \cdot x_k$; en aucun temps elle n'a besoin de connaître l'état des paires sources. La méthode de hashing généralisée demeure donc un PP1.

Si on analyse le cas où W est un état de Werner, on a que la méthode de hashing ($k = 1$) donne un rendement positif si $S(W_F) < 1$, i.e. lorsque $F > 0.8107$. La méthode de hashing généralisée permet d'abaisser cette borne jusqu'à 0.8096 pour $k = 5$. Pour des valeurs de F supérieures disons à 0.82, la méthode de hashing procure cependant le meilleur rendement, les autres variantes ($k \geq 2$) sacrifiant trop de paires comparé au gain réalisé sur l'entropie.

On peut se demander si l'ensemble d'opérations effectuées sur les blocs de taille k est optimal. Autrement dit, existe-t-il un autre ensemble d'opérations (autre que les BXORs en séquence) qui permettrait d'abaisser davantage l'entropie des paires sources ? Par plusieurs simulations, certaines aléatoires et d'autres déterministes, nous avons tenté d'obtenir, toujours sans succès, un tel groupe d'opérations. Par exemple, pour le cas $k = 9$, on peut imaginer la configuration de la figure 2.8.

L'intérêt d'une telle configuration est que, pour $k = 3$, on sait que l'on peut réaliser une baisse d'entropie avec la méthode de hashing généralisée. Alors, en réappliquant la même méthode aux trois paires sources, peut-on réduire de nouveau l'entropie ? La réponse est non et cela n'est pas très surprenant. En effet, les paires sources issues du premier groupe d'opérations ne sont plus décrites par un état de Werner, mais plutôt,

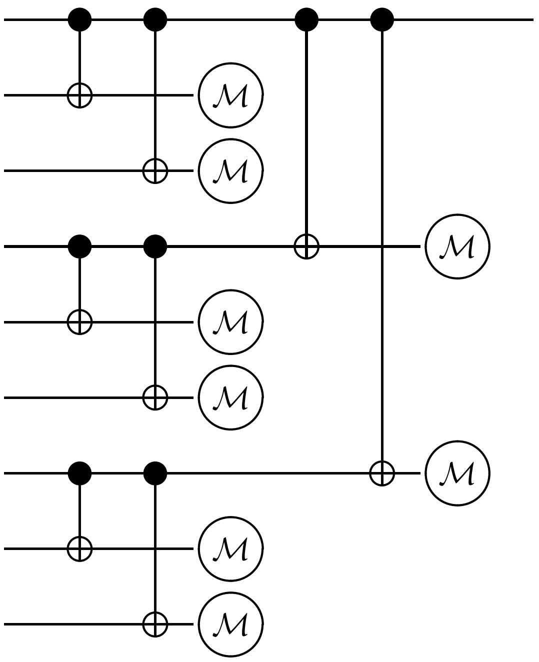

FIG. 2.8: Circuit proposé pour réduire l'entropie d'un mélange d'états de Bell. Il s'agit en fait de la double application du circuit de la figure 2.7 avec $k = 3$. La mesure d'un bit amplitude est désignée par $\mathcal{M}$.

comme on l'a vu, par une collection d'états $\mathcal{E} = \{p_i, W_i\}$. Toutefois, il demeure encore possible que le groupe d'opérations découvert par Shor et Smolin ne soit pas optimal.

À première vue, l'intérêt de la méthode de hashing généralisée ne semble pas trop important puisqu'elle ne fait qu'abaisser le seuil atteint par la méthode de hashing d'un faible pourcentage (0.1 %). Cependant, ce résulat est d'une grande importance théorique ; il démontre qu'un PP1 n'a pas besoin de déterminer le syndrôme d'erreur avec exactitude pour obtenir un rendement positif. Nous allons revenir sur ce point dans le chapitre suivant.

Chapitre 3

Protocoles de purification et codes correcteurs quantiques

Jusqu'à présent nous avons considéré des protocoles de purification qui agissaient sur des paires de qubits tirés d'un mélange statistique. Cela est normal puisque nous avions caractérisé le bruit (pouvant survenir d'une transmission par un canal de dépolarisation) par un mélange statistique. Dans ce contexte, nous avons aussi établi plusieurs définitions dans lesquelles le nombre de paires initiales devait tendre vers l'infini. Cependant, la notion de purification n'est pas limitée à ce contexte. D'ailleurs, les protocoles de purification que nous avons analysés peuvent tout aussi bien être appliqués à des blocs finis de paires de qubits. La recherche d'un comportement asymptotique servait seulement à faire le lien avec nos définitions de rendement et d'intrication de distillation.

Dans ce chapitre nous allons considérer un nouveau modèle d'erreur dont l'analyse se prête mieux au cas des blocs finis. Ce modèle d'erreur est très populaire dans le domaine des codes correcteurs quantiques. Le scénario typique est le suivant : Alice possède n singlets purs qu'elle veut utiliser à des fins de téléportation avec Bob. Elle envoie donc une particule sur deux de chaque paires à Bob. Lors de la transmission, pas plus de t de ces n particules interagissent avec l'environnement (i.e. subissent des erreurs). Le but pour Alice et Bob est de retirer de ce bloc $m < n$ paires parfaitement intriquées, i.e. avec $F = 1$ exactement. Voici un résumé des différences avec l'approche du chapitre précédent :

– Il y a encore un ensemble $\mathcal{L}$ comprenant les séquences d'états de Bell pouvant survenir à la suite de la transmission par le canal bruité. Cependant, au lieu de caractériser le bruit par un mélange statistique, on le définit par une promesse que le nombre d'erreurs soit inférieur ou égal à t. Pour simplifier, on supposse que t ne dépend pas de n.

– L'ensemble $\mathcal{L}$ possède donc une taille finie donnée par l'expression :

$$\mathcal{T} = \sum_{p=0}^{t} 3^p \begin{pmatrix} n \\ p \end{pmatrix}.$$

Chaque membre de l'ensemble $\mathcal{L}$ est indexé par i $(1 \leq i \leq \mathcal{T})$ et définit un syndrome d'erreur particulier. Le nombre 3 dans l'expression correspond au nombre possible de types d'erreur : il peut survenir soit une erreur de phase $(\Phi^+ \rightarrow \Phi^-)$, soit une erreur d'amplitude $(\Phi^+ \rightarrow \Psi^+)$ ou soit les deux $(\Phi^+ \rightarrow \Psi^-)$. Il a été montré que corriger ces trois types d'erreur est suffisant pour corriger une erreur arbitraire.

– Comme mentionné plus haut, le but du protocole est d'obtenir m singlets de fidélité parfaite, i.e. $F = 1$ exactement, à partir des n paires de départ.

Pour arriver à ce but, nous allons utiliser une approche semblable à celle de la méthode de hashing : en utilisant des opérations unitaires locales, Alice et Bob visent à transformer l'ensemble $\mathcal{L}$ des syndromes d'erreur en un autre ensemble pour lequel le résultat d'une mesure sur $n - m$ paires permet à Bob d'identifier l'état des m autres. Cette approche, que nous allons formaliser de façon plus précise dans la prochaine section, exclut la possibilité d'utiliser la communication bidirectionnelle. La raison en est que l'application première de tels protocoles est la construction de codes correcteurs quantiques. Or, on sait que seuls les protocoles de purification unidirectionnels servent à cette fin. Un fait encore plus fort, démontré dans [4], est que ces deux classes (codes correcteurs quantiques et PP1) sont équivalentes. Pour cette raison, nous allons parlé de codes correcteurs au lieu de PP1 dans ce qui suit. Cependant, nous devons se rappeler qu'il existe une transformation pour passer de l'un à l'autre.

3.1 Condition nécessaire et suffisante à la réalisation de codes correcteurs quantiques

Dans cette section, nous allons chercher à établir une condition nécessaire et suffisante à la réalisation de codes correcteurs quantiques. Pour ce faire, nous allons considérer, par une approche générale, les protocoles de purification unidirectionnels.

Initialement, Alice et Bob partagent n paires de qubits qui définissent un syndrome d'erreur particulier. Chaque syndrome d'erreur est une suite d'états de Bell (par exemple $|\Phi^+\rangle|\Phi^-\rangle|\Psi^+\rangle$) qui peut être représentée par une chaîne de $2n$ bits (001001). On dénote $x^{(i)}$ la chaîne de $2n$ bits qui représente le $i^{ième}$ syndrome d'erreur. De plus, y_k dénote le $k^{ième}$ bit du mot y.

Dans un premier temps, Alice et Bob appliquent sur les n paires à leur disposition, qui forment un certain syndrome d'erreur $x^{(i)}$, une séquence d'opérations tirées du taleau 2.1. Puisque ces opérations unitaires transforment les états de Bell en d'autres états de Bell, l'effet d'une telle séquence d'opérations est d'appliquer une certaine fonction booléenne L_U à $x^{(i)}$, obtenant ainsi une chaîne $w^{(i)}$:

$$w^{(i)} = L_U(x^{(i)}).$$

On note que L_U est contrainte à être une fonction booléenne linéaire et réversible.

La prochaine étape du protocole consiste à mesurer le bit amplitude des $n - m$ dernières paires. Pour chaque syndrome d'erreur $x^{(i)}$, les résultats de ces mesures peuvent être exprimés sous forme d'une autre chaîne booléenne $v^{(i)}$ de longueur $n - m$.

Cette chaîne est définie par :

$$v_k^{(i)} = w_{2m+2k}^{(i)} \qquad k = 1, 2, \ldots, n-m.$$

L'état des particules non mesurées est simplement le mot tronqué $w'^{(i)}$ de longueur $2m$:

$$w'^{(i)} = (w_1 w_2 \cdots w_{2m})^{(i)}.$$

Pour atteindre le résultat voulu, à savoir l'obtention de m paires dans l'état $|\Phi^+\rangle$, Bob doit appliquer une rotation finale U_f sur $w'^{(i)}$ pour le ramener dans l'état $00\cdots0$. Cela est possible si et seulement si Bob connait avec certitude $w'^{(i)}$. La seule information que Bob possède pour déduire $w'^{(i)}$ sont les résultats des mesures regroupés dans le mot $v^{(i)}$. Il est clair que s'il existe deux syndromes d'erreur différents menant à des w' distincts mais produisant les mêmes résultats lors des mesures, Bob ne pourra pas déduire avec certitude w' à partir des résulats des mesures. De façon mathématique, la condition nécessaire et suffisante au succès de la correction d'erreurs s'écrit donc :

$$\forall i, j \qquad w'^{(i)} \neq w'^{(j)} \Longrightarrow v^{(i)} \neq v^{(j)}. \tag{3.1}$$

Il est clair que si chaque syndrome d'erreur est associé à un v distinct, la condition 3.1 sera remplie. Cette nouvelle condition, qui permet d'identifier complètement le syndrome d'erreur à partir du résultat des mesures, s'écrit sous une forme similaire à la condition précédente :

$$\forall i, j \qquad i \neq j \Longrightarrow v^{(i)} \neq v^{(j)}. \tag{3.2}$$

Cependant, cette condition suffisante ne semble pas nécessaire au succès de la correction d'erreur. Si elle l'était, on aurait une restriction sur la taille de l'ensemble $\mathcal{L}$ des syndromes d'erreur. Celle-ci ne pourrait excéder le nombre de mesures possibles :

$$\mathcal{T} = \sum_{p=0}^{t} 3^p \begin{pmatrix} n \\ p \end{pmatrix} \leq 2^{n-m}. \tag{3.3}$$

La condition 3.1 ouvre la possibilité que cette borne puisse être surpassée. Par exemple, si la transformation L_U pouvait être la fonction booléenne qui à chaque $x^{(i)}$ retourne la valeur $w^{(i)} = 00\cdots0$, alors aucune mesure ne serait nécessaire ($m = n$) et la borne serait violée. Cependant, L_U est fortement contrainte à être une fonction linéaire réversible ce qui invalide l'exemple précédent dans lequel L_U était clairement non réversible. Dans la prochaine section, on montre un exemple valide qui viole la borne 3.3.

3.2 Exemple de code correcteur qui n'identifie pas complètement le syndrome d'erreur

On peut adapter l'équation 3.3 dans le contexte du chapitre précédent, i.e. avec un modèle d'erreur basé sur les mélanges statistiques. On a vu dans la section 2.6 sur la méthode de hashing qu'on pouvait se restreindre seulement aux syndromes d'erreur compris dans $\mathcal{L}$, un ensemble dont la cardinalité est égal à :

$$\mathcal{T} = |\mathcal{L}| \simeq 2^{nS(W)}.$$

(On ne peut pas diminuer davantage le nombre de syndromes d'erreur.) De plus, on sait que le nombre de particules sacrifiées lors de ce protocole est à peu de chose près égal à $nS(W)$. La méthode de hashing atteint donc, à toute fin pratique, la borne 3.3.

Maintenant, considèrons la méthode de hashing modifié avec $k = 5$. On sait que l'état de Werner $W_{.8096}$ est purifiable par cette méthode. Avec cet état comme point de départ, on peut vérifier que la collection d'états $\mathcal{E} = \{p_i, W_i\}$ issue du premier stade d'opérations (i.e. l'application des quatre BXORs en séquence suivi des quatre mesures) est d'entropie égale à 1. Le nombre de particules sacrifiées lors du protocole global est donc égale à :

$$\left(\frac{k-1}{k}\right) \cdot n + \frac{n}{k} \cdot S(\mathcal{E}) = n.$$

Le premier terme indique qu'une fraction $\frac{k-1}{k}$ des particules sont mesurées lors du premier stade du protocole, tandis que le second terme provient du fait que le reste des particules ($\frac{n}{k}$) sont ensuite soumises à la méthode hashing simple ($k = 1$), sacrifiant ainsi $\frac{n}{k} \cdot S(\mathcal{E})$ particules. Puisque $S(\mathcal{E}) = 1$, le nombre total de particules perdues se réduit simplement à n. D'autre part, on a que :

$$\mathcal{T} = 2^{nS(W_{.8096})} > 2^n \qquad \text{car } S(W_{.8096}) > 1.$$

Cette équation montre que la borne 3.3 est violée ! De cela, on conclut qu'en général la méthode de hashing généralisée n'identifie pas complètement le syndrome d'erreur. Mais plus important encore, puisqu'il est impossible de sacrifier plus de n particules, l'état $W_{.8096}$ n'est pas purifiable par un protocole qui identifie complètement le syndrome d'erreur, et donc, la méthode de hashing généralisée ne peut être substituée par aucun protocole de purification qui satisfait la condition 3.2. Il y a donc une réelle distinction entre la condition 3.1 et la condition 3.2 : la première rend compte de codes correcteurs plus puissants que la seconde.

En ce moment, la méthode de hashing généralisée est le seul exemple qui viole la borne 3.3. Il existe des codes correcteurs qui n'identifient pas complètement le syndrome d'erreur mais, dans la majorité des cas, on peut les substituer par d'autres qui le font et qui fournissent un rendement égal. Par exemple, la méthode de hashing est un protocole de purification qui ne satisfait pas la condition 3.2. Cependant, il existe un protocole similaire, appelé “the breeding method”, qui est décrit dans [4] et qui satisfait la condition 3.2 tout en accomplissant le même rendement que la méthode de hashing. Il n'existe pas de tel protocole de remplacement pour la méthode de hashing généralisée.

Dans la prochaine section, nous allons présenter les recherches de codes correcteurs que nous avons effectuées pour de faibles valeurs des paramètres t et m ($t \leq 2, m \leq 3$). Entre autre, on cherche à savoir si certains de ces codes violent la borne 3.3.

3.3 Algorithmes Monte Carlo pour trouver des codes correcteurs

Dans l'approche générale exposée à la section 3.1, Alice et Bob doivent soumettre leur syndrome d'erreur $x^{(i)}$ à une séquence d'opérations unitaires de la table 2.1. En particulier, ils peuvent choisir de faire :

1. un ou-exclusif bilatéral (BXOR) ;
2. une rotation bilatérale de $\pi/2$ (B_y) ;
3. une rotation unilatérale de π (σ_z) ;
4. une opération composée $\sigma_x B_x$.

Il se trouve que ces quatre opérations peuvent remplacer l'ensemble complet des opérations de la table 2.1. En effet, on peut montrer que chaque opération de cette table peut être réalisée par une composition particulière des quatre opérations ci-dessus. Par exemple, $B_z = \sigma_z(\sigma_x B_x) B_y(\sigma_x B_x)$.

Partant de ce résultat, nous avons implanté un algorithme de type Monte Carlo afin de trouver de codes correcteurs. L'algorithme, plutôt simple, est décrit dans [4]. On commence d'abord par énumérer tous les syndromes d'erreur possibles. Ensuite, on choisit au hasard une des quatre opérations de base et sélectionne, encore de manière aléatoire, un état de Bell (ou une paire d'états de Bell dans le cas du BXOR) sur lequel on applique l'opération choisie. Après avoir sauvegardé l'opération effectuée dans une liste, on vérifie si l'ensemble des $w^{(i)}$ résultants satisfait la condition 3.1. Si la réponse est non, on répète la procédure en ajoutant une nouvelle opération aléatoire. Si la réponse est oui, la liste d'opérations constituent alors un bon code correcteur.

L'effet de cet algorithme est de mélanger aléatoirement les $\mathcal{T}$ syndromes initiaux. Autrement dit, après l'application d'un grand nombre d'opérations aléatoires choisies par le programme, les $\mathcal{T}$ syndromes d'erreur, représentés par des chaînes de $2n$ bits, sont transformés en chaînes "quasi" aléatoires. Les chaînes $v^{(i)}$ sont alors bien distribuées, ce qui augmente les chances d'obtenir un code correcteur. En réalité, l'idée de l'algorithme est emprunté à la méthode de hashing : faire en sorte que chaque mesure retire le maximum d'information (soit un bit chacune) sur le syndrome d'erreur initial. Vu cette similitude, nous avons implanté un second algorithme basé sur la méthode de hashing. En premier lieu, on choisit au hasard une chaîne de bits s de longueur $2n$. On calcule ensuite la parité du syndrome initial en prenant s comme sous-ensemble de bits. On recommence le procédé jusqu'à temps d'avoir complété les $n - m$ mesures (à raison d'une mesure à chaque itération).

Quoique que les deux algorithmes se ressemblent en substance, l'algorithme de hashing permet de systématiser davantage la réalisation de codes correcteurs. En effet, comme démontré à la section sur la méthode de hashing, la mesure de la parité d'un sous-ensemble arbitraire d'une chaîne de bits retire le maximum d'information, soit un bit. De plus, le nombre d'opérations élémentaires constituant les codes correcteurs produits par l'algorithme de hashing est en général beaucoup plus petit que celui de ceux produits par l'algorithme aléatoire.

Nous avons effectué des recherches pour de petites valeurs de t et m. En fixant la valeur de ces deux paramètres, nous avons tenté de trouver la plus petite valeur de n qui permettait d'obtenir un code correcteur. On espérait ainsi obtenir un code correcteur violant la borne 3.3. Mais avant d'effectuer les recherches, nous avons considéré un résultat important de Knill et Laflamme [10] qui donne une condition nécessaire à l'existence de codes correcteurs. Ce résultat est énoncé par un théorème :

Théorème 1 *Un code correcteur quantique à n qubits qui corrige t erreurs en restorant m singlets doit satisfaire $n \geq 4t + m$.*

La preuve de ce théorème demande des connaissances particulières sur la théorie des codes correcteurs quantiques ; le lecteur qui s'y intéresse devra donc se référer à [10]. Le tableau 3.1 résume le théorème pour les cas particuliers qui nous intéressent.

L'avant-dernière colonne indique pour quelles valeurs de n l'équation 3.3 est violée. La dernière colonne indique les cas intéressants à enquêter, i.e. ceux qui violent la borne

t	m	Condition pour qu'un code correcteur existe	Condition pour que la borne 3.3 soit violée	Cas susceptibles de violer la borne 3.3
1	1	$n \geq 5$	$n \leq 4$	aucun
1	2	$n \geq 6$	$n \leq 6$	$n = 6$
1	3	$n \geq 7$	$n \leq 7$	$n = 7$
2	1	$n \geq 9$	$n \leq 9$	$n = 9$
2	2	$n \geq 10$	$n \leq 11$	$n = 10$ ou $n = 11$

TAB. 3.1: Résumé du théorème 1 pour quelques valeurs de t et m.

3.3 et qui satisfont la condition nécessaire sur l'existence de codes correcteurs.

3.4 Résultats des recherches

Pour le cas $t = 1$ et $m = 1$, plusieurs solutions ont été trouvées lorsque $n = 5$. En fait, nous obtenons une nouvelle solution pratiquement à chaque fois qu'on exécute le programme. Le nombre d'opérations élémentaires qui constituent ces codes correcteurs tournent autour de vingt (lorsqu'on utilise l'algorithme de hashing). Nous avons effectué une recherche exhaustive sur les sous-ensembles de parité pour trouver le code correcteur avec le plus petit nombre de portes élémentaires. Nous avons obtenu un code correcteur avec dix opérations élémentaires ce qui surpasse le résultat trouvé dans [4] (eux avaient trouvé onze). Le circuit qui correspond à ce code correcteur est illustré à l'annexe D. En accord avec le théorème 1, aucun code correcteur n'a été trouvé pour $n < 5$.

Pour le cas $t = 1$ et $m = 2$, la borne inférieure donnée par le théorème 1 est $n = 6$. De plus, on peut déduire à partir du résultat précédent qu'il existe un code correcteur pour $n = 10$. Il suffit de diviser le bloc de dix en deux blocs de cinq et d'appliquer sur chacun d'eux le code correcteur avec $t = 1, m = 1$ et $n = 5$. Après quelques recherches,

t	m	Plus petite valeur de n trouvé	Nombre d'opérations élémentaires	Sous-ensembles de parité qui décrivent le code correcteur
1	1	5	10	0101010001 01001001 100101 0011
1	2	8	25	0110001011010001 00010111000101 001001000101 0001100001 01001001 110001
1	3	9	46	000110100000111001 1011000010000101 01011111110101 101010110001 1110110011 01101011
2	1	15	147	voir Annexe D
2	2	-	-	-

TAB. 3.2: Résulats des recherches.

notre programme a trouvé un code correcteur pour $n = 8$. En dépit du grand nombre de recherches, quelques espoirs demeurent pour $n = 7$. Cependant tout laisse à croire qu'il n'existe pas de bons codes correcteurs pour $n = 6$ (pour le cas $t = 1$ et $m = 2$ bien sûr).

Nous avons aussi effectué des recherches pour le cas $t = 1$ et $m = 3$. Nous avons obtenu un code correcteur pour $n = 9$. Il est intéressant de noter que le premier code correcteur découvert permettait de corriger une erreur sur neuf qubits en restorant un singlet [13]. Le code correcteur que nous avons trouvé en restore trois dans la même situation.

Quant au cas $t = 2$ et $m = 1$, nous pensions avoir des chances d'obtenir un code correcteur pour $n = 9$. Fait plutôt étonnant, nous n'avons pas trouvé mieux que $n = 15$. En raison de la complexité grandissante du problème, nous avons décidé d'arrêter là nos recherches. Les résultats les plus importants sont résumés dans le tableau 3.2. Les codes correcteurs qui y figurent identifient tous complètement le syndrome d'erreur, i.e. satisfont la condition suffisante (mais non nécessaire) 3.2. Nous avons trouvé des codes correcteurs qui n'identifient pas complètement le syndrome d'erreur pour les cas ($t = 1$, $m = 2$, $n = 8$) et ($t = 2$, $m = 1$, $n = 15$), mais ils possèdent plus de portes élémentaires (32 et 178 respectivement) que ceux présentés dans le tableau 3.2. Nous n'avons pas trouvé de codes correcteurs qui n'identifiaient pas complètement le syndrome d'erreur pour les cas ($t = 1$, $m = 1$, $n = 5$) et ($t = 1$, $m = 3$, $n = 9$).

3.5 Analyse des résultats

Pour évaluer si un circuit constitue un code correcteur, on doit vérifier la condition 3.1. Soit $\mathcal{C}$ l'ensemble suivant :

$$\mathcal{C} = \{(i,j) \mid i \neq j \wedge v^{(i)} = v^{(j)} \wedge w'^{(i)} \neq w'^{(j)}\}.$$

La cardinalité de cet ensemble est appelé le nombre de conflits. La condition 3.1 est satisfaite si et seulement si le nombre de conflits est nul. Le but de nos algorithmes Monte Carlo est donc de réduire le nombre de conflits initiaux à zéro. Lorsque cela est réussi, nous obtenons un code correcteur.

Comme on a déjà vu dans la section sur la méthode hashing, la probabilité que deux syndromes distincts i et j produisent un conflit satisfait :

$$\mathrm{P}((i,j) \in \mathcal{C} \mid i \neq j) = \mathrm{P}(v^{(i)} = v^{(j)} \wedge w'^{(i)} \neq w'^{(j)}) \leq \left(\frac{1}{2}\right)^{n-m}.$$

Si on fait l'approximation que les syndromes d'erreur deviennent parfaitement aléatoires suite aux opérations de hashing, on peut estimer cette probabilité à :

$$\mathrm{P}(v^{(i)} = v^{(j)} \wedge w'^{(i)} \neq w'^{(j)}) = \left(\frac{1}{2}\right)^{n-m} \left(1 - \left(\frac{1}{2}\right)^{2m}\right).$$

Le nouveau terme de droite correspond à la probabilité que l'état des m paires restantes du syndrome d'erreur i soit distinct de celui des m paires restantes du syndrome d'erreur j. Puisqu'il y a $\mathcal{T}(\mathcal{T}-1)/2$ telles paires, le nombre moyen de paires (i,j) qui appartiennent à $\mathcal{C}$ est donc estimé à :

$$\frac{\mathcal{T}(\mathcal{T}-1)}{2} \left(\frac{1}{2}\right)^{n-m} \left(1 - \left(\frac{1}{2}\right)^{2m}\right).$$

Cette valeur, que l'on nomme le nombre moyen de conflits, permet d'évaluer les chances d'obtenir un code correcteur. Plus celle-ci est près de zéro, meilleures sont ces chances. Le tableau 3.3 montrent quelques unes de ces valeurs.

Lors de nos simulations, nous avons mesuré le nombre de conflits après l'application de la méthode de hashing aux syndromes initiaux. Nous avons observé des valeurs très proches de celles indiquées à l'avant-dernière colonne du tableau 3.3. On remarque que pour les cas possédant une vingtaine de syndromes, nous avons obtenu un code correcteur lorsque le nombre moyen de conflits était plus petit que six. Pour les cas avec avec quelques centaines de syndromes, nous avons l'impression que cette valeur ne doit pas dépasser la trentaine pour espérer obtenir un code correcteur. Bien sûr, à mesure que le nombre de syndromes augmente, on peut tolérer un nombre moyen de conflits plus élevé.

Bien qu'il n'y ait pas de relation précise entre le nombre moyen de conflits et la possibilité d'obtenir un code correcteur, il peut s'avérer utile de calculer ce nombre avant d'effectuer des recherches (en particulier pour des grandes valeurs de t et m).

Par exemple, supposons que deux cas possèdent un nombre de syndromes et un nombre moyen de conflits similaires et que l'on trouve un code correcteur pour un de ces cas. Nous croyons alors qu'il est prometteur de faire des receherches dans l'autre cas.

t	m	n	Nombre de syndromes	Nombre initial de conflits	Nombre moyen de conflits	Code correcteur trouvé
1	1	5	16	18	5.6	oui
1	1	6	19	21	4.0	oui
1	1	7	22	24	2.7	oui
1	2	6	19	45	10.0	non
1	2	7	22	51	6.8	non
1	2	8	25	57	4.4	oui
1	2	9	28	63	2.8	oui
1	3	7	22	81	14.2	non
1	3	8	25	90	9.2	non
1	3	9	28	99	5.8	oui
2	1	9	352	2106	181.0	non
2	1	10	436	2760	138.9	non
2	1	11	529	3531	102.3	non
2	1	12	631	4428	72.8	non
2	1	13	742	5460	50.3	non
2	1	14	862	6636	34.0	non
2	1	15	991	7965	22.5	oui
2	1	16	1129	9456	14.6	oui

Tab. 3.3: Nombre de conflits pour quelques cas possibles de codes correcteurs. La dernière colonne indique si effectivement un code correcteur a été trouvé.

Chapitre 4

Conclusions

Dans ce mémoire, nous avons pu constater que les protocoles de purification sont très variés. De plus, ils peuvent être définis dans plusieurs contextes. Celui dans lequel le bruit est caractérisé par un mélange statistique est fondamental. Il mène directement à la définition de l'intrication de distillation pour les mélanges statistiques. Dans la plupart des cas, on n'a pas pu mesurer exactement l'intrication de distillation. Par contre, plusieurs bornes inférieures ont été fournies grâce à des protocoles de purification explicites.

Afin de maximiser le rendement d'un protocole de purification, certaines règles doivent être appliquées. Premièrement, il est essentiel de considérer la structure unique des états qui sont soumis au protocole. En tenant compte de la spécificité du mélange statistique de départ, nous avons pu dériver un critère simple pour optimiser le rendement de la méthode de récurrence. Toute opération non unitaire faite pour ramener l'état de départ sous une certaine forme risque de perdre de l'intrication. Deuxièmement, à la suite de chaque mesure il faut considérer tous les cas possibles. En faisant cela, nous avons obtenu des procédés récursifs de purification pour les méthodes de purification directe et de mesure généralisée.

Malgré le fait que chaque état doive être purifié différemment, on remarque que certains schèmes sont récurrents. Ceci est vrai en particulier pour la méthode de hashing, laquelle a pu être adaptée au modèle d'erreur des codes correcteurs quantiques. Dans ce dernier cas, les protocoles de purification ont servi à déterminer plusieurs codes correcteurs quantiques pour des cas particuliers jamais explorés auparavant. Pour le cas $(t = 1, m = 1, n = 5)$, plusieurs codes correcteurs avaient déjà été identifiés mais dont le nombre de portes élémentaires étaient plus grand que celui qu'on a trouvé. Il semble donc que l'approche par purification soit une voie prometteuse pour l'étude des codes correcteurs quantiques.

Pour l'analyse des protocoles de purification, nous avons surtout utilisé le rendement comme critère de comparaison. Cependant, il existe d'autres critères importants pour évaluer l'efficacité d'un protocole. Un de ces critère est la notion de complexité. Un survol rapide des protocoles que nous avons analysés permet de constater que certains sont plus "réalisables" que d'autres. Par exemple, la méthode de hashing requiert l'énumération par Bob de tous les syndromes possibles. Puisque la quantité de tels syndromes est exponentielle par rapport au nombre de particules, on peut se poser des questions quant à l'efficacité du protocole. D'un autre côté, les PP2 semble beaucoup plus efficaces. La plupart ne demandent qu'un nombre fini d'opérations simples à chaque itération. Dans le cas de la méthode par purification directe, on obtient même un rendement positif apès une seule itération. Une analyse approfondie des protocoles de purificaition sous l'angle de la complexité révèlerait sûrement des résultats très intéressants.

Bibliographie

[1] C.H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres et W.K. Wootters. Teleporting an unkown quantum state by dual classical and Einstein-Podolsky-Rosen channels. Phys. Rev. Lett. **70**, 1895-1898 (1993).

[2] C.H. Bennett, G. Brassard, S. Popescu, B. Schumacher, J.A. Smolin et W.K. Wooters. Purification of noisy entanglement and faithful teleportation via noisy channels. Phys. Rev. Lett. **76**, 722 (1996).

[3] C.H. Bennett, G. Brassard et J.-M. Robert. How to reduce your ennemy's information. Advances in Cryptology : Proceedings of Cryto 85. Springer-Verlag. 468-476.

[4] C.H. Bennett, D.P. DiVincenzo, J.A. Smolin et W.K. Wooters. Mixed state entanglement and quantum error correction. quant-ph/9604024.

[5] C.H. Bennett, H.J. Bernstein, S. Popescu et B. Schumacher. Concentrating partial entanglement by local prerations. Phys. Rev. A **53**, 2046 (1996).

[6] F. Bessette, G. Brassard, L. Salvail et J.A. Smolin. Experimental quantum cryptography. Journal of Cryptology, vol. 5, no. 1, 2-28 (1992).

[7] M. Horodecki, P. Horodecki et R. Horodecki. Separability of mixed states : necessary and sufficient conditions. quant-ph/9605038.

[8] M. Horodecki, P. Horodecki et R. Horodecki. Mixed-state entanglement and distillation : is there a "bound" entanglement in nature ? quant-ph/9801069.

[9] M. Horodecki, P. Horodecki et R. Horodecki, Distillability of inseparable quantum systems, quant-ph/9607009.

[10] E. Knill et R. Laflamme. A theory of quantum error-correcting codes. quant-ph/9604034.

[11] A. Peres. Separability criterion for density matrices. quant-ph/9604005.

[12] P.W. Shor et J.A. Smolin. Quantum error-correction codes do not need to completely reveal the error syndrome. quant-ph/9604006.

[13] P.W. Shor. Scheme for reducing decoherence in quantum memory. Phys. Rev. A **52**, 2493 (1995).

Annexe A

Rotations bilatérales aléatoires pour rendre un état sous forme de Werner

Dans cette annexe, on montre comment, par l'application de rotations bilatérales aléatoires choisies parmi un ensemble fini, Alice et Bob peuvent rendre n'importe quel état ρ sous forme de Werner tout en conservant la fraction d'intrication maximale. Sans perte de généralité, on suppose que $\langle \Psi^- | \rho | \Psi^- \rangle = f(\rho) = F$. En effet, sinon, soit $|e\rangle$ l'état parfaitement intriqué tel que $\langle e|\rho|e\rangle = f(\rho) = F$, en utilisant la décomposition de Schmidt on peut écrire :

$$|e\rangle = \sum_i \sqrt{\lambda_i} |\alpha_i\rangle \otimes |\beta_i\rangle,$$

où les λ_i sont les valeurs propres de $\mathrm{Tr}_\mathrm{A} |e\rangle\langle e|$ et les $\{|\alpha_i\rangle\}$ ainsi que les $\{|\beta_i\rangle\}$ forment des bases dans les systèmes d'Alice et Bob respectivement. Or, puisque

$$E(|e\rangle\langle e|) = S(\mathrm{Tr}_\mathrm{A} |e\rangle\langle e|) = \mathrm{H}(\lambda_1) = 1,$$

il en découle que $\lambda_1 = \lambda_2 = 1/2$ et donc, on obtient que :

$$|e\rangle = \frac{1}{\sqrt{2}}(|\alpha_1\rangle \otimes |\beta_1\rangle + |\alpha_2\rangle \otimes |\beta_2\rangle).$$

Soit U_1 et U_2 deux transformations unitaires qui ont pour effet :

$$\begin{array}{ll} |\alpha_1\rangle \xrightarrow{U_1} |0\rangle & \quad |\beta_1\rangle \xrightarrow{U_2} |1\rangle \\ |\alpha_2\rangle \xrightarrow{U_1} |1\rangle & \quad |\beta_2\rangle \xrightarrow{U_2} -|0\rangle \end{array}$$

Alors, on a que :

$$\begin{aligned} U_1 \otimes U_2\,|e\rangle &= \frac{1}{\sqrt{2}}(U_1|\alpha_1\rangle \otimes U_2|\beta_1\rangle + U_1|\alpha_2\rangle \otimes U_2|\beta_2\rangle) \\ &= \frac{1}{\sqrt{2}}(|0\rangle \otimes |1\rangle - |1\rangle \otimes |0\rangle) \\ &= |\Psi^-\rangle. \end{aligned}$$

Donc, si Alice et Bob appliquent $U_1 \otimes U_2$ à leurs particules, ils obtiennent ρ' qui satisfait :

$$\langle\Psi^-|\,\rho'|\Psi^-\rangle = \langle\Psi^-|(U_1 \otimes U_2)\rho(U_1 \otimes U_2)^\dagger|\Psi^-\rangle = \langle e|\rho|e\rangle = F,$$

tel que voulu.

Ensuite, avec probabilité uniforme, une sélection aléatoire est faite sur un groupe de N opérations unitaires $\{U_i\}$. Chacune de ces opérations représente l'application d'une rotation bilatérale R_i particulière, i.e. $U_i = R_i \otimes R_i$. Si le choix aléatoire n'est pas connu par Alice et Bob (ou s'ils l'oublient !), alors la matrice de densité ρ est transformée comme suit :

$$M = \frac{1}{N}\sum_{i=1}^{N} U_i\,\rho\,U_i^\dagger.$$

Avant d'introduire ce groupe de N opérations unitaires $\{U_i\}$, nous avons besoin d'apporter une petite modification au tableau 1. Pour alléger la lecture, nous avions omis les changements de phase des états de Bell dans le tableau 2.1. Cela ne change rien aux calculs précédents car nous faisions toujours affaire à des mélanges statistiques diagonaux dans la base de Bell. Cependant, dans ce cas-ci, ρ n'est pas diagonal et nous avons besoin de connaître les changements de phase. Le tableau A.1 illustre ces changements.

		source			
		Ψ^-	Φ^-	Φ^+	Ψ^+
	I	Ψ^-	Φ^-	Φ^+	Ψ^+
Rotations Bilatérales :	B_x	Ψ^-	Φ^-	$i\Psi^+$	$i\Phi^+$
	B_y	Ψ^-	$-\Psi^+$	Φ^+	Φ^-
	B_z	Ψ^-	$i\Phi^+$	$i\Phi^-$	Ψ^+
	B_x^2	Ψ^-	Φ^-	$-\Phi^+$	$-\Psi^+$
	B_y^2	Ψ^-	$-\Phi^-$	Φ^+	$-\Psi^+$
	B_z^2	Ψ^-	$-\Phi^-$	$-\Phi^+$	Ψ^+

TAB. A.1: Ajout des changements de phase des états de Bell dans une partie du tableau 2.1.

La première étape consiste à rendre ρ diagonal dans la base de Bell (c'est souvent la seule chose que l'on requiert). Pour ce faire, un groupe de quatre rotations bilatérales est suffisant :

$$\{U_i\} = \{R_i \otimes R_i\} \qquad \text{avec} \qquad \{R_i\} = \begin{cases} I \\ B_x^2 \\ B_y^2 \\ B_z^2 \end{cases}$$

Si on veut rendre ce nouvel état M sous forme de Werner, on peut appliquer une seconde ronde avec, cette fois, trois rotations bilatérales :

$$\{U_i\} = \{R_i \otimes R_i\} \qquad \text{avec} \qquad \{R_i\} = \begin{cases} B_x \\ B_y \\ B_z \end{cases}$$

Au lieu de passer par deux rondes pour rendre l'état ρ sous forme de Werner ($\rho \longrightarrow W \longrightarrow W_F$), il est possible d'utiliser une seule ronde mais cela requiert douze rotations bilatérales [4].

Annexe B

Quelques propriétés de la fonction g

Soit :

$$g(p_{00}, p_{10}) = \underbrace{\frac{(p_{00}^2 + p_{10}^2)}{((p_{00} + p_{10})^2 + (1 - p_{00} - p_{10})^2)}}_{(*)},$$

on a que :

$$\begin{aligned}
\frac{\partial g}{\partial p_{10}} &= \frac{2p_{10}}{(*)} - \frac{1}{(*)^2}((p_{00}^2 + p_{10}^2)(2(p_{00} + p_{10}) - 2(1 - p_{00} - p_{10}))) \\
&= \frac{1}{(*)^2}(2p_{10}(2p_{00}^2 + 4p_{00}p_{10} + 2p_{10}^2 + 1 - 2p_{00} - 2p_{10}) - (p_{00}^2 + p_{10}^2)(4p_{00} + 4p_{10} - 2)) \\
&= \frac{1}{(*)^2}(4p_{00}p_{10}^2 + 2p_{10} - 4p_{00}p_{10} - 2p_{10}^2 - 4p_{00}^3 + 2p_{00}^2) \\
&= \frac{1}{(*)^2}(4p_{00}p_{10}(p_{10} - 1) + 2p_{10}(1 - p_{10}) + 2p_{00}^2(1 - 2p_{00})) \\
&= \frac{1}{(*)^2}(2p_{10}(1 - 2p_{10})(1 - p_{00}) + 2p_{00}^2(1 - 2p_{00})) \\
&= \frac{1}{(*)^2}(1 - 2p_{00})(2p_{00}^2 + 2p_{10}(1 - p_{10})).
\end{aligned}$$

Puisque $p_{00} > 1/2$, alors $(1 - 2p_{00}) < 0$. Les autres termes étant strictement positif, on obtient que $\frac{\partial g}{\partial p_{10}} < 0$. Les extrema de la fonction sont donc atteints sur les bornes du domaine, i.e. lorsque $p_{10} = 0$ et $p_{10} = 1 - p_{00}$. La valeur maximale de la fonction est :

$$g_{max}(p_{00}) = g(p_{00}, 0) = \frac{p_{00}^2}{p_{00}^2 + (1 - p_{00})^2},$$

et la valeur minimale :

$$g_{min}(p_{00}) = g(p_{00}, 1-p_{00}) = p_{00}^2 + (1-p_{00})^2.$$

On peut voir facilement que $g_{min}(p_{00}) < p_{00} < g_{max}(p_{00})$. (En effet, puisque $g_{max}(p_{00})g_{min}(p_{00}) = p_{00}^2$ et $0 < g_{min}(p_{00}) < g_{max}(p_{00})$, il est impossible qu'il en soit autrement). Cela implique qu'il existe une valeur p_{eqv} telle que $g(p_{00}, p_{eqv}) = p_{00}$, i.e. :

$$p_{00} = \frac{(p_{00}^2 + p_{eqv}^2)}{(p_{00}+p_{eqv})^2 + (1-p_{00}-p_{eqv})^2}.$$

En manipulant cette équation, on obtient :

$$\begin{aligned}
2p_{00}^3 + 4p_{00}^2 p_{eqv} + 2p_{00}p_{eqv}^2 + p_{00} - 2p_{00}^2 - 2p_{00}p_{eqv} &= p_{00}^2 + p_{eqv}^2 \\
(2p_{00}-1)p_{eqv}^2 + (4p_{00}^2 - 2p_{00})p_{eqv} + (2p_{00}^3 - 3p_{00}^2 + p_{00}) &= 0 \\
(2p_{00}-1)p_{eqv}^2 + 2p_{00}(2p_{00}-1)p_{eqv} + p_{00}(2p_{00}-1)(p_{00}-1) &= 0.
\end{aligned}$$

Puisque $2p_{00} - 1 \neq 0$, on peut diviser $2p_{00} - 1$ pour obtenir l'équation quadratique suivante :

$$p_{eqv}^2 + 2p_{00}p_{eqv} + p_{00}(1-p_{00}) = 0,$$

dont les solutions sont :

$$\begin{aligned}
p_{eqv} &= \frac{-2p_{00} \pm \sqrt{4p_{00}^2 - 4p_{00}(p_{00}-1)}}{2} \\
p_{eqv} &= \pm\sqrt{p_{00}} - p_{00}.
\end{aligned}$$

La seule solution valide est $p_{eqv} = \sqrt{p_{00}} - p_{00} = \sqrt{p_{00}}(1-\sqrt{p_{00}})$.

Annexe C

Comment retirer la parité d'un sous-ensemble arbitraire d'une séquence d'états de Bell inconnue

Alice et Bob partagent au départ une séquence d'états de Bell inconnue x. Pour un sous-ensemble arbitraire s, ils veulent apprendre la valeur de $s \cdot x$. Il faut d'abord choisir une paire dans laquelle la valeur de $s \cdot x$ sera stockée dans le bit amplitude. Cette paire, appelée paire destination, est simplement la paire qui correspond aux derniers bits non nuls de s. Par exemple, si $s = 11011000$, la paire destination sera la troisième paire. Les paires associées à 00 (comme la dernière dans l'exemple) n'ont aucun effet sur la parité recherchée et sont donc simplement ignorées.

La première étape consiste à mettre la parité recherchée pour chaque paire dans le bit amplitude de celle-ci. Il y a trois cas à considérer dépendant de l'indice associé à la paire. Le tableau C.1 indique quoi faire dans chacun des trois cas.

Une fois cela fait, il faut calculer la parité des bits amplitudes de toutes ces paires (excluant celles associées à 00) et envoyer ce résultat dans le bit amplitude de la paire destination. Ceci peut être accompli par plusieurs application du BXOR. L'effet d'un

indice associé à la paire	opérations locales
01	rien
10	B_y
11	$\sigma_x B_x$

TAB. C.1: Comment mettre la parité dans le bit amplitude.

BXOR sur les états de Bell se résume de la façon suivante :

$$\begin{array}{ccccc} \text{source} & \text{cible} & & \text{source} & \text{cible} \\ s_1 s_2 & c_1 c_2 & \xrightarrow{\text{BXOR}} & (s_1 \oplus c_1) s_2 & c_1 (s_2 \oplus c_2) \end{array}$$

où $s_1, s_2, c_1, c_2 \in \{0, 1\}$. On voit que le bit amplitude de la paire cible devient la parité des bits amplitudes des deux paires initiales ($s_2 \oplus c_2$). Pour accumuler la parité désirée dans le bit amplitude de la paire destination, il suffit donc d'appliquer des BXORs en choisissant comme paire cible la paire destination et comme paire source les autres paires tour à tour.

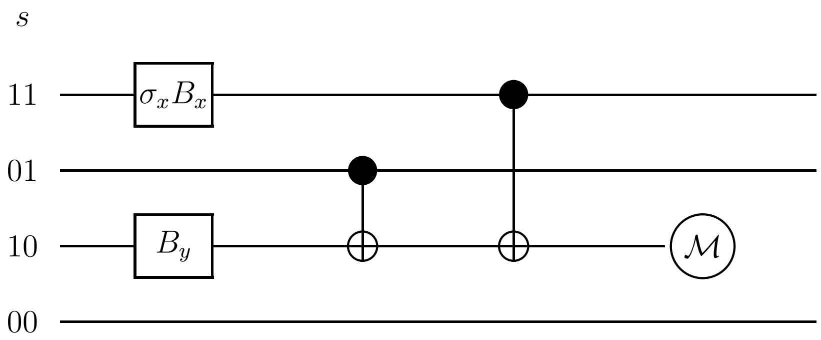

FIG. C.1: Circuit qui calcule la parité du sous-ensemble $s = 11011000$ d'une séquence de quatre états de Bell inconnue. La parité est obtenue par la mesure $\mathcal{M}$.

Annexe D

Codes correcteurs pour les cas $t = 1$, $m = 1$, $n = 5$ et $t = 2$, $m = 1$, $n = 15$.

Voici le circuit qui correspond à un code correcteur pour le cas $t = 1$, $m = 1, n = 5$.

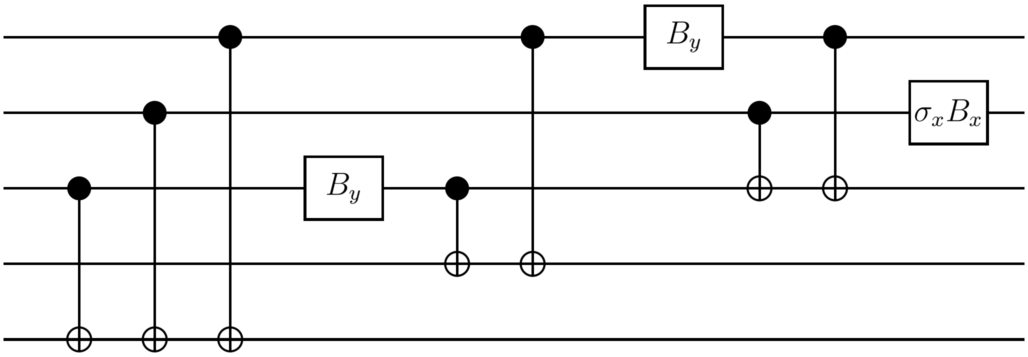

FIG. D.1: Code correcteur pour le cas $t = 1, m = 1, n = 5$. On remarque que les sous-ensemble de parité sont $s_0 = 0101010001$, $s_1 = 01001001$, $s_2 = 100101$ et $s_3 = 0011$. La particule restorée est sur le fil du haut.

Ce circuit possède dix portes élémentaires dont sept ou-excusif bilatéraux (rappelons que le meilleur code correcteur connu avait onze portes [4]). Tout comme les autres codes correcteurs trouvés, il identifie complètement le syndrome d'erreur. Dans le contexte classsique des codes correcteurs quantiques, ce circuit correspont au circuit de décodage. Le circuit d'encodage est le circuit inverse.

Quant au cas $t = 2$, $m = 1$ et $n = 15$, nous avons fait moins de recherches pour minimiser le nombre de porte élémentaires. Celui que nous présentons identifie complètement le syndrome d'erreur et possède 147 opérations élémentaires. Il est décrit par les quatorze sous-ensembles de parité suivants :

$$
\begin{aligned}
s_0 &= 001011111111011101000011001011 \\
s_1 &= 0001010110100111110001010001 \\
s_2 &= 01001011110111010011111001 \\
s_3 &= 110010111001111010001101 \\
s_4 &= 1111011110100011010001 \\
s_5 &= 01000110100001100001 \\
s_6 &= 100110110111101011 \\
s_7 &= 0011000010011001 \\
s_8 &= 00010011101011 \\
s_9 &= 001010111111 \\
s_{10} &= 1101111011 \\
s_{11} &= 10010101 \\
s_{12} &= 100111 \\
s_{13} &= 0111
\end{aligned}
$$

Pour des raisons évidentes nous avons décidé de ne pas illustrer le circuit. Cependant, la recette pour le faire est très simple ; il suffit de mettre bout à bout les quatorze circuits qui calculent la parité de chaque sous-ensemble (voir l'annexe C).