

Chapitre 3

Récepteur optimal

3.1 Introduction

Le signal obtenu en sortie de canal, $r(t)$, résultant de la transmission d'un signal $s_i(t)$ parmi N possibles est une version déformée, déphasée ou retardée et perturbée par un bruit, du signal émis.

Le problème du récepteur (pris ici au sens large) est de *démoduler* le signal reçu en un vecteur \mathbf{r} , et de *décider* de l'affectation de ce vecteur à l'un des mots de l'alphabet de départ (décision).

Figure 3.1 – Structure d'une chaîne d'émission-réception

La décision doit être prise au vu, ou avec les connaissances des probabilités d'émission et de transition, de contraintes sur les probabilités de fausse alarme (Pfa) ou de mauvaise détection (Pmd). Il s'agit donc d'élaborer une *règle de décision*, et une *structure* du récepteur, auquel on devra associer des *performances*. Sauf indication contraire, nous supposons dans la suite qu'un symbole est émis toutes les T secondes, le canal est parfait (les signaux ne sont pas distordus), le bruit est additif, et gaussien. Dans la plus grande partie de ce chapitre, nous supposons que les paramètres des signaux reçus — et en particulier la fréquence et la phase, sont connus ou estimés au niveau du récepteur (voir le chapitre sur la synchronisation) Les récepteurs correspondant sont alors appelés *récepteurs cohérents*.

Nous débuterons en 3.2 par l'introduction de quelques outils mathématiques, notamment sur les problèmes de représentation des signaux et présenterons la notion de filtre adapté. Nous aborderons ensuite en 3.3 les fondements de la théorie de la détection. Les stratégies bayésienne et de Neyman-Pearson permettent alors d'élaborer des règles de décision, et finalement les opérations à effectuer au niveau du récepteur, c'est-à-dire conduisent à la structure du récepteur. Nous indiquerons que ces approches peuvent être équivalentes à l'élaboration d'un récepteur minimisant la probabilité d'erreur. Nous expliquerons en 3.4 comment on peut étendre la méthodologie présentée en 3.3 à la détection de signaux à temps continu, qui sont les signaux observés en pratique en entrée d'un récepteur. Nous aborderons alors en 3.5 le problème de la communication binaire en présence de bruit blanc et établirons alors la structure du récepteur et les performances des modulations binaires, en terme de probabilité d'erreur. En utilisant la même démarche, nous étendrons ensuite ces résultats au cas de communications M -aires, section 3.6, où il s'agit de détecter un symbole parmi M possibles. À nouveau, nous donnerons la structure du récepteur optimal puis nous discuterons des performances dans quelques cas particuliers. Jusque là, nous aurons supposé, implicitement, que les signaux sont sans mémoire. Dans le cas où les symboles sont interdépendants, le récepteur peut tirer parti de cette dépendance, ce qui conduit à détecter une séquence de symboles. La détection au sens du maximum de vraisemblance d'une séquence de symboles sera abordée en 3.7 et nous permettra notamment de présenter l'algorithme de Viterbi. Enfin, nous aborderons en 3.8 le cas d'une phase inconnue en réception, qui conduit à un récepteur non cohérent, puis nous terminerons en traitant le cas de la réception d'un signal issu d'un canal de Rayleigh.

3.2 Outils

On décrit ici différentes méthodes de caractérisation et de représentation des signaux. L'idée sous-jacente est de trouver une bonne représentation pour le problème traité, qui rend parfois la solution de celui-ci très simple. D'autre part, on présentera la notion de filtre adapté, filtre qui est construit afin de maximiser le rapport signal-à-bruit. Ce type de filtre est rencontré dans la plupart des récepteurs.

3.2.1 Représentation de signaux déterministes

On considère un signal $x(t)$ défini sur $[0, T]$. On peut représenter $x(t)$ sur une base de fonctions $\Phi_i(t)$ par

$$x(t) = \sum_{i=1}^{\infty} x_i \Phi_i(t),$$

où les $\Phi_i(t)$ sont un ensemble de fonctions orthonormées. On peut par exemple choisir

$$\Phi_i(t) = \left(\frac{1}{T}\right)^{\frac{1}{2}} \quad \Phi_{2m}(t) = \left(\frac{2}{T}\right)^{\frac{1}{2}} \cos\left(\frac{2\pi}{T}mt\right)$$

$$\Phi_{2m+1}(t) = \left(\frac{2}{T}\right)^{\frac{1}{2}} \sin\left(\frac{2\pi}{T}mt\right)$$

Lorsque l'on tronque le développement à un ordre N , l'erreur de représentation est $e_N(t) = x(t) - \sum_{i=1}^N x_i \Phi_i(t)$. La minimisation de l'énergie de l'erreur $\int_0^T e_N(t)^2 dt$ conduit alors à choisir les coefficients x_i comme

$$x_i = \int_0^T x(t) \Phi_i(t) dt.$$

Lorsque l'énergie de l'erreur tend vers 0 avec N , l'ensemble des $\Phi_i(t)$ représente un ensemble orthonormé complet.

Deux méthodes sont possibles pour générer les coefficients x_i . La première méthode consiste à multiplier $x(t)$ par chacune des $\Phi_i(t)$ puis à intégrer sur une durée T . Il s'agit du produit scalaire entre $x(t)$ et $\Phi_i(t)$. On parle également d'une opération de corrélation (x_i résulte de l'intercorrélation entre $x(t)$ et $\Phi_i(t)$, prise pour un retard nul).

Figure 3.2 – Calcul des coefficients par corrélation.

La seconde méthode consiste à passer le signal $x(t)$ dans un ensemble de filtres de réponse impulsionnelle $h_i(t) = \Phi_i(T - t)$ et à conserver la sortie au temps T .

$$\int_0^T x(t) h_i(T - t) dt = \int_0^T x(t) \Phi_i(t) dt = x_i$$

Figure 3.3 – Calcul des coefficients par filtrage.

Ce type de représentation permet de considérer un signal comme un point d'un espace N -dimensionnel (N éventuellement infini).

3.2.2 Représentation des signaux aléatoires

Nous avons vu dans la section précédente 3.2.1 qu'il est possible de représenter un signal déterministe, par un développement en série. On peut étendre cette idée aux processus aléatoires sur $[0, T]$:

$$x(t) = \lim_{N \rightarrow \infty} \sum_{i=1}^N x_i \Phi_i(t) \quad t \in [0, T]$$

avec

$$x_i = \int_0^T x(t) \Phi_i(t) dt$$

La limite utilisée devra être une limite en moyenne quadratique, ce qui signifie que

$$\lim_{N \rightarrow \infty} E \left[\left(x(t) - \sum_{i=1}^N x_i \Phi_i(t) \right)^2 \right] = 0, \quad t \in [0, T]$$

Les coefficients x_i sont des variables aléatoires. En recherchant une décomposition qui conduise à des coefficients *décorrélés*, c'est-à-dire tels que

$$E \left[(x_i - m_i) (x_j - m_j) \right] = \lambda_i \delta_{ij},$$

avec $m_i = E[x_i]$, on obtient la condition suivante sur les fonctions de base $\Phi_i(t)$

$$\lambda_i \Phi_i(t) = \int_0^T C_x(t, u) \Phi_i(u) du$$

où le noyau $C_x(t, u)$ de cette équation intégrale est une fonction de covariance, définie non négative et symétrique : $C_x(t, u) = E_x \left[(X(t) - m_x(t)) (X(u) - m_x(u)) \right]$. Les fonctions de base $\Phi_i(t)$ sont déterminées en résolvant l'équation intégrale précédente et l'on peut alors développer un signal aléatoire sur cette base de fonctions, ce développement, appelé développement de Kahunen-Loève, étant optimal en moyenne quadratique.

3.2.3 Filtrage adapté

Considérons simplement pour le moment le modèle

$$y(t) = x(t) + b(t).$$

L'approche habituelle consiste à rechercher à minimiser l'effet du bruit d'observation. On cherche alors à construire un filtre $h(t)$ tel que le rapport signal-à-bruit en sortie soit maximal, à un instant T , appelé instant de décision. Cet instant T devra être défini pour que l'observation ait été effectuée et que le filtre ait agi.

Notons $z(t)$ la sortie du filtre de réponse impulsionnelle $h(t)$. On a alors

$$z(t) = (h * y)(t) = (h * x)(t) + (h * b)(t).$$

On écrit ainsi la sortie comme la somme de la sortie non-bruitée et de la contribution du bruit. Le rapport signal-à-bruit vaut ainsi

$$\rho(t) = \frac{|(h * x)(t)|^2}{E[|(h * B)(t, \omega)|^2]},$$

où le numérateur représente la puissance instantanée de la sortie non bruitée et le dénominateur la puissance liée au bruit. On évalue ce rapport signal-à-bruit à l'instant de décision T .

$$\rho(T) = \frac{|(h * x)(T)|^2}{E[|(h * B)(T, \omega)|^2]}.$$

En développant les produits de convolution, on obtient

$$\rho(T) = \frac{|\int_{-\infty}^{+\infty} h(u)x(T-u)du|^2}{E\left[|\int_{-\infty}^{+\infty} h(u)b(T-u)du|^2\right]}.$$

En ce qui concerne tout d'abord le dénominateur, il s'agit là de la puissance d'un signal aléatoire à la sortie d'un filtre, et l'on a donc

$$E[|(h * b)(t)|^2] = \int_{-\infty}^{+\infty} S_{bb}(f)df = \sigma^2 \int_{-\infty}^{+\infty} |H(f)|^2 df,$$

où σ^2 est la puissance du bruit d'entrée.

L'inégalité de Schwartz permet de majorer le numérateur :

$$\left| \int_{-\infty}^{+\infty} h(u)x(T-u)du \right|^2 \leq \int_{-\infty}^{+\infty} |h(u)|^2 du \int_{-\infty}^{+\infty} |x^*(T-u)|^2 du,$$

avec égalité lorsque les vecteurs $h(u)$ et $x^*(T-u)$ sont colinéaires. L'égalité de Parseval-Plancherel, qui exprime la conservation du produit scalaire entraîne que

$$\int_{-\infty}^{+\infty} |h(u)|^2 du = \int_{-\infty}^{+\infty} |H(f)|^2 df.$$

On en déduit donc que le rapport signal-à-bruit est majoré selon

$$\rho(T) \leq \frac{\int_{-\infty}^{+\infty} |x^*(T-u)|^2 du}{\sigma^2} = \frac{E_x}{\sigma^2},$$

où E_x est l'énergie du signal $x(t)$. L'égalité est atteinte lorsque $h(u)$ et $x^*(T-u)$ sont colinéaires, c'est-à-dire

$$h(u) = kx^*(T-u),$$

où k est une constante arbitraire. Le filtre optimal maximisant le rapport signal-à-bruit en sortie, à l'instant T , est ainsi le filtre dont la réponse impulsionnelle est la copie retournée et décalée dans le temps du signal que l'on cherche à retrouver. En ce sens, le filtre est adapté au signal.

La relation de filtrage de $y(t, \omega)$ avec une « copie retournée » équivaut en fait à effectuer une intercorrélacion (au sens déterministe). En effet,

$$\begin{aligned} z(t) &= \int_{-\infty}^{+\infty} h(u)y(t-u)du = \int_{-\infty}^{+\infty} x^*(T-u)y(t-u)du \\ &= \int_{-\infty}^{+\infty} x^*(T+v)y(t+v)dv = \int_{-\infty}^{+\infty} x^*(T-t+v)y(v)dv, \end{aligned}$$

soit

$$z(t) = R_{yx}(T-t).$$

Le récepteur optimal consiste donc à calculer l'intercorrélacion entre le signal reçu $y(t)$ et le signal attendu $x(t)$. On parle alors souvent de récepteur à corrélacion.

Dans le domaine fréquentiel, la fonction de transfert du filtre adapté, la transformée de Fourier de $h(u)$ vaut simplement

$$H(f) = \text{TF}\{h(u)\} = \text{TF}\{kx^*(T-u)\} = kX(f)^* e^{-j2\pi fT},$$

c'est-à-dire le conjugué de la transformée de Fourier du signal d'intérêt, à un facteur et une phase près. Ainsi, la portion non bruitée du signal de sortie est

$$(h * x)(t) = \text{TF}^{-1}\{kX(f)^* e^{-j2\pi fT} X(f)\} = k\text{TF}^{-1}\{|X(f)|^2 e^{-j2\pi fT}\},$$

et à l'instant T ,

$$(h * x)(T) = k \int_{-\infty}^{+\infty} |X(f)|^2 df = kE_x.$$

3.3 Fondements

3.3.1 Modélisation

Le problème central est de choisir l'un des mots de l'alphabet d'émission (l'ensemble des mots codes c_i issus du codeur de canal) à partir du signal reçu r . La règle de décision habituelle consiste à effectuer une partition de l'ensemble des mots de réception.

Figure 3.4 – Partition de l'ensemble des signaux reçus

On note Z l'ensemble de tous les mots r susceptibles d'être reçus, et on note Z_i l'ensemble de tous les mots r reçus tel que c_i soit décidé. C'est l'ensemble *décodeur* associé à c_i . Le problème est alors de trouver la meilleure partition possible, au sens d'un critère à définir, de l'ensemble des signaux reçus Z .

Dans le schéma précédent, il y aura erreur dès que le signal reçu r n'appartient pas à Z_i , alors que c_i a été émis :

$$\begin{aligned} P(\text{erreur} \mid c_i) &= P(r \in \overline{Z}_i \mid c_i), \\ &= 1 - P(r \in Z_i \mid c_i). \end{aligned}$$

En moyenne, la probabilité d'erreur au décodage vaut

$$P(\text{erreur}) = \sum_{i=1}^N P(c_i) P(\text{erreur} \mid c_i)$$

$$\text{soit } P(\text{erreur}) = \sum_{i=1}^N \sum_{j \neq i} P(\text{décider } c_j \mid c_i \text{ émis}).$$

Lorsque les symboles émis sont équiprobables, on obtient

$$P(\text{erreur}) = \frac{1}{N} \sum_{i=1}^N P(\text{erreur} \mid c_i)$$

Construction de la partition :

On distingue en général deux règles, selon que l'on considère les probabilités *a priori* ou non (ou que l'on prend une loi uniforme pour la source). La probabilité *a posteriori* est obtenue en appliquant la règle de Bayes :

$$P(c_i \mid r) = \frac{P(c_i) P(r \mid c_i)}{P(r)}$$

La probabilité $P(c_i \mid r)$ est la probabilité *a posteriori*, c'est-à-dire la probabilité de c_i connaissant l'observation r . La probabilité $P(r \mid c_i)$ est la vraisemblance de l'observation, c'est-à-dire la probabilité d'observer r lorsque c_i est émis.

1. règle du maximum *a posteriori*

$$Z_i = \left\{ r \mid P(c_i \mid r) = \max_K P(c_K \mid r) \right\}$$

cette partition dépend à la fois de la loi d'entrée $P(c_i)$ et des probabilités de transition $P(r \mid c_i)$, (c'est à dire du canal).

2. règle du maximum de vraisemblance

$$Z_i = \left\{ r \mid P(r \mid c_i) = \max_K P(r \mid c_K) \right\}$$

Lorsque $P(c_i)$ est uniforme, maximiser $P(c_K \mid r)$ est équivalent à maximiser $P(r \mid c_K)$, et les deux règles sont alors équivalentes.

On considère les quatre espaces suivants et leurs relations :

Figure 3.5 – Espaces de signaux pour une chaîne de transmission.

Ces différents espaces sont des espaces de fonctions/variables certaines ou aléatoires, de dimension finie ou infinie.

Les passages de M à S et de R à D sont généralement des applications certaines, alors que le canal (passage de S à R) est aléatoire et caractérisé par une distribution de probabilité conditionnelle.

3.3.2 Stratégie bayésienne-Notions de coût

On définit un « coût » $C(s, D)$, qui représente la pénalité ou la perte associée à l'événement conjoint (s émis, Décision D prise). Le coût attaché à l'événement conjoint représente donc en quelque sorte le prix à payer pour chacune des décisions. Afin de privilégier les bonnes décisions, on attachera alors un coût important aux mauvaises décisions et un coût nul, ou négatif aux bonnes décisions. Notons que certaines mauvaises décisions peuvent être « plus graves » que d'autres ; on pourra ordonner ces décisions en jouant sur les coûts associés. Lorsque les coûts sont fixés, on pourra chercher à minimiser le coût moyen d'une décision, c'est-à-dire la moyenne des coûts pondérée par les probabilités de chacun des événements conjoints : c'est la stratégie bayésienne.

La stratégie bayésienne consiste à minimiser le coût moyen. On parle alors de la minimisation du risque bayésien.

$$R = E_{D,s} [C(s, D)].$$

Examinons tout d'abord comment ceci se traduit dans le cas binaire : la source peut émettre deux messages s_0 et s_1 , et il s'agit de décider quel message a été émis au vu de l'observation r . La stratégie doit alors amener à trouver la meilleure partition de l'ensemble des observations possibles Z en deux sous ensembles, auxquels on associe des décisions D_0 et D_1 .

Figure 3.6 – Partition des observations et décisions.

On note C_{ij} le coût associé à la décision D_i (on décide que s_i est émis), lorsque l'hypothèse H_j est vraie (s_j a été émis).

La stratégie bayésienne consiste à trouver la partition Z_0, Z_1 tel que le risque bayésien R soit minimal. Ici, le risque bayésien s'exprime alors, en fonction des probabilités de transition et des régions de décisions, comme :

$$R = P(H_0) C_{10} + P(H_1) C_{11} + \int_{Z_0} P(H_1) (C_{01} - C_{11}) p(r | H_1) - P(H_0) (C_{10} - C_{00}) p(r | H_0) dr$$

Les deux premiers termes de R sont des termes constants. L'intégrale représente le coût contrôlé par les points assignés à Z_0 . En supposant que le coût associé à une mauvaise décision est plus important que celui associé à une bonne décision, on note que l'intégrande est constitué par la différence de deux termes positifs.

Comme on veut trouver une partition de Z qui permette de minimiser le risque bayésien, il faut que tous les points r tels que

$$P(H_0) (C_{10} - C_{00}) p(r | H_0) > P(H_1) (C_{01} - C_{11}) p(r | H_1)$$

soient assignés à l'ensemble Z_0 (et donc à la décision D_0), car ils contribuent alors à rendre l'intégrale négative. En suivant la même idée, tous les points r tels que l'inégalité précédente ne soit pas vérifiée doivent être exclus de Z_0 , c'est-à-dire affectés à Z_1 (décision D_1). Ceci nous permet donc de définir la partition Z_0/Z_1 qui minimise le risque bayésien. Ceci s'écrit

$$P(H_0) (C_{10} - C_{00}) p(r | H_0) \begin{matrix} \rangle \\ \langle \end{matrix} \begin{matrix} H_0 \\ H_1 \end{matrix} P(H_1) (C_{01} - C_{11}) p(r | H_1)$$

Plus classiquement, on écrit le *test du rapport de vraisemblance (Likelihood Ratio Test)*

$$\frac{p(r | H_1)}{p(r | H_0)} \begin{matrix} \rangle \\ \langle \end{matrix} \begin{matrix} H_1 \\ H_0 \end{matrix} \frac{P(H_0)}{P(H_1)} \cdot \frac{(C_{10} - C_{00})}{(C_{01} - C_{11})}$$

ou encore

$$\Lambda(r) \begin{matrix} \rangle \\ \langle \\ H_1 \end{matrix} \eta.$$

Il est important de noter que

- le rapport de vraisemblance $\Lambda(r)$ ne dépend que des probabilités de transition $p(r|H_i)$, c'est-à-dire du canal, et non des coûts et des probabilités *a priori*,
- le rapport de vraisemblance, comme rapport de deux distributions, est une *quantité monodimensionnelle*,
- le seuil η ne dépend que des coûts et des probabilités *a priori*,
- le rapport de vraisemblance, comme rapport de deux fonctions d'une variable aléatoire est lui-même une variable aléatoire. Les performances du test sont alors liées aux caractéristiques statistiques de $\Lambda(r)$.

L'implantation du détecteur consiste donc à calculer le rapport de vraisemblance $\Lambda(r)$, puis à comparer ce rapport à un seuil η .

Figure 3.7 – Structure d'un détecteur à maximum de vraisemblance.

On utilise souvent le logarithme du rapport de vraisemblance, $l(r)$ et le test devient alors

$$l(r) = \log(\Lambda(r)) \begin{matrix} \rangle \\ \langle \\ H_0 \end{matrix} \log \eta.$$

En général, on prend simplement $C_{00} = C_{11} = 0$, pour les bonnes décisions, et $C_{01} = C_{10} = 1$ pour les mauvaises décisions. Si en outre les deux hypothèses sont équiprobables, le test peut s'écrire

$$p(r|H_1) \begin{matrix} \rangle \\ \langle \\ H_0 \end{matrix} p(r|H_0),$$

qui est simplement la *règle du maximum de vraisemblance*.

Lien avec la probabilité d'erreur

Lorsque

$$\begin{cases} C_{00} = C_{11} = 0 \\ C_{01} = C_{10} = 1 \end{cases}$$

et si les probabilités *a priori* sont équiprobables, alors le risque est simplement

$$\begin{aligned} R &= C_{01} P(D_0|H_1) + C_{10} P(D_1|H_0) \\ &= P(D_0|H_1) + P(D_1|H_0) = P(\text{erreur}). \end{aligned}$$

Ainsi, dans ce cas particulier, la *minimisation du risque bayésien* R est équivalente à la *minimisation de la probabilité d'erreur*.

Exemple : détection d'un signal constant dans du bruit blanc gaussien.

Le problème est le suivant : sous H_1 , la sortie de la source est une constante m ; sous H_0 , la sortie de la source est zéro. Avant l'observation, cette sortie est perturbée par un bruit blanc additif

gaussien centré et de variance σ^2 . Le signal reçu est échantillonné sur une durée T et on obtient ainsi N échantillons.

$$\begin{aligned} H_0 &: r(t) = b(t) & \mathbf{r} = \mathbf{b} \\ H_1 &: r(t) = m + b(t) & \mathbf{r} = m + \mathbf{b} \end{aligned}$$

La distribution de probabilité de b est

$$p(\mathbf{b}) = \prod_{i=1}^N p(b_i) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{b_i^2}{2\sigma^2}\right)$$

Par conséquent,

$$p(\mathbf{r}|H_0) = p(\mathbf{b}) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{r_i^2}{2\sigma^2}\right)$$

$$p(\mathbf{r}|H_1) = p(\mathbf{b} + m) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(r_i - m)^2}{2\sigma^2}\right)$$

Le test du rapport de vraisemblance fournit alors

$$\Lambda(\mathbf{r}) = \frac{p(\mathbf{r}|H_1)}{p(\mathbf{r}|H_0)} = \prod_{i=1}^N \exp\left(-\frac{1}{2\sigma^2} [(r_i - m)^2 - r_i^2]\right)$$

En prenant le logarithme, il vient

$$\log \Lambda(\mathbf{r}) \underset{H_0}{\overset{H_1}{\rangle}} \log(\eta),$$

soit

$$\sum_{i=1}^N r_i \underset{H_0}{\overset{H_1}{\rangle}} \frac{\sigma^2}{m} \log(\eta) + \frac{Nm}{2}$$

Le récepteur ne fait donc qu'additionner les observations et les compare à un seuil... Lorsque $\eta=1$, le test est simplement :

$$\frac{1}{N} \sum_{i=1}^N r_i \underset{H_0}{\overset{H_1}{\rangle}} \frac{m}{2},$$

c'est-à-dire qu'on calcule la moyenne arithmétique des observations, et qu'on décide que le signal m est présent si cette moyenne arithmétique est supérieure à $m/2$, et absent dans le cas contraire. Le test fournit alors un résultat remarquablement conforme à l'intuition.

3.3.3 Stratégie de Neyman-Pearson

La stratégie de Neyman-Pearson est essentiellement définie dans le cas binaire. On considère deux types d'erreur :

- L'erreur de fausse alarme : $\alpha = P_{fa} = P(D_1|H_0) = \int_{Z_1} p(r|H_0) dr$,
- L'erreur de mauvaise détection : $\beta = P_{md} = P(D_0|H_1) = \int_{Z_0} p(r|H_1) dr$,

La probabilité $P(D_1|H_1) = 1 - \beta$ est appelée la puissance du test. La stratégie consiste à se fixer une probabilité maximale de fausse alarme, et, cette probabilité étant fixée, à maximiser la puissance du test, c'est-à-dire à minimiser la probabilité de mauvaise détection.

Ceci conduit à un test du type rapport de vraisemblance,

$$\frac{p(r|H_1)}{p(r|H_0)} \underset{H_0}{\overset{H_1}{>}} \lambda.$$

où le paramètre de Lagrange λ est déterminé afin de satisfaire la contrainte :

$$P_{fa} = \int_{\lambda}^{\infty} p(\Lambda|H_0) d\Lambda = P_r(\Lambda > \lambda|H_0)$$

La stratégie de Neyman-Pearson conduit donc à la même structure de récepteur que la stratégie bayésienne : seul change le seuil de décision, qui est ici fixé par la probabilité de fausse alarme.

Exemple : détection d'un signal constant dans un bruit additif blanc gaussien (suite)
Nous avons obtenu le test suivant :

$$\sum_{i=1}^N r_i \underset{H_0}{\overset{H_1}{>}} \frac{\sigma^2}{m} \log(\eta) + \frac{Nm}{2}$$

Dans la stratégie de Neyman-Pearson, seul change ici le seuil η , $\eta = \lambda$, qui est fixé par la probabilité de fausse alarme.

Comme somme de variables gaussiennes indépendantes et de même variance, $\sum r_i$ est elle même une variable gaussienne, de variance $N\sigma^2$. On pose maintenant $l(\mathbf{r}) = \frac{1}{\sigma\sqrt{N}} \sum_{i=1}^N r_i$. Il s'agit là d'une variable gaussienne normée (de variance unité). Le test devient alors

$$l(\mathbf{r}) \underset{H_0}{\overset{H_1}{>}} \gamma = \frac{\sigma}{\sqrt{Nm}} \log(\lambda) + \frac{\sqrt{Nm}}{2\sigma},$$

Sous H_0 , chaque r_i étant centré, $l(\mathbf{r})$ est de moyenne nulle. Sous H_1 , les r_i sont de moyenne m et la moyenne de $l(\mathbf{r})$ vaut $\frac{m\sqrt{N}}{\sigma}$. La probabilité de fausse alarme est donc

$$P_{fa} = \int_{\gamma}^{\infty} p(l|H_0) dl = \frac{1}{\sqrt{2\pi}} \int_{\gamma}^{\infty} \exp\left(-\frac{l^2}{2}\right) dl = Q(\gamma),$$

où l'on a utilisé la fonction Q définie par

$$Q(x) \triangleq \frac{1}{\sqrt{2\pi}} \int_x^{+\infty} e^{-\frac{y^2}{2}} dy.$$

On rencontre également la fonction erf, (« *error function* »),

$$\text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-y^2} dy,$$

et sa fonction complémentaire erfc

$$\text{erfc}(x) = 1 - \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_x^{+\infty} e^{-y^2} dy,$$

avec les relations suivantes :

$$\begin{aligned} Q(x) &= \frac{1}{2} \text{erfc}\left(\frac{x}{\sqrt{2}}\right), \\ \text{erfc}(x) &= 2Q(\sqrt{2}x). \end{aligned}$$

Les fonctions Q , erf et erfc sont tabulées.

À probabilité de fausse alarme fixée, on déduit des tables la valeur de γ , et par suite celle de λ .

La probabilité de détection vaut

$$P_d = \int_{\gamma}^{+\infty} p(l|H_1) dl = \int_{\gamma}^{+\infty} \frac{1}{\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(l - \frac{m\sqrt{N}}{\sigma} \right)^2 \right] dl = Q \left(\gamma - \frac{m\sqrt{N}}{\sigma} \right),$$

et la probabilité de mauvaise détection vaut $P_{md} = 1 - P_d$.

3.3.4 Observation vectorielle - Cas gaussien général

On considère maintenant le problème de détection binaire

$$\begin{aligned} H_0 &: \mathbf{r} \sim N(\mathbf{m}_0, \mathbf{\Gamma}_0) \\ H_1 &: \mathbf{r} \sim N(\mathbf{m}_1, \mathbf{\Gamma}_1) \end{aligned}$$

où il s'agit de décider entre deux signaux gaussiens de moyennes respectives \mathbf{m}_0 et \mathbf{m}_1 , et de matrices de covariances $\mathbf{\Gamma}_0$ et $\mathbf{\Gamma}_1$.

Le rapport de vraisemblance est, comme précédemment

$$\Lambda(\mathbf{r}) = \frac{p(\mathbf{r}|H_1)}{p(\mathbf{r}|H_0)}$$

or $p(\mathbf{r}|H_i) = \frac{1}{\sqrt{(2\pi)^N \det \mathbf{\Gamma}_i}} \exp -\frac{1}{2} (\mathbf{r} - \mathbf{m}_i)^t \mathbf{\Gamma}_i^{-1} (\mathbf{r} - \mathbf{m}_i)$, où det désigne le déterminant. Alors

$$\begin{aligned} l(\mathbf{r}) = \log \Lambda(\mathbf{r}) &= \frac{1}{2} \left[(\mathbf{r} - \mathbf{m}_0)^t \mathbf{\Gamma}_0^{-1} (\mathbf{r} - \mathbf{m}_0) - (\mathbf{r} - \mathbf{m}_1)^t \mathbf{\Gamma}_1^{-1} (\mathbf{r} - \mathbf{m}_1) \right] \\ &\quad \left. \begin{array}{l} H_1 \\ \rangle \\ \langle \\ H_0 \end{array} \right\} \log(\eta) + \frac{1}{2} \log \left[\frac{\det \mathbf{\Gamma}_1}{\det \mathbf{\Gamma}_0} \right]. \end{aligned}$$

Nous pouvons maintenant examiner le problème de la détection de deux signaux certains, connus, \mathbf{s}_0 et \mathbf{s}_1 , à partir d'une observation bruitée.

$$\begin{aligned} H_0 &: \mathbf{r} = \mathbf{s}_0 + \mathbf{b} \quad N(\mathbf{s}_0, \mathbf{\Gamma}_b) \\ H_1 &: \mathbf{r} = \mathbf{s}_1 + \mathbf{b} \quad N(\mathbf{s}_1, \mathbf{\Gamma}_b) \end{aligned}$$

Compte tenu des résultats précédents, on obtient alors

$$\begin{aligned} l(\mathbf{r}) &= \frac{1}{2} \left[(\mathbf{r} - \mathbf{s}_0)^t \mathbf{\Gamma}_b^{-1} (\mathbf{r} - \mathbf{s}_0) - (\mathbf{r} - \mathbf{s}_1)^t \mathbf{\Gamma}_b^{-1} (\mathbf{r} - \mathbf{s}_1) \right] \\ &= \frac{1}{2} \left[2\mathbf{r}^t \mathbf{\Gamma}_b^{-1} (\mathbf{s}_1 - \mathbf{s}_0) - \mathbf{s}_1^t \mathbf{\Gamma}_b^{-1} \mathbf{s}_1 + \mathbf{s}_0^t \mathbf{\Gamma}_b^{-1} \mathbf{s}_0 \right] \left. \begin{array}{l} \rangle \\ \langle \end{array} \right\} \log \eta \end{aligned}$$

Il vient donc

$$\mathbf{r}^t \mathbf{\Gamma}_b^{-1} (\mathbf{s}_1 - \mathbf{s}_0) \left. \begin{array}{l} H_1 \\ \rangle \\ \langle \\ H_0 \end{array} \right\} \log \eta + \frac{1}{2} \left(\mathbf{s}_1^t \mathbf{\Gamma}_b^{-1} \mathbf{s}_1 - \mathbf{s}_0^t \mathbf{\Gamma}_b^{-1} \mathbf{s}_0 \right).$$

Cas particuliers :

Si le bruit d'observation est blanc, on a $\mathbf{\Gamma}_b = \sigma^2 \mathbf{1}$, où $\mathbf{1}$ est la matrice identité. Si d'autre part on a $\eta = 1$, alors

$$\mathbf{r}^t (\mathbf{s}_1 - \mathbf{s}_0) \left. \begin{array}{l} H_1 \\ \rangle \\ \langle \\ H_0 \end{array} \right\} \frac{1}{2} \left(\mathbf{s}_1^t \mathbf{s}_1 - \mathbf{s}_0^t \mathbf{s}_0 \right).$$

Si les deux signaux possèdent la même énergie $E_s = \mathbf{s}_1^t \mathbf{s}_1 = \mathbf{s}_0^t \mathbf{s}_0$, on obtient le simple test

$$\mathbf{r}^t \mathbf{s}_1 \underset{H_0}{\overset{H_1}{>}} \mathbf{r}^t \mathbf{s}_0,$$

qui consiste à comparer les produits scalaires entre les observations \mathbf{r} et les deux signaux attendus \mathbf{s}_0 et \mathbf{s}_1 .

Enfin, dans le cas où l'on cherche à détecter un seul signal \mathbf{s}_1 dans du bruit blanc ($\mathbf{s}_0 = \mathbf{0}$), le test obtenu consiste à comparer le résultat du produit scalaire à la moitié de l'énergie

$$\mathbf{r}^t \mathbf{s}_1 \underset{H_0}{\overset{H_1}{>}} \frac{1}{2} \mathbf{s}_1^t \mathbf{s}_1.$$

3.3.5 Test à M hypothèses

Nous avons présenté précédemment le cas de la décision entre deux hypothèses, ce qui correspond finalement à l'étude de signaux binaires. On peut généraliser ceci au test à M hypothèses, c'est-à-dire par exemple au problème d'affectation d'un symbole reçu à un point d'une constellation. On a ainsi un problème de détection d'un signal parmi M :

$$\begin{aligned} H_0 : & \quad r = s_0 + b \\ H_1 : & \quad r = s_1 + b \\ & \quad \vdots \\ & \quad \vdots \\ & \quad \vdots \\ H_{M-1} : & \quad r = s_{M-1} + b \end{aligned}$$

À ces hypothèses correspondent M^2 éventualités (D_i, H_j) , auxquelles on associe les coûts C_{ij} . La minimisation du coût moyen, c'est-à-dire du risque bayésien, conduit à diviser l'espace d'observation Z en M sous-espaces Z_i , tels que $r \in Z_i \Leftrightarrow$ Décision D_i . L'étude précise conduit ici simplement au fait qu'il faut tester les hypothèses deux à deux. En définissant les rapports de vraisemblance

$$\Lambda_{1|0}(r) \stackrel{D}{=} \frac{p_{r|H_1}}{p_{r|H_0}}, \quad \Lambda_{2|0}(r) \stackrel{D}{=} \frac{p_{r|H_2}}{p_{r|H_0}}, \quad \dots \quad \Lambda_{M-1|0}(r) \stackrel{D}{=} \frac{p_{r|H_{M-1}}}{p_{r|H_0}},$$

on a accès à l'ensemble des rapports de vraisemblance, puisque

$$\Lambda_{i|j}(r) \stackrel{D}{=} \frac{p_{r|H_i}}{p_{r|H_j}} = \frac{\Lambda_{i|0}}{\Lambda_{j|0}}.$$

On effectue ensuite tous les tests

$$\Lambda_{i|j}(r) \underset{H_j \text{ (ou } H_k, k \neq i, j)}{\overset{H_i \text{ (ou } H_k, k \neq i, j)}{>}} \frac{p_j}{p_i},$$

qui permettent de tester une hypothèse (mais une seule) contre une autre. En collectant l'ensemble des résultats, on obtiendra alors la décision.

Pour illustrer ceci, considérons simplement le cas d'une constellation à quatre états $\{m_0, m_1, m_2, m_3\} = \{1 + j, -1 + j, -1 - j, 1 - j\}$, c'est-à-dire qu'un symbole émis a_k est susceptible de prendre l'une

quelconque de ces quatre valeurs (hypothèses). Le signal reçu, r , est entâché par un bruit additif b , que l'on supposera gaussien de variance σ^2 . Dans ce cas, l'un des tests s'écrit simplement

$$\Lambda_{i|j}(r) = e^{\frac{1}{2\sigma^2}((r-m_j)^2 - (r-m_i)^2)} \begin{matrix} H_j \text{ (ou } H_k, k \neq i, j) \\ \langle \\ H_i \text{ (ou } H_k, k \neq i, j) \end{matrix} \frac{p_j}{p_i},$$

ce qui fournit simplement, en prenant le log et en supposant que les probabilités *a priori* sont uniformes ($p_i = p_j$),

$$(r - m_j)^2 \begin{matrix} H_j \text{ (ou } H_k, k \neq i, j) \\ \langle \\ H_i \text{ (ou } H_k, k \neq i, j) \end{matrix} (r - m_i)^2.$$

Ainsi, le test revient simplement à sélectionner le point de la constellation qui est le plus proche, au sens d'une distance euclidienne habituelle, du signal reçu... Si l'on reprend l'exemple de la constellation à 4 états, tester m_0 contre m_1 revient à affecter les points du demi-plan droit à m_0 préférentiellement à m_1 , tester m_0 contre m_2 revient à affecter les points du demi-plan supérieur à m_0 préférentiellement à m_2 , et tester m_0 contre m_3 revient à affecter les points du demi-plan supérieur à la diagonale principale à m_0 préférentiellement à m_3 . Finalement, ceci revient alors à affecter les points à m_0 s'ils appartiennent au quart de plan supérieur droit. Par symétrie, la stratégie de décision consiste alors tout simplement à affecter la valeur reçue à l'état du quart de plan correspondant.

Bien entendu, les résultats précédents sont plus généraux et s'appliquent à des constellations ou à des tests plus élaborés, en prenant en compte d'autres statistiques de bruit.

3.4 Détection de signaux à temps continu

Nous n'avons considéré pour le moment que la détection de constantes, ou de signaux à temps discrets, représentés sous la forme de vecteurs. L'information est véhiculée, le long du canal, par des signaux analogiques à temps continu, et il convient donc de s'intéresser au problème de la détection des signaux à temps continu. Ceci nous amènera en fait à présenter la structure générale d'un récepteur. Pour passer à l'analyse de signaux à temps continu, sans perdre l'investissement théorique précédent, nous allons utiliser les méthodes de représentation des signaux aléatoires présentées au début de ce chapitre, qui nous permettront de nous ramener au cas vectoriel. Pour introduire cette démarche, nous débiterons par le problème de détection d'un signal dans du bruit.

3.4.1 Représentation

Les deux hypothèses sont

$$\begin{aligned} H_0 &: r(t) = b(t) \\ H_1 &: r(t) = s(t) + b(t) \end{aligned}$$

où $s(t)$ est un signal certain, d'énergie $E_s = \int_0^T |s(t)|^2 dt$, et $b(t)$ un signal aléatoire, centré, gaussien, de fonction de corrélation $\Gamma_b(t, u)$.

On choisit comme base de représentation les $\Phi_i(t)$ (fonctions propres) de $\Gamma_b(t, u)$, définis par $\int_0^T \Phi_i(t) \Gamma_b(t, u) dt = \lambda_i \Phi_i(u)$ pour $u \in [0, T]$, et où les coefficients b_i du développement de $b(t)$ sont décorrélés : $E[b_i b_j] = \lambda_i \delta_{ij}$. L'utilisation de la transformée de Kahunen-Loève permet ici de traiter sans difficulté le cas d'un bruit coloré.

Dans ces conditions, les signaux d'intérêt peuvent s'exprimer à l'aide des développements en série

suivants :

$$\begin{cases} s(t) = \sum_{i=1}^{\infty} s_i \Phi_i(t) & \text{avec } s_i = \int_0^T s(t) \Phi_i(t) dt, \\ b(t) = \sum_{i=1}^{\infty} b_i \Phi_i(t) & \text{avec } b_i = \int_0^T b(t) \Phi_i(t) dt, \\ r(t) = \sum_{i=1}^{\infty} r_i \Phi_i(t) & \text{avec } r_i = \int_0^T r(t) \Phi_i(t) dt. \end{cases}$$

Le bruit $b(t)$ étant gaussien, on peut connaître $r|H_0$ et $r|H_1$...

Par propriétés du développement de Kahunen-Loève, et en tenant compte du fait que $b(t)$ est par hypothèse gaussien centré, les variables aléatoires b_i sont gaussiennes dans leur ensemble, centrées et décorréelées. On déduit de ceci les distributions de $r_i|H_0$ et $r_i|H_1$:

- les $r_i|H_0$ sont des variables aléatoires gaussiennes centrées, non corrélées,
- les $r_i|H_1$ sont gaussiennes, non corrélées, de valeurs moyennes $E[r_i|H_1] = s_i$

3.4.2 Troncature

On peut tronquer les développements à l'ordre N (on sait que le développement de KL converge en moyenne quadratique). On obtient alors les développements suivants :

$$r_N(t) = \sum_{i=1}^N r_i \Phi_i(t) \quad s_N(t) = \sum_{i=1}^N s_i \Phi_i(t) \quad b_N(t) = \sum_{i=1}^N b_i \Phi_i(t)$$

Sur la base $\{\Phi_i\}$, $i = 1..N$, les différents signaux pourront simplement être représentés par les composantes du développement, c'est-à-dire par un vecteur.

Sous l'hypothèse H_0 : $r(t)$ est représenté par $r_N(t)$, i.e. par les r_i pour $i = 1$ à N ,

$$\mathbf{r}|H_0 = [r_1, r_2, \dots, r_N]^T = [b_1, b_2, \dots, b_N]^T,$$

et sous H_1 :

$$\mathbf{r}|H_1 = [r_1, r_2, \dots, r_N]^T = [s_1, s_2, \dots, s_N]^T + [b_1, b_2, \dots, b_N]^T = \mathbf{s} + \mathbf{b}.$$

3.4.3 Stratégie bayésienne

Nous pouvons maintenant appliquer les résultats obtenus précédemment : le problème est simplement la détection d'un signal certain \mathbf{s} dans un bruit gaussien \mathbf{b} , avec $p_{\mathbf{r}|H_0}$ une densité gaussienne centrée de matrice de corrélation Γ_b , et $p_{\mathbf{r}|H_1}$ gaussienne de moyenne \mathbf{s} et de même matrice de corrélation.

La matrice de corrélation est diagonale, par propriété de décorrélation du développement de Kahunen-Loève :

$$\Gamma_b = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_N \end{bmatrix}.$$

On a alors :

$$l(\mathbf{r}) = r^T \Gamma_b^{-1} \mathbf{s} \Bigg|_{H_0}^{H_1} \log \eta + \frac{1}{2} \mathbf{s}^T \Gamma_b^{-1} \mathbf{s}$$

or, la matrice diagonale Γ_b est particulièrement simple à inverser, et

$$\Gamma_b^{-1} = \begin{bmatrix} \frac{1}{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{\lambda_N} \end{bmatrix}.$$

On obtient alors finalement

$$l(\mathbf{r}) = \sum_{i=1}^N \frac{r_i s_i}{\lambda_i} \left\langle \begin{matrix} H_1 \\ H_0 \end{matrix} \right\rangle \log \eta + \frac{1}{2} \sum_{i=1}^N \frac{s_i^2}{\lambda_i}$$

3.4.4 Retour au continu

Comme on sait que le développement de Kahunen-Loève converge, on peut faire tendre N vers l'infini.

On ne peut en général pas revenir à une expression simple en fonction des signaux à temps continu initiaux, et à l'exception notable du bruit additif blanc gaussien, le traitement doit passer soit par un blanchiment préalable, soit par l'emploi effectif d'une décomposition de type Kahunen-Loève. D'importants résultats sont cependant obtenus lorsque le bruit additif est blanc gaussien.

Cas du bruit blanc

Dans le cas du bruit blanc, toutes les valeurs propres sont égales, et on a $\lambda_i = \frac{N_0}{2} \forall i$. Le test du log-rapport de vraisemblance s'écrit alors

$$l = \frac{2}{N_0} \sum_{i=1}^{\infty} r_i s_i \left\langle \begin{matrix} H_1 \\ H_0 \end{matrix} \right\rangle \log \eta + \frac{1}{N_0} \sum_{i=1}^{\infty} s_i^2,$$

ou

$$l' = \sum r_i s_i \left\langle \begin{matrix} H_1 \\ H_0 \end{matrix} \right\rangle \log \eta + \sum s_i^2.$$

En remplaçant r_i par son expression,

$$\begin{aligned} l' &= \sum_{i=1}^{\infty} \int_0^T r(t) \Phi_i(t) s_i dt = \int_0^T r(t) \sum s_i \Phi_i(t) dt \\ &= \int_0^T r(t) s(t) dt, \end{aligned}$$

on arrive alors à :

$$\int_0^T r(t) s(t) dt \left\langle \begin{matrix} H_1 \\ H_0 \end{matrix} \right\rangle \frac{N_0}{2} \log \eta + \frac{1}{2} \int_0^T s(t) s(t) dt.$$

Le récepteur optimal consiste ainsi à calculer le *produit scalaire* entre $s(t)$ et $r(t)$, c'est-à-dire aussi l'intercorrélacion entre $r(t)$ et $s(t)$, pour un retard nul, et à comparer le résultat à un seuil qui fait intervenir l'énergie du signal.

Figure 3.8 – Récepteur optimal pour la détection d'un signal noyé dans du bruit.

Cette opération peut également s'interpréter comme un filtrage particulier, un *filtrage adapté*, présenté dans la section 3.2.3, au signal attendu $s(t)$, avec échantillonnage et prise de décision au temps T . On parlera alors de manière équivalente de récepteur à corrélation ou de récepteur à filtre adapté.

Pour caractériser le récepteur, il nous faut maintenant chercher à établir quelles sont les performances du récepteur optimal, c'est-à-dire les probabilités d'erreur, de détection, de fausse alarme, etc... Pour celà, il suffit de connaître les deux distributions de probabilité $p_{l|H_0}(l)$ et $p_{l|H_1}(l)$. La variable l étant gaussienne, il suffit alors de connaître la moyenne et la variance. On a

$$l|H_0 = \int_0^T b(t) s(t) dt \quad \text{soit} \quad E[l|H_0] = 0$$

$$l|H_1 = \int_0^T s(t)s(t) + s(t)b(t) dt \quad \text{soit} \quad E[l|H_1] = E_s$$

puis

$$\begin{aligned} \text{Var}[l|H_0] &= \text{Var}[l|H_1] = E[l^2|H_0] \\ &= E\left[\int_0^T b(t)s(t)b(u)s(u) dt du\right] \\ &= \frac{N_0}{2} \int_0^T (s(t))^2 dt = \frac{N_0}{2} E_s, \end{aligned}$$

en utilisant le fait que $b(t)$ est un bruit blanc de fonction d'autocorrélation $R_{bb}(\tau) = \frac{N_0}{2}\delta(t - \tau)$.

On définit l'indice de performance comme

$$d^2 = \frac{\{E[l|H_1] - E[l|H_0]\}^2}{\text{Var}(l|H_0)} = \frac{2E_s^2}{N_0E_s} = \frac{2E_s}{N_0}.$$

Il s'agit là aussi de la valeur du rapport signal-à-bruit en sortie d'un filtre adapté. Les variables $l|H_0$ et $l|H_1$ étant respectivement deux variables gaussiennes,

$$l|H_0 = N\left(0, \frac{N_0}{2}E_s\right),$$

$$l|H_1 = N\left(E_s, \frac{N_0}{2}E_s\right),$$

la probabilité de fausse alarme s'exprime comme

$$P_{fa} = \frac{1}{\sqrt{2\pi}} \int_{\sqrt{\frac{N_0}{2E_s}} \log \eta + \frac{1}{2} \sqrt{\frac{2E_s}{N_0}}}^{+\infty} \exp\left\{-\frac{l'^2}{2}\right\} dl' = Q\left(\sqrt{\frac{N_0}{2E_s}} \log \eta + \frac{1}{2} \sqrt{\frac{2E_s}{N_0}}\right)$$

On trouvera enfin :

$$\begin{aligned} P_{fa} &= Q\left[\log \eta \sqrt{\frac{N_0}{2E_s}} + \frac{1}{2} \sqrt{\frac{2E_s}{N_0}}\right] \\ P_d &= Q\left[\log \eta \sqrt{\frac{N_0}{2E_s}} - \frac{1}{2} \sqrt{\frac{2E_s}{N_0}}\right] \end{aligned}$$

On pourra noter que les probabilités P_{fa} et P_d ne dépendent que du rapport signal-à-bruit $\frac{2E_s}{N_0}$. D'autre part, la minimisation du risque bayésien dans le cas du signal blanc gaussien revient à maximiser le rapport signal-à-bruit et à retrouver la notion de filtre adapté.

3.5 Communication binaire en présence de bruit blanc.

La démarche ayant été conduite dans le cas de la détection d'un signal à temps continu noyé dans du bruit, nous pouvons maintenant aborder le problème de la détection de deux signaux certains $s_0(t)$ et $s_1(t)$, connus, noyés dans un bruit blanc $b(t)$. Ce problème comprend la détection des « niveaux » dans un train binaire, mais aussi la détection d'une fréquence parmi deux (BFSK) ou d'un état de phase parmi deux (BPSK). Évidemment, nous étendrons ensuite ces résultats au cas M hypothèses. Le problème est cependant pour le moment de déterminer, au vu de l'observation $r(t)$, laquelle des deux hypothèses suivantes est vraie.

$$\begin{aligned} H_0 &: r(t) = s_0(t) + b(t) \\ H_1 &: r(t) = s_1(t) + b(t) \end{aligned} \quad \left\{ \begin{array}{l} t \in [0, T] \end{array} \right.$$

On supposera ici que s_0 et s_1 ont même énergie E_s (ceci n'est pas limitatif, et on traiterai de façon analogue la détection de deux signaux d'énergies différentes) :

$$\int_0^T |s_0|^2 dt = |s_1(t)|^2 dt = E_s.$$

On ne suppose pas nécessairement que les signaux soient orthogonaux, et on définit le coefficient de corrélation, pour des signaux certains réels, comme

$$\rho = \frac{1}{E_s} \int_0^T s_0(t) s_1(t) dt \in [0, 1]$$

Afin de pouvoir traiter ce problème de détection/décision, nous devons, comme précédemment, rechercher une "bonne" base de représentation, sur laquelle on sait mener facilement les calculs. La construction de la base de décomposition repose ici sur une procédure d'orthonormalisation de Gram-Schmidt :

$$\Phi(t) = \begin{bmatrix} \Phi_1(t) = \frac{s_0(t)}{\sqrt{E_s}} \\ \Phi_2(t) = \frac{s_1(t) - \rho s_0(t)}{\sqrt{E_s(1-\rho^2)}} \\ \Phi_3(t) = \dots \\ \vdots \end{bmatrix},$$

qui permet de construire des fonctions de base Φ_i orthonormales, en enlevant la partie corrélée, puis en normalisant l'énergie.

Après cette décomposition, seules les composantes r_1 et r_2 du développement de $r(t)$ portent l'information, puisque les autres composantes sont définies comme la projection de $r(t)$ sur des fonctions $\Phi_i, i \neq 1, 2$ qui sont justement orthogonales aux signaux d'intérêt $s_0(t)$ et $s_1(t)$:

$$r(t) = \sum_i r_i \Phi_i(t), \quad \text{avec} \quad r_i = \int_0^T r(t) \Phi_i(t) dt.$$

Sous les deux hypothèses, $\mathbf{r} = [r_1, r_2, \dots, r_N]^T$ est un vecteur gaussien, qui sera caractérisé par la connaissance des moyennes et variances. En notant que les composantes sont décorrélatées (par l'orthonormalisation et parce que le bruit additif est blanc), il suffira de connaître, pour l'hypothèse H_0 par exemple,

$$\begin{cases} E[r_1 | H_0] \\ E[r_2 | H_0] \\ \text{Var}[r_{1,2} | H_0] \end{cases}$$

et de même pour l'hypothèse H_1 . Il suffit ensuite d'écrire le test du rapport de vraisemblance (en ayant tronqué la décomposition à un ordre N), puis en faisant tendre N vers l'infini, à trouver l'expression "continue" du test. Tous calculs faits, le résultat est ici

$$l = \int_0^T r(t) \cdot [s_1(t) - s_0(t)] dt \underset{H_0}{\overset{H_1}{>}} \frac{N_0}{2} \log \eta.$$

Ce dernier résultat est indépendant du fait que les signaux recherchés soient orthogonaux ou non : finalement, que les signaux soient orthogonaux ou non, le récepteur est le même, et il faut corrélérer $r(t)$ avec $s_1(t)$ d'une part, et $r(t)$ avec $s_0(t)$ d'autre part. De manière équivalente, on construit deux filtres adaptés, pour les signaux attendus $s_0(t)$ et $s_1(t)$. Que les signaux soit ou non orthogonaux, la structure du récepteur optimal est inchangée, mais les performances sont affectées par l'éventuelle corrélation des signaux.

Dans le cas où $\eta = 1$, le test devient

$$l = \int_0^T r(t) s_1(t) dt \underset{H_0}{\overset{H_1}{>}} \int_0^T r(t) s_2(t) dt$$

Figure 3.9 – Récepteur optimal dans le cas binaire.

et le récepteur optimal a la structure représentée sur la figure 3.9 :

Le choix des signaux influe sur les performances du récepteur. Les performances peuvent alors fournir un guide pour le choix des signaux, c'est-à-dire finalement pour le choix d'une méthode de modulation, en gardant toutefois à l'esprit que d'autres critères de choix entrent en jeu, comme la bande occupée, la complexité de réalisation, etc...

Comme précédemment, on introduit l'indice de performance

$$d^2 = \frac{[E[l|H_1] - E[l|H_0]]^2}{\text{Var}[l|H_0]}.$$

Sous chacune des deux hypothèses, le log rapport de vraisemblance (qui est une variable gaussienne, comme intégrale d'une variable gaussienne), vaut :

$$\begin{aligned} l|H_0 &= \int_0^T [s_0(t) + b(t)] [s_1(t) - s_0(t)] dt, \\ l|H_1 &= \int_0^T [s_1(t) + b(t)] [s_1(t) - s_0(t)] dt. \end{aligned}$$

On en déduit les moyennes :

$$\begin{aligned} E[l|H_0] &= \rho E_s - E_s = E_s(\rho - 1), \\ E[l|H_1] &= E_s(1 - \rho). \end{aligned}$$

Les variances valent quant-à-elles

$$\begin{aligned} \text{Var}[l|H_0] = \text{Var}(l|H_1) &= E \left\{ \int \int b(t) (s_1(t) - s_0(t)) b(u) (s_1(u) - s_0(u)) dt du \right\} \\ &= N_0 E_s (1 - \rho) \end{aligned}$$

Finalement, l'indice de performance s'exprime en fonction du rapport signal-à-bruit et du coefficient de corrélation :

$$d^2 = \frac{[2E_s(1 - \rho)]^2}{N_0 E_s (1 - \rho)} = \frac{4E_s}{N_0} (1 - \rho).$$

On peut ensuite facilement, cf 3.3.3, calculer les différentes probabilités, pour obtenir :

$$\begin{aligned} P_{fa} &= Q \left[\frac{\log \eta}{d} + \frac{d}{2} \right], \\ P_d &= Q \left[\frac{\log \eta}{d} - \frac{d}{2} \right], \end{aligned}$$

et dans le cas où $\eta=1$, la probabilité d'erreur vaut

$$\text{Pr(erreur)} = Q \left[\frac{d}{2} \right] = Q \left[\sqrt{\frac{E_s}{N_0}} (1 - \rho) \right], \quad (1)$$

et le meilleur choix, qui minimise la probabilité d'erreur, la probabilité de fausse alarme et maximise la probabilité de détection, consiste à rechercher deux signaux $s_0(t)$ et $s_1(t)$ tels que $\rho = -1$.

Interprétation en terme de distance :

Dans le cas de la détection binaire avec bruit blanc additif gaussien, on peut facilement interpréter la probabilité d'erreur comme étant liée à la distance euclidienne entre les deux signaux d'intérêt $s_0(t)$ et $s_1(t)$, normalisée par la variance du bruit additif. En effet, si les probabilités d'émission des deux signaux sont égales, la frontière de décision est constituée par la bissectrice perpendiculaire à la droite rejoignant les points s_0 et s_1 , coordonnées de $s_0(t)$ et $s_1(t)$ dans la base de représentation.

Ainsi, lorsque le signal $s_0(t)$ (ou $s_1(t)$) est émis, une erreur est commise lorsque le signal reçu est plus proche de $s_1(t)$ (réciproquement $s_0(t)$) que de $s_0(t)$ (réciproquement $s_1(t)$). Comme le bruit est blanc gaussien de variance $N_0/2$, la probabilité de cet événement est donnée par

$$\Pr(\text{erreur} | s_0 \text{ émis}) = \Pr(\|r - s_1\| < \|r - s_0\| | s_0 \text{ émis}) = \frac{1}{\sqrt{2\pi N_0/2}} \int_{d_*/2}^{+\infty} \exp\left(\frac{-u^2}{2N_0/2}\right) du,$$

où d_* désigne la distance euclidienne entre s_0 et s_1 . Les probabilités d'erreur si $s_0(t)$ et si $s_1(t)$ sont égales, et comme les deux messages sont équiprobables, la probabilité d'erreur totale est égale à l'expression précédente. En effectuant le changement de variable $v = u/\sqrt{N_0/2}$, on obtient alors

$$\Pr(\text{erreur}) = \frac{1}{\sqrt{2\pi}} \int_{\frac{1}{2} \frac{d_*}{\sqrt{N_0/2}}}^{+\infty} \exp\left(\frac{-v^2}{2}\right) dv,$$

et par définition de la fonction Q ,

$$\Pr(\text{erreur}) = Q\left(\frac{1}{2} \frac{d_*}{\sqrt{N_0/2}}\right) = Q\left(\frac{d_*}{\sqrt{2N_0}}\right), \quad (2)$$

ou le terme apparaissant dans la fonction Q est bien la demi-distance $d_*/2$, normalisée par l'écart type du bruit $\sqrt{N_0/2}$. On peut enfin facilement calculer la valeur de d_* :

$$d_*^2 = \int_0^T (s_1(t) - s_0(t))^2 dt = 2E_s(1 - \rho),$$

et l'on retrouve alors l'expression (1).

Exemples :

Considérons deux modulations classiques dans le cas binaire : la modulation de fréquence (FSK) et la modulation de phase (PSK).

- Dans le cas de la modulation FSK, l'information est codée à l'aide de deux fréquences : on associe au niveau 1 une fréquence ν_1 et au niveau 0 une fréquence ν_0 , de sorte que

$$s_1(t) = \sqrt{\frac{2E_s}{T}} \cos(2\pi\nu_1 t),$$

$$s_0(t) = \sqrt{\frac{2E_s}{T}} \cos(2\pi\nu_0 t).$$

Les deux signaux sont orthogonaux si $T(\nu_1 - \nu_0)$ est soit entier, soit $\gg 1$, et dans ce cas $\rho = 0$. L'indice de performances vaut alors $d^2 = 4E_s/N_0$, et la probabilité d'erreur

$$\Pr(\text{erreur}) = Q\left[\sqrt{\frac{E_s}{N_0}}\right]. \quad (3)$$

- Dans la cas de la PSK, l'information est codée en phase. On peut obtenir un coefficient de corrélation de -1 en prenant un déphasage de π , avec

$$s_1(t) = -\sqrt{\frac{2E_s}{T}} \cos(2\pi\nu_0 t) \quad (\Phi_1 = \pi)$$

$$s_0(t) = \sqrt{\frac{2E_s}{T}} \cos(2\pi\nu_0 t) \quad (\Phi_0 = 0)$$

L'indice de performance vaut alors $d^2 = 8E_s/N_0$, et la probabilité d'erreur est

$$\Pr(\text{erreur}) = Q \left[\sqrt{\frac{2E_s}{N_0}} \right]. \quad (4)$$

Notons que cette PSK est ici équivalente à une modulation ASK dans laquelle on prendrait une amplitude 1 pour représenter le signal (le "1" logique) et une amplitude -1 pour le "0" logique.

Dans le cas binaire, avec bruit blanc gaussien additif, si le critère de choix est la probabilité d'erreur, le meilleur choix est ainsi obtenu pour une modulation PSK avec un déphasage de π .

3.6 Communication M-aire (M hypothèses)

En utilisant la même démarche que précédemment, on peut étendre les résultats au cas M -hypothèses. Ce sera par exemple le cas d'une modulation QAM à M états, ou d'un codage en fréquence (M fréquences pour coder M messages). On considère ainsi

$$H_i : r_i(t) = s_i(t) + b(t) \quad t \in [0, T],$$

où les signaux s_i sont certains. Comme dans le cas binaire, on obtient qu'il faut comparer les produits scalaires de signal reçu $r(t)$ avec chacun des signaux sources $s_i(t)$,

$$l_i = \int_0^T r(t)s_i(t)dt$$

et prendre une décision en faveur du signal (de l'hypothèse) associé au plus grand produit scalaire. Le récepteur optimal bayésien consiste donc à calculer l'intercorrélacion, ou le filtre adapté, sur M voies et à retenir l'hypothèse correspondant à la plus grande valeur de sortie, voir la figure 3.10.

Figure 3.10 – Structure du récepteur optimal bayésien : autocorrélacion ou filtrage adapté sur chacune des voies puis comparaison.

L'évaluation de la probabilité d'erreur est quant-à-elle un peu plus ardue. Nous traiterons tout d'abord le cas de signaux orthogonaux, qui peuvent par exemple correspondre à une M -FSK (avec des fréquences orthogonales), puis nous traiterons le cas des M -PSK, et enfin nous donnerons le résultat pour des modulations QAM.

Probabilité d'erreur pour des signaux orthogonaux

En supposant que les différentes hypothèses H_i sont équiprobables et que les coûts sont tels que $C_{ij} = 1$, $C_{ii} = 0$, la probabilité d'erreur s'écrit

$$\begin{aligned} \Pr(\text{erreur}) &= \sum_{i \neq j} \sum p_j p(\text{Décider } H_i | H_j \text{ réalisé}) \\ &= \sum_j p_j \sum_{i \neq j} p(\text{Décider } H_i | H_j \text{ réalisé}) = \sum_j p_j p(\text{erreur} | H_j). \end{aligned}$$

Les différentes hypothèses étant équiprobables, on a $p_j = \frac{1}{M}$, et il est assez évident, par symétrie du problème, que $p(e) = p(e | H_j)$ quelque soit j . Considérons par exemple l'hypothèse H_1 :

$$\begin{aligned} p(e | H_1) &= 1 - p(l_1 > l_j \forall j \neq 1 / H_1) \\ &= 1 - p(l_1 > l_2) \cdot p(l_1 > l_3) \dots p(l_1 > l_M) \\ &= 1 - \prod_{j \neq 1} p(l_1 > l_j | H_1) \end{aligned}$$

La probabilité de non erreur si H_1 est la probabilité $p(l_1 \text{ supérieur à tous les } l_j, l_j \neq l_1 | H_1)$, soit

$$P(\text{non erreur sous } H_1) = \int_{-\infty}^{+\infty} p(l_1 | H_1) \int_{-\infty}^{l_1} p(l_2 | H_1) \int_{-\infty}^{l_1} \int_{-\infty}^{l_1} \dots \int_{-\infty}^{l_1} p(l_M | H_1) dl_1 dl_2 \dots dl_M \quad (5)$$

(M-1) fois

Ici les produits scalaires l_1, l_j conduisent à

$$\left. \begin{aligned} l_1 &= \int_0^T r(t) s_1(t) dt \\ l_j &= \int_0^T r(t) s_j(t) dt \end{aligned} \right| \rightarrow \begin{aligned} l_{1/H_1} &= \int_0^T [s_1(t) + b(t)] s_1(t) dt, \\ l_{j/H} &= \int_0^T [s_1(t) + b(t)] s_j(t) dt. \end{aligned}$$

En utilisant l'orthogonalité des signaux, les moyennes sont données par

$$\begin{aligned} E[l_j | H_1] &= 0 \quad \text{pour } j \neq 1 \\ &= E_s \quad \text{pour } j = 1. \end{aligned}$$

Afin d'évaluer la probabilité de non erreur, ou d'erreur, si H_1 , il nous faut connaître $p(l_j | H_1)$, $j = 1 \dots M$. Or les $l_j | H_1$, pour $j = 1 \dots M$ sont des variables gaussiennes dont les variances valent :

$$\begin{aligned} \text{Var } l_1 | H_1 &= \frac{N_0}{2} E_s, \\ \text{Var } l_j | H_1 &= \frac{N_0}{2} E_s. \end{aligned}$$

Toutes les variables $l_j | H_1$ pour $j \neq 1$ sont ainsi des variables de même loi, et pour alléger un peu les relations, on pose $\sigma^2 = \frac{N_0}{2} E_s$. Ainsi, la probabilité (5) s'écrit

$$\begin{aligned} P(\text{non erreur} | H_1) &= \int_{-\infty}^{+\infty} \int_{-\infty}^{l_1} \int_{-\infty}^{l_1} p(l_1/H_1) p(l_2/H_1) \dots p(l_M/H_1) dl \\ &= \int_{-\infty}^{+\infty} p(l_1/H_1) \left[\int_{-\infty}^{l_1} p(l_j/H_1) dl_j \right]^{M-1}, \\ &\quad \begin{array}{cc} \uparrow & \uparrow \\ \text{gaussienne} & \text{gaussienne} \\ \text{non centrée} & \text{centrée} \end{array} \\ &= \int_{-\infty}^{+\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{l_1 - E_s}{\sigma^2}\right)^2} \left[\int_{-\infty}^{l_1} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\frac{l_j^2}{\sigma^2}} dl_j \right]^{M-1} dl_1. \end{aligned}$$

Finalement, en posant $y = l_j/\sigma$ et $x = l_1/\sigma$, on obtient la probabilité d'erreur :

$$\text{Pr(erreur)} = 1 - \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2}\left(x - \sqrt{\frac{2E_s}{N_0}}\right)\right)^2 \left[\int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \right]^{M-1} dx \quad (6)$$

où on a posé $x = l_1/\sigma$. Cette relation n'admet pas de solution analytique, et doit être évaluée numériquement.

Borne d'union

On peut cependant obtenir une borne supérieure simple pour cette probabilité d'erreur. Donnons tout d'abord une formulation générale de cette borne, pour le problème de la détection parmi M signaux, non nécessairement orthogonaux. Rappelons que la probabilité d'erreur pour la détection entre deux signaux $s_i(t)$ et $s_j(t)$ de même énergie E_s est donnée par (1), et si l'on note d_{ij} l'indice de performance et ρ_{ij} le coefficient de corrélation correspondant, on a

$$\text{Pr(erreur)} = Q\left[\frac{d_{ij}}{2}\right] = Q\left[\sqrt{\frac{E_s}{N_0}}(1 - \rho_{ij})\right].$$

On peut voir le détecteur pour les M signaux comme effectuant $M - 1$ décisions binaires entre la voie contenant le signal, disons $s_i(t)$, et les autres voies. On aura alors erreur sur la détection si l'une au

moins des décisions binaires est erronée. Notons A_k , pour $k = 1..M - 1$ l'événement correspondant à une décision binaire erronée. Alors la probabilité d'erreur pour la détection du signal $s_i(t)$ parmi M est égale à la probabilité de l'union (du « ou » logique) des $M - 1$ événements. Or on sait que la probabilité de l'union de $M - 1$ événements est toujours inférieure à la somme des probabilités élémentaires : $\Pr(\bigcup_{k=1}^{M-1} A_k) \leq \sum_{k=1}^{M-1} \Pr(A_k)$. Ainsi,

$$\Pr(\text{erreur} | s_i(t) \text{ émis}) \leq \sum_{\substack{j=1 \\ j \neq i}}^M Q \left[\frac{d_{ij}}{2} \right] \quad \text{pour } i = 1, 2, \dots, M.$$

La probabilité d'erreur symbole est alors la moyenne de cette expression pour les M symboles, et

$$\begin{aligned} \Pr(\text{erreur}) &= \sum_{i=1}^M p_i \Pr(\text{erreur} | s_i(t) \text{ émis}) \\ &\leq \sum_{i=1}^M \sum_{\substack{j=1 \\ j \neq i}}^M p_i Q \left[\frac{d_{ij}}{2} \right], \end{aligned} \quad (7)$$

où les p_i sont les probabilités d'émission des différents signaux $s_i(t)$. Notons (ceci nous sera utile pour établir la probabilité d'erreur pour les M -PSK), que dans le cas d'une constellation circulaire symétrique, $\Pr(\text{erreur} | s_i(t) \text{ émis})$ est la même pour tout i , et

$$\Pr(\text{erreur}) \leq \sum_{\substack{j=1 \\ j \neq i}}^M Q \left(\frac{d_{ij}}{2} \right) = \sum_{\substack{j=1 \\ j \neq i}}^M Q \left(\frac{d_{*ij}}{\sqrt{2N_0}} \right) \quad \text{pour tout } i, \quad (8)$$

où d_{*ij} est la distance euclidienne entre les signaux $s_i(t)$ et $s_j(t)$, comme nous l'avons utilisée en (2).

Revenons maintenant au cas des signaux orthogonaux de même énergie. Dans ce cas, $\rho_{ij} = 0$, et tous les indices d_{ij} sont égaux. À partir de (7), on obtient alors pour (6) la borne supérieure simple suivante :

$$\Pr(\text{erreur}) \leq (M - 1)Q \left(\frac{d}{2} \right) = (M - 1)Q \left(\sqrt{\frac{E_s}{N_0}} \right).$$

Application : transmission de digits binaires.

À titre d'application, considérons le cas d'une transmission de digits binaires, et étudions l'évolution de la probabilité d'erreur lorsque M croît. Les niveaux 0, 1 émis par la source sont équiprobables et émis tous les T . La puissance disponible à l'émission est donnée, et fixée à P . On forme ensuite des symboles en associant plusieurs digits binaires. On considèrera par ailleurs que les signaux $s_i(t)$ associés à chacun des symboles sont orthogonaux.

1. Transmission binaire s_0, s_1 , l'énergie par signal est $E_s = TP$,
2. pour $M = 4$ (groupes de 2) $E_s = 2TP$,
3. pour $M = 8$ (groupes de 3 digits binaires) $E_s = 3TP$,
4. pour $M = 2^K$ (groupes de $K = \log_2 M$ digits binaires) $E_s = K TP = \log_2(M) TP$.

Ainsi, l'énergie associée à chacun des signaux émis est liée à la puissance disponible et à la durée d'émission de la source. La probabilité d'erreur est donnée par la relation (6), avec $\sigma^2 = \frac{N_0}{2} E_s = \frac{N_0}{2} \log_2(M) TP$; et il vient donc

$$\Pr(\text{erreur}) = 1 - \int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi \log_2(M)}} \exp -\frac{1}{2} \left(x - \sqrt{\frac{2TP \log_2(M)}{N_0}} \right)^2 \left[\int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} dy \right]^{M-1} dx,$$

et après un changement de variable,

$$\Pr(\text{erreur}) = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \exp\left(-\frac{y^2}{2}\right) \left[\int_{-\infty}^{y+\sqrt{\frac{2TP \log_2 M}{N_0}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \right]^{M-1} dy$$

En fonction de M , on peut enfin montrer que

$$\lim_{M \rightarrow \infty} (M-1) \log \left[\int_{-\infty}^{y+\sqrt{\frac{2TP \log_2 M}{N_0}}} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt \right] = \begin{cases} -\infty & \text{si } \frac{TP}{N_0} < \log_2, \\ 0 & \text{sinon.} \end{cases}$$

On en déduit que $\Pr(\text{erreur}) \rightarrow 1$ si $\frac{TP}{N_0} < \log_2$, et que $\Pr(\text{erreur}) \rightarrow 0$ si $\frac{TP}{N_0} > \log_2$.

Ainsi, en dessous d'un certain débit $1/T$, on peut garantir une probabilité d'erreur aussi faible que l'on veut. Par contre, lorsque le taux d'émission est supérieur à ce seuil, il n'est pas possible de transmettre sans erreur, et la probabilité d'erreur tend même vers 1. Ceci n'est autre que le second théorème de Shannon, qui indique qu'on peut espérer transmettre de l'information *sans erreur* en dessous d'un certain débit maximum, et qu'*a contrario* la probabilité d'erreur devient très importante, quelques soient les moyens mis en œuvre, dès que l'on dépasse ce seuil.

Probabilité d'erreur pour une modulation M -PSK

Dans la section précédente, nous avons étudié la probabilité d'erreur pour la transmission de M signaux orthogonaux. Évidemment, les signaux utilisés dans une modulation PSK ne sont pas orthogonaux entre eux, et il nous faudrait donc reprendre le calcul dans ce cas de figure. Pour des corrélations arbitraires ρ_{ij} entre les signaux $s_i(t)$ et $s_j(t)$, pour $i, j = 1..M$, on peut établir, voir [1, pages 381-383], que la probabilité d'erreur s'exprime comme

$$\Pr(\text{erreur}) = 1 - \frac{1}{M} \exp\left(-\frac{E_s}{N_0}\right) \int_{-\infty}^{+\infty} \exp\left(x \frac{2E_s}{N_0}\right) \times \left[\underbrace{\frac{d}{dx} \int_{-\infty}^x \cdots \int_{-\infty}^x}_{(M-1) \text{ fois}} \frac{1}{\sqrt{(2\pi)^M \det(\rho)}} \exp\left(-\frac{1}{2} \mathbf{y}^T \rho^{-1} \mathbf{y}\right) d\mathbf{y} \right] dx,$$

ce qui n'est pas précisément une expression très sympathique à manipuler. On peut cependant facilement trouver une borne supérieure, en utilisant la borne d'union présentée dans la section précédente, dans laquelle on utilise l'expression (1) de la probabilité d'erreur pour la détection binaire de deux signaux de corrélation quelconque. Dans la mesure où cette dernière probabilité d'erreur est elle-même majorée par la probabilité d'erreur pour les deux signaux de corrélation maximum, on obtient alors

$$\Pr(\text{erreur}) \leq (M-1)Q \left[\sqrt{\frac{E_s}{N_0} (1 - \rho_{\max})} \right],$$

où ρ_{\max} est le plus grand des coefficients de corrélation.

Dans le cas des modulations PSK, on peut obtenir une expression simple en utilisant la borne d'union établie en (8) et reproduite ci-dessous :

$$\Pr(\text{erreur}) \leq \sum_{\substack{j=1 \\ j \neq i}}^M Q\left(\frac{d_{ij}}{2}\right) = \sum_{\substack{j=1 \\ j \neq i}}^M Q\left(\frac{d_{*ij}}{\sqrt{2N_0}}\right) \quad \text{pour tout } i, \quad (9)$$

La constellation d'une M -PSK est circulaire, et si les signaux sont d'énergie E_s , la distance entre un point de la constellation et l'un de ses deux plus proches voisins est donnée par

$$d_* = 2\sqrt{E_s} \sin\left(\frac{\pi}{M}\right).$$

Si l'on suppose que le rapport signal-à-bruit est suffisamment grand pour que les erreurs soient essentiellement dûes à la confusion entre deux voisins, la somme en (9) se réduit à deux termes, et il reste simplement

$$\Pr(\text{erreur}) \simeq 2Q\left(\sqrt{\frac{2E_s}{N_0}} \sin\left(\frac{\pi}{M}\right)\right) = 2Q\left(\sqrt{\frac{2kE_b}{N_0}} \sin\left(\frac{\pi}{M}\right)\right),$$

avec E_b l'énergie bit et $k = \log_2 M$. Cette approximation de la probabilité d'erreur se révèle bonne pour tout $M \geq 4$, très proche de la vraie probabilité d'erreur si le rapport signal-à-bruit est élevé.

Pour une modulation QAM, on obtient la probabilité d'erreur suivante, cf [3, pages 369-373] :

$$\Pr(\text{erreur}) \leq 1 - \left[1 - 2Q\left(\sqrt{\frac{3\bar{E}_s}{(M-1)N_0}}\right)\right]^2,$$

où \bar{E}_s est l'énergie symbole moyenne. Pour les probabilités d'erreur correspondant à d'autres schémas de modulations, on pourra se reporter à [2] ou [3].

3.7 Détection à maximum de vraisemblance d'une séquence de symboles

Lorsque les signaux sont *sans mémoire*, les détecteurs précédents, qui procèdent symbole par symbole, sont optimaux au sens de la probabilité d'erreur. Cependant, lorsque le signal présente une mémoire, c'est-à-dire lorsque les symboles codés successifs sont interdépendants, le récepteur devrait tirer parti de cette dépendance. Par exemple, pour une modulation à phase continue (CPM), les phases successives ne sont pas indépendantes. Ainsi, pour une modulation MSK, qui est une modulation à phase continue d'indice $1/2$, on ne peut passer d'une phase nulle qu'à une phase de $\pi/2$ ou de $3\pi/2$. En quelque sorte, les hypothèses de détection à considérer à chaque instant *dépendent de l'état antérieur*. On peut bien entendu traiter la détection des signaux avec mémoire symbole par symbole sans se préoccuper de la dépendance intersymboles, ce qui mènerait pour la MSK à décider à chaque instant entre les 4 états de phase $\{0, \pi/2, \pi, 3\pi/2\}$, mais il est clair que la prise en compte de la dépendance intersymboles doit permettre d'améliorer les performances. On cherche ainsi à détecter une *séquence de symboles*.

Comme dans la présentation de [2], considérons un signal NRZ encodé différenciellement, c'est-à-dire que le signal de sortie de l'encodeur, b_k , ne change d'état que lorsque l'entrée a_k de l'encodeur est un « 1 ». Ainsi, on a $b_k = a_k + b_{k-1}$, où l'addition est une addition modulo 2. Notons S_0 et S_1 les deux états possibles de b_k , auxquels on pourra associer deux niveaux d'amplitude, deux signaux différents $s_0(t)$ et $s_1(t)$, etc. La *figure 3.11* donne ainsi le diagramme en treillis pour un NRZ différentiel, où les changements d'état sont provoqués par $a_k = 1$.

Figure 3.11 – Diagramme en treillis pour un NRZ différentiel

Supposons ici que l'on associe les niveaux $s_0 = -s_1 = A$ aux deux états. Pour le k^{e} symbole, l'observation est alors de la forme

$$r_k = \pm A + b_k,$$

en sortie du filtre adapté de réception, avec b_k un bruit blanc additif gaussien de variance σ_b^2 . Les densités de probabilité conditionnelles aux deux hypothèses s'écrivent alors

$$p(r_k | s_{0,1}) = \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp \left\{ -\frac{(r_k \mp A)^2}{2\sigma_b^2} \right\}.$$

Considérons une séquence de symboles de longueur K , et notons $\mathbf{s}^{(n)}$ la n^{e} parmi les 2^K séquences possibles. Alors, la densité de probabilité conditionnelle de la séquence d'observation r_1, r_2, \dots, r_K si la séquence $\mathbf{s}^{(n)}$ a été émise, est

$$p(r_1, r_2, \dots, r_K | \mathbf{s}^{(n)}) = \prod_{k=1}^K \frac{1}{\sqrt{2\pi\sigma_b^2}} \exp \left\{ -\frac{(r_k - s_k^{(n)})^2}{2\sigma_b^2} \right\} = \left(\frac{1}{\sqrt{2\pi\sigma_b^2}} \right)^K \exp \left\{ -\frac{1}{2\sigma_b^2} \sum_{k=1}^K (r_k - s_k^{(n)})^2 \right\},$$

où $s_k^{(n)} = \pm A$. Ainsi, on peut rechercher la séquence $\mathbf{s}^{(n)}$ telle que la densité conditionnelle précédente, la *vraisemblance*, soit maximale. On obtient alors un détecteur de séquence au maximum de vraisemblance. En prenant le logarithme, on voit que la recherche de la solution au maximum de vraisemblance correspond à la minimisation de la *métrique*

$$\sum_{k=1}^K (r_k - s_k^{(n)})^2.$$

Il s'agit donc de retenir la séquence associée à la métrique (ici une distance euclidienne) la plus faible, parmi les 2^K séquences possibles. Il est possible de réduire cette complexité en utilisant une recherche séquentielle des séquences sur le treillis. L'algorithme de recherche correspondant s'appelle l'algorithme de Viterbi. Cet algorithme peut être employé dès que l'on a une recherche de solution au maximum de vraisemblance pour une séquence de données non indépendantes. On le rencontre ainsi également dans le contexte de l'égalisation (voir le chapitre sur l'égalisation) et pour le décodage des codes convolutifs (voir le chapitre sur le codage canal).

Dans le cas de notre exemple du NRZ différentiel, chaque noeud du treillis, correspondant à un instant de décision – multiple de la période symbole, est atteint par deux chemins « entrants », et deux chemins partent de chaque noeud. Pour les deux chemins entrants, on peut calculer la métrique cumulée. Considérons ainsi que l'on soit à l'étape i et que l'on considère un noeud sur l'état S_0 . Alors, pour le premier chemin atteignant $S_{0,i}$, on a $D_1(i) = D_1(i-1) + (r_i - A)^2$, où $D_1(i-1)$ est la métrique du premier chemin entrant ; et pour le second chemin, on a $D_2(i) = D_2(i-1) + (r_i - A)^2$.

L'algorithme de Viterbi consiste alors à comparer ces deux métriques au noeud i et à ne retenir que le chemin correspondant à la métrique la plus faible. En effet, le chemin écarté correspond forcément à une séquence (partielle) de vraisemblance plus faible, et la prolongation de ce chemin amènera toujours à une vraisemblance plus faible que la prolongation de l'autre chemin. Le chemin conservé est appelé chemin *survivant*. Les mêmes opérations sont effectuées pour les chemins entrant dans $S_{1,i}$, le noeud sur S_1 à l'étape i , et l'on obtient un survivant. Les deux survivants sont alors prolongés dans le treillis jusqu'aux états suivants à l'étape $i+1$. Ainsi, le nombre de chemins à explorer, parmi les 2^K , est divisé par deux à chaque étape, ce qui explique la popularité de cet algorithme. Ce processus est poursuivi jusqu'à $i = K$, et l'on retient alors le chemin de métrique la plus faible, c'est-à-dire de vraisemblance la plus grande. Enfin, en notant les transitions sur le chemin retenu, on reconstruit la séquence de symboles.

Notons qu'en général, on s'arrange pour débiter d'un état connu et terminer le treillis sur un état connu. Cela correspond par exemple à avoir ajouté quelques bits à 0 en début et en fin de séquence (bits de terminaison – *tail bits*), qui imposent l'état S_0 en début et fin de séquence, et permet d'éviter des ambiguïtés. Lorsque deux chemins entrant en un noeud ont même métrique, ces chemins ne sont pas distinguables (sont équivalents) en terme de vraisemblance, et on peut tout simplement n'en retenir qu'un par tirage aléatoire (pile ou face).

D'après ce qui précède, on voit qu'il faut attendre la fin du treillis avant de prendre une décision sur la séquence, et par conséquent sur les symboles individuels. Il s'en suit donc un retard de décision qui est considérable. En pratique, si on appelle L la mémoire du signal, c'est-à-dire le nombre de symboles précédents dont dépend le symbole courant ($L = 1$ pour le NRZ différentiel), on s'aperçoit qu'avec une probabilité très élevée, toutes les séquences survivantes en i sont identiques jusqu'à $i - 5L$. On peut donc prendre les décisions avec un retard de $5L$, sans attendre la fin du treillis.

Revenons maintenant au cas des modulations à phase continue. Ces modulations sont caractérisées par un nombre fini d'états de phase, le nombre exact d'état dépendant à la fois de l'indice de modulation et de la largeur (en nombre de symboles) de l'impulsion de mise en forme. Les transitions entre les différents états de phase sont gouvernées par un treillis de phase. On peut montrer que la log-vraisemblance de l'observation $r(t)$ conditionnellement à une séquence de symboles $\mathbf{s} = [s_1, s_2, \dots, s_K]$, est proportionnelle à

$$M_K(\mathbf{s}) = \int_{-\infty}^{(K+1)T} r(t) \cos[2\pi f_0 t + \phi(t; \mathbf{s})] dt = M_{K-1}(\mathbf{s}) + \int_{KT}^{(K+1)T} r(t) \cos[2\pi f_0 t + \theta(t; \mathbf{s}) + \theta_K] dt,$$

où $\theta(t; \mathbf{s})$ dépend de la séquence de symboles et de l'impulsion quand θ_K ne dépend que de la séquence de symboles précédents, et où $M_{K-1}(\mathbf{s})$ est la métrique jusqu'en KT . Compte tenu de ce qui précède, on voit que l'on peut appliquer l'algorithme de Viterbi afin de reconstituer la séquence des états de phase et par suite la séquence de symboles. Pour cela, il faut utiliser la métrique précédente, c'est-à-dire en fin de compte calculer un produit scalaire avec une fonction des états de phase.

Pour terminer cette section, on notera que l'algorithme de Viterbi permet d'obtenir d'excellentes performances, mais que son emploi est typiquement limité par le nombre d'états. En effet, si on a une mémoire de longueur L avec une modulation binaire, on obtient 2^L états, ce qui devient prohibitif dès que L est supérieur à 9 ou 10.

3.8 Détection de signaux de phase inconnue, ou à paramètres inconnus.

Les signaux reçus sont non seulement bruités, mais il peuvent être *imparfaitement connus* : la phase, le rythme ou la fréquence des signaux reçus peuvent être mal connus, en raison de la propagation ou de la désynchronisation de l'émetteur et du récepteur. Il s'agit donc de détecter la présence de signaux qui sont eux même imparfaitement caractérisés : on considèrera par exemple le cas d'une détection binaire

$$\begin{aligned} r(t) &= s_0(t, \theta) + b(t) \\ r(t) &= s_1(t, \theta) + b(t), \end{aligned}$$

où le paramètre θ est inconnu.

Comme classiquement, on formule le rapport de vraisemblance

$$\Lambda(\mathbf{r}) = \lim_{N \rightarrow \infty} \frac{p(\mathbf{r} | H_1)}{p(\mathbf{r} | H_0)},$$

où l'on définit les distributions $p(\mathbf{r} | H_0)$ et $p(\mathbf{r} | H_1)$ comme les distributions marginales

$$\begin{cases} p(\mathbf{r} | H_0) = \int_{\theta} p(\mathbf{r} | H_0, \theta) p(\theta | H_0) d\theta, \\ p(\mathbf{r} | H_1) = \int_{\theta} p(\mathbf{r} | H_1, \theta) p(\theta | H_1) d\theta. \end{cases}$$

On dit qu'on intègre le paramètre de « nuisance » θ hors du problème. Le problème est alors résolu, si l'on se donne des les distributions *a priori* $p(\theta | H_0)$ et $p(\theta | H_1)$, et si les distributions marginales sont calculables.

Une autre possibilité souvent utilisée est de construire un *rapport de vraisemblance conditionnel* :

$$\Lambda_{\theta}(\mathbf{r}) = \frac{p(\mathbf{r} | H_1, \theta)}{p(\mathbf{r} | H_0, \theta)}$$

On retient alors dans un premier temps la valeur de θ comme l'argument du maximum du rapport de vraisemblance : $\theta^* = \operatorname{argmax}_{\theta} \Lambda_{\theta}(\mathbf{r})$, puis il reste à résoudre un problème classique, avec $\theta = \theta^*$ connu. Le problème a alors été scindé en deux étapes : d'abord une étape d'*estimation* du paramètre, puis une étape de détection.

Récepteurs pour un signal à phase inconnue

Pour illustrer le problème de la détection d'un signal à paramètres inconnus et fournir le récepteur correspondant à un cas particulier important, considérons maintenant le problème de la détection d'un signal sinusoïdal dont la phase est inconnue.

$$\begin{aligned} H_1 \quad r(t) &= \cos(2\pi f_0 t + \Phi(t) + \theta) + b(t) \\ H_0 \quad r(t) &= b(t) \end{aligned}$$

Après décomposition sur une bonne base, on peut écrire le rapport de vraisemblance, puis faire tendre N vers l'infini pour obtenir l'expression continue. Cette démarche ayant déjà été répétée à plusieurs reprises, on peut donner ici directement le résultat. On a en outre introduit une loi *a priori* $p_{\theta}(\theta)$ sur le paramètre θ , et on a exprimé les lois marginales.

$$\Lambda(r) = \int_{-\pi}^{\pi} p_{\theta}(\theta) d\theta \exp \left[\frac{2}{N_0} \int_0^T r(t) s_0(t, \theta) dt - \frac{4}{N_0} \int_0^T s_0(t, \theta)^2 dt \right]$$

Le signal $s_0(t, \theta) = \cos(2\pi f_0 t + \Phi(t) + \theta)$ peut être exprimé comme

$$\cos(\omega t + \Phi(t) + \theta) = \cos(\omega t + \Phi(t)) \cos \theta - \sin(\omega t + \Phi(t)) \sin \theta$$

On peut alors décomposer $r(t) s_0(t, \theta)$ sous la forme $L_C \cos \theta - L_S \sin \theta$, avec

$$\begin{aligned} L_C &= \int_0^T r(t) \cos[\omega t + \Phi(t)] dt \\ L_S &= \int_0^T r(t) \sin[\omega t + \Phi(t)] dt \end{aligned}$$

et le rapport de vraisemblance devient alors (à une constante près) :

$$\Lambda'(r) = \int_{-\pi}^{\pi} p_{\theta}(\theta) \exp \left[\frac{2}{N_0} (L_C \cos \theta - L_S \sin \theta) \right] d\theta.$$

Il nous reste maintenant à spécifier la distribution *a priori*. On peut prendre une famille de densités indexée par un paramètre L_m :

$$p_{\theta}(\theta, L_m) = \frac{\exp(L_m \cos \theta)}{2\pi I_0(L_m)}$$

où I_0 est la fonction de Bessel de 1ère espèce. Pour cette famille paramétrée, on a en particulier

$$\begin{aligned} L_m = 0 &\quad \rightarrow p_{\theta}(\theta) = \frac{1}{2\pi} \\ L_m \rightarrow \infty &\quad \rightarrow p_{\theta}(\theta) = \delta(\theta) \end{aligned}$$

Le rapport de vraisemblance devient alors

$$\begin{aligned} \Lambda'(r(t)) &= \int_{-\pi}^{\pi} \exp \left[\left(L_m + \frac{2L_C}{N_0} \right) \cos \theta - \frac{2L_S}{N_0} \sin \theta \right] \frac{1}{2\pi I_0(L_m)} d\theta \\ &= \frac{1}{I_0(L_m)} \cdot I_0 \left[\left\{ \left(L_m + \frac{2L_C}{N_0} \right)^2 + \left(\frac{2L_S}{N_0} \right)^2 \right\}^{\frac{1}{2}} \right] \end{aligned}$$

et

$$\log \Lambda'(r) = \log I_0 \left[\left\{ \left(L_m + \frac{2L_C}{N_0} \right)^2 + (L_s)^2 \right\}^{\frac{1}{2}} \right] \begin{matrix} > \\ < \\ < \end{matrix} \log \eta + \frac{2E_s}{N_0} + \log I_0(L_m).$$

En tenant compte du fait que

$$I_0(x) \approx 1 + \frac{x^2}{4} \quad (x \ll 1),$$

c'est-à-dire $\log(I_0(x)) \approx \frac{x^2}{4}$,

il vient

$$\left(L_C + \frac{N_0}{2} L_m \right)^2 + L_s^2 \begin{matrix} > \\ < \end{matrix} \gamma,$$

et on en déduit le récepteur optimal : un schéma de démodulation non cohérente typique, qui consiste à démoduler par deux signaux en quadrature, à intégrer puis à élever au carré afin de faire disparaître la dépendance en la phase inconnue : Dans ce dernier schéma, on peut remplacer la partie démodulation

Figure 3.12 – Récepteur optimal non cohérent dans le cas de la détection d'un signal dans du bruit.

et intégration (le calcul de corrélation) par un filtre adapté suivi d'un échantillonnage au temps T pour obtenir un récepteur équivalent à filtres adaptés.

Ce type de résultat s'étend à des modulations à plus d'états ; dans le cas d'une modulation M-FSK avec une incertitude sur la phase de réception, on trouverait ainsi un récepteur constitué par M voies où l'on calcule les produits scalaires du signal reçu avec les composantes en quadrature de la fréquence attendue, intégration et élévation au carré ; c'est-à-dire simplement le récepteur élémentaire précédent, répliqué pour les M fréquences potentielles. Enfin, la décision est prise en faveur de la voie fournissant la plus grande valeur.

Lorsque $L_m = 0$, le test se réduit à

$$L_C^2 + L_s^2 \begin{matrix} > \\ < \end{matrix} \gamma,$$

et l'on doit calculer $L_C^2 + L_s^2$, ou, de manière équivalente, $\sqrt{L_C^2 + L_s^2}$. Ceci peut être réalisé en employant un filtre passe-bande isolant le signal recherché suivi par un détecteur d'enveloppe. En effet, si l'on prend $h(t) = \cos(2\pi f_0 t + \psi(t))$, alors le signal de sortie vaut

$$\begin{aligned} (h * r)(t) &= \int_0^T h(t - \tau) r(\tau) d\tau = \int_0^T \cos(2\pi f_0(t - \tau) + \psi(t - \tau)) r(\tau) d\tau \\ &= L_C \cos(2\pi f_0 t + \psi(t) - \Phi(t)) - L_s \sin(2\pi f_0 t + \psi(t) - \Phi(t)). \end{aligned}$$

L'enveloppe de la sortie du filtre est alors proportionnelle à la racine carrée du test. Il suffit donc d'utiliser un détecteur d'enveloppe après la démodulation (un seuil à zéro suivi par un filtrage passe-bas par exemple). L'enveloppe étant insensible à la phase du signal reçu, ce dispositif permet ainsi la détection. Si le signal émis comportait en outre une modulation d'amplitude $a(t)$, c'est-à-dire qu'il serait de la forme $a(t) \cos(2\pi f_0 t + \phi(t))$, alors on devrait utiliser un filtre passe-bande de réponse impulsionnelle $h(t) = b(t) \cos(2\pi f_0 t + \psi(t))$, cf [1, pages 341-343] et dans ce cas l'enveloppe de sortie est la racine du test $\sqrt{L_C^2 + L_s^2}$ à l'instant T , si on choisit $b(t) = a(T - t)$ et $\psi(t) = -\Phi(T - t)$. Ceci peut alors être interprété comme la combinaison d'une démodulation d'une part et un filtrage adapté d'autre part. Cette opération, suivie par la détection d'enveloppe et échantillonnage à l'instant T est parfois appelée *filtre adapté non cohérent*.

Récepteur pour un signal à amplitude et phase inconnue

Pour conclure ce chapitre, nous nous intéresserons au cas d'un signal issu d'un canal de Rayleigh. En l'absence de bruit d'observation, le signal recueilli a la forme

$$r(t) = a(t) \cos(2\pi f_0 t + \theta(t)),$$

où l'enveloppe $a(t)$ varie continûment et a une densité de probabilité de Rayleigh, et où la phase est distribuée uniformément entre 0 et 2π . Lorsque l'on a un « fading » lent, on peut considérer que $a(t)$ et $\theta(t)$ sont constants sur chaque durée symbole, et si on s'intéresse simplement à un problème de détection, on a

$$\begin{aligned} H_1 \quad r(t) &= a \cos(2\pi f_0 t + \theta) + b(t), \\ H_0 \quad r(t) &= b(t), \end{aligned}$$

où a est une variable aléatoire de Rayleigh et θ une variable aléatoire uniforme. On peut écrire la partie signal en fonction des composantes en quadrature :

$$a \cos(2\pi f_0 t + \theta) = a_1 \cos(2\pi f_0 t) + a_2 \sin(2\pi f_0 t) = a_1 s_1(t) + a_2 s_2(t),$$

où a_1 et a_2 sont deux variables aléatoires gaussiennes indépendantes, centrées et de même variance σ^2 . De plus, les deux composantes sont orthogonales. On obtient alors pour le rapport de vraisemblance

$$\begin{aligned} \Lambda(r(t)) &= \int \int p_{a_1}(A_1) p_{a_2}(A_2) \exp\left(\frac{2}{N_0} \int_0^T r(t) (A_1 s_1(t) + A_2 s_2(t)) dt\right. \\ &\quad \left. - \frac{1}{N_0} \int_0^T \sum_{i=1}^2 \sum_{j=1}^2 A_i A_j s_i(t) s_j(t) dt\right) dA_1 dA_2. \end{aligned}$$

En définissant alors

$$L_i = \int_0^T r(t) s_i(t) dt,$$

et en utilisant l'orthogonalité entre s_1 et s_2 , on obtient

$$L_1^2 \left(\frac{\sigma_{a_1}^2}{\sigma_{a_1}^2 + N_0/2} \right) + L_2^2 \left(\frac{\sigma_{a_2}^2}{\sigma_{a_2}^2 + N_0/2} \right) > \gamma.$$

En se souvenant enfin que $\sigma_{a_1}^2 = \sigma_{a_2}^2 = \sigma^2$, et en notant que L_1 et L_2 ne sont rien d'autre que les fonctions L_C et L_s introduites précédemment, on obtient finalement

$$L_C^2 + L_s^2 > \gamma',$$

qui est le même test que celui obtenu pour le cas d'une phase aléatoire, et l'on adoptera donc les mêmes structures de récepteur. Bien entendu, les performances seront dégradées par rapport au cas où seule la phase est inconnue, mais l'étude de cette question sort du cadre de ce texte, et l'on pourra se reporter à [1].

3.9 Conclusions

Dans ce chapitre, nous avons tâché de présenter des outils et une méthodologie pour comprendre ou concevoir un récepteur de communication numérique. Nous avons notamment donné la démarche qui permet d'établir une structure de récepteur et les moyens de qualifier celui-ci en calculant ses performances. Ces points ont été développés dans quelques cas particuliers. Évidemment, nous n'avons pu aborder l'ensemble des schémas de modulation et des scénarios caractérisant une chaîne de transmission ; mais nous pensons que notre présentation constitue une bonne base pour une exploration

plus approfondie, débutant par exemple par les ouvrages cités en bibliographie. En outre, il faut garder à l'esprit que les récepteurs présentés ici correspondent presque à un cas idéal, puisque en pratique le problème de la réception est souvent compliqué par la présence d'interférence entre symboles, qui nécessitera un dispositif d'égalisation, et par des problèmes de synchronisation entre l'émetteur et le récepteur. Nous abordons maintenant ces derniers points dans les deux chapitres suivants.

Bibliographie

- [1] Van Trees H. L. – *Detection, Estimation and Modulation Theory*, New York, Wiley, 1968.
- [2] Proakis J. G. – *Digital Communications*, 4^e édition, New York, Mc Graw-Hill, 2000.
- [3] Haykin S. – *Communication Systems*, 4^e édition, New York, John Wiley and Sons, 2001.