Cours n°7 : codes polaires

Scribe : Charles Moury

Comment construire des codes qui permettent d'atteindre la capacité ?
En 2008, Erdal Arikan propose cette construction de codes qui atteignent la capacité pour un grand nombre de canaux.

Propriétés de l'information mutuelle :

Rappel : $(X,Y) \sim p(x,y)$
$I(X,Y) = H(Y) - H(Y \mid X)$
$\quad = H(X) - H(X \mid Y)$

Définition : (X,Y,Z)
$I(X;Y \mid Z) = H(Y \mid Z) - H(Y \mid X,Z)$

Proposition :

- $H(X_1,\dots,X_n) = \sum_{i=1}^{n} H(X_i \mid X^{i-1})$ où $X^{i-1} \triangleq (X_1,\dots,X_{i-1})$
- $I(X_1,\dots,X_n;Y) = \sum_{i=1}^{n} I(X_i;Y \mid Y^{i-1})$

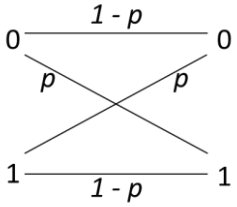
Preuve :

- $H(X_1,\dots,X_n) = -\mathbb{E}_{(X_1,\dots,X_n)}[log(p(X_1,\dots,X_n))]$
  $et\ log(p(X_1,\dots,X_n)) = log\left(\prod_{i=1}^{n} p(X_i \mid X^{i-1})\right)$
  $$donc\ H(X_1,\dots,X_n) = -\sum_{i=1}^{n} \mathbb{E}_{(X_1,\dots,X_n)}\left[log\left(p(X_i \mid X^{i-1})\right)\right]$$
  $$\triangleq \sum_{i=1}^{n} H(X_i \mid X^{i-1})$$
- $I(X_1,\dots,X_n;Y) = H(X^n) - H(X^n \mid Y)$
  $$= \sum_{i=1}^{n}\left[H(X_i \mid X^{i-1}) - H(X_i \mid Y;X^{i-1})\right]$$
  $et\ H(X_i \mid X^{i-1}) - H(X_i \mid Y,X^{i-1}) = I(X_i;Y \mid X^{i-1})$

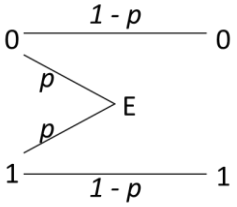
Définition : Un canal $Q(.|.)$ à entrée binaire est à sortie symétrique s'il existe une permutation $\pi : y^n \to y^n$ telle que :

- $\pi^2 = 1$
- $\forall y,\ Q(y \mid 0) = Q(\pi(y) \mid 1)$

BSC :



BEC :



Remarque : La capacité d'un tel canal est atteinte pour la distribution uniforme en entrée :

$$argmax_X I(X,Y) = Bern\left(\frac{1}{2}\right)$$

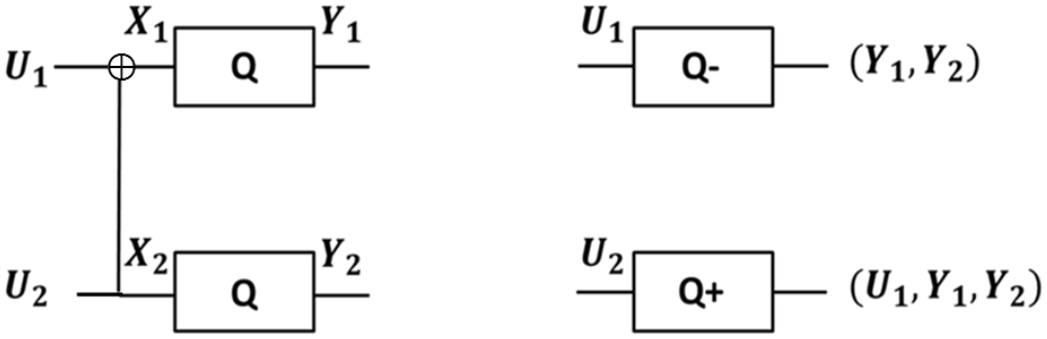
Théorème : (E. Arikan, 2008)

Soit Q un canal à entrée binaire et à sortie symétrique de capacité $I_0 = I(Q)$.
Soient $0 \leq R \leq I_0$ et $0 < \varepsilon < 1$.
Alors il existe un code polaire de longueur $N = 2^n, n \geq 1$, de taux R et tel que ${P_e}^N \leq \varepsilon$.

Remarque : ${P_e}^N \sim O\left(e^{-\sqrt{N}}\right)$, la convergence asymptotique des codes polaires est meilleure que celle des codes LDPC.
Néanmoins la constante de domination est très grande et pour des mots code de petite taile un code LDPC reste plus efficace qu'un code polaire.

Qu'est-ce que la polarisation ?
Elle transforme une famille de canaux moyens en une famille de canaux soit très bons, soit très mauvais.

Voici la transformation fondamentale des codes polaires :



« Combinaison de canaux » « Séparation de canaux »

$$Q \longrightarrow (Q^-, Q^+)$$

$$U_1, U_2 \sim Bern\left(\frac{1}{2}\right) \ i.i.d.$$

<u>Proposition :</u>

1. $I(U_1, U_2 ; Y_1, Y_2) = I(X_1, X_2 ; Y_1, Y_2)$
2. $I(X_1, X_2 ; Y_1, Y_2) = 2I_0$
3. $I(U_1, U_2 ; Y_1, Y_2) = I(U_1; Y_1, Y_2) + I(U_2; U_1, Y_1, Y_2)$

<u>Preuve :</u>

2. $I(X_1, X_2 ; Y_1, Y_2) \overset{\text{def}}{=} H(Y_1, Y_2) - H(Y_1, Y_2 \mid X_1, X_2)$
$= H(Y_1) + H(Y_2) - H(Y_1 \mid X_1, X_2) - H(Y_2 \mid X_1, X_2, Y_1)$
$= I(X_1; Y_1) + I(X_2; Y_2)$

et $I(X_1; Y_1) = I(X_2; Y_2) = I_0$

donc $I(X_1, X_2 ; Y_1, Y_2) = 2I_0$

3. $I(U_1, U_2 ; Y_1, Y_2) = I(U_1 ; Y_1, Y_2) + I(U_2 ; Y_1, Y_2 \mid U_1)$

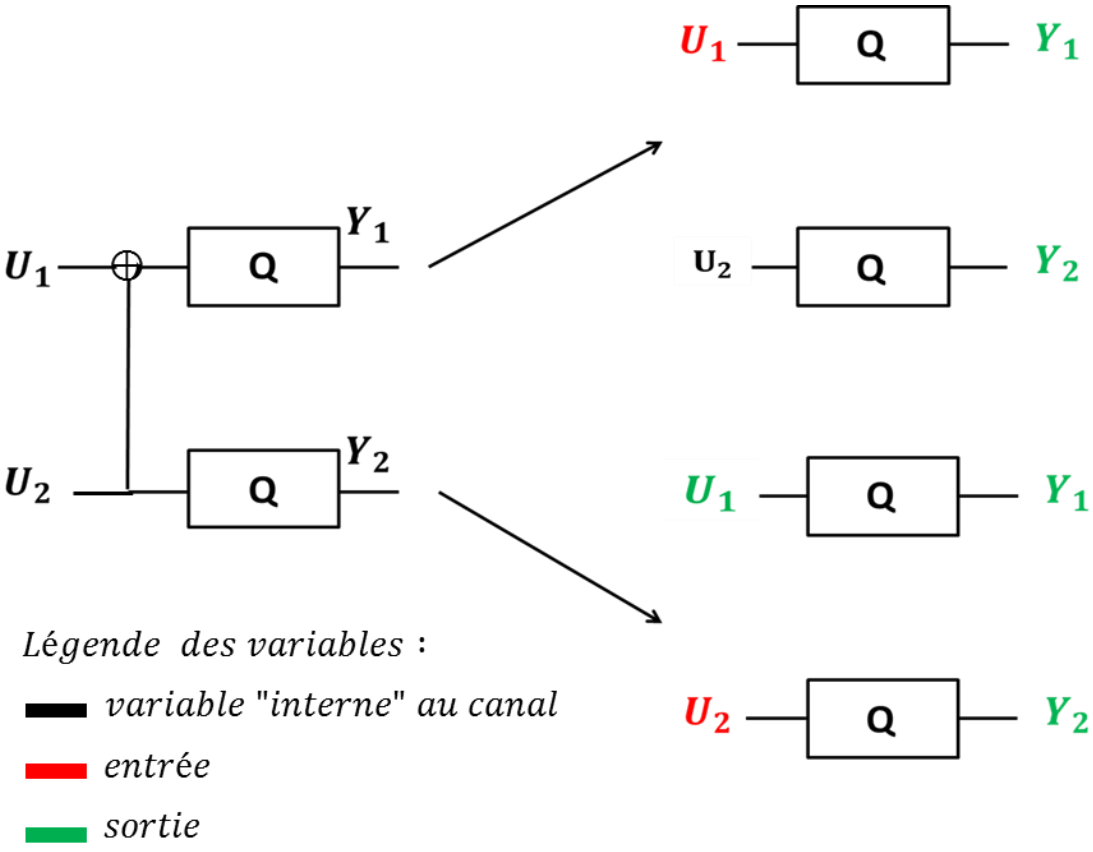
or $I(U_2 ; Y_1, Y_2 \mid U_1) = H(U_2 \mid U_1) - H(U_2 \mid U_1, Y_1, Y_2)$

et $H(U_2 \mid U_1) = H(U_2)$

donc $I(U_1, U_2 ; Y_1, Y_2) = I(U_1; Y_1, Y_2) + I(U_2; U_1, Y_1, Y_2)$

Retour sur la séparation des canaux :
Remarque : U_2 correspond à une randomisation au niveau du bruit, c'est une variable interne de Q^-.

$Q^-: U_1 \to (Y_1, Y_2)$
$Q^+: U_2 \to (U_1, Y_1, Y_2)$



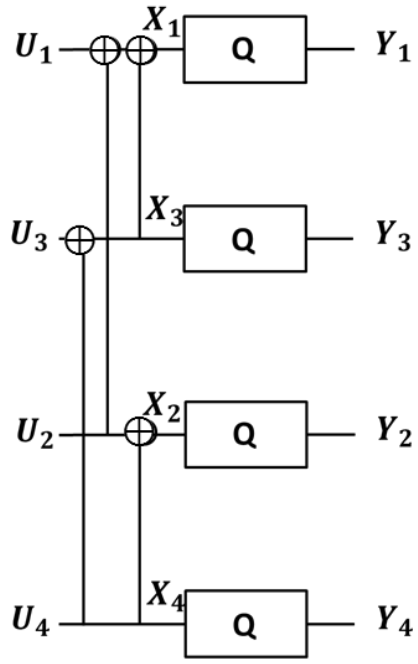
$I(U_1, U_2\,; Y_1, Y_2) = I(U_1; Y_1, Y_2) + I(U_2; U_1, Y_1, Y_2) = I(Q^+) + I(Q^-) = 2I_0$
$Donc \dfrac{I(Q^+) + I(Q^-)}{2} = I_0$

$I(Q^+) = I\big(U_2\,; U_1, Y_1, Y_2\big) \triangleq I(U_2; Y_2) + \cdots = I_0 + \cdots \; donc \; I(Q^+) \geq I_0$
$De\ même\ I(Q^-) \leq I_0$
$D'où\ \ I(Q^-) \leq I_0 \leq I(Q^+)$

Itération de la transformation fondamentale :

$Q \to (Q^-, Q^+) \to (Q^{--}, Q^{-+}, Q^{+-}, Q^{++}) \to etc.$

2ième itération : $(Q^-, Q^+) \to (Q^{--}, Q^{-+}, Q^{+-}, Q^{++})$

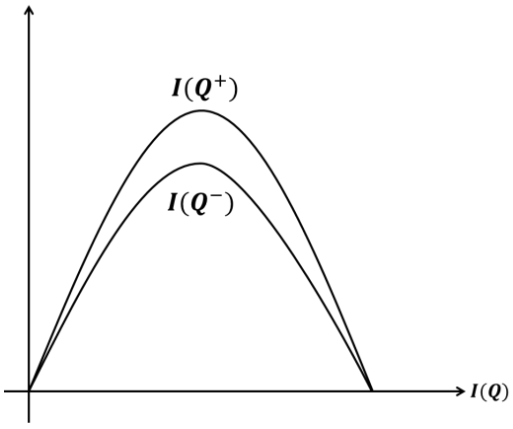


$$\equiv (\mathbf{X_1}, \mathbf{X_2}, \mathbf{X_3}, \mathbf{X_4}) = (\mathbf{U_1}, \mathbf{U_2}, \mathbf{U_3}, \mathbf{U_4}) \begin{pmatrix} \mathbf{1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{1} & \mathbf{0} & \mathbf{1} & \mathbf{0} \\ \mathbf{1} & \mathbf{1} & \mathbf{1} & \mathbf{1} \end{pmatrix}$$

Cas général : $(X_1, \dots, X_{2^n}) = (U_1, \dots, U_{2^n}) \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}^{\otimes n}$

nième itération : $\left\{Q^{b_1, b_2, \dots, b_n};\ b_i \in \{+, -\}\right\}$

Lemme (du progrès garanti) :
$\forall \varepsilon, 0 < \varepsilon < 1, \exists \delta > 0$ *tel que si* $I(Q) \in (\varepsilon, 1-\varepsilon)$,
Alors $|I(Q^+) - I(Q^-)| > \delta$ *et* $\varepsilon(\delta) \underset{\delta \to 0}{\longrightarrow} 0$



Théorème :

Soit $\varepsilon > 0$.

On définit :

$$K_n{}^{+} = \frac{1}{2^n} \#\{Q^{b_1,b_2,\dots,b_n}, I(Q^{b_1,b_2,\dots,b_n}) \geq 1 - \varepsilon\}$$

$$K_n{}^{-} = \frac{1}{2^n} \#\{Q^{b_1,b_2,\dots,b_n}, I(Q^{b_1,b_2,\dots,b_n}) \leq \varepsilon\}$$

$$K_n{}^{0} = \frac{1}{2^n} \#\{Q^{b_1,b_2,\dots,b_n}, \varepsilon < I(Q^{b_1,b_2,\dots,b_n}) < 1 - \varepsilon\}$$

Alors :

$$K_n{}^{+} \xrightarrow[n\to\infty]{} I_0$$

$$K_n{}^{-} \xrightarrow[n\to\infty]{} 1 - I_0$$

$$K_n{}^{0} \xrightarrow[n\to\infty]{} 0$$

Preuve : voir note sur la polarisation

Codage/Décodage :

- Fixer $\varepsilon > 0$.
- Choisir $N = 2^n$ grand.
- Ordonner les canaux : $1 \geq I(Q_1) \dots \geq I(Q_N)$
- Considérer les canaux : $I(Q_i) \geq 1 - \varepsilon$ et envoyer de l'info sur les bits d'entrée correspondants et « 0 » sur les bits d'entrée des mauvais canaux.

Décodage successif :

$$\widehat{U_1} = \frac{Q(Y_1, \dots, Y_{2^n} \mid U_1 = 0)}{Q(Y_1, \dots, Y_{2^n} \mid U_1 = 1)} = L_1$$

$$\begin{cases} si : U_1 \text{ gelé } \rightarrow \widehat{U_1} = 0 \\ sinon : \begin{cases} si\ L_1 > 1\ alors\ \widehat{U_1} = 0 \\ si\ L_1 \leq 1\ alors\ \widehat{U_1} = 1 \end{cases} \end{cases}$$

$$\widehat{U_2} = \frac{Q(Y_1, \dots, Y_{2^n}, \widehat{U_1} \mid U_2 = 0)}{Q(Y_1, \dots, Y_{2^n}, \widehat{U_1} \mid U_2 = 1)} = L_2$$

$$\begin{cases} si : U_2 \text{ gelé } \rightarrow \widehat{U_2} = 0 \\ sinon : \begin{cases} si\ L_2 > 1\ alors\ \widehat{U_2} = 0 \\ si\ L_2 \leq 1\ alors\ \widehat{U_2} = 1 \end{cases} \end{cases}$$

Complexité : $O\big(Nlog(N)\big)$