COM 105 - Résumé du cours

1 Codage correcteur d'erreur

Def (Le modèle de transmission).

$$\xrightarrow{D_1,\dots,D_k} \text{Émetteur } f \xrightarrow{X_1,\dots,X_n} \text{Canal} \xrightarrow{Y_1,\dots,Y_n} \text{Récepteur } g \xrightarrow{\hat{D}_1,\dots,\hat{D}_k}$$

Voc. • Bits d'information : $\mathbf{D} = (D_1, \dots, D_k)$, représentent les données à transmettre, supposés aléatoires et donc i.i.d. $\mathcal{B}\left(\frac{1}{2}\right)$.

- Transmission en bloc : les k bits d'information sont envoyés sur un bloc de n utilisations du canal.
- Émetteur : associe $\mathbf{X} = (X_1, \dots, X_kn)$ à $\mathbf{D} = (D_1, \dots, D_k)$, supposé déterministe et avec f injective.
- Récepteur : associe $\hat{\mathbf{D}} = \left(\hat{D}_1, \dots, \hat{D}_k\right)$ à $\mathbf{Y} = (Y_1, \dots, Y_kn)$, supposé déterministe.
- Erreur : cas où $\left(\hat{D}_1, \dots, \hat{D}_k\right) \neq (D_1, \dots, D_k)$.

1.1 Les canaux

Def (Canaux discrets sans mémoire (**DMC**)). Un DMC est complètement caractérisé par le triplet $\left(\mathcal{X}, \mathcal{Y}, \mathbf{P}_{X|Y}(\cdot \mid \cdot)\right)$ où

- $\mathcal{X}$ est un alphabet fini contenant toutes les valeurs possibles à l'entrée du DMC,
- $\mathcal{Y}$ est un alphabet fini contenant toutes les valeurs possibles à la sortie du DMC,
- $\mathbf{P}_{X|Y}(\cdot \mid \cdot)$ est une loi de probabilité conditionnelle, dite loi de transition, décrivant comment une sortie Y_t est obtenue à partir d'une entrée x_t.

Def (Canal binaire symétrique (**BSC**)). On a $\mathcal{X} = \mathcal{Y} = \{0, 1\}$ et $\forall t \in [\![1\,;n]\!]$, $Y_t = x_t$ avec une probabilité $p \in [0\,;1]$ et $Y_t \neq x_t$ avec une probabilité $1 - p$. On peut toujours supposer $p < \frac{1}{2}$.

Def (Canal binaire à effacement (**BEC**)). On a $\mathcal{X} = \{0, 1\}$ et $\mathcal{Y} = \{0, 1, \Delta\}$ où Δ représente un effacement. Pour tout $t \in [\![1\,;n]\!]$, $Y_t = x_t$ avec une probabilité $\epsilon \in [0\,;1]$ et $Y_t = \Delta$ avec une probabilité $1 - \epsilon$.

1.2 Codage par des codes en bloc

Def. Un **code en bloc** $\mathcal{C}$ de longueur n sur un alphabet $\mathcal{X}$ est un sous-ensemble de $\mathcal{X}^n$, c'est l'ensemble d'arrivée de f. Les éléments de $\mathcal{C}$ sont appelés les mots de code de $\mathcal{C}$.

Def. Le **rendement** (binaire) d'un code en bloc $\mathcal{C}$ de longueur n, aussi appelé taux de codage, est $R = \frac{\log_2(|\mathcal{C}|)}{n}$ où $|\mathcal{C}|$ est le nombre de mots du code $|\mathcal{C}|$.

1.3 Distances

Def. Poids de Hamming pour $\mathbf{x} = (x_1, \dots, x_n) : w_H(\mathbf{x}) = \mathrm{Card}\left(\{x_i \neq 0\}\right)$.

Def. La **distance de Hamming** entre deux mots $\mathbf{x}$ et $\hat{\mathbf{x}}$ est donnée par $d_H(\mathbf{x}, \hat{\mathbf{x}}) = w_H(\mathbf{x} - \hat{\mathbf{x}})$.

Prop. *La distance de Hamming est bien une distance (symétrie, positivité et inégalité triangulaire).*

Def. La **distance minimale** du code en bloc $\mathcal{C}$ est $d_{\min}(\mathcal{C}) = \min_{\mathbf{c} \neq \hat{\mathbf{c}}} d_H(\mathbf{c}, \hat{\mathbf{c}})$.

1.4 Décodage

On décompose la fonction de décodage en deux étapes : $g = g_2 \circ g_1$, g_1 trouve pour toute observation $\mathbf{Y}$ le mot de code $\hat{\mathbf{c}} \in \mathcal{C}$ qui paraît le plus probable et $g_2 = f^{-1}$ produit la suite des bits détectés $\hat{D}_1, \dots, \hat{D}_k$ qui est associée à $\hat{\mathbf{c}}$.

Def. Soit g_1 fixée. La région de décision associée à $\mathbf{c} \in \mathcal{C}$ est $\Omega_{\mathbf{c}} := \{\mathbf{y} \in \mathcal{Y}^n \mid g_1(\mathbf{y}) = \mathbf{c}\}$. Ces régions forment une partition de $\mathcal{Y}^n$.

On a $P_e := \mathbf{P}(\hat{\mathbf{C}} \neq \mathbf{C})$ la probabilité d'erreur et $P_c := \mathbf{P}(\hat{\mathbf{C}} = \mathbf{C})$ la probabilité de succès.

Prop (Optimalité de la règle de maximum vraisemblance). *Si les mots de code sont tous émis avec la même probabilité, minimiser P_e revient à choisir le mot de code qui maximise la vraisemblance, c'est-à-dire, à choisir les $\Omega_{\mathbf{c}}$ tel que* $(\mathbf{y} \in \Omega_{\mathbf{c}}) \iff (\mathbf{P}(\mathbf{Y} = \mathbf{y} \mid \mathbf{C} = \mathbf{c}) = \max_{\tilde{\mathbf{c}} \in \mathcal{C}} \mathbf{P}(\mathbf{Y} = \mathbf{y} \mid \mathbf{C} = \tilde{\mathbf{c}}))$.

Prop (Règle du voisin le plus proche). *Dans le cas d'un BSC on a* $\mathbf{P}(\mathbf{Y} = \mathbf{y} \mid \mathbf{C} = \mathbf{c}) = (1-p)^n \left(\frac{p}{1-p}\right)^{d_H(\mathbf{y},\mathbf{c})}$, $\forall p \in \left[0\,;\frac{1}{2}\right]$ *donc minimiser P_e revient à trouver le mot de code $\mathbf{c} \in \mathcal{C}$ qui minimise $d_H(\mathbf{y}, \mathbf{c})$.*

Voc. On dit qu'un code en bloc $\mathcal{C}$ corrige t erreurs si il existe un décodeur qui permet de corriger toutes les configurations d'erreurs dont le nombre est inférieur ou égal à t.

Prop (Capacité de correction d'un code). *Le décodeur décide toujours du bon mot $\hat{\mathbf{c}} = \mathbf{c}$ lorsque $2d_H(\mathbf{c}, \mathbf{y}) < d_{\min}$. Donc le code peut corriger $t = \left\lfloor \frac{d_{\min}-1}{2} \right\rfloor$ erreurs.*

Lorsque l'on fait de la détection d'erreur, on a $g_1 : \mathcal{Y}^n \to \mathcal{C} \cup \Delta$. La question est alors : est-ce que le mot reçu est bien égal au mot envoyé ? Dans le cas où $d_H(\mathbf{c}, \mathbf{y}) = l \geqslant 1$ et le décodeur produit Δ, on dit que le décodeur a détecté l erreurs.

Voc. Dans le cas d'un BEC, on dit qu'un code en bloc $\mathcal{C}$ détecte t erreurs si il existe un décodeur qui permet de corriger toutes les configurations d'erreurs dont le nombre est inférieur ou égal à t.

Prop (Capacité de détection d'un code). *Un code en bloc est capable de détecter $t' = d_{\min} - 1$ erreurs.*

Il suffit pour cette détection de déclarer Δ dès que $\mathbf{y} \notin \mathcal{C}$.

1.5 Codes linéaires en bloc

Def. Un **code en bloc linéaire** binaire de longueur n est un sous-espace vectoriel de $\mathbb{F}_2^n$.

Def. La dimension k d'un code en bloc linéaire est sa dimension en tant que ss-ev de $\mathbb{F}_2^n$.

On peut alors simplifier l'expression du rendement et de la distance minimale :

$$R = \frac{\log_2(|\mathcal{C}|)}{n} = \frac{k}{n} \qquad \text{et} \qquad d_{\min}(\mathcal{C}) = \min_{\mathbf{c} \in \mathcal{C}, \mathbf{c} \neq 0} w_H(\mathbf{c}) \,.$$

Not. Un code linéaire $\mathcal{C}$ de longueur n, de dimension k et de distance minimale $d_{\min}$ sera noté $(n, k, d_{\min})$.

Def. Un codeur linéaire associe au bits $d_1, \ldots, d_k$ la valeur $\sum_{i=1}^n d_i \mathbf{e}_i$ où $\mathcal{B} = (\mathbf{e}_1, \ldots, \mathbf{e}_k)$ est une base du code.

Def. On appelle **matrice génératrice** du code $\mathcal{C}$ toute matrice G à k lignes et n colonnes de la forme $G = \begin{bmatrix} \mathbf{e}_1 \\ \vdots \\ \mathbf{e}_k \end{bmatrix}$.

Tout mot de code $\mathbf{c} \in \mathcal{C}$ peut s'écrire alors $\mathbf{c} = \mathbf{d} \cdot G$ où $\mathbf{d}$ est le mot d'information.

Deux codes $\mathcal{C}$ et $\tilde{\mathcal{C}}$ sont dits équivalents si et seulement si $\exists \sigma \in \mathfrak{S}_n, \forall \mathbf{c} \in \mathcal{C}, \exists \tilde{]} \in \tilde{\mathcal{C}}, (\tilde{c}_1, \ldots, \tilde{c}_n) = (c_{\sigma(1)}, \ldots, c_{\sigma(n)})$. Deux opérations sont permises sur G pour trouver une autre matrice génératrice pour le même code (ou un code équivalent) : combinaisons linéaires de lignes et permutations de colonnes.

Def. On appelle **matrice génératrice systématique** du code $\mathcal{C}$ toute matrice obtenue à la sortie du pivot de Gauss appliqué à une matrice génératrice G quelconque de $\mathcal{C}$. Elle est sous la forme $G_s = [I_k \parallel P]$ où P dépend du code $\mathcal{C}$.

On a alors $\mathbf{c} = \mathbf{d} \cdot G_s = [\mathbf{d} \quad \mathbf{d} \cdot P]$. Ainsi les k premiers bits sont les bits d'information alors que les $(n-k)$ bits restants dépendent de $\mathbf{d}$ et du code et sont appeles bits de parité.

Ex. On appelle code de parité binaire de longueur n un code binaire de longueur n dont les mots sont tous les n-uplets binaires de poids de Hamming pair. Sa dimension est $n-1$ et $d_{\min} = 2$. Sa matrice génératrice systématique est $G_s = \begin{pmatrix} 1 & 0 & \ldots & 0 & 1 \\ 0 & 1 & & 0 & 1 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & 1 \end{pmatrix}$.

Def. Deux mots $\mathbf{x}$ et $\tilde{\mathbf{x}}$ sont dits orthogonaux si $\mathbf{x} \cdot {}^t\tilde{\mathbf{x}} = \sum_{i=1}^n x_i \tilde{x}_i = 0$. À la différence de l'espace euclidien, tout mot de poids de Hamming pair est orthogonal à lui-même.

Def. Le **code dual** de $\mathcal{C}$ sur $\mathcal{X}$ est $\mathcal{C}^\perp := \left\{ \mathbf{x} \in \mathcal{X}^n \mid \forall \mathbf{c} \in \mathcal{C}, \mathbf{x} \cdot {}^t\mathbf{c} = 0 \right\}$.

Def. Une **matrice de contrôle de parité** de $\mathcal{C}$ est toute matrice H qui est matrice génératrice de $\mathcal{C}^\perp$. Ainsi H est une matrice à $n-k$ lignes et n colonnes de rang $n-k$.

Th. *Soit G une matrice génératrice de $\mathcal{C}$. Toute matrice $H \in \mathfrak{M}_{n-k,n}$ de rang $n-k$ qui vérifie $G \cdot {}^tH = 0$ est une matrice de contrôle de parité de $\mathcal{C}$.*

On en déduit la matrice de contrôle de parité systématique $H_s = [-{}^tP \parallel I_{n-k}]$.

Th. *Pour tout code linéaire en bloc, $d_{\min}$ est égale au plus petit nombre de colonnes dépendantes de H.*

Th (**Borne de Singleton**). *Pour tout code linéaire en bloc $(n, k, d_{\min})$, on a $d_{\min} \leqslant n - k + 1$.*

Def. Soit $\mathbf{y} \in \mathcal{Y}^n$. On appelle **syndrome** la quantité $\mathbf{s} = \mathbf{y} \cdot {}^tH \in \mathfrak{M}_{1,n-k}$.

Prop. *On a $\mathbf{y} \in \mathcal{C}$ si et seulement si $\mathbf{s} = 0$.*

Algorithme de décodage par syndrome

1. Calculer le syndrome $\mathbf{s} = \mathbf{y} \cdot {}^tH$.
2. Si $\mathbf{s} = 0$ alors on déclare $\hat{\mathbf{c}} = \mathbf{y}$ et l'algorithme se termine.

3. Vérifier si ${}^t\mathbf{y}$ est égal à une colonne de H. Si ${}^t\mathbf{y} = \mathbf{h}_i$, déclarer $\mathbf{c} = (y_1, \dots, y_{i-1}, 1 - y_i, y_{i+1}, \dots, y_n)$ et l'algorithme se termine. S'il existe plusieurs i, en choisir un au hasard.
4. Vérifier si ${}^t\mathbf{y}$ est égal à la somme de deux colonnes de H. Si ${}^t\mathbf{y} = \mathbf{h}_i + \mathbf{h}_j$, déclarer $\mathbf{c}$ en inversant y_i et y_j et l'algorithme se termine. S'il existe plusieurs paires, en choisir un au hasard.
5. Continuer ainsi de suite.

Cet algorithme utilisé sur un canal BSC peut corriger $\left\lfloor \frac{d_{\min}-1}{2} \right\rfloor$ erreurs.

Def. Soit $m \geqslant 3$ entier. Un **code de Hamming binaire** est un code de longueur $2^m - 1$ et de dimension $2^m - m - 1$. Sa matrice de contrôle de parité contient, en tant que colonnes, tous les m-uplets binaires non nuls (il y en a bien $2^m - 1$).

1.6 Performances

Probabilité d'erreur par mot : $P_{e,\text{mot}} \leqslant \sum_{i=t+1}^{n} \binom{n}{i} p^i (1-p)^{n-i}$ en considérant que au moins toutes les configurations dont le nombre d'erreurs est inférieur ou égal à t sont corrigées. On peut donc approcher la probabilité d'erreur par bit décodé (en supposant que bits d'information et bits de parits auront la même probabilité d'erreur) par $P_b \simeq \frac{d_{\min}}{n} \binom{t+1}{n} p^{t+1} (1-p)^{n-(t-1)} \overset{p \ll 1}{\simeq} \frac{d_{\min}}{n} \binom{t+1}{n} p^{t+1}$.

2 Théorie de l'information

On montrera ici que la probabilité d'erreur peut être rendue artificiellement faible pour peu que R ne dépasse pas un certain seuil et sous l'hypothèse que k et n tendent vers l'infini.

2.1 Entropie et information mutuelle

Def. Soit X une v.a. sur $\mathcal{X}$ fini avec loi de probabilité $\mathbf{P}_X$. Son **entropie** est $H(X) := -\sum_{x \in \mathcal{X}} \mathbf{P}_X(x) \log_2(\mathbf{P}_X(x))$ avec, par convention, $0 \cdot \log_2(0) = 0$.

L'entropie permet de capter le degré d'incertitude contenue dans une variable aléatoire. Elle ne dépend pas des valeurs prises, mais seulement des probabilités associées.

Th (Valeurs extrêmes de l'entropie). *Pour toute v.a. X sur $\mathcal{X}$ fini, on a*

$$0 \leqslant H(X) \leqslant \log_2(|\mathcal{X}|) .$$

En outre $H(X) = 0$ si et seulement si X est déterministe et $H(X) = \log_2(|\mathcal{X}|) \iff X \sim \mathcal{U}(\mathcal{X})$.

Def. La fonction d'**entropie binaire** est définie par $H_b(p) := -p \log_2(p) - (1-p) \log_2(1-p)$ où $p \in [0\,;1]$. Donc $H_b(p) = H(X)$ si $X \sim \mathcal{B}(p)$.

Def. Soit X et Y deux v.a. sur $\mathcal{X}$ et $\mathcal{Y}$ discrets avec $\mathbf{P}_{XY}$ comme loi de probabilité conjointe. Leur **entropie conjointe** est définie comme $H(X,Y) := -\sum_{x \in \mathcal{X}, y \in \mathcal{Y}} \mathbf{P}_{XY}(x,y) \log_2 (\mathbf{P}_{XY}(x,y))$.

L'entropie conjointe est symétrique et on retrouve $H(X,X) = H(X)$.

Th (Valeurs extrêmes de l'entropie conjointe). *Pour toute paire de v.a. X et Y sur $\mathcal{X}$ et $\mathcal{Y}$ discrets, on a*

$$\max\{H(X), H(Y)\} \leqslant H(X,Y) \leqslant H(X) + H(Y) .$$

En outre $H(X,Y) = H(X) \iff Y = g(X)$ avec g quelconque, et $H(X,Y) = H(X) + H(Y) \iff X \perp\!\!\!\perp Y$.

Def. L'**entropie conditionnelle** de X sachant Y est
$H(x \mid Y) := \sum_{y \in \mathcal{Y}} \mathbf{P}_Y(y) H(X \mid Y = y) = -\sum_{x \in \mathcal{X}} \sum_{y \in \mathcal{Y}} \mathbf{P}_{XY}(x,y) \log_2(\mathbf{P}_{X|Y}(x \mid y))$.

Prop (Valeurs extrêmes de l'entropie conditionnelle). *Pour toute paire de v.a. X et Y sur $\mathcal{X}$ et $\mathcal{Y}$, on a $0 \leqslant H(X \mid Y) \leqslant H(X)$. En outre $H(X \mid Y) = 0 \iff X = f(Y)$ avec f une fonction quelconque, et $H(X,Y) = H(X) \iff X \perp\!\!\!\perp Y$.*

Prop (Règle de chaînage (*chain rule*)). *$H(X,Y) = H(Y) + H(X \mid Y) = H(X) + H(Y \mid X)$ pour n'importe quel X et Y.*

Def. L'**information mutuelle** de X et Y est $I(X;Y) := H(X) + H(Y) - H(X,Y) = H(X) - H(X \mid Y) = H(Y) - H(Y \mid X)$.

Cette information mutuelle (car symétrique) permet de quantifier l'information commune entre X et Y.

Prop (Valeurs extrêmes de l'information mutuelle). *Pour toute paire de v.a. X et Y sur $\mathcal{X}$ et $\mathcal{Y}$ finis, on a $0 \leqslant I(X;Y) \leqslant \min\{H(X), H(Y)\}$. En outre $I(X;Y) = 0 \iff X \perp\!\!\!\perp Y$, et $I(X;Y) = H(X) \iff X = f(Y)$ avec f une fonction quelconque.*

2.2 Définition et théorème de la capacité pour le DMC

Not. On note $f^{(n)}$ et $g^{(n)}$ les fonctions de codage et décodage pour indiquer la taille des blocs. On les inclut dans la définition des codes, qui sont alors spécifiés par $(n, k, f^{(n)}, g^{(n)})$.

On considère les probabilités d'erreur en bloc $P_e^{(n)} := \mathbf{P}\left((D_1, \dots, D_k) \neq (\hat{D}_1, \dots, \hat{D}_k)\right)$.

Def. Un taux $R > 0$ est dit atteignable sur un DMC $(\mathcal{X}, \mathcal{Y}, \mathbf{P}_{Y|X})$ s'il existe une suite $\left(n, k = \lfloor nR \rfloor, f^{(n)}, g^{(n)}\right)_{n \in \mathbf{N}^*}$ de codes, telle que $P_e^{(n)} \underset{n\to\infty}{\longrightarrow} 0$.

Def. La **capacité** C d'un DMC $(\mathcal{X}, \mathcal{Y}, \mathbf{P}_{Y|X})$ est $C = \max_{\mathbf{P}_X} I(X;Y)$ où la maximisation se fait sur toutes les lois de probabilité de X et où $Y \sim \mathbf{P}_{X|Y}(\cdot \mid X)$. Donc, dans cette expression, la paire (X, Y) suit la loi de probabilité $\mathbf{P}_{XY}(x, y) = \mathbf{P}_X(x)\mathbf{P}_{Y|X}(x \mid Y)$.

Th (Théorème de Shannon de la capacité). *Pour un DMC $(\mathcal{X}, \mathcal{Y}, \mathbf{P}_{Y|X})$: tous les débits $0 < R < C$ sont atteignables et aucun débit $R > C$ ne l'est.*

Prop. *La capacité d'un* $\mathrm{BSC}(p)$ *est égale à* $C_{\mathrm{BSC}(p)} = 1 - H_{\mathrm{b}}(p)$.

Prop. *La capacité d'un* $\mathrm{BEC}(\epsilon)$ *est égale à* $C_{\mathrm{BEC}(\epsilon)} = 1 - \epsilon$.

3 Modulations numériques

3.1 Canal de propagation

Hyp. Le bruit $b(t)$ est i.i.d. gaussien de moyenne nulle, de fonction d'autocorrélation $r_{bb}(\tau) := \mathbf{E}(b(t+\tau)b(t))$ et satisfait $r_{bb}(\tau) = \frac{N_0}{2}$.

Prop. *Soit $x(t)$ le signal émis et $y(t)$ le signal reçu. Le canal multi-trajets conduit à $y(t) = c_p(t) \star x(t) + b(t)$.*

Lorsque $c_p(t) = \delta(t)$ le canal est appelé **canal gaussien**, car seul le bruit gaussien vient perturber la transmission. Dans ce cas $y(t) = x(t) + b(t)$. C'est notamment vrai dans les cas suivants :

- Faisceaux hertziens : entre antenne fixes avec une visibilité directe entre elles $\to$ antenne émettrice directive orienté vers l'antenne de réception $\to$ ni dispersion, ni écho.
- Liaisons satellitaires : en première approximation l'onde ne subit pas d'obstacle entre l'émission par le satellite et la réception par une antenne parabolique.
- Réseaux câblés co-axiaux : produisent très peu de multitrajets.

Lorsque $c_p(t) \neq \delta(t)$, le canal est appelé **canal sélectif en fréquence**, car alors $Y(f) = C_p(f)X(f) + B(f)$ (en prenant les TF) $\to C_p$ n'est plus constante et donc agit différemment selon les fréquences.

3.2 Description de l'émetteur

Def. Transmission d'un signal provenant d'un code correcteur d'erreur, mathématiquement on a :

$$x(t) = \sum_{n=0}^{N-1} s_n g(t - nT_s)$$

avec $g(t)$ le filtre d'émission, $\{s_n\}_n$ la suite de symboles s'exprimant en fonction des données et T_s le temps-symbole.

Def. On note $\mathcal{M}$ l'ensemble des valeurs possibles pour chaque symbole s_n et $M = \mathrm{Card}(\mathcal{M})$ le **nombre de valeurs possibles pour chaque symbole**.

Une fois M fixé, on appelle **constellation** la manière dont sont répartis les M valeurs possibles des symboles sur l'axe des réels. Voici les deux constellations les plus utilisées :

— M-OOK : $\{0; A; 2A; ...; (M-2)A; (M-1)A\}$, les valeurs sont espacées de A.

— M-PAM : $\{-(M-1)A; -(M-3)A; ...; -A; A; ...; (M-3)A; (M-1)A\}$, les valeurs sont espacées de $2A$.

Hyp. Soit $\{s_n\}_{n=0,\dots,N-1}$ et $\mathcal{M} = \left\{s^{(m)}\right\}_{m=0,\cdots,M-1}$ rangé par ordre croissant. On considère que :

- La suite $\{s_n\}_{n=0,\dots,N-1}$ est i.i.d.
- Chaque s_n prend une valeur dans $\mathcal{M}$ de façon équiprobable : $\forall m, n, \mathbf{P}\left(s_n = s^{(m)}\right) = \frac{1}{M}$.

On définit la **moyenne symbole** m_s, la **variance symbole** σ_s^2 et **l'énergie symbole** E_s. Comme la suite des symboles est i.i.d elles ne dépendent pas de n et on a :

$$\forall n, \quad m_s = \frac{1}{M}\sum_{m=0}^{M-1} s^{(m)} \qquad \sigma_s^2 = \frac{1}{M}\sum_{m=0}^{M-1} (s^{(m)} - m_s)^2 \qquad E_s = \frac{1}{M}\sum_{m=0}^{M-1} {s^{(m)}}^2 = m_s^2 + \sigma_s^2$$

Prop. *Résultats à connaître (pour aller plus vite) :*

	m_s	σ_s^2	E_s
OOK	$\frac{A(M-1)}{2}$	$\frac{A^2(M^2-1)}{12}$	$\frac{A^2(2M^2-3M+1)}{6}$
PAM	0	$\frac{A^2(M^2-1)}{3}$	$\frac{A^2(M^2-1)}{3}$

Prop. *L'énergie consommée pour émettre un bit d'information (**énergie bit**) s'écrit :*

$$E_b = \frac{1}{\log_2(M)} \left(E_s E_g + m_s \sum_{n \neq 0} h_n \right)$$

*Avec $E_g = \int g(t)^2 dt$ l'énergie du filtre d'émission et $h_n = h(nT_s)$ où $h(t) = g(-t) \star g(t)$. Sauf indication contraire, le filtre choisi amènera toujours à $E_g = 1$ et $m_s = 0$ **ou** $h_n = 0$ pour tout $n \neq 0$. Ainsi, on retiendra :*

$$E_b = \frac{E_s}{\log_2(M)}$$

3.3 Description du récepteur

Prop. *La suite optimale au sens de la probabilité d'erreur est la suivante :*

$$z_n = z(nT_s) \quad \textit{avec} \quad z(t) = g(-t) * y(t)\,.$$

Prop. *Dans le contexte d'un canal gaussien on a $z_n = h_n \star s_n + w_n$.*

Def (Filtre de Nyquist, Filtre en racine de Nyquist).

- Un filtre de réponse impulsionnelle $l(t)$ est dit de Nyquist, si et seulement si : $l_n = l(nT_s) = \begin{cases} \neq 0 & \text{pour} \quad n = 0 \\ 0 & \text{pour} \quad n \neq 0 \end{cases}$
- Un filtre est dit en racine de Nyquist si et seulement si le filtre $l(-t) \star l(t)$ est un filtre de Nyquist. Autrement dit, le filtre convolué à son filtre adapté est de Nyquist.

Prop (Canal gaussien à temps discret). *Une fois la contrainte de filtre de Nyquist vérifiée par $h(t)$, l'équation $z_n = h_n \star s_n + w_n$ se simplifie en :*

$$z_n = s_n + w_n$$

Avec w_n un bruit blanc gaussien de variance $\frac{N_0}{2}$.

Prop. *La largeur de bande, notée B, de tout filtre de Nyquist ou en racine de Nyquist vérifie $B \geqslant \dfrac{1}{T_s}$.*

3.4 Détecteur optimal

Prop. *Soit $\{s_n\}_n$ une suite de symboles et le canal gaussien à temps discret donnée plus haut ($z_n = s_n + w_n$). Alors le détecteur optimal obtient le symbole $\hat{s}_n$ de la manière suivante :*

$$\hat{s}_n = \begin{cases} s^{(0)} & \textit{si} \quad z_n \in]-\infty, t^{(0)}] \\ s^{(m)} & \textit{si} \quad z_n \in]t^{(m-1)}, t^{(m)}] \textit{ pour } m \in \{1, \cdots, M-2\} \\ s^{(M-1)} & \textit{si} \quad z_n \in]t^{(M-2)}, +\infty[\end{cases}$$

Avec, pour $m \in \{0, \cdots, M-2\}$, les seuils suivants : $t^{(m)} = \dfrac{s^{(m)} + s^{(m+1)}}{2}$.

3.5 Performances

Prop. *Si l'étiquetage permet d'avoir seulement un bit de différent entre deux symboles adjacents, alors on a cette relation :*

$$P_b = \frac{1}{\log_2(M)} P_e$$

Avec P_b et P_e les probabilités d'erreur bit et symbole respectives. (NdR : En réalité, le symbole $=$ dans l'équation plus haut est un $\approx$. Cependant, pour les applications en COM105, on a bien un $=$.)

Prop. *Si l'hypothèse sur l'émetteur est vérifiée, la constellation 2-PAM admet les performances suivantes :*

$$P_b = P_e = Q\left(\sqrt{2\frac{E_b}{N_0}}\right) \quad \textit{avec} \quad Q := x \mapsto \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-u^2/2}\, \mathrm{d}u\,.$$

Rem. On a $P_{\text{dB}} = 10 \cdot \log_{10}(P)$.