

# Information mutuelle pour la construction d'un détecteur quantique optimal

Le problème de la détection d'état quantique porte sur un ensemble de  $m$  états quantiques représentés par les opérateurs densité  $\{\rho_i, 1 \leq i \leq m\}$  munis des probabilités préalables  $\{p_i \geq 0, 1 \leq i \leq m\}$ . L'objectif est d'obtenir un ensemble de  $m$  opérateurs de mesure  $\{\Pi_j, 1 \leq j \leq m\}$  permettant de mesurer le mieux possible par rapport aux probabilités les états d'entrée qui nous arrivent.

Les opérateurs  $\rho_i$  et  $\Pi_i$  sont des matrices Hermitiennes semi-définies positives, de la forme 
$$\begin{bmatrix} a & b + ic \\ b - ic & d \end{bmatrix}.$$

Plusieurs critères ont été proposés à optimiser afin de construire ces détecteurs optimaux. D'une part, on a la possibilité de travailler sur la minimisation de l'erreur de mesure (l'erreur moyenne ou l'erreur quadratique). D'autre part, et c'est ce sur quoi nous avons travaillé, on peut considérer le critère de l'information mutuelle comme critère à maximiser. Ce critère indique la dépendance de deux variables aléatoires entre elles, il permet dans notre cas de bien caractériser la quantité d'information qu'on peut retirer des états d'entrée en ayant les opérateurs de mesure

L'information mutuelle de deux variables aléatoires  $X$  et  $Y$  est donnée par :

$$I(X; Y) = \sum_{y \in Y} \sum_{x \in X} p_{(X,Y)}(x, y) \log \left( \frac{p_{(X,Y)}(x, y)}{p_X(x)p_Y(y)} \right), \quad (1)$$

mais peut aussi être écrite en fonction des entropies des variables aléatoires :

$$I(X;Y) = H(X) - H(X|Y) \quad (2)$$

$$= H(Y) - H(Y|X) \quad (3)$$

$$= H(X) + H(Y) - H(X,Y). \quad (4)$$

Avec  $H(X)$  entropie marginale de  $X$ ,  $H(Y)$  entropie marginale de  $Y$ ,  $H(X|Y)$  entropie conditionnelle de  $X$  sachant  $Y$  et enfin  $H(X,Y)$  entropie jointe de  $X$  et  $Y$ . On peut utiliser indifféremment  $\log_2$ ,  $\log_{10}$  ou  $\ln$  pour le logarithme, le changement étant à une constante près.

Dans le cas classique, les entropies marginales, conditionnelles et jointes sont définies par :

$$H(X) = - \sum_{x \in X} p(x) \log(p(x)), \quad (5)$$

$$H(Y) = - \sum_{y \in Y} p(y) \log(p(y)), \quad (6)$$

$$H(X,Y) = - \sum_{x \in X} \sum_{y \in Y} p(x,y) \log(p(x,y)), \quad (7)$$

$$H(Y|X) = - \sum_{x \in X, y \in Y} p(x,y) \log\left(\frac{p(x,y)}{p(x)}\right) \quad (8)$$

Dans le cas quantique, les formules restent les mêmes, mais on exprime les probabilités des variables en fonction des valeurs des états quantiques d'entrée.

En fonction d'un état d'entrée  $\rho_i$  de probabilité préalable  $p_i$ , et d'un opérateur de sortie  $\Pi_i$ , on peut définir leur probabilité jointe :

$$p(X = \rho_i, Y = \Pi_i) = p_i \operatorname{tr}(\rho_i \Pi_i). \quad (9)$$

On en déduit les probabilités marginales :

$$p(X = \rho_i) = \sum_j p_i \operatorname{tr}(\rho_i \Pi_j) \quad (10)$$

$$p(Y = \Pi_j) = \sum_i p_i \operatorname{tr}(\rho_i \Pi_j), \quad (11)$$

Et les probabilités conditionnelles :

$$P(Y = \Pi_j | X = \rho_i) = \frac{\text{tr}(\rho_i \Pi_j)}{\sum_k \text{tr}(\rho_i \Pi_k)} \quad (12)$$

L'information mutuelle pour notre problème peut donc être ré-écrite de la façon suivante, en utilisant  $\alpha_{ij} = \text{tr}(\rho_i \Pi_j)$  :

$$\begin{aligned} I(\rho; \Pi) &= H(\rho) + H(\Pi) - H(\rho, \Pi) \\ &= - \sum_{i=1}^m \left( \sum_{j=1}^m \alpha_{ij} \right) \log \left( \sum_{j=1}^m \alpha_{ij} \right) - \sum_{i=1}^m \left( \sum_{j=1}^m \alpha_{ji} \right) \log \left( \sum_{j=1}^m \alpha_{ji} \right) + \sum_{i=1}^m \sum_{j=1}^m \alpha_{ij} \log(\alpha_{ij}) \end{aligned} \quad (13)$$

On peut aussi exprimer l'information mutuelle en fonction de l'entropie conditionnelle, mais il est plus efficace d'utiliser celle donnée à l'équation 13 pour la résolution numérique.

Le problème se formule comme un problème de maximization de l'information mutuelle : on cherche à maximiser l'information qu'on peut obtenir sur  $\rho_i$  quand on a les opérateurs de mesure  $\Pi_i$  :

$$\max_{\Pi} I(\rho, \Pi) \quad (14)$$

tel que :

$$\Pi_j \succeq 0 \quad 1 \leq j \leq m \quad (15)$$

$$\sum_{j=1}^m \Pi_j = I \quad (16)$$

La contrainte 15 correspond à la semi-définition positive des opérateurs de mesure  $\Pi_j$ . Enfin, la contrainte 16 permet d'obtenir des opérateurs de mesure cohérents pour que les probabilités de mesure  $p(j) = \text{tr}(\rho \Pi_j)$  soient positives et se somment à 1.

On est en présence d'une fonction convexe, et les contraintes engendrent un ensemble admissible convexe. C'est le cas idéal lors d'une minimisation, mais le problème est une maximisation, donc concave, on ne peut donc pas juste faire une descente de gradient pour le résoudre. On peut utiliser un certain nombre de méthodes approximatives, nous utilisons le calcul par intervalle afin d'obtenir un résultat sûr dans un intervalle.

La définition du problème permet de résoudre notamment les cas immédiats des opérateurs de densité orthogonaux, mais la résolution devient très lente lorsqu'on passe à d'autres cas non orthogonaux. On rajoute des conditions au problème pour accélérer la résolution.

Le premier élément à simplifier est l'expression de l'entropie marginale de  $X = \rho_i$ . En effet, nous l'avons exprimé en fonction de la trace de la multiplication matricielle, mais on peut reprendre la définition donnée lors du cas classique qui indique que  $H(X) = - \sum_{x \in X} p(x) \log(p(x))$ . Le problème nous indique que nous connaissons les probabilités préalables des états d'entrée, on peut donc directement exprimer cette entropie en fonction de ces données et donc sans les variables de sortie  $\Pi_i$ .

Ensuite, on sait que les opérateurs de mesure se somment à l'identité. Cela signifie d'une part qu'on peut passer d'un problème à  $m$  matrices à un problème à  $m - 1$  matrices pour  $m \geq 2$ . Les matrices étant carrées de dimension  $n$ , on passe de  $m \times n^2$  variables à  $(m - 1) \times n^2$  variables, ce qui est non négligeable.

Enfin, le problème et les contraintes sont symétriques, on peut intervertir les  $\Pi_j$  sans influencer le résultat de la fonction coût. Cela nous permet de couper le problème au moins en deux pour réduire à nouveau le temps de calcul. Du fait de la somme à l'identité, on peut ajouter en condition que  $\Pi_{1,1} \leq \frac{1}{m}$  pour  $m$  opérateurs de mesure, puis  $\Pi_{2,1} \leq \frac{1}{m-1}$ , etc.

Pour la résolution de ce problème, utilisons la librairie **ibex** permettant de faire du calcul par intervalle, et possède entre autres un outil d'optimisation, **ibexopt**. Le problème est formulé avec un langage dédié, **minibex**. Nous avons eu besoin de définir un opérateur additionnel à ceux présents, l'opérateur **xlog** permettant d'effectuer l'opération  $x \times \log(x)$  en redéfinissant  $0 \times \log(0) = 0$  pour que les intervalles ne tombent pas à l'infini quand ils contiennent 0. De plus, **minibex** ne considère que des problèmes de minimisation, on ré-écrit le problème en prenant la fonction coût opposée :  $\max f(x) \Leftrightarrow \min -f(x)$ .

Le premier test effectué est sur le cas de deux états d'entrée  $|\psi_1\rangle = |0\rangle$  et  $|\psi_2\rangle = |1\rangle$  ayant pour probabilité respective  $p_1 = 0.1$  et  $p_2 = 0.9$ . Le résultat théorique est connu : les états étant orthogonaux, on doit obtenir les opérateurs de mesure égaux aux opérateurs densité d'entrée. On obtient bien avec **ibex** les opérateurs suivant :

$$\Pi_1 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix},$$

qui correspondent bien à deux opérateurs de mesure orthogonaux. Dans ce cas, l'information mutuelle est comprise dans l'intervalle  $I(\rho, \Pi) \in [0.3250, 0.3254]$ ,

avec un temps de calcul de 24 millisecondes.

Le deuxième cas qu'on peut présenter est le suivant :  $|\psi_1\rangle = |0\rangle$  et  $|\psi_1\rangle = |+\rangle$  avec comme probabilité respectives  $p_1 = 0.1$  et  $p_2 = 0.9$ . Le résultat théorique n'est pas donné, et on obtient avec **ibex** le résultat suivant :

$$\Pi_1 = \begin{bmatrix} 0.446 & 0.497 \\ 0.497 & 0.554 \end{bmatrix}, \quad \Pi_2 = \begin{bmatrix} 0.554 & -0.497 \\ -0.497 & 0.446 \end{bmatrix},$$

avec une information mutuelle comprise dans l'intervalle  $I(\rho, \Pi) \in [0.1348, 0.1349]$ , et un temps de calcul de 300 secondes.