ACCQ202 23 Octobre 2015

Cours 5

Enseignant: A. Tchamkerten *Scribes: R. Aymeric and R. Yehia*

1 Preuve du théorème de codage de canal

1.1 Réciproque

Soit $W \sim Unif[1, 2, ..., 2^{nR}]$. Alors

$$\begin{aligned} nR &= H(W) \\ &= H(W) - H(W|Y^n) + H(W|Y^n) \\ &= I(W; Y^n) + H(W|Y^n) \end{aligned}$$

Or d'après le Data Processing Inequality et le lemme de Fano on a respectivement :

$$I(W; Y^n) \leq I(X^n; Y^n)$$

et

$$H(W|Y^n) \leq 1 + nR \cdot P_n(\hat{W} \neq W)$$

d'où

$$nR \leq I(X^n; Y^n) + 1 + nR \cdot P_n(\hat{W} \neq W)$$

Ce qui donne

$$nR \leq n \cdot C + 1 + nR \cdot P_n(\hat{W} \neq W)$$

$$\Rightarrow R \leq C + \frac{1}{n} + R \cdot P_n(\hat{W} \neq W)$$

Donc si $\{(2^{nR}, n)\}_{n \geq 1}$ est tel que $P_n(\hat{W} \neq W) \underset{n \to +\infty}{\longrightarrow} 0$ alors:

$$R \leq C$$

On a montré que pour tout code dont l'erreur tend vers zéro le rendement est au plus égale à la capacité du canal.

1.2 Partie directe

A montrer: pour tout $R < C$ il existe $\{(2^{nR}, n)\}_{n\geq 1}$ tel que

$$P_n(\hat{W} \neq W) \underset{n\to+\infty}{\longrightarrow} 0$$

Pour $Q(y|x)$ donné, $x \in \S$, $y \in \mathcal{Y}$, et pour $p_x(x)$ quelconque sur $\mathcal{X}$, on génère le code aléatoirement:

$$\mathcal{C} = \begin{pmatrix} x_1(1) & x_2(1) & ... & x_n(1) \\ x_1(2) & x_2(2) & ... & x_n(2) \\ ... & ... & ... & ... \\ x_1(2^{nR}) & x_2(2^{nR}) & ... & x_n(2^{nR}) \end{pmatrix}$$

où les x_i sont des instances tirées aléatoiremement selon $p_X(x)$. La probabilité d'un code particulier est donc $P(\mathcal{C}) = \prod_{i=1}^{n} \prod_{j=1}^{2^{nR}} p_X(x_i(j))$

Une fois $\mathcal{C}$ généré, il est révélé à l'émetteur et au récepteur.

Décodeur: Il va choisir déclarer W tel que $x^n(W)$ est typique avec le vecteur reçu Y^n. Si plusieurs messages sont typiques avec Y^n , le décodeur choisit un de ces messages alléatoirement. S'il n'y en a aucun, le décodeur déclare un message au hasard.

Calculons la probabilité d'erreur moyennée sur l'ensemble des codes:

$$\overline{P_{\mathcal{C}}(\hat{W} \neq W)} = \sum_{\mathcal{C}} P(\mathcal{C}) \cdot P(\hat{W} \neq W|\mathcal{C})$$

$$\overline{P_{\mathcal{C}}(\hat{W} \neq W)} = \sum_{\mathcal{C}} P(\mathcal{C}) \cdot \sum_{i=1}^{2^{nR}} P(\hat{W} \neq i|W = i, \mathcal{C}) \cdot \frac{1}{2^{nR}}$$

$$\overline{P_{\mathcal{C}}(\hat{W} \neq W)} = \frac{1}{2^{nR}} \cdot \sum_{i=1}^{2^{nR}} \sum_{\mathcal{C}} P(\mathcal{C}) \cdot P(\hat{W} \neq i|W = i, \mathcal{C})$$

On a de plus $\sum_{\mathcal{C}} P(\mathcal{C}) \cdot P(\hat{W} \neq i|W = i, \mathcal{C}) = \overline{P_{\mathcal{C}}(\hat{W} \neq 1|W = 1)}$ par symétrie du codage et du décodage.

On a donc

$$\overline{P_{\mathcal{C}}(\hat{W} \neq W)} = \overline{P_{\mathcal{C}}(\hat{W} \neq 1 | W = 1)}$$

Soit l'évènement $E_i = \{(X^n(i), Y^n) \in \tilde{\mathcal{A}}_\epsilon^n\}$

$$\overline{P_{\mathcal{C}}(\hat{W} \neq 1 | W = 1)} = P(E_1^c \cup E_2 \cup ... \cup E_{2^{nR}})$$

$$\overline{P_{\mathcal{C}}(\hat{W} \neq 1 | W = 1)} \leq P(E_1^c) + \sum_{i=2}^{2^{nR}} P(E_i)$$

Avec n suffisament grand

$$\overline{P_{\mathcal{C}}(\hat{W} \neq 1 | W = 1)} \leq \epsilon + \sum_{i=2}^{2^{nR}} 2^{-n(I(X;Y)-3\epsilon)}$$

$$\overline{P_{\mathcal{C}}(\hat{W} \neq 1 | W = 1)} \leq \epsilon + 2^{nR} \cdot 2^{-n(I(X;Y)-3\epsilon)}$$

Donc pour $R < I(X;Y) - 3\epsilon$ et n suffisamment grand on a

$$\overline{P_{\mathcal{C}}(\hat{W} \neq 1 | W = 1)} \leq 2\epsilon$$

Dernire tape: choisir p_X pour maximisant $I(X;Y)$. Dans ce cas on a pour $R < C - 3\epsilon$ et n suffisamment grand

$$\overline{P_{\mathcal{C}}(\hat{W} \neq 1 | W = 1)} \leq 2\epsilon$$

La moyenne étant inférieure à 2ϵ il existe un code $\mathcal{C}$ tel que $P(\hat{W} \neq W | \mathcal{C}) \leq 2\epsilon$ (argument probabiliste).

Jusqu'ici la probabilité d'erreur est moyennée sur les messages d'entrée et cela suffit pour conclure la preuve du théorème. Cependant on peut également étendre la conclusion au cas où la probabilité dérreur est non plus moyennée sur les mots code mais est définie comme étant la probabilité d'erreur maximale sur les messages. On note que si on élimine la moitié des mots de code de $\mathcal{C}$ dont la probabilité d'erreur est la plus haute on obtient $\mathcal{C}'$ tel que la probabilité d'erreur maximale satisfait

$$max_w P(\hat{W} \neq w | W = w, \mathcal{C}') \leq 2\epsilon \cdot 2 = 2\epsilon$$

Le taux de ce code est $\frac{nR-1}{n} = R - \frac{1}{n} = R'$ donc quasiment pas de perte pour n suffisamment grand.

On déduit que même sous le critère plus contraignant de probabilité d'erreur maximale tout taux $R < C$ est atteignable.