Introduction aux capacités de canaux

Léo Gagnon

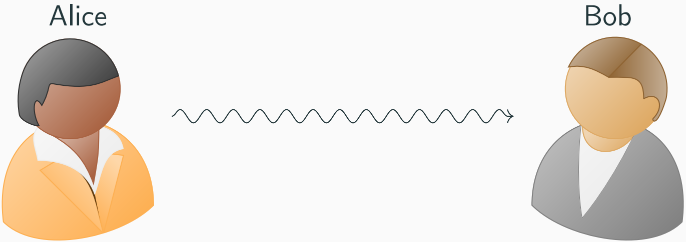
June 25, 2020

Université de Montréal

Structure de la présentation

Canal bruité classique

Canal bruité classique : Mise en contexte



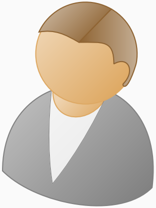
Canal bruité classique : Mise en contexte

Alice

$m \to \mathcal{E}(m)$

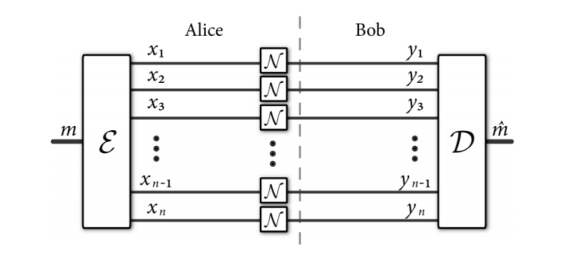
$$\mathcal{E}(m) \to \mathcal{N}(\mathcal{E}(m))$$

Bob



$\mathcal{N}(\mathcal{E}(m)) \to \hat{m}$

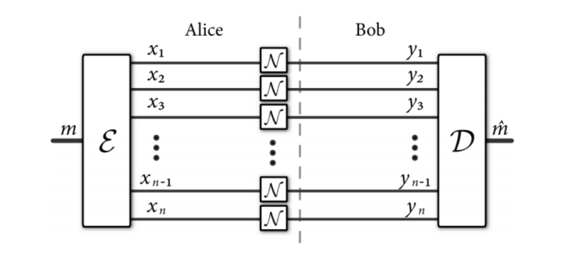
Canal bruité classique : Description formelle



Hypothèses :

- Tout les M messages sont équiprobables.
- L'entrée et la sortie du canal sont représentés par des V.A.D.
- Les canaux $\mathcal{N}$: $p_{Y|X}(y|x)$ sont i.i.d.

Canal bruité classique : Description formelle

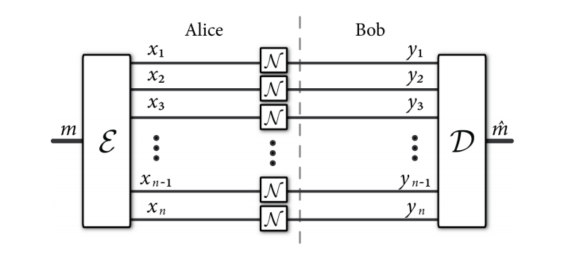


Hypothèses :

- Tout les M messages sont équiprobables.
- On peux écrire la distribution de probabilité conditionnelle comme

$$p_{Y^n|X^n}(y^n|x^n) = \prod_{i=1}^{n} p_{Y|X}(y_i|x_i)$$

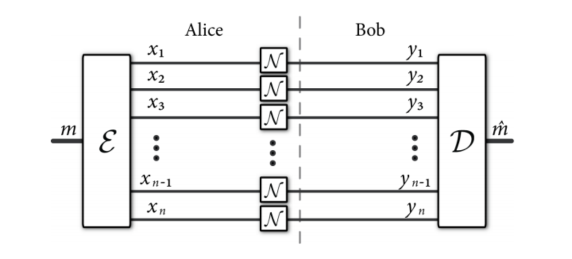
Canal bruité classique : Description formelle



Stratégie :

1. Alice et Bob choisissent un code $C \equiv \{x^n(m)\}_{m \in [M]}$
2. Alice transforme son message m en $x^n(m)$
3. Elle envoie $x^n(m)$ à Bob avec n utilisations du canal $\mathcal{N}$.
4. Bob reçoit y^n et détermine quel mot de code x^n est le plus probable.

Canal bruité classique : Description formelle

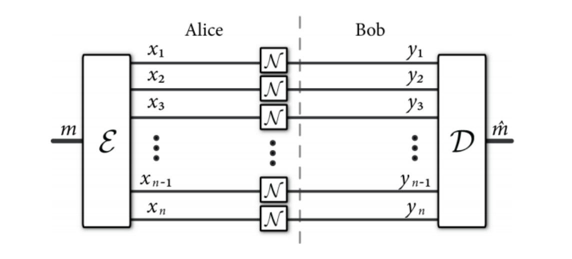


Stratégie :

1. Alice et Bob choisissent un code $C \equiv \{x^n(m)\}_{m\in[M]}$
2. Alice transforme son message m en $x^n(m)$
3. Elle envoie $x^n(m)$ à Bob avec n utilisations du canal $\mathcal{N}$.
4. Bob résout

$$\max_{y^n} = p_{Y^n|X^n}(y^n|x^n)$$

Canal bruité classique : Description formelle



Taux de transmission

$$R \equiv \frac{\#\text{bits du message}}{\#\text{utilisations du canal}} = \frac{\log_2(M)}{n}$$

Canal bruité classique : Description d'un code

Définition

Un (n, R, ε)-*code* C pour un canal bruité $\mathcal{N}$ est défini par deux fonctions :

- Encodage : $E^n : M \to X^n$
- Décodage : $D^n : Y^n \to M$

Le taux de transmission est de $R \equiv \frac{\log_2(M)}{n}$ et on a

$$p_e^*(\mathcal{C}) \equiv \max_m \Pr\{D^n(\mathcal{N}^n(E^n(m)) \neq m)\} \leq \varepsilon$$

Canal bruité classique : Taux atteignable

Définition

Un taux de transmission R est *atteignable* pour un canal $\mathcal{N}$ s'il existe un $(n, R - \delta, \varepsilon)$-code pour tout $\varepsilon \in (0, 1)$, $\delta > 0$ lorsque n est assez grand.

Canal bruité classique : Capacité d'un canal

Définition

La *capacité* d'un canal $\mathcal{N}$ est définie comme

$$C(\mathcal{N}) \equiv \sup_R \{R \text{ est un taux atteignable pour le canal } \mathcal{N}\}$$

Théorème de capacité de Shannon : Énoncé

Théorème

Soit $\mathcal{N} = p_{Y|X}$ un canal bruité. Alors

$$C(\mathcal{N}) = I(\mathcal{N}) \equiv \max_{C} I(X;Y)$$

où $I(X;Y)$ est l'information mutuelle.

Théorème de capacité de Shannon : Typicalité classique

Définition

L'entropie d'échantillon $\overline{H}(x^n)$ d'un échantillon de X^n est définie comme suit :

$$\overline{H}(x^n) \equiv -\frac{1}{n}\log_2(p_{X^n}(x^n))$$

Théorème de capacité de Shannon : Typicalité classique

Définition

Une chaîne x^n est δ-*typique* si

$$|\overline{H}(x^n) - H(X)| \leq \delta$$

Théorème de capacité de Shannon : Typicalité classique

Définition

Une chaîne x^n est δ-*typique* si

$$|\overline{H}(x^n) - H(X)| \leq \delta$$

Exemple

Soit $X = \{(10\%, 0), (90\%, 1)\}$ tel que $H(X) = 0.469$.

$$\overline{H}(11101111) \approx 0.54$$
$$\overline{H}(11111111) \approx 0.15$$
$$\overline{H}(10101000) \approx 2.13$$
$$\overline{H}(1^{18}0^2) \approx 0.469$$

Théorème de capacité de Shannon : Typicalité classique

Définition

Une chaîne x^n est δ-*typique* si

$$|\overline{H}(x^n) - H(X)| \leq \delta$$

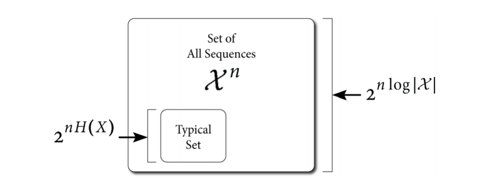
Définition

On note $T_\delta^{X^n}$ l'ensemble des chaînes δ-typiques de longeur n.

Théorème de capacité de Shannon : Typicalité classique

Lois des grands nombres $\implies$ Propriétés

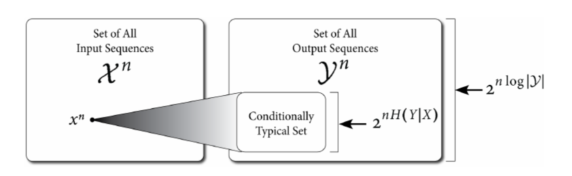
1. $\Pr\{X^n \in T_\delta^{X^n}\} \geq 1 - \varepsilon$ pour tout $\varepsilon \in (0,1)$ et $\delta > 0$ lorsque n est grand.
2. $|T_\delta^{X^n}| \leq 2^{n(H(X)+\delta)}$ pour tout $\delta > 0$
3. Soit $x^n, y^n \in T_\delta^{X^n}$, alors $p_{X^n}(x^n) \approx_\delta p_{X^n}(y^n)$



Théorème de capacité de Shannon : Typicalité classique

Lois des grands nombres $\Longrightarrow$ Propriétés

1. $\Pr\{X^n \in T_\delta^{X^n}\} \geq 1 - \varepsilon$ pour tout $\varepsilon \in (0,1)$ et $\delta > 0$ lorsque n est grand.
2. $|T_\delta^{X^n}| \leq 2^{n(H(X)+\delta)}$ pour tout $\delta > 0$
3. Soit $x^n, y^n \in T_\delta^{X^n}$, alors $p_{X^n}(x^n) \approx_\delta p_{X^n}(y^n)$



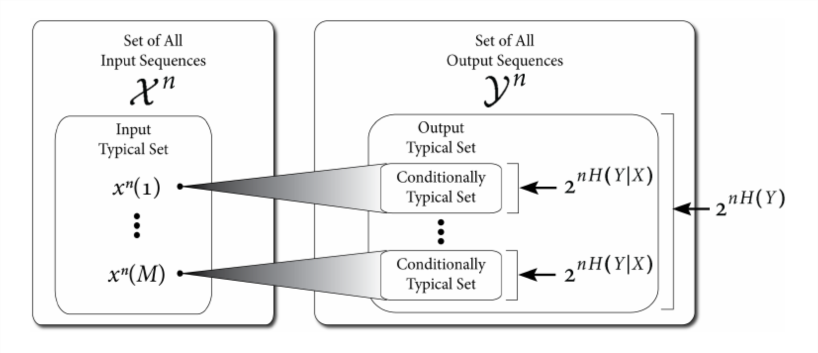
Théorème de capacité de Shannon : Protocole

Soit $\mathcal{N} : p_{Y|X}(y|x)$ un canal bruité et $p_X(x)$ la distribution maximisant $I(X;Y)$.

1. Alice et bob s'entendent sur un code aléatoire $\mathcal{C} = \{x^n(m)\}_{m \in [M]}$ en échantillionant m fois X^n.
2. Alice envoie $x^n(m)$ à Bob à travers $\mathcal{N}$. Bob reçoit y^n.
3. Bob vérifie si $y^n \in T^{Y^n}$. Ensuite, il vérifie s'il existe un message unique m tel que $y^n \in T^{Y^n|x^n(m)}$. Si les deux tests sont positifs, il déclare m comme étant le message d'Alice.

Le taux de transmission sera déterminé par le nombre de $T^{Y^n|x^n(m)}$ qu'on peux "packer" dans T^{Y^n}.

Théorème de capacité de Shannon : Protocole



Théorème de capacité de Shannon : Idée de la preuve ($\exists$)

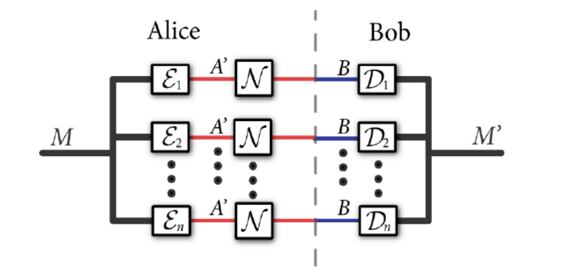
- Pour que Bob puisse décoder les messages d'Alice avec grande probabilité, les ensembles typiques conditionnels ne doivent pas trop "overlap".
- Intuitivement, Bob pourra décoder correctement si l'ensemble typique d'arrivée est divisé en M sous-ensembles de taille $2^{nH(Y|X)}$.
- Le nombres de messages distinguables M que Alice peux envoyer est donc au maximum

$$M = 2^{nR} = \frac{2^{nH(Y)}}{2^{nH(Y|X)}} = 2^{n(H(Y)-H(Y|X))} = 2^{nI(X;Y)}$$

- Les propriétés des ensembles typiques garantissent que $\mathcal{N}(x^n(m))$ appartient bien à $T_\delta^{Y^n|x^n(m)}$.

Capacité (classique) d'un canal bruité quantique

Approche naïve

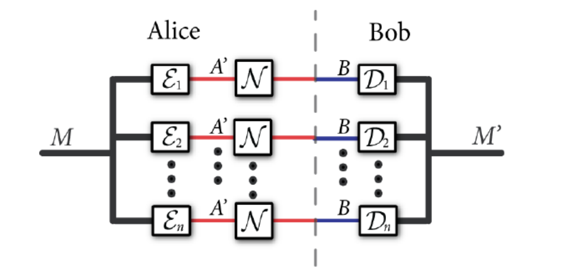


Modifications :

Alice et Bob s'ententent sur un ensemble $\{\rho_x\}$ d'opérateurs densité qui serviront d'input. Les mots de code sont maintenant de la forme

$$\rho_{x^n(m)} \equiv \rho_{x_1(m)} \otimes \rho_{x_2(m)} \cdots \otimes \rho_{x_n(m)}$$

Approche naïve

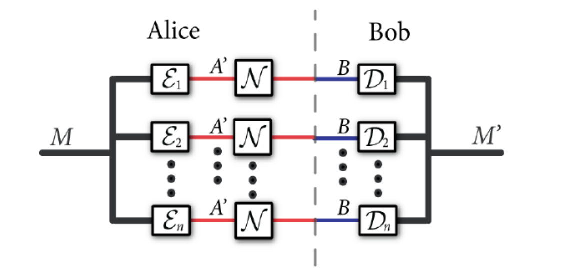


Modifications :

Bob mesure chaque $\mathcal{N}(\rho_{x_i(m)})$ avec un POVM $\{\Lambda_y\}$, induisant une distribution de probabilité conditionnelle

$$p_{Y|X}(y|x) \equiv \text{Tr}\{\Lambda_y \mathcal{N}(\rho_x)\}$$

Approche naïve



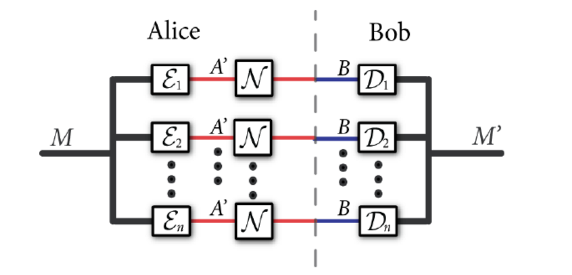
Modifications :

Le meilleur taux auquel ils peuvent communiquer est donc

$$I_{acc}(\mathcal{N}) \equiv \max_{p_X(x), \rho_x, \Lambda} I(X;Y)$$

où X, Y sont des V.A.D.

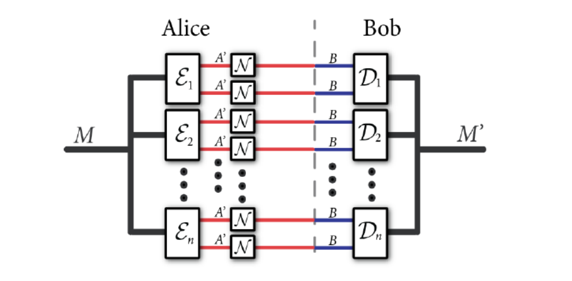
Approche naïve



Problème :

Cette façon de faire est essentiellement classique : pas d'intrication.

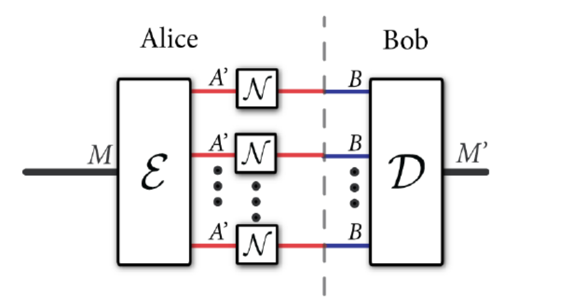
Approche moins naïve



Taux atteignable :

$$\frac{1}{2} I_{acc}(\mathcal{N} \otimes \mathcal{N}) \geq I_{acc}(\mathcal{N})$$

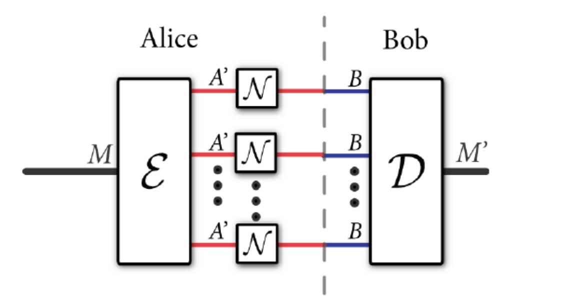
Approche moins naïve



Taux atteignable :

$$I_{reg}(\mathcal{N}) \equiv \lim_{k\to\infty} \frac{1}{k} I_{acc}(\mathcal{N}^{\otimes k}) \geq I_{acc}(\mathcal{N})$$

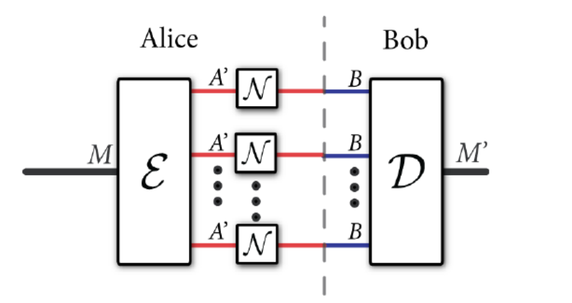
Approche moins naïve



Problème :

L'optimisation de cette quantité est pratiquement impossible.

Approche moins naïve



Problème :

L'optimisation de cette quantité est pratiquement impossible.

Solution :

Utiliser la borne du théorème d'Holevo ($I(X;Y) \leq \chi(\mathcal{N})$)

Information d'Holevo (Rappel?)

Définition

Soit $\mathcal{E} = \{(p_X(x), \rho_B^x)\}$ une source quantique. *L'information d'Holevo* $\chi(\mathcal{E})$ est une mesure de l'information classique accessible à propos de x sachant ρ_B^x :

$$\chi(\mathcal{E}) \equiv I(X;B)_\sigma$$

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

Information d'Holevo (Rappel?)

Définition

L'information d'Holevo $\chi(\mathcal{N})$ d'un canal quantique $\mathcal{N}$ est donnée par

$$\chi(\mathcal{N}) \equiv \max_{\rho_{XA}} I(X;B)_{\rho_{XB}}$$

où $\rho_{XA} = \sum_x p_X(x) |x\rangle\langle x|_X \otimes \rho_A^x$
et $\rho_{XB} = \sum_x p_X(x) |x\rangle\langle x|_X \otimes \mathcal{N}(\rho_A^x)_B$

Théorème d'Holevo : Énoncé

Théorème

Considérons les variables suivantes :

- $\mathcal{N} : A \to B$ un canal quantique;
- $C = \{x^n(m)\}_{m\in[M]}$ où les mots de code sont déterminés par $p_X(x)$;
- $\{\Lambda_y\}$ un POVM et Y la V.A.D. correspondant au résultat.

Alors

$$I(X;Y) \leq \chi(\mathcal{N})$$

Théorème d'Holevo : Idée de la preuve

Preuve

Soit $\rho_{XB} = \sum_x p_X(x) |x\rangle\langle x|_X \otimes \mathcal{N}(\rho_A^x)_B$ l'état *cq* définissant une communication sur le canal et soit $\mathcal{F} : B \to B \otimes Q$ le super-opérateur qui mesure B avec $\{\Lambda_y\}$ et qui place le résultat dans le registre Q. Alors on a

$$(\mathbb{I}_X \otimes \mathcal{F})\rho_{XB} = \sum_{x,y} p_X(x) |x\rangle\langle x|_X \otimes \sqrt{\Lambda_y}\mathcal{N}(\rho_A^x)_B \sqrt{\Lambda_y}^* \otimes |y\rangle\langle y|_Q$$

On a alors que

$$\chi(\mathcal{N}) \geq I(X;B)_\rho \geq I(X;BQ)_{(\mathbb{I}_X \otimes \mathcal{F})\rho} \geq I(X;Q)_{(\mathbb{I}_X \otimes \mathcal{F})\rho} = I(X;Y)$$

Théorème d'Holevo : Corollaire

Corollaire

La capacité classique pour la communication sur un canal quantique $\mathcal{N}$ est bornée supérieurement de la façon suivante :

$$
\begin{aligned}
C(\mathcal{N}) &= I_{reg}(\mathcal{N}) \\
&= \lim_{k\to\infty} \frac{1}{k} I_{acc}(\mathcal{N}^{\otimes k}) \\
&\leq \lim_{k\to\infty} \frac{1}{k} \chi(\mathcal{N}^{\otimes k}) \\
&= \chi_{reg}(\mathcal{N})
\end{aligned}
$$

Théorème de Holevo-Schumacher-Westmoreland (HSW)

Théorème

La capacité classique d'un canal quantique $\mathcal{N}$ est donné par la régularisation de l'information d'Holevo du canal :

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

Théorème HSW : Structure de la preuve

Structure semblable à la preuve du théorème de Shannon

- Ensembles typiques $\leftrightarrow$ Espace typique
- Entropie d'échantillion ($\overline{H}(x^n)$) $\leftrightarrow$ "Packing lemma"

Typicalité quantique

Soit $\mathcal{E} = \{(p_X(x), |\psi_x\rangle)\}_{x\in\mathcal{X}}$ une source quantique et

$$\rho_A = \sum_{x\in\mathcal{X}} p_X(x)\,|x\rangle\!\langle x|_A$$

sa matrice densité. Remarquons alors que

$$\rho_{A^n} = \sum_{x^n\in\mathcal{X}^n} p_{X^n}(x^n)\,|x^n\rangle\!\langle x^n|_{A^n}$$

donc on peux utiliser la typicalité classique sur les vecteurs propres.

Typicalité quantique : Sous-espace typique

Définition

Le sous-espace δ-*typique* $T^{\delta}_{A^n}$ est engendré par les états pures $|x^n\rangle_{A^n}$ où x^n est δ-typique :

$$T^{\delta}_{A^n} \equiv \mathrm{span}\{|x^n\rangle_{A^n} : x^n \in T^{X^n}_{\delta}\}$$

Définition

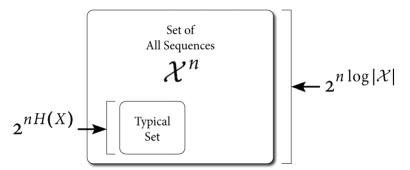
Notons $\Pi^{\delta}_{A^n}$ le projecteur sur $T^{\delta}_{A^n}$:

$$\Pi^{\delta}_{A^n} = \sum_{x^n \in T^{X^n}_{\delta}} |x^n\rangle\langle x^n|_{A^n}$$

Typicalité quantique : Propriétés

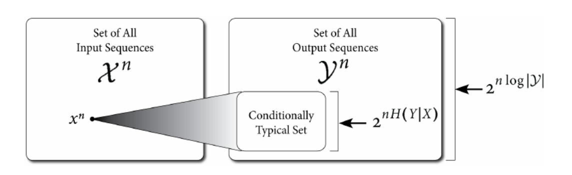
Lois des grands nombres $\implies$ Propriétés

1. $\mathrm{Tr}\{\Pi^{\delta}_{A^n}\rho_{A^n}\} \geq 1 - \epsilon$ pour tout $\varepsilon \in (0,1)$ et $\delta > 0$ lorsque n est grand.
2. $\mathrm{Tr}\{\Pi^{\delta}_{A^n}\} \leq 2^{n(H(A)+c\delta)}$ pour tout $\delta > 0$ lorsque n est grand.
3. Soit $|x^n\rangle, |y^n\rangle \in T^{\delta}_{A^n}$, alors $\mathrm{Tr}\{|x^n\rangle\langle x^n|\,\rho_{A^n}\} \approx_{\delta} \mathrm{Tr}\{|y^n\rangle\langle y^n|\,\rho_{A^n}\}$



Lois des grands nombres $\implies$ Propriétés

1. $\mathrm{Tr}\{\Pi^{\delta}_{A^n}\rho_{A^n}\} \geq 1-\epsilon$ pour tout $\varepsilon \in (0,1)$ et $\delta > 0$ lorsque n est grand.
2. $\mathrm{Tr}\{\Pi^{\delta}_{A^n}\} \leq 2^{n(H(A)+c\delta)}$ pour tout $\delta > 0$ lorsque n est grand.
3. Soit $|x^n\rangle, |y^n\rangle \in T^{\delta}_{A^n}$, alors
   $\mathrm{Tr}\{|x^n\rangle\langle x^n|\,\rho_{A^n}\} \approx_{\delta} \mathrm{Tr}\{|y^n\rangle\langle y^n|\,\rho_{A^n}\}$



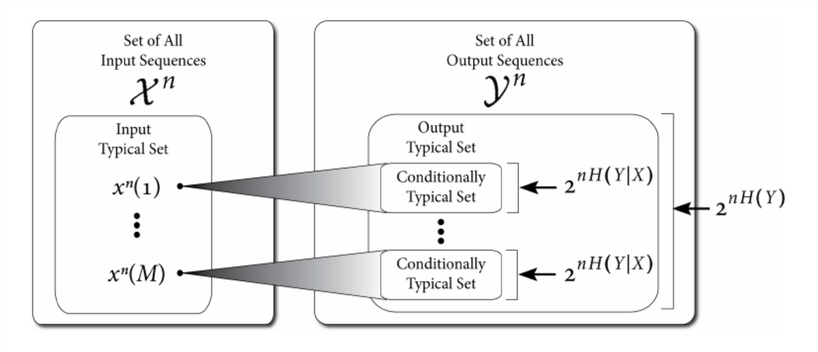
Théorème HSW : Protocole

Soit $\mathcal{N} : A \to B$ un canal bruité et X une V.A.D.

1. Alice et bob s'entendent sur un code aléatoire $\mathcal{C} = \{x^n(m)\}_{m\in[M]}$ en échantillionant m fois X^n et sur un emsembles d'opérateurs densité $\{\rho^x\}$.
2. Alice envoie $\rho^{x^n(m)} = \rho^{x_1(m)} \otimes \ldots \otimes \rho^{x_n(m)}$ à Bob à travers $\mathcal{N}$. Bob reçoit $\sigma^{x^n(m)} = \mathcal{N}^{\otimes n}(\rho^{x^n(m)})$.
3. Bob utilise le POVM $\{\Lambda_y\}$ du "packing lemma" pour déterminer avec grande probabilité dans quel sous-espace typique $T^{B^n|x^n(m)}$ se trouve $\sigma^{x^n(m)}$.

Le taux de transmission sera déterminé par le nombre de $T^{B^n|x^n(m)}$ qu'on peux "packer" dans T^{B^n}.

Théorème HSW : Protocole



"Packing lemma"

Théorème

Soient X une V.A.D. avec comme distribution $p_X(x)$ et $\{p_X(x), \sigma_x\}_{x\in\mathcal{X}}$ une distribution d'état quantique. Supposons qu'il un projecteur $\Pi = \sum_{x\in\mathcal{X}} \Pi_x$ ayant les propriétés suivantes :

- $\mathrm{Tr}\{\Pi\sigma_x\} \geq 1 - \varepsilon$
- $\mathrm{Tr}\{\Pi_x\sigma_x\} \geq 1 - \varepsilon$
- $\mathrm{Tr}\{\Pi_x\} \leq d$
- $\Pi\sigma\Pi \leq \frac{1}{D}\Pi$ ($\Pi\sigma\Pi$ environ l'état complètement mixte)

Alors soit $\mathcal{C} = \{C_m\}_{m\in\mathcal{M}}$ un code aléatoire. Il existe un POVM $\{\Lambda_m\}_{m\in\mathcal{M}}$ qui distingue presque toujours les états $\{\sigma_{C_m}\}$:

$$\mathbb{E}_{\mathcal{C}}\left\{\frac{1}{|\mathcal{M}|}\sum_{m\in\mathcal{M}}\mathrm{Tr}\{\Lambda_m\sigma_{C_m}\}\right\} \geq 1 - 2(\varepsilon + 2\sqrt{\varepsilon}) - 4|\mathcal{M}|\frac{d}{D}$$

"Packing lemma" : construction du POVM

On veux construire un POVM qui distingue les sous-espaces décrits par les Π_x.

Solution :

"Packing lemma" : construction du POVM

On veux construire un POVM qui distingue les sous-espaces décrits par les Π_x.

Solution : Pretty good measurement! (☹)

"Packing lemma" : construction du POVM

On veux construire un POVM qui distingue les sous-espaces décrits par les Π_x.

Solution :

Un bon candidat est le POVM décrit par les opérateurs suivants :

$$\Lambda_m \equiv \left(\sum_{i=1}^{|\mathcal{M}|} \Upsilon_{c_i}\right)^{-1/2} \Upsilon_{c_m} \left(\sum_{i=1}^{|\mathcal{M}|} \Upsilon_{c_i}\right)^{-1/2}$$

où $\Upsilon_x \equiv \Pi \Pi_x \Pi$ s'assure simplement que $\text{supp}\{\Upsilon_x\} \subseteq \text{supp}\{\Pi\}$

"Packing lemma" : Analyse de l'erreur

$$
\begin{aligned}
p_e(m, \mathcal{C}) &= \mathrm{Tr}\{(\mathbb{I} - \Lambda_m)\sigma_{c_m}\} \\
&\leq 2\,\mathrm{Tr}\{(\mathbb{I} - \Upsilon_{c_m})\sigma_{C_m}\} + 4 \sum_{m' \neq m}^{|\mathcal{M}|} \mathrm{Tr}\{\Upsilon_{C_{m'}}\sigma_{C_m}\}
\end{aligned}
$$

"Packing lemma" : Analyse de l'erreur

$$
\begin{aligned}
p_e(m, \mathcal{C}) &= \mathrm{Tr}\{(\mathbb{I} - \Lambda_m)\sigma_{c_m}\} \\
&\leq 2\,\mathrm{Tr}\{(\mathbb{I} - \Upsilon_{c_m})\sigma_{C_m}\} + 4 \sum_{m' \neq m}^{|\mathcal{M}|} \mathrm{Tr}\{\Upsilon_{C_{m'}}\sigma_{C_m}\} \\
&\leq 2(\varepsilon + 2\sqrt{\varepsilon}) + 4 \sum_{m' \neq m}^{|\mathcal{M}|} \mathrm{Tr}\{\Upsilon_{C_{m'}}\sigma_{C_m}\}
\end{aligned}
$$

"Packing lemma" : Analyse de l'erreur

$$\overline{p}_e(\mathcal{C}) = \frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|}\mathrm{Tr}\{(\mathbb{I} - \Lambda_m)\sigma_{C_m}\}$$

$$\leq 2(\varepsilon + 2\sqrt{\varepsilon}) + \frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|}\sum_{m'\neq m}^{|\mathcal{M}|}\mathrm{Tr}\{\Upsilon_{C_{m'}}\sigma_{C_m}\}$$

"Packing lemma" : Analyse de l'erreur

$$\mathbb{E}_{\mathcal{C}}\{\overline{p}_e(\mathcal{C})\} = \mathbb{E}_{\mathcal{C}}\left\{\frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|}\text{Tr}\{(\mathbb{I}-\Lambda_m)\sigma_{C_m}\}\right\}$$

$$\leq 2(\varepsilon + 2\sqrt{\varepsilon}) + \frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|}\sum_{m'\neq m}^{|\mathcal{M}|}\mathbb{E}_{\mathcal{C}}\{\text{Tr}\{\Upsilon_{C_{m'}}\sigma_{C_m}\}\}$$

"Packing lemma" : Analyse de l'erreur

$$
\begin{aligned}
\mathbb{E}_{\mathcal{C}}\{\mathrm{Tr}\{\Upsilon_{C_{m'}}\sigma_{C_m}\}\} &= \mathbb{E}_{\mathcal{C}}\{\mathrm{Tr}\{\Pi\Pi_{C_{m'}}\Pi\sigma_{C_m}\}\} \\
&= \mathbb{E}_{\mathcal{C}}\{\mathrm{Tr}\{\Pi_{C_{m'}}\Pi\sigma_{C_m}\Pi\}\} \\
&= \mathbb{E}_{C_m,C_{m'}}\{\mathrm{Tr}\{\Pi_{C_{m'}}\Pi\sigma_{C_m}\Pi\}\} \\
&= \mathrm{Tr}\{\mathbb{E}_{C_{m'}}\{\Pi_{C_{m'}}\}\Pi\mathbb{E}_{C_m}\{\sigma_{C_m}\}\Pi\} \\
&= \mathrm{Tr}\{\mathbb{E}_{C_{m'}}\{\Pi_{C_{m'}}\}\Pi\sigma\Pi\} \\
&\leq \frac{1}{D}\,\mathrm{Tr}\{\mathbb{E}_{C_{m'}}\{\Pi_{C_{m'}}\}\Pi\} \\
&\leq \frac{1}{D}\,\mathrm{Tr}\{\mathbb{E}_{C_{m'}}\{\Pi_{C_{m'}}\}\} \\
&= \frac{1}{D}\mathbb{E}_{C_{m'}}\{\mathrm{Tr}\{\Pi_{C_{m'}}\}\} \\
&\leq \frac{d}{D}
\end{aligned}
$$

"Packing lemma" : Analyse de l'erreur

$$\mathbb{E}_{\mathcal{C}}\{\overline{p}_e(\mathcal{C})\} = \mathbb{E}_{\mathcal{C}}\left\{\frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|}\mathrm{Tr}\{(\mathbb{I}-\Lambda_m)\sigma_{C_m}\}\right\}$$

$$\leq 2(\varepsilon + 2\sqrt{\varepsilon}) + \frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|}\sum_{m'\neq m}^{|\mathcal{M}|}\mathbb{E}_{\mathcal{C}}\{\mathrm{Tr}\{\Upsilon_{C_{m'}}\sigma_{C_m}\}\}$$

"Packing lemma" : Analyse de l'erreur

$$
\begin{aligned}
\mathbb{E}_{\mathcal{C}}\{\overline{p}_e(\mathcal{C})\} &= \mathbb{E}_{\mathcal{C}}\left\{\frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|}\mathrm{Tr}\{(\mathbb{I}-\Lambda_m)\sigma_{C_m}\}\right\} \\
&\leq 2(\varepsilon + 2\sqrt{\varepsilon}) + \frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|}\sum_{m'\neq m}^{|\mathcal{M}|}\mathbb{E}_{\mathcal{C}}\{\mathrm{Tr}\{\Upsilon_{C_{m'}}\sigma_{C_m}\}\} \\
&\leq 2(\varepsilon + 2\sqrt{\varepsilon}) + \frac{1}{|\mathcal{M}|}\sum_{m=1}^{|\mathcal{M}|}\sum_{m'\neq m}^{|\mathcal{M}|}\frac{d}{D} \\
&\leq 2(\varepsilon + 2\sqrt{\varepsilon}) + 4|\mathcal{M}|\frac{d}{D}
\end{aligned}
$$

"Packing lemma" : Retouche finale

Le théorème qu'on vient de prouver montre que l'espérance de la probabilité d'erreur est petite. Les deux arguments suivant prouvent l'existance d'un code $\mathcal{C}_0$ avec $p^*(\mathcal{C}_0)$ petit :

1. $\mathbb{E}_{\mathcal{C}}\{\overline{p}_e(\mathcal{C})\} \leq B \implies (\exists \mathcal{C}_0)[\overline{p}_e(\mathcal{C}_0) \leq B]$
2. $\overline{p}_e(\mathcal{C}_0) \leq B \implies p_e(m) \leq B$ pour au moins la moitié des messages m. On peut donc construire un code $\mathcal{C}_0'$ avec ces messages. Le taux de transmission est assimptotiquement le même puisque $\frac{2^{nR}}{2} = 2^{nR-1} = 2^{n(R-\frac{1}{n})}$. On a alors $p^*(\mathcal{C}_0') \leq B$.

Théorème HSW : Protocole

Soit $\mathcal{N} : A \to B$ un canal bruité et X une V.A.D.

1. Alice et bob s'entendent sur un code aléatoire $\mathcal{C} = \{x^n(m)\}_{m \in [M]}$ en échantillionant m fois X^n et sur un emsembles d'opérateurs densité $\{\rho^x\}$.
2. Alice envoie $\rho^{x^n(m)} = \rho^{x_1(m)} \otimes \ldots \otimes \rho^{x_n(m)}$ à Bob à travers $\mathcal{N}$. Bob reçoit $\sigma^{x^n(m)} = \mathcal{N}^{\otimes n}(\rho^{x^n(m)})$.
3. Bob utilise le POVM $\{\Lambda_y\}$ du "packing lemma" pour déterminer avec grande probabilité dans quel sous-espace typique $T^{B^n|x^n(m)}$ se trouve $\sigma^{x^n(m)}$.

Le taux de transmission sera déterminé par le nombre de $T^{B^n|x^n(m)}$ qu'on peux "packer" dans T^{B^n}.

"Packing lemma"

Théorème

Soient X une V.A.D. avec comme distribution $p_X(x)$ et $\{p_X(x), \sigma_x\}_{x \in \mathcal{X}}$ une distribution d'état quantique. Supposons qu'il un projecteur $\Pi = \sum_{x \in \mathcal{X}} \Pi_x$ ayant les propriétés suivantes :

- $\mathrm{Tr}\{\Pi \sigma_x\} \geq 1 - \varepsilon$
- $\mathrm{Tr}\{\Pi_x \sigma_x\} \geq 1 - \varepsilon$
- $\mathrm{Tr}\{\Pi_x\} \leq d$
- $\Pi \sigma \Pi \leq \frac{1}{D} \Pi$ ($\Pi \sigma \Pi$ environ l'état complètement mixte)

Alors soit $\mathcal{C} = \{C_m\}_{m \in \mathcal{M}}$ un code aléatoire. Il existe un POVM $\{\Lambda_m\}_{m \in \mathcal{M}}$ qui distingue presque toujours les états $\{\sigma_{C_m}\}$:

$$\mathbb{E}_{\mathcal{C}} \left\{ \frac{1}{|\mathcal{M}|} \sum_{m \in \mathcal{M}} \mathrm{Tr}\{\Lambda_m \sigma_{C_m}\} \right\} \geq 1 - 2(\varepsilon + 2\sqrt{\varepsilon}) - 4|\mathcal{M}| \frac{d}{D}$$

Théorème HSW : Application du "Packing lemma"

- L'ensemble d'état est $\{p_{X^n}(x^n), \mathcal{N}^{\otimes n}(\rho^{x^n(m)})\}_{m\in\mathcal{M}}$.
- Le projecteur sur le code est $\Pi^{\delta}_{B^n}$.
- Les projecteurs sur les mots de codes sont $\Pi^{\delta}_{B^n|x^n}$
- $d = 2^{H(B|X)+c\delta}$ et $D = 2^{H(B)+c'\delta}$

On a alors

- $\mathrm{Tr}\{\Pi^{\delta}_{B^n}\mathcal{N}^{\otimes n}(\rho^{x^n(m)})\} \geq 1-\varepsilon$
- $\mathrm{Tr}\{\Pi^{\delta}_{B^n|x^n}\mathcal{N}^{\otimes n}(\rho^{x^n(m)})\} \geq 1-\varepsilon$
- $\mathrm{Tr}\{\Pi^{\delta}_{B^n|x^n}\} \leq 2^{n(H(B|X)+c\delta)}$
- $\Pi^{\delta}_{B^n}\mathbb{E}_{X^n}\{\mathcal{N}^{\otimes n}(\rho^{x^n(m)})\Pi^{\delta}_{B^n}\} \leq \frac{1}{1-\epsilon}\cdot\frac{1}{2^{n(H(B)+c'\delta)}}\Pi^{\delta}_{B^n}$

Ainsi, $\mathcal{M}$ peut avoir comme taille

$$|\mathcal{M}| \approx \frac{D}{d} = \frac{2^{nH(B)}}{2^{nH(B|X)}} = 2^{n(H(B)-H(B|X))} = 2^{nI(X;B)}$$