PROJET DE SIMULATION

Communication en présence de bruit avec retour

3 juillet 2016

Jocelyn Beauchesne
Gabriel Misrachi

ÉCOLE POLYTECHNIQUE
UNIVERSITÉ PARIS-SACLAY

Préliminaires théoriques

1. (a) On pose $\forall x \in R_+^*,\ g(x) = x \cdot ln(x)$.
Alors,
$$\forall x \in R_+^*,\ g'(x) = ln(x) + 1 \text{ et } g''(x) = \frac{1}{x} > 0 \tag{1}$$
D'où g est convexe et le tableau de variation suivant :

x	0		$\frac{1}{e}$		1		$+\infty$
$g'(x)$		$-$	0	$+$		$+$	
$g(x)$	$\Vert\ 0$	$\searrow$	$-\frac{1}{e}$	$\nearrow$	0	$\nearrow$	$+\infty$

On pose alors
$$A = \left\{ y \in R \text{ tel que } \frac{f_X(y)}{f_Y(y)} < 1 \right\} \tag{2}$$
De sorte qu'en écrivant
$$D(f_X||f_Y) = \int_{R-A} \frac{f_X(y)}{f_Y(y)} ln\left(\frac{f_X(y)}{f_Y(y)}\right) \cdot f_Y(y)dy + \int_A \frac{f_X(y)}{f_Y(y)} ln\left(\frac{f_X(y)}{f_Y(y)}\right) \cdot f_Y(y)dy \tag{3}$$
On ait :
$$\begin{aligned} 0 < \mathbb{E}^- &= -\int_A \frac{f_X(y)}{f_Y(y)} ln\left(\frac{f_X(y)}{f_Y(y)}\right) \cdot f_Y(y)dy \\ &< \int_A \frac{1}{e} \cdot f_Y(y)dy \\ &< \frac{1}{e} \end{aligned}$$
D'où $D(f_X||f_Y)$ est bien définie (éventuellement $= +\infty$).

(b) Puisque d'après (1), g est convexe, on peut appliqué l'inégalité de Jensen intégrale à $D(f_X||f_Y)$:
$$\begin{aligned} D(f_X||f_Y) &= \mathbb{E}\left(g\left(\frac{f_X(y)}{f_Y(y)}\right)\right) \\ &\geq g\left(\mathbb{E}\left(\frac{f_X(y)}{f_Y(y)}\right)\right) \\ &\geq g(1) \\ &\geq 0 \end{aligned}$$
Qui plus est, le cas d'égalité de l'inégalité de Jensen dan le cas strictement convexe nous indique que $U = \frac{f_X(Y)}{f_Y(Y)}$ est constante presque partout. Or U est continue et son espérance vaut 1. Ainsi, en cas d'égalité, U = 1. C'est à dire $f_X = f_Y$. Autrement dit, X et Y suivent la même loi.

(c) Prenons Y suivant une loi $\mathcal{N}(0, \sigma^2)$ où $\sigma^2 = Var(X)$. On a donc :

$$\forall y \in \mathbb{R},\ f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} \tag{4}$$

Puis :

$$\begin{aligned} D(f_X||f_Y) &= \int f_X(y) \left[ln\left(f_X(y)\right) - ln\left(\frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{y^2}{2\sigma^2}} \right) \right] dy \\ &= -h(X) + \int ln\left(\sqrt{2\pi\sigma^2}\right) f_X(y)dy + \frac{1}{2\sigma^2} \int y^2 f_X(y)dy \\ &= -h(X) + ln\left(\sqrt{2\pi\sigma^2}\right) + \frac{1}{2\sigma^2}\mathbb{E}(X^2) \end{aligned}$$

Or, comme X est centrée, $\mathbb{E}(X^2) = Var(X) = \sigma^2$ et :

$$D(f_X||f_Y) = -h(X) + ln\left(\sqrt{2\pi}\right) + ln\left(\sqrt{\sigma^2}\right) + \frac{1}{2} \tag{5}$$

Finalement, par positivité de la divergence de Kullback obtenue à la question précédente on obtient bien :

$$h(X) \leq c + ln\left(\sqrt{\mathbb{E}(X^2)}\right) \text{ où } c = \frac{1}{2} + ln\left(\sqrt{2\pi}\right) \tag{6}$$

D'après la question précédente, on a égalité si et seulement si X et Y sont de même loi. C'est à dire si et seulement si : $X \hookrightarrow \mathcal{N}(0, \sigma^2)$.

(d) Supposons désormais $|X| < A$ presque sûrement. Il en découle :

$$\begin{aligned} \mathbb{P}_X\left(\subset \mathbb{R} - [-A, A]\right) = 0 &\Longleftrightarrow \int_{\mathbb{R}-[-A,A]} f_X(u)du = 0 \\ &\Longleftrightarrow f_X|_{\mathbb{R}-[-A,A]} = 0 \text{ p.p.} \end{aligned}$$

Quitte à choisir un représentant continu de f_X dans L^1 on peut supposer que la borne supérieure de f_X est atteinte sur $[-A, A]$. On choisit alors Y telle que $Y \hookrightarrow \mathcal{U}\left([-A, A]\right)$. La divergence de Kullback s'écrit dans ce cas :

$$\begin{aligned} D(f_X||f_Y) &= \int_{-A}^{A} f_X(y) ln\left(\frac{f_X(y)}{f_Y(y)}\right) dy \\ &= \int_{-A}^{A} f_X(y) ln\left(f_X(y)\right) dy - \int_{-A}^{A} f_X(y) ln\left(f_Y(y)\right) dy \\ &= -h(X) + ln(2A) \\ &= -h(X) + ln(2) + ln(A) \end{aligned}$$

A nouveau grâce à la question 1(b) on obtient :

$$h(X) \leq ln(2) + ln(A) \tag{7}$$

Avec égalité si et seulement si X suit une loi uniforme sur $[-A, A]$.

(e) Pour montrer la définition de $I(X;Y)$ il suffit de reconduire le raisonnement de la question 1(a) avec cette fois-ci :

$$A = \left\{(x,y) \in R^2 \text{ tels que } \frac{f_{X,Y}(x,y)}{f_X(x)f_Y(y)} < 1\right\} \tag{8}$$

Comme en 1(a), une inégalité de Jensen permet de montrer :

$$I(X;Y) \geq 0 \tag{9}$$

En ce qui concerne la symétrie, on a :

$$\begin{aligned} I(X;Y) &= \int f_{X,Y}(u,v) ln\left(\frac{f_{X,Y}(u,v)}{f_X(u)f_Y(v)}\right) dudv \\ &= \int f_{Y,X}(\alpha,\beta) ln\left(\frac{f_{Y,X}(\alpha,\beta)}{f_X(\alpha)f_Y(\beta)}\right) d\alpha d\beta \\ &= I(Y;X) \end{aligned}$$

Par ailleurs :

$$\begin{aligned} h(X) - I(X;Y) &= -\int f_X(u) ln\left(f_X(u)\right) du - \int f_{X,Y}(\alpha,\beta) ln\left(\frac{f_{X,Y}(\alpha,\beta)}{f_X(\alpha)f_Y(\beta)}\right) d\alpha d\beta \\ &= -\int f_X(u) ln\left(f_X(u)\right) du + \int ln(f_X(\alpha)) \left(\int f_{X,Y}(\alpha,\beta) d\beta\right) d\alpha \\ &\quad - \int f_{X,Y}(\alpha,\beta) ln\left(\frac{f_{X,Y}(\alpha,\beta)}{f_Y(\beta)}\right) d\alpha d\beta \\ &= \underbrace{-\int f_X(u) ln\left(f_X(u)\right) du + \int ln(f_X(\alpha)) f_X(\alpha) d\alpha}_{=0} \\ &\quad - \int f_{X,Y}(\alpha,\beta) ln\left(\frac{f_{X,Y}(\alpha,\beta)}{f_Y(\beta)}\right) d\alpha d\beta \end{aligned}$$

Finalement,

$$h(X|Y) = -\int f_{X,Y}(\alpha,\beta) ln\left(\frac{f_{X,Y}(\alpha,\beta)}{f_Y(\beta)}\right) d\alpha d\beta \tag{10}$$

2. (a) Notons $(a_i)_{i=1}^n$ les coefficients de la matrice A qui ici est un vecteur ligne. n étant la taille du vecteur colonne $X = (X_1, \ldots, X_n)$. Considérons alors le couple $(X, Y - AX)$.
D'après l'énoncé (X,Y) est gaussien, donc toute combinaison linéaire de X et Y est gaussienne. On en conclut que $(X, Y - AX)$ est gaussien.
Or, d'après un résultat du cours sur les vecteurs gaussiens on sait qu'il suffit que la matrice de covariance du couple$(X, Y - AX)$ soit diagonale pour qu'on ait l'indépendance entre X et $Y - AX$. Cela se traduit par un système linéaire de n équations à n inconnues comme suit :

$$\begin{aligned} \forall i \in [1,n],\ Cov\left(X_i, Y - \sum_{j=1}^{n} a_j X_j\right) &= 0 \\ \iff Cov(X_i, Y) &= \sum_{j=1}^{n} a_j Cov(X_i, X_j) \end{aligned}$$

Cependant, il nous est pour l'instant impossible d'affirmer que ce système a bien une solution. Afin de résoudre ce problème nous allons prendre un point de vue « Hilbertien ».
Considérons $E = Vect(X_1, \ldots, X_n)$ comme un sous-espace vectoriel de L^2 que l'on a préalablement équipé de son produit scalaire classique : $< U, V >= \mathbb{E}(UV)$.
Puisque nous travaillons avec des variables centrées, le produit scalaire devient la covariance : $< U, V >= Cov(U, V)$.
Ainsi, il nous suffirait de trouver $Z \in E$ tel que $\forall i \; Cov(X_i, Y - Z) = 0$. En effet, $Z \in E$ signifie que Z est une combinaison linéaire des X_i. Donc, que l'on dispose de $M \in \mathcal{M}_n(\mathbb{R})$ telle que $Z = MX$.
Or, d'après le début de l'étude, le résultat sur les covariances suffit puisque l'on travaille avec des vecteurs gaussiens.
En réécrivant la covariance comme un produit scalaire :

$$\forall i \;\; < X_i, Y - Z >= 0 \tag{11}$$

On reconnait une des propriété du projeté orthogonal. Or il se trouve que E, en tant que sous-espace vectoriel de dimension finie est à la fois convexe et fermé et que $(L^2, < ., . >)$ est un Hilbert. Le théorème de projection sur un convexe fermé nous permet alors de prouver l'existence d'un tel Z et donc l'existence d'une matrice telle que demandée dans l'énoncé.

(b) Soit f telle que $Y - f(X)$ soit indépendante de X. Nous travaillons toujours avec des variables centrées donc on a $\mathbb{E}(Y - f(X)) = 0$.
Par indépendance avec X nous pouvons alors écrire :

$$\mathbb{E}(Y - f(X) \mid X) = 0$$

ce qui, par linéarité, donne :

$$\mathbb{E}(f(X) \mid X) = \mathbb{E}(Y \mid X)$$

Or $\forall x \in X(\Omega)$,

$$\mathbb{E}(f(X) \mid X = x) = f(x)$$

D'où $\mathbb{E}(f(X) \mid X) = f(X)$

Ainsi, une condition nécessaire sur f pour $Y - f(X)$ soit indépendante est $f(X) = \mathbb{E}(Y \mid X)$. Ayant prouvé l'existence d'une telle fonction f à la question précédente, cette condition est aussi suffisante.

(c) Notons $c(X) = \mathbb{E}(Y \mid X)$ et montrons que pour toute fonction $u : \mathbb{R} \longrightarrow \mathbb{R}$,

$$\mathbb{E}\left((Y - u(X))^2\right) \geq \mathbb{E}\left((Y - c(X))^2\right) \tag{12}$$

On a :

$$\begin{aligned}\mathbb{E}\left((Y - u(X))^2\right) &= \mathbb{E}\left((Y - m(X) + m(X) - u(X))^2\right) \\ &= \mathbb{E}\left((Y - m(X))^2\right) + \mathbb{E}\left((m(X) - u(X))^2\right) + 2\mathbb{E}\left((Y - m(X))(m(X) - u(X))\right)\end{aligned}$$

Calculons $\mathbb{E}\left((Y - m(X))(m(X) - u(X))\right)$ par conditionnement :

$$\mathbb{E}\left((Y - m(X))(m(X) - u(X))\right) = \mathbb{E}\left[\mathbb{E}\left((Y - m(X))(m(X) - u(X))\right) \mid X\right]$$

Or, $\forall x \in X(\Omega), \forall g$ telles que $g(X) \in L^2, \forall$ v.a. $Z \in L^2$:

$$\mathbb{E}(g(X)Z \mid X = x) = g(x)\mathbb{E}(Z \mid X)$$

Ainsi,

$$\mathbb{E}(g(X)Z \mid X) = g(X)\mathbb{E}(Z \mid X)$$

En posant $g(X) = m(X) - u(X)$, on a :

$$\begin{aligned}
\mathbb{E}\left[\mathbb{E}\left((Y - m(X))(m(X) - u(X))\right) \mid X\right] &= \mathbb{E}\left[(m(X) - u(X)) \mid \mathbb{E}\left((Y - m(X)) \mid X\right)\right] \\
&= \mathbb{E}\left[(m(X) - u(X))(\mathbb{E}[Y \mid X] - \mathbb{E}[m(X)|X])\right] \\
\text{Avec 2 (a) il vient : } &= \mathbb{E}\left[(m(X) - u(X))(\mathbb{E}[Y \mid X] - m(X))\right]
\end{aligned}$$

Or, $m(X) = \mathbb{E}[Y \mid X]$, donc :

$$\begin{aligned}
\mathbb{E}[(Y - u(X))^2] &= \mathbb{E}[(Y - m(X))^2] + \underbrace{\mathbb{E}[(m(X) - u(X))^2]}_{\geq 0} \\
&\geq \mathbb{E}[(Y - m(X))^2]
\end{aligned}$$

L'égalité force $\mathbb{E}[(m(X) - u(X))^2] = 0$ soit $m(X) = u(X)$ presque sûrement. Soit $m(X) = u(X)$ quitte à changer de représentant dans la classé d'équivalence de $m(X)$ dans L^2.
Notons $\delta = \mathbb{E}[(Y - \mathbb{E}[X \mid Y])^2]$, l'erreur quadratique moyenne, et calculons :

$$\begin{aligned}
\delta &= \mathbb{E}[(Y - \mathbb{E}[X \mid Y])^2] \\
&= \mathbb{E}[(Y - AX)^2] \\
&= \mathbb{E}(Y^2) + \mathbb{E}((AX)^2) - 2\mathbb{E}[Y(AX)] \\
&= V(Y) + V(AX) - Z\sum_i a_i Cov(Y, X_i) \\
&= V(Y) + \sum_j a_j^2 V(X_j) + \sum_{p \neq q} a_p a_q Cov(X_p, X_q) - 2\sum_i a_i Cov(Y, X_i) \\
&= V(Y) + \sum_i (V(X_i) - 2Cov(Y, X_i)) + \sum_{p \neq q} Cov(X_p, X_q)
\end{aligned}$$

(d) En gardant les notations, on a montré à la 2(b) que :

$$\mathbb{E}(Y \mid X) = AX \tag{13}$$

Or, ici $Y = f(X)$ et $\mathbb{E}(f(X) \mid X) = f(X)$.
Finalement, on obtient bien la linéarité de $f(X) = AX$:

$$\begin{aligned}
f : &\mathbb{R}^n \to \mathbb{R} \\
&x \longmapsto Ax
\end{aligned}$$

MODÈLE DE COMMUNICATION

3. (a) On applique le résultat de la question 1 (c) à la variable aléatoire $U - V$:

$$\begin{aligned} D &= \mathbb{E}\left((U-V)^2\right) \\ &\geq exp\left(2 \cdot (h(U-V) - c)\right) \end{aligned}$$

(b) On a $I(U - V; Y) \geq 0$ d'où on en déduit :

$$\begin{aligned} h(U-V) - h(U-V|Y) &\geq 0 \\ h(U-V) &\geq h(U-V|Y) \\ &\geq h(U|Y) \end{aligned}$$

Par croissance l'exponentielle il vient :

$$\begin{aligned} D &\geq exp\left(2 \cdot (h(U|Y) - c)\right) \\ &\geq exp\left(2 \cdot (h(U) - I(U;Y) - c)\right) \end{aligned}$$

Or U suit une loi normale de variance σ^2. D'où d'après la question 1(c),

$$h(U) = c + \frac{1}{2} ln(\sigma^2) \tag{14}$$

Ainsi,

$$\begin{aligned} D &\geq exp\left(2 \cdot (h(U) - I(U;Y) - c)\right) \\ &\geq \sigma^2 \cdot exp(-2 \cdot I(U;Y)) \end{aligned}$$

Soit finalement :

$$\frac{\sigma^2}{D} \leq exp(2 \cdot I(U;Y)) \tag{15}$$

(c) Sachant que que $Y_i = X_i + Z_i$ et que Z_i est indépendant de X_i et ne dépend pas des variables où $j > i$. Or $X_1 = C_i(U)$ d'où la connaissance de x_1 est directement donné par celle de $U = u$ et il n'y a pas plus d'information à tirer de l'égalité $U = u$ pour Y_1. Ainsi :

$$f_{Y|U=u}(y) = f_{Y_1|X_1=x_1}(y_1) \cdot f_{(Y_2,...,Y_n)|U=u \text{ et } X_1=x_1}(y) \tag{16}$$

Puis de proche en proche avec la remarque précédente, on obtient $X_i = C_i(u, Y_1 = y_1, ..., Y_{i-1} = y_{i-1})$ d'où :

$$f_{Y|U=u}(y) = \Pi_{i=1}^{n} f_{Y_i|X_i=x_i}(y_i) \tag{17}$$

(d) En appliquant moins le logarithme à l'expression ci-dessus, on obtient le calcul suivant :

$$\begin{aligned}I(U;Y) &= I(Y;U) \\ &= h(Y) - h(Y|U) \\ &= h(Y) - \sum_{i=1}^{n} h(Y_i|X_i) \\ &= h(Y) - \sum_{i=1}^{n} h(Y_i) + \sum_{i=1}^{n} I(X_i, Y_i)\end{aligned}$$

Or,

$$f_{(Y_1,\ldots,Y_n)} = f_{Y_1} \cdots f_{Y_n|(Y_1,\ldots,Y_n)} \tag{18}$$

D'où par passage au logarithme dans l'expression précédente :

$$h(Y) = \sum_{i=1}^{n} h(Y_i|Y_1,\ldots,Y_{i-1}) \tag{19}$$

Mais on a toujours $h(Y_i) \geq h(Y_i|Y_1,\ldots,Y_{i-1})$ par positivité de $I(Y_i;Y_1,\ldots,Y_{i-1})$. Finalement :

$$\begin{aligned}h(Y) &= \sum_{i=1}^{n} h(Y_i|Y_1,\ldots,Y_{i-1}) \\ &\leq \sum_{i=1}^{n} h(Y_i) \\ 0 &\leq \sum_{i=1}^{n} h(Y_i) - h(Y)\end{aligned}$$

Ainsi, on en déduit en réinjectant :

$$I(U;Y) \leq \sum_{i=1}^{n} I(X_i;Y_i) \tag{20}$$

(e) Puis, par un calcul :

$$\begin{aligned}I(X_i;Y_i) &= h(Y_i) - h(Y_i|X_i) \\ &= h(Y_i) - h(X_i + Z_i|X_i) \\ &= h(Y_i) - h(Z_i|X_i) \\ &= h(Y_i) - h(Z_i)\end{aligned}$$

(f) On sait que :

$$\frac{\sigma^2}{D} \leq exp(2 \cdot I(U;Y)) \tag{21}$$

Ainsi avec les deux résultats précédents :

$$\frac{\sigma^2}{D} \leq exp\left(2 \cdot \sum_{i=1}^{n} h(Y_i) - h(Z_i)\right) \tag{22}$$

Or les (Z_i) suivent des lois normales de variance N, on applique alors 1) à Y_i et Z_i :

$$\begin{aligned}\sum_{i=1}^{n} h(Y_i) - h(Z_i) &\leq \sum_{i=1}^{n} c + \frac{1}{2}ln(\mathbb{E}(Y_i^2)) - \left(c + \frac{1}{2}ln(\mathbb{E}(Z_i^2)\right) \\ &\leq \sum_{i=1}^{n} \frac{1}{2}ln(\mathbb{E}(X_i^2) + \mathbb{E}(Z_i^2) + 2\mathbb{E}(X_iZ_i)) - \frac{1}{2}ln(\mathbb{E}(Z_i^2))\end{aligned}$$

Grace à l'indépendance de X_i et Z_i on a : $\mathbb{E}(X_iZ_i) = \mathbb{E}(X_i) \cdot \mathbb{E}(Z_i) = 0$

$$\begin{aligned}\sum_{i=1}^{n} h(Y_i) - h(Z_i) &\leq \sum_{i=1}^{n} \frac{1}{2}ln\left(\frac{\mathbb{E}(X_i^2) + N}{N}\right) \\ &\leq \frac{n}{2}\sum_{i=1}^{n} \frac{1}{n}ln\left(\frac{\mathbb{E}(X_i^2) + N}{N}\right)\end{aligned}$$

Par concavité du logarithme et l'inégalité de Jensen on en déduit que :

$$\begin{aligned}\sum_{i=1}^{n} h(Y_i) - h(Z_i) &\leq \frac{n}{2}ln\left(\sum_{i=1}^{n} \frac{1}{n}\frac{\mathbb{E}(X_i^2) + N}{N}\right) \\ &\leq \frac{n}{2}ln\left(1 + \frac{P}{N}\right)\end{aligned}$$

Finalement, en réinjectant le résultat précédent dans l'inégalité suivante :

$$\frac{\sigma^2}{D} \leq exp\left(2 \cdot \sum_{i=1}^{n} h(Y_i) - h(Z_i)\right) \tag{23}$$

$$\leq \left(1 + \frac{P}{N}\right)^n \tag{24}$$

Schéma d'Elias

4. (a) Le caractère optimale de l'inégalité, soit l'égalité, impose l'égalité dans toutes les inégalités intermédiaires. En l'occurrence :

$$\begin{aligned} & h(U-V) = h(U-V|Y) \\ \Longleftrightarrow & h(U-V) - h(U-V) + I(U-V;Y) = 0 \\ \Longleftrightarrow & I(U-V;Y) = 0 \end{aligned}$$

Ce qui implique l'indépendance de $(U-V)$ avec les $Y_1, \ldots, Y_n$.
Ainsi que :

$$\forall i \in [2, n], \; h(Y_i) = h(Y_i \mid Y_1, \ldots, Y_{i-1}) \tag{25}$$

D'où $I(Y_i; Y_1, \ldots, Y_{i-1}) = 0$ et Y_i est indépendant des $Y_1, \ldots, Y_{i-1}$.
De plus, nous savions déjà que Z_i est indépendant des $Y_1, \ldots, Y_{i-1}$, on en déduit donc que $X_i = Y_i - Z_i$ est également indépendant des $Y_1, \ldots, Y_{i-1}$.

(b) On sait déjà que $U \hookrightarrow \mathcal{N}(0, \sigma^2)$ et $Z_i \hookrightarrow \mathcal{N}(0, N)$ $\forall i \in [1, n]$.
Pour Y_i, l'égalité supposée par l'énoncé :

$$\frac{\sigma^2}{D} = \left(1 + \frac{P}{N}\right)^n$$

Une fois imposée dans le raisonnement de 3(f) force :

$$h(Y_i) = c + \frac{1}{2} ln(E[Y_i^2])$$

D'après 1.(c) on en déduit :

$$\forall i \in [1, n], \; Y_i \hookrightarrow \mathcal{N}(0, Var[Y_i]) \tag{26}$$

De plus, $X_i = Y_i - Z_i$ d'où X_i suit également une loi normale comme combinaison linéaire de gaussiennes.

$$\forall i \in [1, n], \; X_i \hookrightarrow \mathcal{N}(0, Var[X_i]) \tag{27}$$

D'après la question 4 (a) nous savons également que $U - V \perp Y$ soit $U - \mathcal{D}(Y) \perp Y$. Le résultat de 2(a) et 2(b) nous prouve alors l'existence d'une matrice A telle que

$$\mathcal{D}(Y) = AY = \sum_{i=1}^{n} a_i Y_i$$

Finalement, V est également une gaussienne comme combinaison linéaire de gaussiennes :

$$V \hookrightarrow \mathcal{N}(0, Var(V)) \tag{28}$$

(c) Nous travaillons à présent sous l'hypothèse que l'ensemble des variables du problème forment un vecteur gaussien. En particulier,

$$\forall i \in [1, n], \; (U, Y_1, \ldots, Y_{i-1}, X_i) \text{ est gaussien.}$$

De plus, X_i est une variable aléatoire de dimension 1. On se retrouve alors dans le cadre de la question 2). En appliquant le résultat obtenu en 2(d) on montre que C_i est linéaire pour tout i dans $[1, n]$.
Soient $b_0, \ldots, b_{i-1}$ les réels tels que :

$$X_i = b_0 U + \sum_{j=1}^{i-1} b_j Y_j$$

Alors, on dispose de

$$\begin{aligned} g :& \mathbb{R}^{i-1} \to \mathbb{R} \\ & (y_1, \ldots, yi-1) \longmapsto -\frac{1}{b_0} \sum_{j=1}^{i-1} b_j Y_j \end{aligned}$$

telle que :

$$\frac{1}{b_0} X_i = U - g(Y_1, \ldots, Y_{i-1}) \tag{29}$$

Mais X_i est indépendante de $Y_1, \ldots, Y_{i-1}$. D'où d'après 2(b) on a :

$$g(Y_1, \ldots, Y_{i-1}) = \mathbb{E}(U \mid Y_1, \ldots, Y_{i-1}) \tag{30}$$

Ainsi,

$$\begin{aligned} X_i &= b_0 \left(U - \mathbb{E}(U \mid Y_1, \ldots, Y_{i-1})\right) \\ &= b_0 U_i \end{aligned}$$

En prenant la variance de part et d'autre on trouve :

$$b_0^2 = \frac{Var(X_i)}{Var(U_i)} \text{ d'où } b_0 = \frac{\sqrt{Var(X_i)}}{\sigma_i} \tag{31}$$

Ce que l'on peut faire moyennant l'hypothèse qu'on ne travaille pas avec des variables constantes p.s.
En admettant alors que

$$Var(X_i) = P$$

On en déduit :

$$X_i = \frac{\sqrt{P}}{\sigma_i} U_i \tag{32}$$

Poursuivons. D'après les questions 2(a) et 2(b) nous savons que nous disposons d'un réel b tel que :

$$\mathbb{E}(U_i \mid Y_i) = bY_i$$

En prenant, cette fois-ci, la covariance, on obtient :

$$\begin{aligned} Cov(\mathbb{E}(U_i \mid Y_i), Y_i) &= bVar(Y_i) \\ \text{d'où } b &= \frac{Cov(\mathbb{E}(U_i \mid Y_i), Y_i}{Var(Y_i} \end{aligned}$$

Or, nous savons d'après 2 que $U_i - \mathbb{E}(U_i \mid Y_i) \perp Y_i$, en particulier :

$$\begin{aligned} &Cov(U_i - \mathbb{E}(U_i \mid Y_i), Y_i) = 0 \\ \iff &Cov(U_i, Y - i) = Cov(\mathbb{E}(U_i \mid Y_i), Y_i) \end{aligned}$$

De plus

$$Y_i = X_i + Z_i \tag{33}$$

$$\text{et } U_i = \frac{\sigma_i}{\sqrt{P}} X_i \tag{34}$$

D'où :

$$\begin{aligned} b &= \frac{\sigma_i}{\sqrt{P}} \frac{Cov(X_i, X_i + Z_i)}{Var(X_i + Z_i)} \\ \text{or } X_i \perp Z_i : &= \frac{\sigma_i}{\sqrt{P}} \frac{Cov(X_i, X_i) + \overbrace{Cov(X_i, Z_i)}^{=0}}{Var(X_i) + Var(Z_i)} \\ &= \frac{\sigma_i}{\sqrt{P}} \frac{P}{P+N} \\ &= \frac{\sigma_i \sqrt{P}}{P+N} \end{aligned}$$

Soit :

$$\mathbb{E}(U_i \mid Y_i) = \frac{\sigma_i \sqrt{P}}{P+N} Y_i \tag{35}$$

(d) On sait que $\mathcal{D}(Y)$ est telle que l'erreur quadratique moyenne $D = \mathbb{E}[(U-V)^2$ soit minimale. Ainsi, d'après les question 2(c) et 2 (b), $\mathcal{D}$ est linéaire en les Y_i. On admet alors que :

$$\mathcal{D}(Y) = \sum_{i=1}^{n} \mathbb{E}(U_i \mid Y_i) \tag{36}$$

Puis :

$$\begin{aligned} \sigma_{i+1}^2 &= Cov(U_i - \frac{\sigma_i \sqrt{P}}{P+N} Y_i, U_i - \frac{\sigma_i \sqrt{P}}{P+N} Y_i) \\ &= \sigma_i^2 + \left(\frac{\sigma_i \sqrt{P}}{P+N} \right)^2 Cov(Y_i, Y_i) - 2 \frac{\sigma_i \sqrt{P}}{P+N} Cov(U_i, Y_i) \\ &= \sigma_i^2 + \left(\frac{\sigma_i \sqrt{P}}{P+N} \right)^2 (P+N) - 2 \frac{\sigma_i^2}{P+N} Cov(X_i, Y_i) \\ &= \sigma_i^2 \cdot \left(1 + - \frac{P}{P+N} \right) \end{aligned}$$

Ainsi,

$$\sigma_i^2 = \left(\frac{N}{P+N} \right)^{i-1} \cdot \sigma^2 \tag{37}$$

5. Voici le code python de l'implémentation du schéma d'Elias.

```python
from random import gauss
from math import sqrt,log10
import matplotlib.pyplot as plt


P = 0.5
N = 1
sigma = 1
nmax = 100
realisations = 1000

D = [0]*nmax

for n in range(0,nmax):#On fait varier la valeur de n de 0 a 99
    for j in range(0,realisations):#On fait 1000 realisations source-canal
        Source = gauss(0,sigma**2) #On tire U
        sigmaI = sigma
        U = Source
        V = 0 #Ce que le destinataire recoit a l'issue de l'algorithme

        for i in range(0,n+1):#+1 pour que n+1 varie de 1 a 100
            X = (sqrt(P)/sigmaI)*U
            Y = X + gauss(0,N**2) #Ajout du bruit

            E = (sigmaI*sqrt(P)/(P+N))*Y #L'espereance conditionnelle
            U = U - E
            sigmaI = sqrt(N/(P+N))*sigmaI
            V += E

        D[n] += (Source-V)**2
    D[n] = 10*log10(sigma**2/(D[n]/realisations))

#Affichage
x = range(1,nmax+1)
plt.plot(x,D)
plt.show()
```

On obtient alors le résultat suivant pour le rapport signal-à-bruit en décibel : Ce qui est cohérent avec

$$10 * log_{10}\left(\frac{\sigma^2}{D}\right) = n \cdot 10 * log_{10}\left(1 + \frac{P}{N}\right) \tag{38}$$

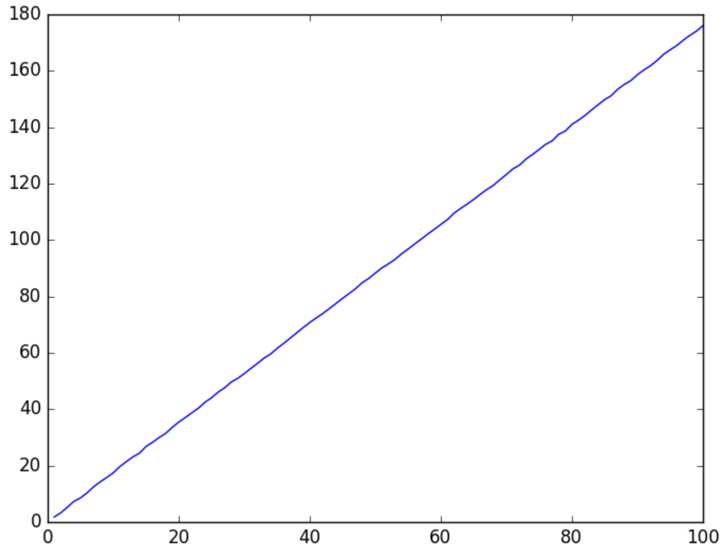


Figure 1 – Rapport signal-à-bruits en décibels en fonction de n

Schéma de Schalkwijk

7. Après des recherches personnelles infructueuses nous avons trouvé [1] où le résultat est obtenu des pages 26 à 28. Nous avons donc :

$$P_e(n) \leq \sqrt{\frac{2}{\pi}} exp\left(-\frac{2^{n(C-R)}\sqrt{P}}{2}\right) \tag{39}$$

En posant $C - R = 1$, $P = 1$, on obtient le graphique suivant :

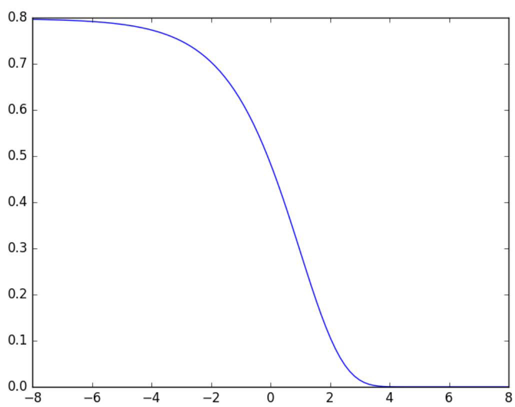


Figure 2 – Graphique de $P_e(x)$

8. La valeur de M de la question 6) n'étant pas donnée, on a choisi de considérer U gaussien et, afin de déterminer si la valeur imparfaite de U obtenue est convenable, on a choisi d'introduire un ϵ et de regarder $P(|U - U_{imparfait}| > \epsilon)$.
   Si on avait considéré U comme dans 6) l'écart entre deux valeurs consécutives est $\frac{2}{M-1}$. Ainsi, afin de déterminer si $U_{imparfait}$ et U sont égales, il suffirait de considérer $\epsilon = \frac{2}{M-1}$. Ou encore :

$$M_\epsilon = \frac{2}{\epsilon} + 1$$

   Ainsi, avec les graphiques ci-dessous et pour les valeurs de M_ϵ correspondantes, on obtient le nombre n minimal d'utilisations du canal afin d'avoir $P_e \simeq 0$.
   Voici le code :

```
from random import gauss
from math import sqrt
import matplotlib.pyplot as plt

def schema(source,n,sigma): #Le schema de transmission d'Elias
    sigmaI = sigma
    U = source
    V = 0 #Ce que le destinataire recoit a l'issue de l'algorithme

    for i in range(0,n):
                X = (sqrt(P)/sigmaI)*U
                Y = X + gauss(0,N**2) #Ajout du bruit

                E = (sigmaI*sqrt(P)/(P+N))*Y #L'espereance conditionnelle
                U = U - E
                sigmaI = sqrt(N/(P+N))*sigmaI
                V += E
    return V

P = 0.5
N = 1
sigma = 1
epsilon = 0.00000001
nmax = 100
realisations = 1000
Pe = [0]*nmax

for n in range(1,nmax+1):

    for i in range(0,realisations):

        U = gauss(0,sigma**2) #On tire la valeur de U
        X = (sqrt(P)/sigma)*U #On transmet X1
        Y = X + gauss(0,N) #La voie de retour transmet Y1
```

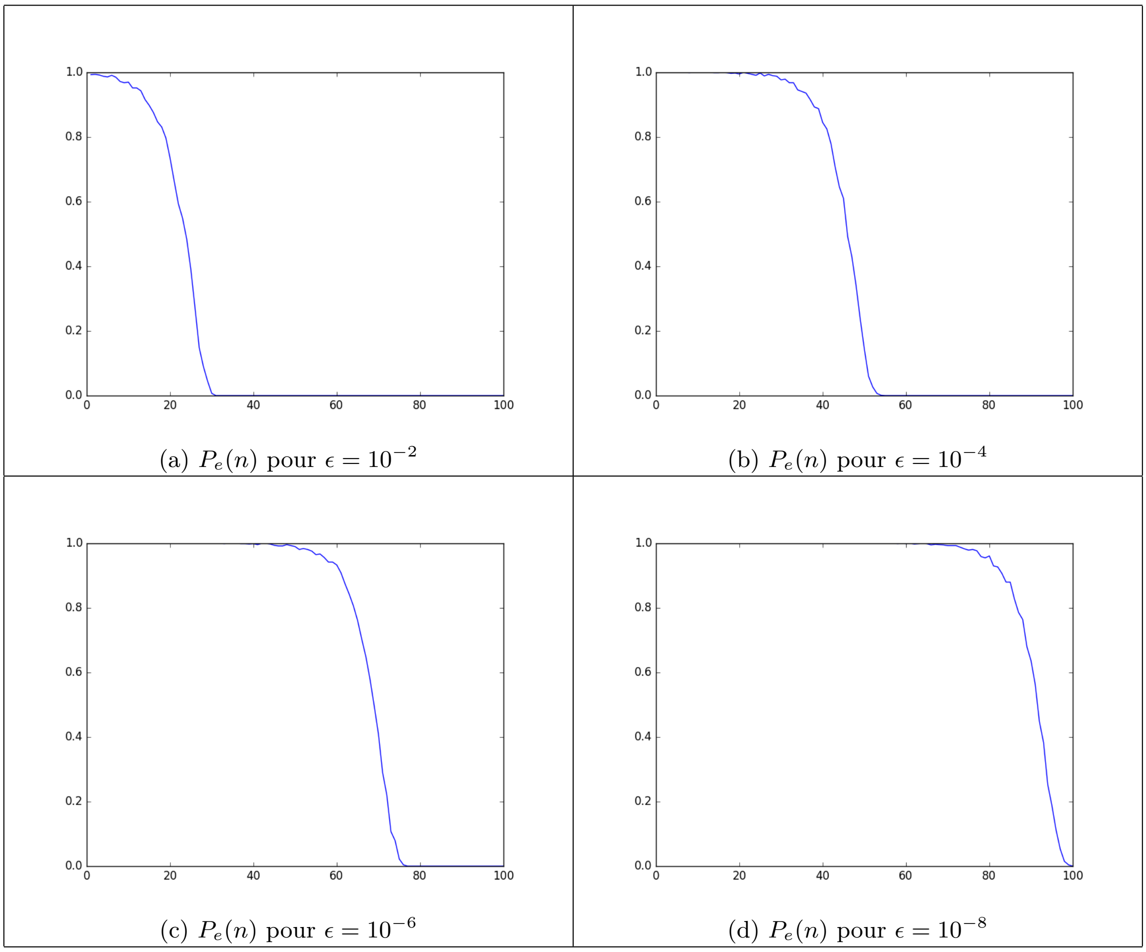
```
        Vz = schema(Y-X,n-1,sigma) #Resultat de la transmission de U2 = Z1
        Upresque = sigma/sqrt(P)*(Y-Vz)

        if (abs(U-Upresque)>epsilon): #Si l'ecart entre U et Upresque est
                                      # superieur a une marge d'erreur
            Pe[n-1] += 1              #On considere que la transmission a echouee


    Pe[n-1] = Pe[n-1]/realisations #On calcule la moyenne

#Affichage
x = range(1,nmax+1)
plt.plot(x,Pe)
plt.show()
```

Et les graphiques :



(a) $P_e(n)$ pour $\epsilon = 10^{-2}$

(b) $P_e(n)$ pour $\epsilon = 10^{-4}$

(c) $P_e(n)$ pour $\epsilon = 10^{-6}$

(d) $P_e(n)$ pour $\epsilon = 10^{-8}$

Ceux-ci sont bien compatibles avec l'expression de $P_e(n)$ donnée par [1].

RÉFÉRENCES

[1] Youlong Wu. *New Ways of Exploiting Feedback in Memoryless Broadcast Channels*. PhD thesis, TELECOM ParisTech Spécialitée " Electronique et Communications ", 2014.