REPUBLIQUE ALGERIENNE DEMOCRATIQUE ET POPULAIRE

MINISTERE DE L'ENSEIGNEMENT SUPERIEUR ET DE LA RECHERCHE SCIENTIFIQUE
UNIVERSITE YAHIA FARES DE MEDEA



Faculté de Technologie

Mémoire de fin d'études de Master

Filière : Télécommunications.

Spécialité : Systèmes des télécommunications.

Détection distribuée dans les réseaux de capteur sans fil (WSN) à travers un canal de rayleigh.

Présenté par :

Mr. BOUDERBALA Mohamed

Mr. HIZI Alla eddine

Proposé et Dirigé par :

Mr. SOUMALI Chaabane

Année universitaire 2020-2021

REMERCIEMENT

Nous remercions Allah tout puissant de m'avoir accordé la volonté et le courage pour réaliser ce mémoire.

Nous remercions vivement et du fonds de cœurs notre promoteur : **M^r^ :** Soumali chaabane, qui nous a orienté dans ce travail.

Pour ses fructueuses remarques, ses encouragements et sa passion pour la recherche. Qu'il trouve ici le témoignage de notre profonde gratitude. Pour l'honneur qu'il nous accordé, Nous tenons également à remercier tous les enseignants de télécom.

Nos vifs remerciements à tous les membres de jury qui nous font l'honneur de juger notre modeste travail.

Nous tenons à exprimer ma très grande gratitude, et notre profonde affection à mes chers parents pour leurs encouragements, leur patience et leur grand soutien, Durant toutes ces années d'études.

Sans oublier mes frères Et mes sœurs pour leurs encouragements, et leurs aides. Nous remercions également tous mes amis et camarades, Pour leurs encouragements Et leur précieux soutien.

DÉDICACES

-Je remercie dieu et grâce à lui que je vous arrivée à ce niveau.

-Je dédie mon travail à mes parents qui m'ont aide dans tous les circonstances pour la bonne réalisation de cette mémoire.

-Comme je la dédie à ma seur. Même je dédie cette mémoire aussi à toute la famille ***Bouderbala****, et à mon binôme* ***alla eddine*** *et à tous mes amis.*

Mohamed

DÉDICACES

Je dédie ce projet :

A ma chère mère,

A mon cher père,

Qui n'ont jamais cessé, de formuler des prières à mon égard, de me soutenir et de

M'épauler pour que je puisse atteindre mes objectifs.

A mes frères, *lahcen, smayda.*

A mes SŒURS *nawel et soundes ,koko ,mkaa*

A mes petits frères, *barhoum, hamoudisghir,*

Douaazina.

Pour ses soutiens moral et leurs conseils précieux tout au long de mes études.

A mon cher grand-père,

Qu'Allah repose son âme en paix.

A ma chère binôme, BouderbalaMohammed

.Pour sa entente et sa sympathie.

A mes chères amies *yahia ,brahim, ishak, ziad , zizou , adel , nasro , siradj ,adem, badr, taher et tous mes amis.*

. Et celui qui l'a oublié, pardonne-moi

Pour leurs aides et supports dans les moments difficiles.

A tous ma famille,

Sans ta présence,

Mon âme est triste et mon coeur est lourd.

Alla eddine

ملخص

في هذا العمل تمت معالجة مشكلة دمج القرارات المرسلة عبر قناة رايلي في شبكة الاستشعار اللاسلكية (WSN) ,

اقترحنا قاعدة قرار تستند الى قياس *نسبة* الاحتمالية التي تتطلب معرفة احصائيات القناة بدلا من معلومات عن حالتها,

توضح لنا أخيرا من خلال تقيمنا للأداء أن هذه القاعدة بها تدهور طفيف مقارنة بقاعدة الدمج بناء على نسبة الاحتمالية .

Résumé

Dans ce travail, le problème de fusion de décisions transmis à travers un canal de Rayleigh dans un réseau de capteur sans fil (WSN) est traité. On propose une règle de décision basée sur le rapport de vraisemblance (LR : likelihood ratio) qui nécessitent uniquement la connaissance des statistiques du canal au lieu du CSI (Channel State Information). L'évaluation des performances montre que cette règle à une légère dégradation par rapport à la règle de fusion basée sur le rapport de vraisemblance utilisant le CSI.

Abstract

In this work, the decision fusion problem transmitted through a Rayleigh channel in a wireless sensor network (WSN) is addressed. We propose a decision rule based on the likelihood ratio (LR) that requires only the knowledge of the channel statistics instead of the CSI (Channel State Information). The performance evaluation shows that this rule has a slight degradation compared to the likelihood ratio based fusion rule using CSI.

Table des matières

Remerciement……..……..…….……I

Dédicace……..……..……..……II

Résumé……..……..……..……IV

Table des matière……..……..……..……..V

Table des figure……..……..……..VIII

Symbole et abréviation……..……..……..IX

CHAPITRE I : Généralités sur les réseaux de capteurs sans fil (WSN).

I.1 Introduction : 3

I.2 Structure d'un réseau de capteur sans fil (WSN) : 3

I.2.1 L'observation d'un phénomène : 3

I.2.2 Les capteurs : 4

I.2.3 Le canal de Rayleigh : 4

I.2.4 Le système MIMO : 4

I.2.4.1 Principe de la technique MIMO 4

I.2.4.2 Modèle de canal MIMO : 5

I.2.4.3 Les différents types de codage MIMO : 6

I.2.4.3.1 Le multiplexage par répartition de fréquence orthogonale (OFDM Orthogonal Frequency Division Multiplexing) 6

I.2.4.3.2 Le multiplexage par division spatiale (SDM Spatial Division Multiplexing) 6

I.2.4.3.3 Le codage spatio-temporel par bloc (STBC Space-Time Bloc Code) 7

I.2.4.4 Capacité de canal 7

I.2.5 Le centre de fusion : 8

I.3 Types de WSN 8

I.3.1 Terrestres WSNs : 8

I.3.2 Souterrains WSNs : 9

I.3.3 Réseaux de capteurs acoustiques sous-marins (UASNs) 10

I.3.4 MultiMedia des WSNs 10

I.3.5 Mobile des WSNs 10

I.4 Domaines d'applications des réseaux de capteurs sans fil : 10

I.4.1 Applications militaires 11

I.4.2 Applications environnementales et agricoles 12

I.4.3 Applications médicales 13

I.4.4 Applications de surveillance 14

I.4.5 Applications industrielles 14
I.4.6 Applications commerciales 14
I.4.7 Applications domestiques 14
I.5 Conclusion : 15

CHAPITRE II : Rappel sur la théorie d'estimation et de détection.

II.1 Introduction : 16
II.2 Théorie de la décision statistique 18
II.2.1 CRITÈRE DE BAYES : 18
II.2.1.1 Test d'hypothèse binaire 18
II.2.1.2 Test d'hypothèse M-aire : 22
II.2.2 CRITÈRE MINIMAX 29
II.2.3 CRITÈRE DE NEYMAN-PEARSON 32
II.2.4 TEST D'HYPOTHÈSE COMPOSITE 33
II.2.5 DÉTECTION SÉQUENTIELLE 33
II.3 Les Paramètres d'Estimation 35
II.3.1 ESTIMATION MAXIMALE DE PROBABILITE 35
II.3.2 TEST GÉNÉRALISÉ DU RATIO DE PROBABILITÉ 35
II.3.3 QUELQUES CRITÈRES POUR DE BON ESTIMATEURS 36
II.3.4 ESTIMATION DE BAYES : 38
II.3.4.1 Estimation de l'erreur quadratique moyenne minimale : 39
II.3.4.2 Valeur absolue moyenne minimale de l'estimation de l'erreur : 40
II.3.4.3 Estimation maximale a posteriori : 40
II.3.5 INÉGALITÉ CRAMER-RAO : 41
II.3.6 MEILLEUR ESTIMATEUR LINÉAIRE NON BIAISÉ 44
II.3.6.1 Estimation de la moyenne quadratique linéaire à un paramètre 44
II.3.6.2 Vecteur aléatoire 45
II.3.6.3 BLEU dans le bruit blanc gaussien : 47
II.3.7 ESTIMATION DES MOINDRES CARRÉS 49
II.4 Conclusions : 51

CHAPITRE III : Résultat et simulation.

III.1 : Introduction : 54
III.2 Formulation du problème : 55
III.2.1 : Model du système à trois couches : 56
III.2.1.1 Couche capteur local : 56
III.2.1.2 Couche canal « Fading » 57

III.2.1.3 Couche centre de fusion :..... 57
III.2.2 Aperçu sur les règles de fusion existantes : 57
III.2.2.1 Règle de décision optimale basée sur le rapport de vraisemblance :..... 57
III.2.2.2 Règle de fusion Chair-Varshney :..... 58
III.2.2.3 Règle de fusion MRC :..... 58
III.2.2.4 Règle de fusion EGC :..... 58
III.3 Rapport de vraisemblance basé sur les statistiques du canal :..... 59
III.4 ANALYSE DE PERFORMANCE : 62
III.5. Simulation et résultats 63
III.6 Conclusion : 65
Conclusion Générale.....66

Table des figures

Chapitre I : Généralités sur les réseaux de capteurs sans fil (WSN).

Figure I.1 : Architecture d'un réseau WSN. 3
Figure I.2 : Principe du MIMO. 5
Figure I.3 : Schémas représentatifs du SISO, MISO, SIMO et MIMO. 8
Figure I.7 : Domaines d'applications. 11
Figure I.8 : Les réseaux de capteurs Application militaire [28]. 12
Figure I.9 : Application environnementale [29]. 13
Figure I.10 : Un réseau de capteurs dans le domaine medical [30]. 13

CHAPITRE II : Rappel sur la théorie d'estimation et de détection.

Figure II.1 : Source de l'hypothèse binaire. 16
Figure II.2 : Région de décision. 17
Figure II.3 : Fonction de densité du cout. 20
Figure II.4 : Espace de décision pour M = 3. 27
Figure II.5 : Escape de décisions pour M = 3 : (a) régions de décision résultante et (b) espace de décision global. 28
Figure II.6 : Espace de décision utilisant le logarithme pour M = 3. 29
Figure II.7 : Fonctions de densité conditionnelle et régions de décision. 30
Figure II.8 : Risque en fonction de P1. 31
Figure II.9 : Fonction de densité de l'estimateur sans biais (θ). 37
Figure II.10 : fonctions de couts: (a) squared error, (b) absolute value of error, and (c) uniform. 38

CHAPITRE III : Résultat et simulation.

Figure III.1 : Modèle de fusion parallèle on présence des chaînes « fading » et « noisy » entre les capteurs locaux et le centre de fusion uk. 56
Figure III.2 : Courbes ROC (Receiver Operating Characteristic) pour règles de fusion LR et LRT-CS. 64
Figure III.3 : Niveau de probabilité de détection en fonction du SNR du canal de Rayleigh 65

Symboles et abbreviations

WSN	Wireless Sensor Network
MANET	Mobile Ad hoc Networks
RCSF	les réseaux sans fils
MIMO	Multiple-input and multiple-output
OFDM	Orthogonal Frequency Division Multiplexing
SDM	Spatial Division Multiplexing
STBC	Space-Time Bloc Code
CSI	Channel state information
SISO	Single-input single-output
SIMO	Single Input, Multiple Output
DSN	Distributed Sensor Network
DARPA	Defense Advanced Research Projects Agency
WATS	Wide Area Tracking System
MVU	Moreno Valley Electric Utility
MMSE	Minimum mean-squared error
BLEU	Best linear unbiased estimator
SNR	Signal-to-noise ratio
EGC	école de gestion et de commerce

Introduction générale :

Un réseau de capteurs sans fil est un réseau ad hoc d'un grand nombre de *nœuds,* qui sont des micro-capteurs capables de recueillir et de transmettre des données d'une manière autonome. La position de ces nœuds n'est pas obligatoirement prédéterminée. Ils peuvent être aléatoirement répartis dans une zone géographique, intitulée *« champ de captage »* correspondant au terrain concerné pour le phénomène capté. En plus d'applications civiles, il existe des applications militaires aux réseaux de capteurs (détection d'intrusions, localisation de combattants, véhicules, armes, etc. sur un champ de bataille, sous l'eau, dans l'espace, au sol…)

Depuis leur création, les réseaux de communication sans fil ont connu un succès sans cesse croissant au sein des communautés scientifiques et industrielles. Grâce à ses divers avantages, cette technologie a pu s'instaurer comme acteur incontournable dans les architectures réseaux actuelles. Le média hertzien offre en effet des propriétés uniques, qui peuvent être résumées en trois points : la facilité du déploiement, l'ubiquité de l'information et le coût réduit d'installation. Au cours de son évolution, le paradigme sans fil a vu naître diverses architectures dérivées, telles que : les réseaux cellulaires, les réseaux locaux sans fils et autres. Durant cette dernière décennie, une architecture nouvelle a vu le jour : les réseaux de capteurs sans fil. Ce type de réseaux résulte d'une fusion de deux pôes de l'informatique moderne : les systèmes embarqués et les communications sans fil. Un réseau de capteurs sans fil (RCSF), ou "Wireless Sensor Network" (WSN), est composé d'un ensemble d'unités de traitements embarquées, appelées "motes", communiquant via des liens sans fil. Le but général d'un WSN est la collecte d'un ensemble de paramètres de l'environnement entourant les motes, telles que la température ou la pression de l'atmosphère, afin de les acheminer vers des points de traitement. Les RCSF sont souvent considérés comme étant les successeurs des réseaux ad hoc. En effet, les RCSF partagent avec les MANET (Mobile Ad hoc NETworks) plusieurs propriétés en commun, telles que l'absence d'infrastructure et les communications sans fil. Mais l'une des différences clés entre les deux architectures est le domaine d'application. Contrairement aux réseaux MANET, qui n'ont pas pu connaître un vrai succès, les RCSF ont su attirer un nombre croissant d'industriels, vu leur réalisme et leur apport concret. En effet, le besoin d'un suivie continu d'un environnement donné est assez courant dans diverses activités de la société.

Chaque jour, un ensemble de nouvelles décisions sont prises sur la base de différentes hypothèses. Le critère approprié est choisi en fonction. Par exemple, nous prenons en compte le domaine de l'ingénierie lorsqu'il y a un problème pour connaître le signal radar reçu, sa qualité

est connue et une décision est prise sur le résultat. Comme dans le système de communication, un groupe de signaux représentés par des zéros et des uns est transmis dans le récepteur, car il est affecté par du bruit ajouté par le support externe lors de sa transmission. Le récepteur ne connaît pas le type de signal 0 ou 1, mais une décision doit être prise sur le signal reçu qui représente l'un d'entre eux. Le signal émis par le récepteur est révélé dans le choix de la règle de décision.

Notre travail est organisé comme suit : dans le premier chapitre, nous donnons un aperçu sur les réseaux de capteurs sans fil (leur structure, des petite définition) après nous donnons les domaines d'applications de celle-ci (application militaire, application environnementales…etc.), et dans le deuxième chapitre, tout d'abord, nous examinerons les différents types de détection (Critère de Bayes, critère Minimax…etc.). Ensuite, nous parlerons sur les différents paramètres d'estimation comme l'estimation maximale de probabilité et l'estimation de Bayes et leur type…etc. Dans le dernier chapitre nous ferons une étude comparative entre le rapport de vraisemblance (LR) et la règle LRT-CS, quelle méthode est la plus performante ? et nous montrerons les résultats

Chapitre I :

Généralités sur les réseaux de capteurs sans fil (WSN).

I.1 Introduction :

Les réseaux de capteurs sans fil sont des réseaux ad hoc constitués par de petits nœuds de détection appelés capteurs. Ils ont une faible vitesse de traitement et des capacités de stockage réduites. Les capteurs peuvent être déployés de façon aléatoire dans une zone à partir de laquelle les utilisateurs souhaitent recueillir des données.

Un réseau de capteurs sans fil (WSN) est un réseau constitué des dizaines à des milliers de ces nœuds de capteurs qui sont densément déployés dans un environnement sans surveillance avec des capacités de détection, de communications sans fil et de calculs (c.-à-d., collecte et diffusion de données environnementales) Ces nœuds sont capables de communications, de détection et de calcul sans fil (logiciel, matériel, algorithmes). Donc, il est évident qu'un WSN est le résultat de la combinaison de techniques de capteurs, de techniques embarquées, d'informations distribuées mécanismes de traitement et de communication.

En effet, les nœuds de capteurs ont la capacité de communiquer soit entre eux, soit directement avec une station de base.

I.2 Structure d'un réseau de capteur sans fil (WSN) :

La structure d'un réseau de capteur sans typique est donnée par le schéma suivant :

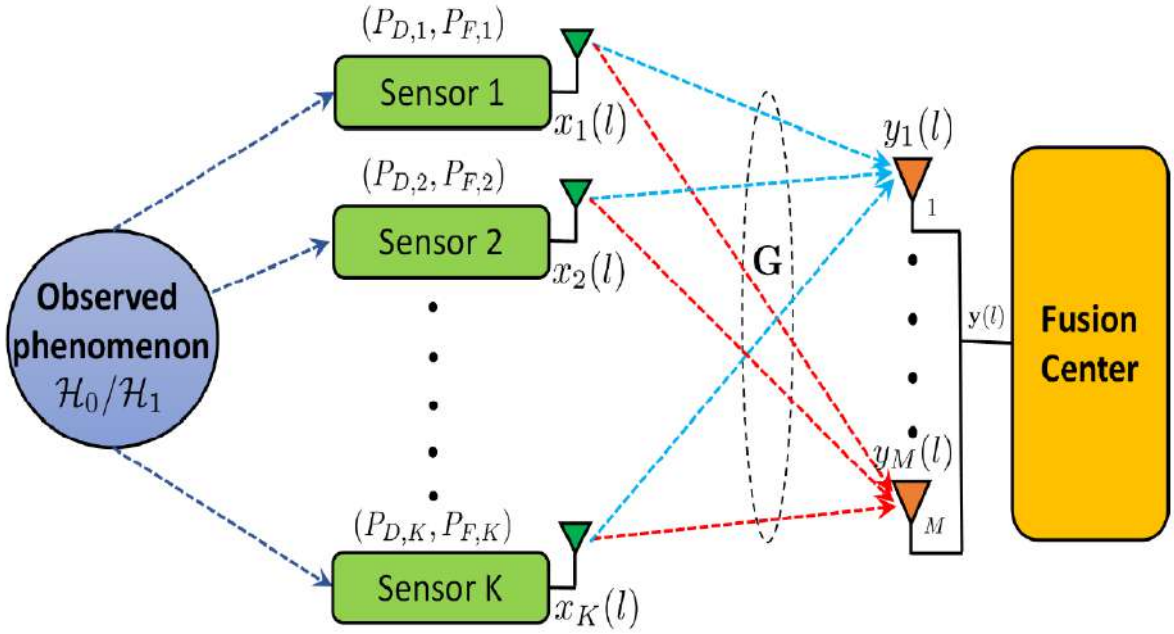


Figure I.1 : Architecture d'un réseau WSN.

I.2.1 L'observation d'un phénomène :

Le phénomène observé peut être une pression, une température, humidité ou autres ou bien une mesure sous-marine, militaire etc.…

I.2.2 Les capteurs :

Un capteur est un dispositif transformant l'état d'une grandeur physique observée (information d'entrée) en une autre grandeur physique de nature électrique qui est convertit en une valeur mesurée dans une représentation analogique ou numérique (utilisable), telle qu'une tension électrique, une hauteur de mercure, une intensité, une accélération, l'humidité. Le capteur se distingue de l'instrument de mesure par le fait qu'il ne s'agit que d'une simple interface entre un processus physique et une information manipulable. Par opposition, l'instrument de mesure est un appareil autonome se suffisant à lui-même, disposant d'un affichage ou d'un système de stockage des données. Le capteur, lui, en est dépourvu. Et nous pouvons dire que les capteurs sont les éléments de base des systèmes d'acquisition de données et que leur mise en œuvre est du domaine de l'instrumentation. Notre capteur peut mesurer un phénomène de pression, de température, d'humidité …etc.

Dans notre cas on a des capteurs qui mesure l'information et l'envoyer directement sans aucun traitement, et on a des capteurs avant l'envoi de l'information, il mesure l'information après il prend une décision partielle avec un processeur local [2].

I.2.3 Le canal de Rayleigh :

Deux images sont tout juste séparées lorsque le maximum central d'une figure coïncide avec le premier minimum de l'autre.

I.2.4 Le système MIMO :

Les systèmes MIMO (multiple input multiple output) sont apparus à la fin des années 1990 grâce à Gerard. J. Foschini [5], ils sont l'un des principaux axes de développement pour augmenter les débits des communications sans fil, nous assistons à un très rapide développement de cette technologie avec des applications déjà envisagées dans les réseaux locaux sans fil et les réseaux de communication de 3ème génération.

La mise en place d'une telle structure permet au système utilisé d'atteindre des débits importants et cela sans changer la largeur de la bande allouée au signal ni sa puissance d'émission. De plus, le fait d'utiliser plus d'une antenne dans les deux côtés du système permet d'apporter de la diversité.

I.2.4.1 Principe de la technique MIMO

Dans les systèmes de communications traditionnels, il n'existe qu'une antenne à l'émission et une antenne à la réception (SISO). Or les futurs services de communications mobiles sans fils demandent plus de transmissions de données (augmentation de la capacité de transmission). Ainsi pour augmenter la capacité des systèmes SISO et satisfaire ces demandes, les bandes passantes de ces systèmes et les puissances à transmettre ont été largement

augmentées. Mais les récents développements ont montré que l'utilisation de plusieurs antennes à l'émission et à la réception permettait d'augmenter le débit de transmission des données et cela sans augmenter ni la bande passante de l'antenne réceptrice du système SISO, ni la puissance du signal à l'émission. Cette technique de transmission s'appelle MIMO (Multiple Input, Multiple Output) (Figure I.2). Un système MIMO tire profit de l'environnement multi-trajets en utilisant les différents canaux de propagation créés par réflexion et/ou par diffraction des ondes pour augmenter la capacité de transmission. D'où l'intérêt aussi d'obtenir des signaux indépendants sur les antennes.

Quand on utilise plusieurs antennes à l'émission, chacune devient une source d'information différente pour les antennes de réception, donc On augmente encore la diversité.

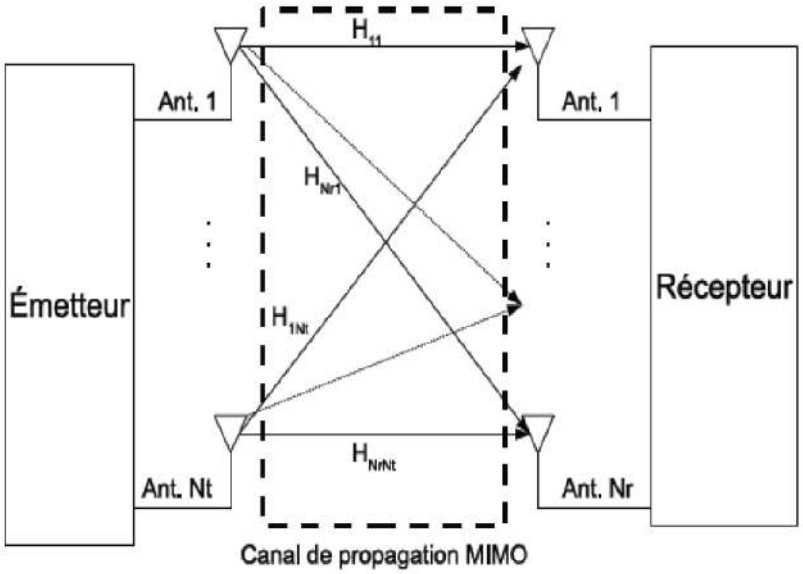


Figure I.2 : Principe du MIMO.

I.2.4.2 Modèle de canal MIMO :

Soit un système MIMO à N_t antennes d'émission et N_r antennes de réception comme illustré sur la figure ci-dessus.

On suppose que la bande de fréquence utilisée pour la transmission est assez étroite pour que le canal soit non sélectif en fréquence. A la réception, chaque antenne reçoit la somme des symboles x_i transmis simultanément par chacune des N_t antennes émettrices. Le signal y_j reçu par la $j^{ième}$ antenne peut d'écrire de manière discrète :

$$y_j = \sum_{i=1}^{Nt} h_{ij} X_i + n_j \qquad \text{(I.1)}$$

Où h_{jj} est le gain (coefficient d'évanouissement) du canal non sélectif en fréquence entre l'antenne d'émission *i* et l'antenne de réception *j*, n_j est le bruit additif qui est modélisé

par des échantillons indépendants et suivant une loi gaussienne centrée de variance $N_0/2$ par dimension réelle. Le rapport signal à bruit (RSB) par antenne de réception est défini par :

$$RSB = Es/N0. \quad \text{(I.2)}$$

L'équation (I.1) peut s'écrire sous une forme matricielle comme suit :

$$Y = Hx + n \quad \text{(I.3)}$$

Soit la matrice du canal MIMO H de dimension $N_r \times N_t$ suivante :

$$H = \begin{matrix} h_{11} & \dots & h_1 N_t \\ \vdots & \ddots & \vdots \\ h_{N_r}1 & \cdots & h_{N_r} N_t \end{matrix} \quad \text{(I.4)}$$

Où y et n sont respectivement les vecteurs de réception et de bruit de dimension $N_r \times 1$.

I.2.4.3 Les différents types de codage MIMO :

Dans un système MIMO, il existe plusieurs méthodes pour effectuer le codage des signaux :

I.2.4.3.1 Le multiplexage par répartition de fréquence orthogonale (OFDM Orthogonal Frequency Division Multiplexing)

Consiste à diviser sur un grand nombre de porteuses, le signal numérique que nous souhaitons transmettre (comme si nous combinons le signal à transmettre sur des émetteurs indépendants et à des fréquences différentes). Pour que les fréquences des porteuses soient les plus proches possibles et ainsi transmettre le maximum d'information sur une portion de fréquences donnée, l'OFDM utilise des porteuses orthogonales entre elles. Les signaux des différentes porteuses se chevauchent mais grâce à l'orthogonalité, n'interfèrent pas entre eux. Ainsi, dans un environnement multi-trajets où certaines fréquences seront détruites à cause des perturbations, le système sera tout de même capable de récupérer l'information perdue sur d'autres fréquences porteuses qui elles n'auront pas été détruites [3].

I.2.4.3.2 Le multiplexage par division spatiale (SDM Spatial Division Multiplexing)

Au cours duquel plusieurs flux de données indépendants (essentiellement des canaux virtuels) sont simultanément multiplexés dans un canal spectral. Le multiplexage SDM peut améliorer le débit de façon significative, car le nombre de données spatiales résolues est plus important. Chaque flux spatial doit disposer de sa propre paire d'antennes de transmission/réception à chaque extrémité du lien radio. Il est important de noter qu'une chaîne de radiofréquences RF et qu'un convertisseur analogique-numérique distincts sont nécessaires pour chaque antenne du système MIMO. Les configurations qui nécessitent plus de deux chaînes d'antennes RF doivent être conçues avec attention pour maintenir des coûts peu élevés tout en répondant aux attentes en matière de performances [4].

I.2.4.3.3 Le codage spatio-temporel par bloc (STBC Space-Time Bloc Code)

Tout comme le SDM permet d'envoyer des signaux différents sur chaque antenne. Le principe du STBC est d'introduire une redondance d'information entre les deux antennes. Le canal STBC comprend *M*N* sous canaux. Chaque sous canal est un canal à évanouissements indépendants ; ce qui fait que le STBC augmente la diversité du canal de transmission et donc la robustesse du récepteur. Cette méthode est très attractive car elle n'exige pas la connaissance de l'état du canal (CSI) même si cela peut réduire la capacité de transmission des données [6].

Le gain de diversité résultant améliore la fiabilité des liaisons sans fil à évanouissements et améliore la qualité de la transmission. Il est à noter que ce type de codage n'améliore pas la capacité de transmission linéairement avec le nombre d'éléments utilisés. Ainsi pour améliorer à la fois la capacité et la qualité, un système MIMO doit être implémenté avec les deux types de codages à savoir le SDM et le STBC [7].

I.2.4.4 Capacité de canal

Nous allons ici comparer les capacités des différents canaux existants (SISO, SIMO, MIMO) (Figure I.3) sans connaissance préalable de l'état du canal CSI. De même, nous allons aussi comparer les limites théoriques données par la capacité de Shannon [8] qui est l'espérance de la capacité E[C] et qui ne peut être obtenue que dans un canal idéal, avec un codage idéal.

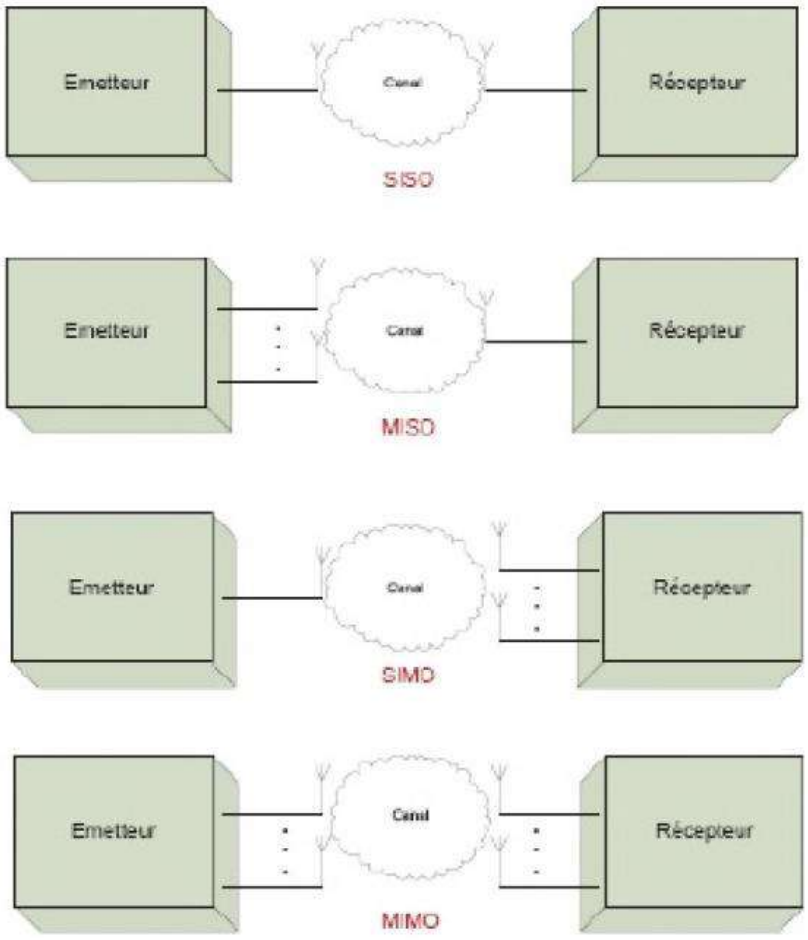


Figure I.3 : Schémas représentatifs du SISO, MISO, SIMO et MIMO.

I.2.5 Le centre de fusion :

Le rôle principal du centre de fusion est de décider et déclarer une valeur optimale selon une règle établi au préalable.

I.3 Types de WSN

Les WSN peuvent être déployés au sol, sous terre et sous l'eau. Cinq types fonctionnels peuvent être distingués, en particulier, terrestres, souterrains, sous-marins, WSN multimédias et mobiles. Ce qui suit fournit les détails de chaque type.

I.3.1 Terrestres WSNs :

Dans WSN non structuré, qui contient une collection dense de nœuds de capteurs. Les nœuds de capteurs peuvent être déployés de manière ad hoc sur le terrain, une fois déployé, le réseau est laissé sans surveillance pour effectuer la surveillance et les rapports les fonctions. Dans un WSN non structuré, la maintenance du réseau telle que la gestion la connectivité et la détection des pannes sont difficiles car il y a tellement de nœuds.

Dans WSN structuré, tout ou partie des nœuds de capteur sont déployés dans un de manière pré-planifiée. L'avantage un réseau structuré est que moins de nœuds peut être déployé avec des coûts de maintenance et de gestion du réseau inférieurs. Moins les nœuds sont déployés de manière bénéfique puisqu'ils sont placés à des emplacements spécifiques pour fournir une couverture tandis que le déploiement ad hoc peut avoir des régions non couvertes. Les nœuds de capteur sont déployés sur le champ de capteur à portée de la transmission distance les uns des autres et à des densités pouvant atteindre 20 nœuds / m3. À forte densité le déploiement de centaines ou de milliers de nœuds de capteurs sur un champ nécessite une maintenance de la topologie en trois phases :

- Phase de pré-déploiement et de déploiement. Les nœuds de capteur peuvent être lancés le champ de déploiement en tant que masse un avion ou d'un obus artillerie, ou placé un par un par un humain ou un robot.
- Phase de post-déploiement. Après le déploiement, la topologie change en raison du changement de position des nœuds de capteur, accessibilité (qui peut être affectée par le brouillage, le bruit, obstacles mobiles, etc.), énergie restante, dysfonctionnement et détails de la tâche.

I.3.2 Souterrains WSNs :

Les WSN souterrains se composent un certain nombre de nœuds de capteurs enfouis sous terre ou dans une grotte ou une mine utilisée pour surveiller les conditions souterraines. Des nœuds de puits supplémentaires sont situés au-dessus du sol pour relayer les informations des nœuds de capteur vers la station de base. Un WSN souterrain est plus cher qu'un WSN terrestre en termes d'équipement, de déploiement et de maintenance. Les nœuds de capteurs souterrains sont coûteux car les pièces équipement appropriées doivent être sélectionné pour assurer une communication fiable à travers le sol, les roches, l'eau et autres contenu minéral. Environnement souterrain fait de la communication sans fil un défi en raison des pertes de signal et des niveaux d'atténuation élevés. Contrairement au terrestre.

WSN, le déploiement d'un WSN souterrain nécessite une planification minutieuse et considérations énergétiques et de coût. L'énergie est une préoccupation importante dans le sous-sol WSN. Comme le WSN terrestre, les nœuds de capteurs souterrains sont équipés d'une puissance de la batterie et une fois déployée dans le sol, il est difficile de recharger ou remplacez la pile d'un nœud de capteur. Comme d'habitude, un objectif clé est d'économiser l'énergie afin d'augmenter la durée de vie du réseau, ce qui peut être réalisé en mettant en œuvre protocole de communication efficace.

I.3.3 Réseaux de capteurs acoustiques sous-marins (UASNs)

La technologie des réseaux de capteurs acoustiques sous-marins (UASN) offre de nouvelles opportunités d'explorer les océans, et par conséquent, il améliore la compréhension de la les problèmes environnementaux, tels que le changement climatique, la vie des animaux marins et variations de la population des récifs coralliens. De plus, les UASN peuvent améliorer la capacité de guerre sous-marine des forces navales puisqu'elles peuvent être utilisées pour surveillance, détection de sous-marins, missions de lutte contre les mines et sans pilote opérations dans les champs ennemis. De plus, surveiller les plates-formes pétrolières avec des UASN peut aider à prendre des mesures préventives pour les catastrophes telles que L'explosion de la plate-forme qui a eu lieu dans le golfe du Mexique en 2010. Dernier point mais non le moindre, tremblement de terre et Les systèmes d'alerte aux tsunamis peuvent également bénéficier de la technologie UASN

I.3.4 MultiMedia des WSNs

Des WSN multimédias ont été proposés pour permettre la surveillance et le suivi des événements sous la forme de multimédia comme la vidéo, l'audio et l'imagerie. Les WSN multimédias se composent d'un certain nombre de nœuds de capteurs à faible coût équipés de caméras et microphones. Ces nœuds de capteurs s'interconnectent les uns aux autres sur un connexion sans fil pour la récupération, le traitement, la corrélation et la compression des données. Ils sont déployés de manière pré planifiée dans l'environnement pour garantir couverture. Les défis du WSN multimédia comprennent :

- Demande de bande passante élevée.
- Consommation d'énergie élevée.
- Fourniture de la qualité de service (QoS).
- Techniques de traitement et de compression des données.
- Conception à couches croisées.

I.3.5 Mobile des WSNs

Les WSN mobiles consistent en un ensemble de nœuds de capteurs qui peuvent se déplacer seuls et interagir avec l'environnement physique. Il y a plusieurs problèmes comparatifs entre les nœuds de capteurs mobiles et statiques

I.4 Domaines d'applications des réseaux de capteurs sans fil :

La miniaturisation des capteurs, le coût de plus en plus faible, la large gamme des types de capteurs disponibles ainsi que le support de communication sans fil utilisé, permettent aux

réseaux de capteurs de se développer dans plusieurs domaines d'application [9] [10]. Ils permettent aussi d'étendre les applications existantes.

Les réseaux de capteurs peuvent se révéler très utiles dans de nombreuses applications lorsqu'il s'agit de collecter et de traiter des informations provenant de l'environnement. Parmi les domaines où ces réseaux peuvent offrir les meilleures contributions, nous citons les domaines : militaire, surveillance, environnemental, médical, domestique, commercial, etc.

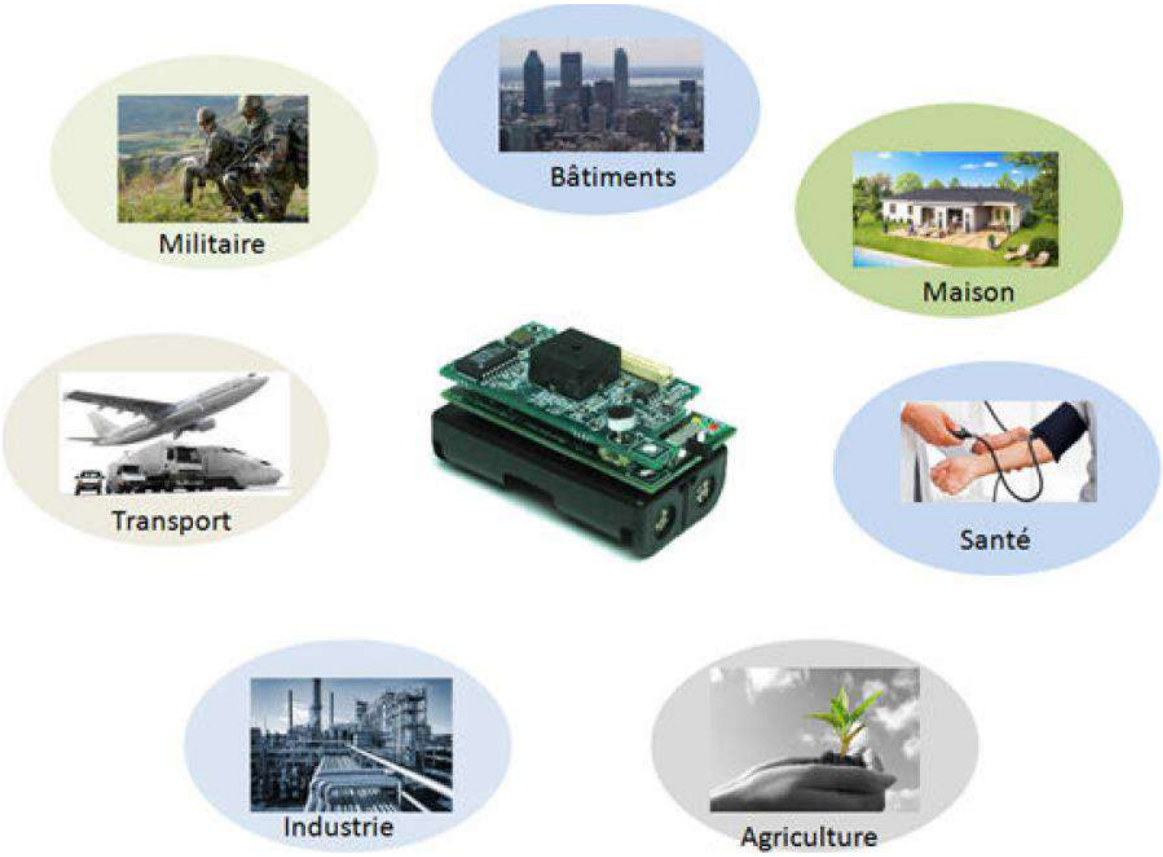


Figure I.4 : Domaines d'applications.

I.4.1 Applications militaires

Le faible coût et le déploiement rapide sont des caractéristiques qui ont rendu les réseaux de capteurs efficaces pour les applications militaires. Plusieurs projets ont été lancés pour aider les unités militaires dans un champ de bataille et protéger les villes contre des attaques, telles que les menaces terroristes. Le projet DSN (Distributed Sensor Network) au DARPA (Defense Advanced Research Projects Agency) [11] était l'un des premiers projets dans les années 80 ayant utilisé les réseaux de capteurs pour rassembler des données distribuées. Les chercheurs du laboratoire national Lawrence Livermore ont mis en place le réseau WATS (Wide Area Tracking System) [12]. Ce réseau est composé de détecteurs des rayons gamma et des neutrons pour détecter et dépister les dispositifs nucléaires. Il est capable d'effectuer la surveillance

constante d'une zone d'intérêt. Il utilise des techniques d'agrégation de données pour les rapporter à un centre intelligent.



Figure I.5 : Les réseaux de capteurs Application militaire [28].

I.4.2 Applications environnementales et agricoles

Le contrôle des paramètres environnementaux par les réseaux de capteurs peut donner naissance à plusieurs applications. Par exemple, le déploiement des thermo-capteurs dans une forêt peut aider à détecter un éventuel début de feu et par suite faciliter la lutte contre les feux de forêt avant leur propagation. Le déploiement des capteurs chimiques dans les milieux urbains peut aider à détecter la pollution et analyser la qualité d'air. De même leur déploiement dans les sites industriels empêche les risques industriels tels que la fuite de produits toxiques (gaz, produits chimiques, éléments radioactifs, pétrole, etc.).

Dans le domaine de l'agriculture, les capteurs peuvent être utilisés pour réagir convenablement aux changements climatiques par exemple le processus d'irrigation lors de la détection de zones sèches dans un champ agricole. Cette expérimentation a été réalisée par Intel Research Laboratory et Agriculture et Agri-Food Canada sur une vigne à British Columbia. L'avantage principal de l'utilisation des RCSF dans l'agriculture est l'absence de câblage ce qui réduit considérablement le coût d'installation. Un autre avantage est la flexibilité de déploiement et la facilité d'entretien. En effet, les capteurs sont autonomes et nécessitent très peu d'intervention humaine sur les champs notamment dans le cas où les protocoles de communication sont tolérants aux pannes et supportent la mobilité des nœuds. Plusieurs projets [13], [14], [15] et [16] et ont utilisé la technologie des RCSF dans l'agriculture [17].



Figure I.6 : Application environnementale [29].

I.4.3 Applications médicales

Dans le domaine de la médecine, les réseaux de capteurs peuvent être utilisés pour assurer une surveillance permanente des organes vitaux de l'être humain grâce à des micro-capteurs qui pourront être avalés ou implantés sous la peau (surveillance de la glycémie, détection de cancers, etc.). Ils peuvent aussi faciliter le diagnostic de quelques maladies en effectuant des mesures physiologiques telles que : la tension artérielle, battements du cœur, etc. à l'aide des capteurs ayant chacun une tâche bien particulière [18]. Les données physiologiques collectées par les capteurs peuvent être stockées pendant une longue durée pour le suivi d'un patient [19]. D'autre part, ces réseaux peuvent détecter des comportements anormaux (chute d'un lit, choc, cri, etc.) chez les personnes dépendantes (handicapées ou âgées).

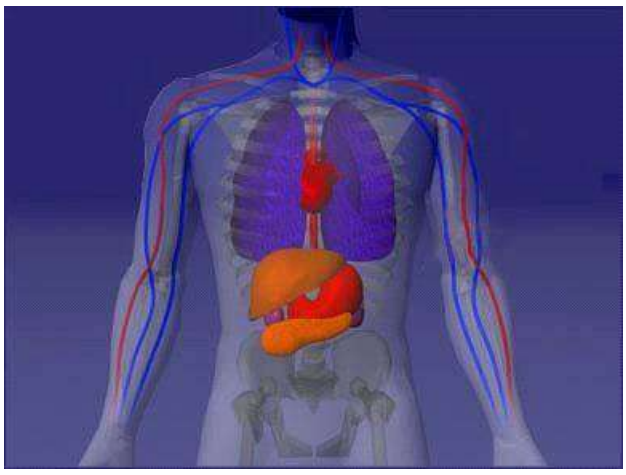


Figure I.7 : Un réseau de capteurs dans le domaine medical [30].

I.4.4 Applications de surveillance

L'application des réseaux de capteurs dans le domaine de la sécurité peut diminuer considérablement les dépenses financières consacrées à la sécurisation des lieux et des êtres humains. Ainsi, l'intégration des capteurs dans de grandes structures telles que les ponts ou les bâtiments aidera à détecter les fissures et les altérations dans la structure suite à un séisme ou au vieillissement de la structure. Le déploiement d'un réseau de capteurs de détection de mouvement peut constituer un système d'alarme qui servira à détecter les intrusions dans une zone de surveillance.

I.4.5 Applications industrielles

Les machines de production peuvent être configurées et pilotées en interconnexion grâce à des systèmes d'acquisition et réseaux de capteurs sans fil, autonomes et communicants. Grâce à ce système, la connaissance précise et en continu de l'état de santé des machines est la garantie d'un maintien en conditions d'exploitation optimale (condition monitoring), d'interventions de maintenance réalisées juste au bon moment, et d'économies conséquentes.

Avec de tels systèmes, les opérationnels ont la possibilité de contrôler, anticiper et surveiller les machines à distance. Chocs, vibrations, déplacements, mouvements, paramètres du contexte environnemental (température, pression, etc..), autant de données stratégiques qu'il convient de collecter en temps réel, au cœur de l'outil de production bien souvent complexe et en milieu contraint.

I.4.6 Applications commerciales

Il est possible d'intégrer des capteurs au processus de stockage et de livraison dans le domaine commercial. Le réseau ainsi formé pourra être utilisé pour connaître la position, l'état et la direction d'un paquet. Il devient alors possible pour un client qui attend la réception d'un colis, d'avoir un avis de livraison en temps réel et de connaître la localisation actuelle du paquet [20]. Pour les entreprises manufacturières, les réseaux de capteurs permettront de suivre le procédé de production à partir des matières premières jusqu'au produit final livré. Grâce aux réseaux de capteurs, les entreprises pourraient offrir une meilleure qualité de service tout en réduisant leurs coûts.

I.4.7 Applications domestiques

Avec le développement technologique, les capteurs peuvent être embarqués dans des appareils, tels que les aspirateurs, les fours à micro-ondes, les réfrigérateurs, les magnétoscopes, etc. Ces capteurs embarqués peuvent interagir entre eux et avec un réseau externe via Internet pour permettre à un utilisateur de contrôler les appareils domestiques localement ou à distance. Le déploiement des capteurs de mouvement et de température dans les futures maisons dites

intelligentes permet d'automatiser plusieurs opérations domestiques telles que : la lumière s'éteint et la musique s'arrête quand la chambre est vide, la climatisation et le chauffage s'ajustent selon les points multiples de mesure, l'alarme est déclenchée par le capteur anti-intrusion quand un étranger veut pénétrer dans la maison [21].

I.5 Conclusion :

Dans ce chapitre nous donnons une généralité sur les réseaux de capteur sans fil WSN (Wireless Sensor network). Ces réseaux sont omniprésents dans divers domaines tels que la santé et le secteur militaire …etc. et sa principale utilisation est pour mesurer un phénomène dont la pression, la température, l'humidité…etc. Donc les réseaux de capteur sans fil sont très utiles dans notre vie quotidienne.

CHAPITRE II :

Rappel sur la théorie d’estimation et de détection.

II.1 Introduction :

Dans notre vie quotidienne, nous sommes constamment amenés à prendre des décisions. Compte tenu de certaines hypothèses, un critère est sélectionné, sur la base duquel une décision doit être prise. Par exemple, en l'ingénierie, lorsqu'il y a un problème de détection de signal radar, le signal renvoyé est observé et une décision est prise quant à la présence ou l'absence d'une cible. Observé et une décision est prise quant à la présence ou l'absence d'une cible. Dans un système de communication numérique, une chaîne de zéros et d'uns peut être transmise sur un support. Un support quelconque. Au niveau du récepteur, les signaux reçus représentant les zéros et les uns sont corrompus dans le support par un bruit additif et par le bruit du récepteur. Le récepteur ne sait pas quel signal représente un Zéro et quel signal représente un Un. Mais doit décider si les signaux reçus représentent des zéros ou des uns. Le processus que le récepteur entreprend pour sélectionner une règle de décision relève de la théorie de la détection des signaux.

Dans la théorie de la détection des signaux, la situation ci-dessus peut être décrite par une source émettant deux sorties possibles à différents instants. Les sorties sont appelées des hypothèses. L'hypothèse nulle H_0 de L'hypothèse nulle H_0 représente un zéro (cible absente) et l'hypothèse alternative H_1 représente un (cible présente).H_1 représente un 1 (cible présente), comme le montre la Figure II.1.

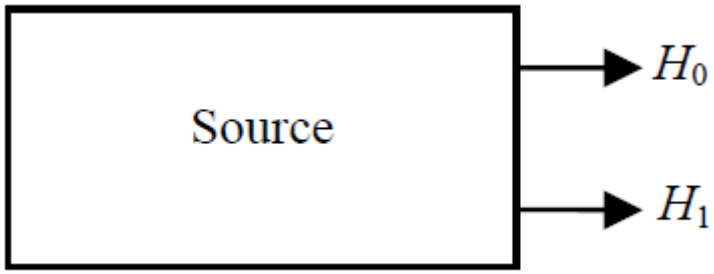


Figure II.8 : Source de l'hypothèse binaire.

Chaque hypothèse correspond à une ou plusieurs observations qui sont représentées par des variables aléatoires. En fonction des valeurs d'observation de ces variables aléatoires, le récepteur décide quelle hypothèse (H_0 ou H_1) est vraie. Supposons que le récepteur doit prendre une décision sur la base d'une seule observation du signal reçu. La gamme de valeurs que prend la variable aléatoire Y constitue l'espace d'observation Z. L'espace d'observation est divisé en deux régions Z_0 et Z_1, de telle sorte que si Y se trouve dans Z_0, le récepteur décide en faveur de

H_0, tandis que si Y se trouve dans Z_1, le récepteur décide en faveur de H_1. L'espace d'observation Z est l'union de Z_0 et Z_1 ; c'est-à-dire que

$$Z = Z_0 \cup Z_1 \qquad \text{(II.1)}$$

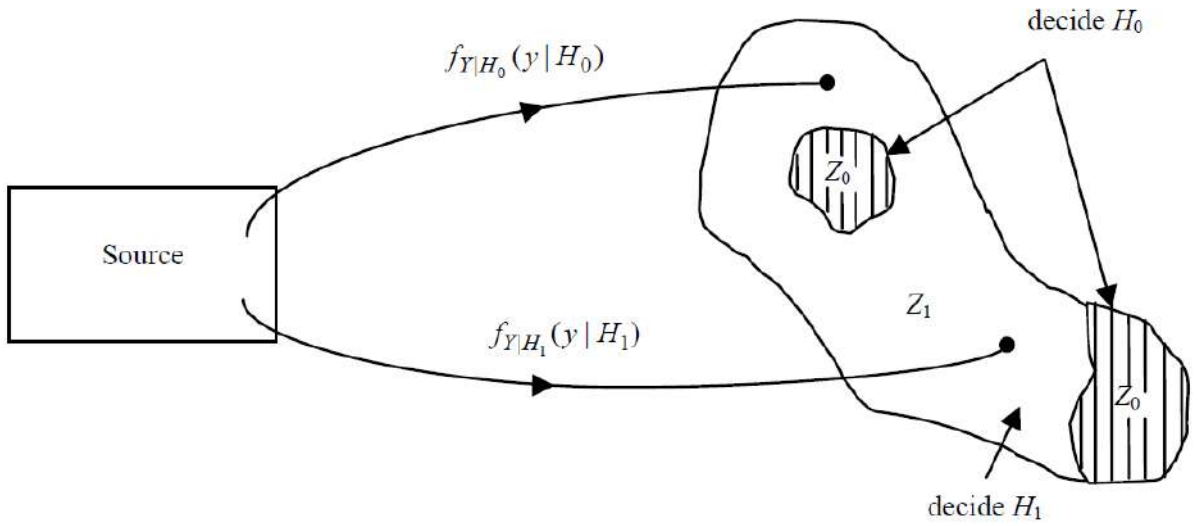


Figure II.9 : Région de décision.

Les fonctions de densité de probabilité de Y correspondant à chaque hypothèse sont $f_Y|_{H_0}(y|H_0)$ et $f_Y|_{H1}(y|H_1)$, où y est une valeur particulière de la variable aléatoire Y.

Chaque fois qu'une décision est prise, sur la base d'un certain critère, pour ce problème de test d'hypothèse binaire, quatre cas possibles peuvent se produire :

1- Décider H_0lorsque H_0 est vrai.

2- Décider H_0lorsque H_1 est vrai.

3- Décider H_1lorsque H_0 est vrai.

4- Décider H_1lorsque H_1 est vrai.

Observons que dans les cas (1) et (4), le récepteur prend une décision correcte, alors que dans les cas (2) et (3), le récepteur fait une erreur. Cas (2) et (3), le récepteur commet une erreur. Dans la nomenclature radar, le cas (2) est appelé miss, le cas (3) une fausse alarme, et le cas (4) une détection.

Dans ce chapitre, nous développons les principes de base nécessaires à la résolution des problèmes de décision. Les observations sont représentées par des variables aléatoires.

Dans les sections suivantes, nous étudions certains des critères utilisés dans la théorie de la décision, ainsi que les conditions dans lesquelles ces critères sont utilisés. Théorie de la décision, et les conditions dans lesquelles ces critères sont utiles.

II.2 Théorie de la décision statistique

II.2.1 CRITÈRE DE BAYES :

II.2.1.1 Test d'hypothèse binaire

L'utilisation du critère de Bayes repose sur deux hypothèses. Premièrement, la probabilité de d'occurrence des deux sorties sources est connue. Ce sont les probabilités a priori $P(H_0)$ et $P(H_1)$. $P(H_0)$ est la probabilité d'occurrence de l'hypothèse H_0, tandis que $P(H_1)$ est la probabilité d'occurrence de l'hypothèse H_1. En désignant les probabilités a priori $P(H_0)$ et $P(H_1)$ par P_0 et P_1 respectivement, et comme l'une ou l'autre desl'hypothèse H_0 ou H_1 se produira toujours, nous avons

$$P0 + P1 = 1 \qquad \text{(II.2)}$$

La deuxième hypothèse est qu'un coût est attribué à chaque décision possible. Le coût est dû au fait qu'une action sera entreprise sur la base de la décision prise.

Les conséquences d'une décision sont différentes des conséquences d'une autre. Par exemple, dans un problème de détection radar, les conséquences d'un échec ne sont pas les mêmes que celles d'une fausse alarme.

Si l'on considère $D_i, i = 0, 1$, où D_0 signifie " décider H_0 " et D_1 signifie " décider H_1 ". Nous pouvons définir$C_{ij}, i, j = 0, 1$, comme étant le coût associé à la décision D_i, étant donné que la véritable hypothèse est H_j.

C'est-à-dire,

$$P\big(\text{Cout enccouru } C_{ij}\big) = P(\text{Décider } D_i \text{ , H j evt vrai}), i, j = 0.1 \qquad \text{(II.3)}$$

En particulier, les coûts pour ce problème de test d'hypothèse binaire sont C_{00} pour le cas (1), C_{01} pour le cas (2), C_{10} pour le cas (3), et C_{11} pour le cas (4). L'objectif du critère de Bayes est de déterminer la règle de décision de sorte que le coût moyen $E[C]$,également appelé risque $\Re$, soit minimisé. L'opération $E\ [C]$ désigne la valeur moyenne. Il On suppose également que le coût d'une mauvaise décision est supérieur au coût d'une bonne décision C'est-à-dire,

$$C_{01} > C_{11} \qquad \text{(II.4)}$$

Et

$$C_{10} > C_{00} \qquad \text{(II.5)}$$

Étant donné $P\ (D_i, H_j)$ la probabilité conjointe que nous décidions D_i, et que l'hypothèse H_j est vraie, le coût moyen est

$$\mathrm{R} = \mathrm{E}[\mathrm{C}] = \mathrm{C}_{00}\mathrm{P}\ (\mathrm{D}_0, \mathrm{H}_0) + \mathrm{C}_{01}\mathrm{P}\ (\mathrm{D}_0, \mathrm{H}_1) + \mathrm{C}_{10}\mathrm{P}\ (\mathrm{D}_1, \mathrm{H}_0) + \mathrm{C}_{11}\mathrm{P}\ (\mathrm{D}_1, \mathrm{H}_1) \quad \text{(II.6)}$$

D'après la règle de Bayes, nous avons

$$P(D_i\ , H_j) = P(D_i \mid H_j)P(H_j) \quad \text{(II.7)}$$

Les fonctions de densité conditionnelle $P(D_i \mid H\ _j\ ), i, j\ =\ 0, 1,$ en termes des régions présentées à la figure II.2, sont

$$P(D_0\ \mid H_0) = P(\text{Décider } H_0\ , H_0\ \ est\ vrai) = \int_{Z_0} f_Y|H_0(y|H_0)dy \quad \text{(II.8)}$$

$$P(D_0\ \mid H_1) = P(\text{Décider } H_0\ , H_1\ \ est\ vrai) = \int_{Z_0} f_Y|H_1(y|H_1)dy \quad \text{(II.9)}$$

$$P(D_0\ \mid H_0) = P(\text{Décider } H_1\ , H_0\ \ est\ vrai) = \int_{Z_1} f_Y|H_0(y|H_0)dy \quad \text{(II.10)}$$

Et

$$P(D_1\ \mid H_1) = P(\text{Décider } H_1\ , H_1\ \ est\ vrai) = \int_{Z_0} f_Y|H_0(y|H_0)dy \quad \text{(II.11)}$$

Les probabilités $P(D_0|H_1)$, $P(D_1|H_0\ )$ et $P(D_1\ , H_1\ )$ représentent les probabilités d'oubli, P_M, la probabilité de fausse alarme, P_F , et la probabilité de détection, P_D, respectivement. Nous observons également que :

$$\mathrm{P_M} = 1 - \mathrm{P_D} \quad \text{(II.12)}$$

Et

$$\mathrm{P}(\mathrm{D}_0|\mathrm{H}_0) = 1 - \mathrm{P_F} \quad \text{(II.13)}$$

Par conséquent, la probabilité d'une décision correcte est donnée par :

$$\begin{aligned} P(Bonne\ décision) = P(c) &= P(D_0, H_0) + P(D_1, H_1) \quad \text{(II.14)} \\ &= \mathrm{P}(\mathrm{D}_0\ \mid \mathrm{H}_0)\mathrm{P}(\mathrm{H}_0) + \mathrm{P}(\mathrm{D}_1\ \mid \mathrm{H}_1)\mathrm{P}(\mathrm{H}_1) \\ &= (1 - \mathrm{P_F})\mathrm{P}_0 + \mathrm{P_D}\mathrm{P}_1 \end{aligned}$$

Et la probabilité d'erreur est donnée par :

$$\begin{aligned} \mathrm{P}(\text{erreur}) = \mathrm{P}(\boldsymbol{\varepsilon}) &= \mathrm{P}(\mathrm{D}_0\ , \mathrm{H}_1) + \mathrm{P}(\mathrm{D}_1\ , \mathrm{H}_0) \quad \text{(II.15)} \\ &= \mathrm{P}(\mathrm{D}_0\ \mid \mathrm{H}_1)\mathrm{P}(\mathrm{H}_1) + \mathrm{P}(\mathrm{D}_1\ \mid \mathrm{H}_0)\mathrm{P}(\mathrm{H}_0) \end{aligned}$$

$$= P_M P_1 + P_F P_0$$

Un graphique de la fonction de densité de probabilité du coût, c (c) , est illustré à la figure II.3. Le coût moyen devient maintenant

$$R = E[C] = C_{00}(1 - P_F)P_0 + C_{01}(1 - P_D)P_1 + C_{10}P_F P_0 + C_{11}P_D P_1 \quad \text{(II.16)}$$

En termes de régions de décision définies en (II.8) à (II.10), le coût moyen est de

$$R = P_0 C_{00} \int_{Z_0} f_Y|_{H_0}(y|H_0)dy + P_1 C_{01} \int_{Z_0} f_Y|_{H_1}(y|H_1)dy + P_0 C_{10} \int_{Z_1} f_Y|_{H_0}(y|H_0)dy + P_1 C_{11} \int_{Z_1} f_Y|_{H_1}(y|H_1)dy \quad \text{(II.17)}$$

En utilisant (II.1) et le fait que

$$\int_Z f_Y|_{H_0}(y|H_0)dy = \int_Z f_Y|_{H_1}(y|H_1)dy = 1 \quad \text{(II.18)}$$

Il s'ensuit que

$$\int_Z f_Y|_{H_j}(y|H_j)dy = 1 - \int_Z f_Y|_{H_j}(y|H_j)dy, \quad j = 0,1 \quad \text{(II.19)}$$

Où $f_Y|_{H_j}(y|H_j), j = 0, 1$, est la fonction de densité de probabilité de Y correspondant à chaque hypothèse. En remplaçant (II.19) dans (II.17), on obtient

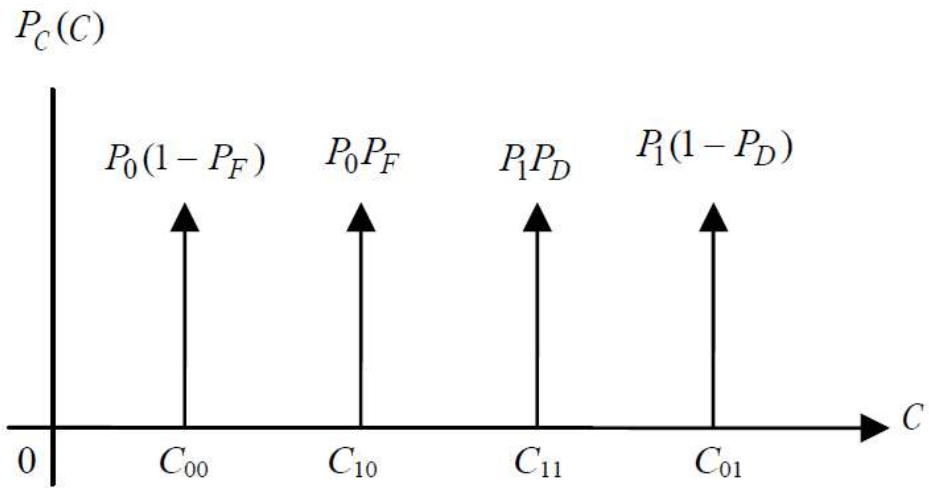


Figure II.10 : Fonction de densité du cout.

$$R = P_0 C_{10} + P_1 C_{11} + \int_{Z_0} \{[P_1(C_{01} - C_{11})f_Y|_{H_1}(y|H_1)] - [P_0(C_{10} - C_{00})f_Y|_{H_0}(y|H_0)]\} dy \quad \text{(II.20)}$$

Nous observons que la quantité $P_0C_{10} + P_1C_{11}$ est constante, indépendamment de la façon dont nous assigner des points dans l'espace d'observation, et que la seule grandeur variable est la

région d'intégration Z_0. De (II.4) et (II.5), les termes entre parenthèses de (II.20) $[P_1(C_{01} - C11)f_y|_{H_1}(y|H_1)$ and $P_0(C_{10} - C_{00})f_y|_{H_0}(y|H_0)]$ sont tous les deux positifs.

Par conséquent, le risque est minimisé en sélectionnant la région de décision Z_0 pour inclure seuls les points de Y pour lesquels le second terme est plus grand, et donc l'intégral est négatif. Plus précisément, nous affectons à la région Z_0 les points pour lesquels

$$[\mathrm{P_1(C_{01} - C_{11})f_Y|_{H_1}(y|H_1)} \text{ and } \mathrm{P_0(C_{10} - C_{00})f_Y|_{H_0}(y|H_0)} \tag{II.21}$$

Toutes les valeurs pour lesquelles le deuxième terme est supérieur seront exclues de Z_0 et affecté à Z_1. Les valeurs pour lesquelles les deux termes sont égaux n'affectent pas le risque, et peut être affecté à Z_0 ou Z_1. Par conséquent, on dit si

$$\mathrm{P_1(C_{01} - C_{11})f_Y|_{H_1}(y|H_1) < P_0(C_{10} - C_{00})f_Y|_{H_0}(y|H_0)} \tag{II.22}$$

Alors nous décidons H_1. Sinon, on décide H_0. Ainsi, la règle de décision résultant du critère de Bayes est

$$\frac{\mathrm{f_Y|_{H_1}(y|H_1)}}{\mathrm{f_Y|_{H_0}(y|H_0)}} \underset{\mathrm{H_0}}{\overset{\mathrm{H_1}}{\gtrless}} \frac{\mathrm{P_0(C_{10} - C_{00})}}{\mathrm{P_1(C_{01} - C_{11})}} \tag{II.23}$$

Le rapport de $f_Y|_{H_0}(y|H_0)$ sur $|f_Y|_{H_1}(y|H_1)$ est appelé le rapport de vraisemblance et est noté $\Lambda(y)$. C'est-à-dire,

$$\mathrm{\Lambda(y) = \frac{f_Y|_{H_1}(y|H_1)}{f_Y|_{H_0}(y|H_0)}} \tag{II.24}$$

Il est à noter que si nous avons K observations, par exemple, K échantillons d'une forme d'onde reçue, $Y_1, Y_2, \ldots\ldots, Y_K$, sur la base de laquelle nous prenons la décision, le rapport de vraisemblance peut être exprimé comme

$$\Lambda(y) = \frac{f_Y|_{H_1}(y|H_1)}{f_Y|_{H_0}(y|H_0)} \tag{II.25}$$

Où Y, le vecteur reçu, est

$$Y^T = [Y_1 \;\; Y_2 \quad \cdots \quad Y^K] \tag{II.26}$$

La statistique de vraisemblance $\Lambda(Y)$ est une variable aléatoire puisqu'elle est fonction de la variable aléatoire Y. Le seuil est

$$\eta = \frac{P_0(C_{10} - C_{00})}{P_1(C_{01} - C_{11})} \tag{II.27}$$

Ainsi, le critère de Bayes, qui minimise le coût moyen, conduit à test du rapport de vraisemblance

$$\Lambda(y) \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{II.28}$$

II.2.1.2 Test d'hypothèse M-aire :

Nous avons considéré le choix entre deux hypothèses, H_0 et H_1. Considérons maintenant le choix d'une hypothèse parmi M hypothèses, $H0, H1, \ldots\ldots, HM-1$, chaque fois qu'une expérience est menée. Le critère de Bayes attribue un coût à chaque alternative. A la $ijème$ alternative, qui est la décision Di étant donné l'hypothèse H_j, le coût $Cij, i, j = 0,1,\ldots,(M-1)$, est assigné. En plus de l'hypothèse $H0, H1, \ldots\ldots, HM-1$, on attribue les probabilités a priori $P_0, P_1, \ldots\ldots\ldots, P_{M-1}$, respectivement. L'objectif est de minimiser le risque défini comme

$$\mathrm{R} = \sum_{\mathrm{j}=1}^{\mathrm{M}-1} \sum_{\mathrm{j}=1}^{\mathrm{M}-1} \mathrm{P_j C_{ij} P(D_i\,|H_j\,)} \tag{II.29}$$

Utilisant le fait que

$$\mathrm{P(D_i\,|H_j\,)} = \int_{\mathrm{Z_i}} \mathrm{f_Y|_{H_j}(y|H_j)dy} \tag{II.30}$$

Le coût moyen devient

$$\mathrm{R} = \sum_{\mathrm{j}=1}^{\mathrm{M}-1} \sum_{\mathrm{j}=1}^{\mathrm{M}-1} \mathrm{P_j C_{ij}} \int_{\mathrm{Z_i}} \mathrm{f_Y|_{H_j}(y|H_j)dy} \tag{II.31}$$

L'espace d'observation Z est maintenant divisé en M sous-espaces,$Z_0, Z_1, \ldots, Z_{M-1}$,tels cette

$$\mathrm{Z = Z_0 \cup\ z_1 \cup\ \ldots \cup\ Z_{M_1}} \tag{II.32}$$

$\Re$ est minimisé,

$$\mathrm{R} = \sum_{\mathrm{i}=0}^{\mathrm{M}-1} \sum_{\substack{\mathrm{j}=0 \\ \mathrm{j}\neq\mathrm{i}}}^{\mathrm{M}-1} \mathrm{P_j(C_{ij} - C_{ij})} \int_{\mathrm{Z_i}} \mathrm{f_Y|_{H_j}(y|H_j)dy} + \sum_{\mathrm{i}=1}^{\mathrm{M}-1} \mathrm{P_j C_{ij}} \int_{\mathrm{Z_i}} \mathrm{f_Y|_{H_i}(y|H_i)dy} \tag{II.33}$$

Du moment que

$$\int_{\mathrm{Z}} \mathrm{f_Y|_{H_j}(y|H_i)dy} = 1 \tag{II.34}$$

Et la superficie

$$z_i = Z - \coprod_{\substack{j=0 \\ j\neq i}}^{M-1} Z_{\ j} \tag{II.35}$$

Substituer dans (II.33), le risque devient

$$R = \sum_{j=1}^{M-1} \int_{Z_i} \sum_{\substack{j=1 \\ j\neq i}}^{M-1} P_j(C_{ij} - C_{jj}) \int_{Z_i} f_Y|_{H_j}(y|H_j)dy + \sum_{i=0}^{M-1} P_j C_{ii} \tag{II.36}$$

En utilisant le même raisonnement que précédemment, nous observons que le second terme de (II.36) est fixe, tandis que le premier terme détermine le coût pour les régions de décision sélectionnées. Par conséquent, la petite valeur intégrale donne la sélection de l'hypothèse pour laquelle

$$I_i(y) = \sum_{\substack{j=0 \\ j\neq i}}^{M-1} P_j(C_{ij} - C_{jj})\, f_Y|_{H_j}(y|H_j) \tag{II.37}$$

Est minimum.

Définition du rapport de vraisemblance $\Lambda i(\boldsymbol{y}), i = 1, 2, \ldots, M-1$, as

$$\Lambda_i(y) = \frac{f_Y|_{H_i}(y|H_i)}{f_Y|_{H_0}(y|H_0)}\ , i = 1,2, \ldots, M-1 \tag{II.38}$$

Et le terme $J_i(Y)$ comme

$$\mathrm{J_i(y)} = \frac{\mathrm{I_i(y)}}{\mathrm{f_Y|_{H_0}(y|H_0)}} = \sum_{\mathrm{j=1}}^{\mathrm{M-1}} \mathrm{P_j}\left(\mathrm{C_{ij}} - \mathrm{C_{jj}}\right)\, \Lambda_\mathrm{j}(\mathrm{y}) \tag{II.39}$$

La règle de décision est de choisir l'hypothèse pour laquelle (II.39) est minimale. Dans les problèmes de communication, il est courant d'avoir les coûts

$$C_{ii} = 0, \quad i = 1,2, \ldots \ldots, M-1 \tag{II.40}$$

Et

$$C_{ij} = 1, \quad i \neq 1 \ \ et\ i,j = 0,1, \ldots \ldots, M-1 \tag{II.41}$$

Dans ce cas, minimiser le risque revient à minimiser la probabilité d'erreur. Après substitution, (II.39) devient

$$I_i(y) = \sum_{j=0}^{M-1} P(H_j) f_Y|_{H_j}(y|H_j) = \sum_{\substack{j=0 \\ i\neq j}}^{M-1} P_j(H_j|Y) f_Y(y) = [1 - P(H_i|Y)] f_Y(y) \tag{II.42}$$

Minimiser (II.42) équivaut à maximiser $(Hi\,|Y\,)$, qui est la probabilité a posteriori de l'hypothèse H_i, étant donné le vecteur d'observation y. Si en plus, les probabilités a priori sont égales

$$\mathrm{P(H_1) = P(H_2) = \cdots = P\left(H_{M-1}\right) = P} \tag{II.43}$$

L'équation (II.42) devient

$$I_i(y) = P\sum_{\substack{j=0 \\ j\neq i}}^{M-1} f_Y|_{H_j}(y|H_j) = p[1 - P(H_i|Y)] \tag{II.44}$$

On observe que dans cette règle de décision, le récepteur calcule les Probabilité à posteriori $P(Hi\,|Y\,), i\; =\; 0, 1, 2, \ldots, M\; - 1$, et se prononce en faveur de l'hypothèse correspondant à la plus grande probabilité a posteriori. Une telle probabilité minimale du récepteur d'erreur est également appelée probabilité maximale a posteriori (MAP). Pour simplifier, soit M = 3 et l'espace d'observation soit $=\; Z0\; U\; Z1\; U\; Z2$.

De (II.31), on obtient

$$\begin{aligned} R = {} & P_0C_{00}\int_{Z_0} f_Y|H_0(y|H_0)dy + P_0C_{10}\int_{Z_1} f_Y|H_0(y|H_0)dy \\ & + P_0C_{20}\int_{Z_2} f_Y|H_0(y|H_0)dy + P_1C_{11}\int_{Z_1} f_Y|H_1(y|H_1) \\ & + P_1C_{01}\int_{Z_0} f_Y|H_1(y|H_1)dy \\ & + P_1C_{21}\int_{Z_2} f_Y|H_1(y|H_1)dy \\ & + P_2C_{22}\int_{Z_2} f_Y|H_2(y|H_2)dy + P_2C_{02}\int_{Z_0} f_Y|H_2(y|H_2)dy \\ & + P_2C_{12}\int_{Z_1} f_Y|H_2(y|H_2)dy \end{aligned} \tag{II.45}$$

Noter que

$$\begin{aligned} \int_{Zi} f_Y|_H(y|H_i)dy &= \int_{Z-\sum_{\substack{j \\ j\neq i}} Z} f_Y|_{H_i}(y|H_i)dy \\ &= 1 - \sum_{\substack{j \\ j\neq 1}} \int_{Z_f} f_Y|_{H_i}(y|H_i)dy, \qquad i,j = 0,1,2 \end{aligned} \tag{II.46}$$

En substituant (II.46) les termes impliquant C_{00}, C_{11}, et C_{22} dans (II.45), nous obtenons

$$R = P_0C_{00} + P_1C_{11} + P_2C_{22} \quad \text{(II.47)}$$
$$+ \int_{Z_0} [P_2(C_{02} - C_{22})f_Y|_{H_2}(y|H_2) + P_1(C_{01} - C_{11})f_Y|_{H_1}(y|H_1)]dy$$
$$+ \int_{Z_1} [P_0(C_{10} - C_{00})f_Y|_{H_0}(y|H_0) + P_2(C_{12} - C_{22})f_Y|_{H_2}(y|H_2)]dy$$
$$+ \int_{Z_2} [P_0(C_{20} - C_{00})f_Y|_{H_0}(y|H_0) + P_1(C_{21} - C_{11})f_Y|_{H_2}(y|H_1)]dy$$

Nous définissons $I_0(y)$, $I_1(y)$ et $I_2(y)$ comme

$$I_0(y) = P_2(C_{02} - C_{22})f_Y|_{H_2}(y|H_2) + P_1(C_{01} - C_{11})f_Y|_{H_2}(y|H_1) \quad \text{(II.48)}$$

$$I_1(y) = P_0(C_{10} - C_{00})f_Y|_{H_0}(y|H_0) + P_2(C_{12} - C_{22})f_Y|_{H_2}(y|H_2) \quad \text{(II.49)}$$

Et

$$I_2(y) = P_0(C_{20} - C_{00})f_Y|_{H_0}(y|H_0) + P_1(C_{21} - C_{11})f_Y|_{H_1}(y|H_1) \quad \text{(II.50)}$$

Pour minimiser $\Re$, nous attribuons des valeurs de Y à la région ayant les plus petits intégral dans (II.47), puisque $I_0(y)$, $I_1(y)$ et $I_2(y)$ sont non négatifs. En conséquence,

$$Z_0 = \{y|I_0(y) < I_1(y) \text{and } I_2(y)\} \quad \text{(II.51)}$$

$$Z_1 = \{y|I_1(y) < I_0(y) and\ I_2(y)\} \quad \text{(II.52)}$$

Et

$$Z_2 = \{y|I_2(y) < I_0(y) and\ I_1(y)\} \quad \text{(II.53)}$$

Où | signifie « tel que ». A partir de (II.38), les rapports de vraisemblance $\Lambda 1\ (y)$ et $\Lambda 2\ (y)$sont

$$\Lambda_1(y) = \frac{f_Y|_{H_1}(y|H_1)}{f_Y|_{H_0}(y|H_0)} \quad \text{(II.54)}$$

Et

$$\Lambda_2(y) = \frac{f_Y|_{H_2}(y|H_2)}{f_Y|_{H_0}(y|H_0)} \quad \text{(II.55)}$$

Afin d'incorporer les rapports de vraisemblance dans la règle de décision, nous utilisons le suivant test équivalent

$$I_0(y) \underset{H_0 or\ H_1}{\overset{H_1 or\ H_2}{\gtrless}} I_1(y) \tag{II.56}$$

$$I_0(y) \underset{H_0 or\ H_1}{\overset{H_1 or\ H_2}{\gtrless}} I_2(y) \tag{II.57}$$

Et

$$\mathrm{I_1(y)} \underset{\mathrm{H_0 or\ H_1}}{\overset{\mathrm{H_0 or\ H_2}}{\gtrless}} \mathrm{I_2(y)} \tag{II.58}$$

En substituant (II.48), (II.49), (II.50) et (II.54), (II.55) dans (II.56), (II.57), (II.58), on obtient le test

$$P_2(C_{01} - C_{11})\Lambda_1(y) \underset{H_0 or\ H_2}{\overset{H_1 or\ H_2}{\gtrless}} P_0(C_{10} - C_{00}) + P_2(C_{12} - C_{02})\boldsymbol{\Lambda_2(y)} \tag{II.59}$$

$$\mathrm{P_2(C_{02} - C_{22})\Lambda_2(y)} \underset{\mathrm{H_0 or\ H_1}}{\overset{\mathrm{H_1 or\ H_2}}{\gtrless}} \mathrm{P_0(C_{20} - C_{00}) + P_1(C_{21} - C_{01})}\boldsymbol{\Lambda_1(y)} \tag{II.60}$$

Et

$$P_2(C_{12} - C_{22})\Lambda_1(y) \underset{H_0 or\ H_1}{\overset{H_0 or\ H_2}{\gtrless}} P_0(C_{20} - C_{10}) + P_1(C_{21} - C_{01})\Lambda_1(y) \tag{II.61}$$

Parce que M = 3, il n'y a que deux rapports de vraisemblance et l'espace de décision est bidimensionnel, comme le montre la figure II.4. Pour les frais

$$\mathrm{C_{00} = C_{11} = C_{22} = 0} \tag{II.62}$$

$$\mathrm{C_{ij} = 1, \quad i \neq j} \tag{II.63}$$

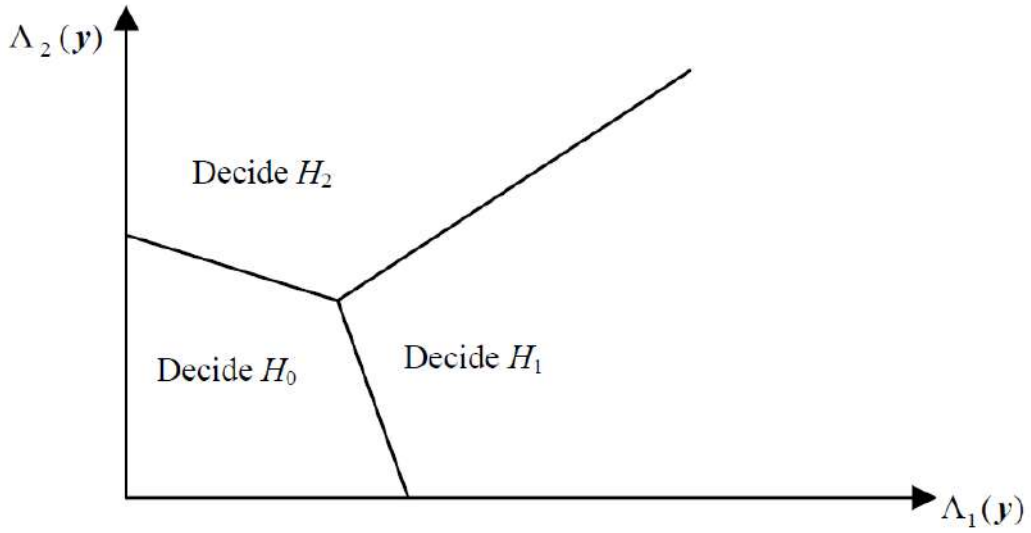


Figure II.11 : Espace de décision pour M = 3.

Il est plus facile d'observer, dans ce cas de M = 3, que minimiser le risque est équivalent à minimiser la probabilité d'erreur, et la règle de décision se réduit à

$$\Lambda_1(y) \underset{H_0 \text{ or } H_2}{\overset{H_1 \text{ or } H_2}{\gtrless}} \frac{P_0}{P_1} \tag{II.64}$$

$$\Lambda_2(y) \underset{H_0 \text{ or } H_1}{\overset{H_1 \text{ or } H_2}{\gtrless}} \frac{P_0}{P_2} \tag{II.65}$$

$$\Lambda_1(y) \underset{H_0 \text{ or } H_2}{\overset{H_1 \text{ or } H_2}{\gtrless}} \frac{P_0}{P_1} \tag{II.66}$$

Et

$$\Lambda_2(y) \underset{H_0 \text{ or } H_1}{\overset{H_1 \text{ or } H_2}{\gtrless}} \frac{P_0}{P_2} \tag{II.67}$$

Les régions de décision résultantes sont illustrées à la figure II.5 (a). La décision globale l'espace est donné dans la figure II.5 (b).

En prenant le logarithme des deux membres de (II.65), (II.66), (II.67) on obtient

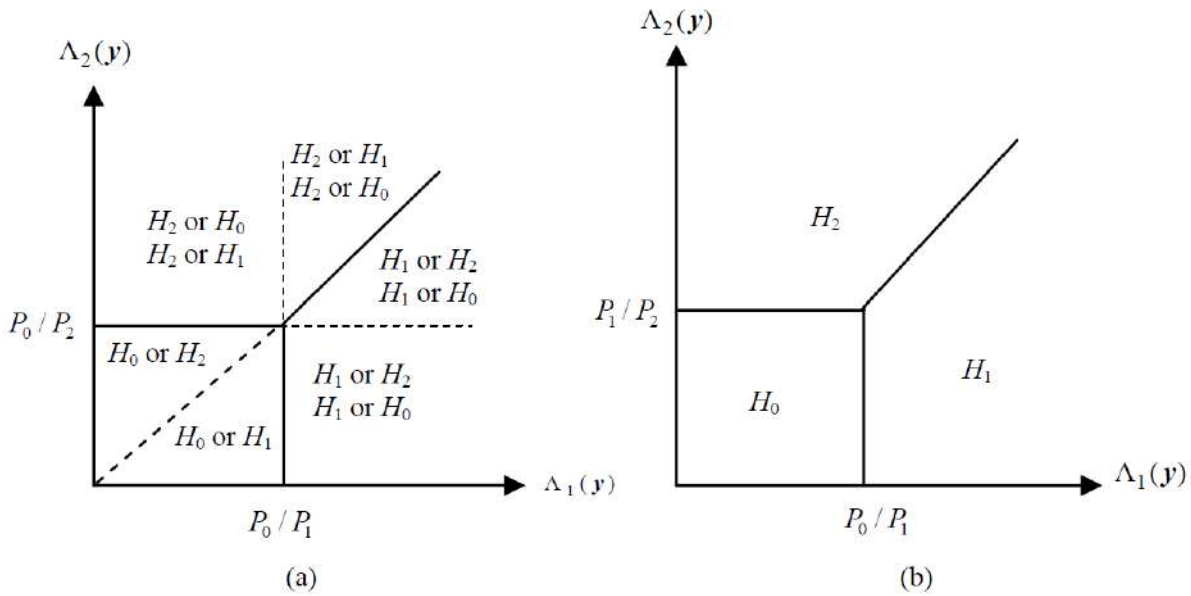


Figure II.12 : Escape de décisions pour M = 3 : (a) régions de décision résultante et (b) espace de décision global.

$$\Lambda_1(y) \underset{H_0 or\ H_2}{\overset{H_1 or\ H_2}{\gtrless}} \frac{P_0}{P_1} \tag{II.68}$$

$$\Lambda_2(y) \underset{H_0 or\ H_1}{\overset{H_1 or\ H_2}{\gtrless}} \frac{P_0}{P_2} \tag{II.69}$$

Et

$$\Lambda_2(y) \underset{H_0 or\ H_1}{\overset{H_0 or\ H_2}{\gtrless}} \frac{P_1}{P_2}\Lambda_1(y) \tag{II.70}$$

L'espace de décision dans le plan ln $\Lambda 1\ (\ y) - ln\ \Lambda 2\ (\ y) -$ est illustré à la Figure II.6. Nous observons que l'espace de décision est maintenant constitué de l'ensemble du plan.

De plus, en remplaçant (II.54), (II.55) dans (II.65), (II.66), (II.67), en divisant par $fY(\ y)$ et en utilisant$P(A \mid B)P(B)$, on obtient la règle de décision suivante

$$\ln\ \Lambda_2(y) \underset{H_0 or\ H_2}{\overset{H_1 or\ H_2}{\gtrless}} \ln\frac{P_1}{P_2} \tag{II.71}$$

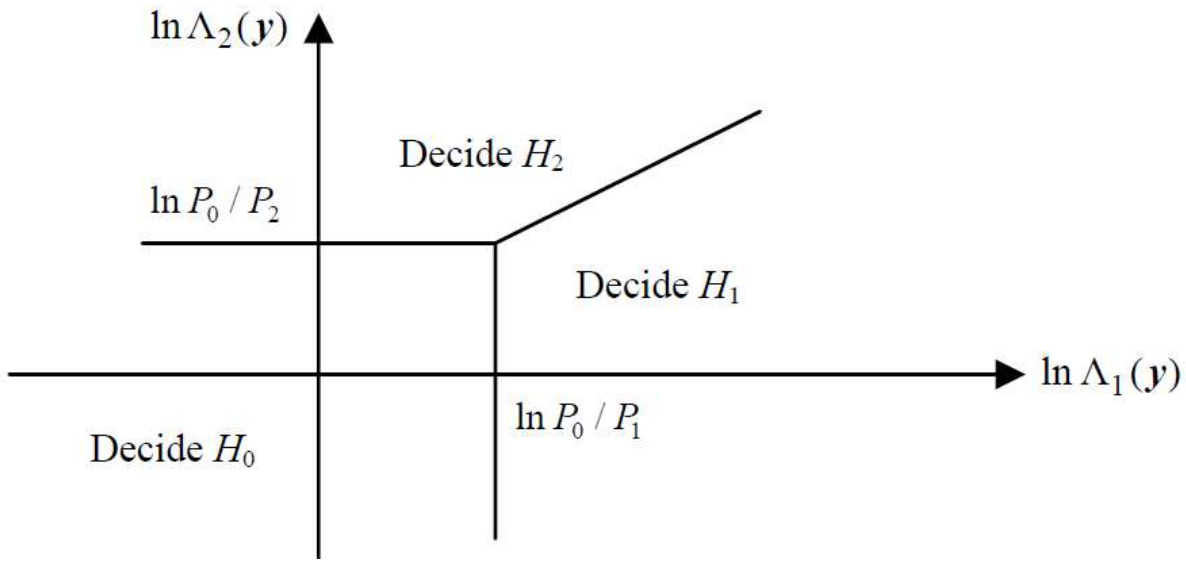


Figure II.13 : Espace de décision utilisant le logarithme pour M = 3.

$$\mathrm{P}(\mathrm{H}_2|\mathrm{y}) \underset{\mathrm{H}_0 \text{or } \mathrm{H}_1}{\overset{\mathrm{H}_1 \text{or } \mathrm{H}_2}{\gtrless}} \mathrm{P}(\mathrm{H}_0|\mathrm{y}) \tag{II.72}$$

Et

$$P(H_2|y) \underset{H_0 \text{or } H_1}{\overset{H_0 \text{or } H_2}{\gtrless}} P(H_1|y) \tag{II.73}$$

II.2.2 CRITÈRE MINIMAX

Le critère de Bayes attribue des coûts aux décisions et suppose la connaissance de la probabilité a priori. Dans de nombreuses situations, il se peut que nous n'ayons pas assez d'informations sur les probabilités a priori et par conséquent, le critère de Bayes ne peut pas être utilisé.

Une approche serait de sélectionner une valeur de P_1, la probabilité a priori de H_1, pourlaquelle le risque est maximal, puis minimiser cette fonction de risque. Ce principe de minimiser le coût moyen maximum pour le P_1 sélectionné est appelé minimax critère.

D'après (II.2), on a

$$P_0 = 1 - P_1 \tag{II.74}$$

En substituant (II.2) dans (II.16), nous obtenons la fonction de risque en termes de P_1 comme

$$R = C_{00}(1 - P_F) + C_{10}P_F + P_1[(C_{11} - C_{00}) + (C_{01} - C_{11})P_M - (C_{10} - C_{00})P_F] \tag{II.75}$$

En supposant une valeur fixe de P_1, $P_1 \in [0,1]$, nous pouvons concevoir un test de Bayes. Celles-ciles régions de décision sont alors déterminées, ainsi que les probabilités de fausse alarme, P_F,et d'oubli, P_M. Le test donne

$$\Lambda(y) \underset{H_0}{\overset{H_1}{\gtrless}} \frac{(1-P_1)(C_{10}-C_{00})}{P_1(C_{01}-C_{11})} \tag{II.76}$$

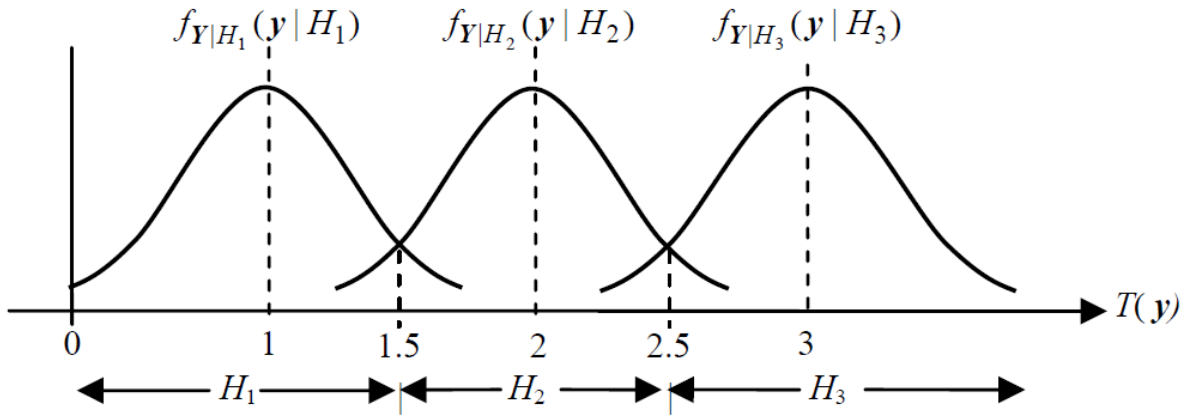


Figure II.14 : Fonctions de densité conditionnelle et régions de décision.

Comme P_1 varie, les régions de décision changent, ce qui entraîne une règle de décision non optimale. Cela entraîne à son tour une variation du coût moyen, qui serait plus important que le Les coûts de Bayes. Les deux cas extrêmes sont lorsque P_1 est égal à zéro ou un. Si P_1 est égal à zéro, alors le seuil est infini et la règle de décision est

$$\Lambda(y) \underset{H_0}{\overset{H_1}{\gtrless}} \infty \tag{II.77}$$

H_0 est toujours vrai. L'observation est Z_0, et la probabilité résultante de faux l'alarme et la probabilité d'oubli sont

$$P_F = \int_{Z_1} f_Y|_{H_2}(y|H_2)]dy = 0 \tag{II.78}$$

Et

$$P_M = \int_{Z_0} f_Y|_{H_1}(y|H_1)]dy = 1 \tag{II.79}$$

En substituant les valeurs de P_1 , P_F et P_M dans (II.75), nous obtenons que le risque est :

$$R = C_{00} \tag{II.80}$$

De même, lorsque $P_1 = 1$, le seuil de (II.76) est nul et la nouvelle règle de décision est

$$\Lambda(y) \underset{H_0}{\overset{H_1}{\gtrless}} 0 \tag{II.81}$$

Puisque $\Lambda(y)$ est non négatif, nous décidons toujours H_1. Par conséquent, P_F = 1 et P_M = 0. Le risque qui en résulte est

$$R = C_{11} \tag{II.82}$$

Si $P_1 = P_1^*$ tel que $P1 * \in (0,1)$, alors le risque en fonction de P_1 est comme indiqué dans Figure II.8. D'après (II.75), on voit que le risque est linéaire en fonction de P_1, et le test de Bayes pour $P_1 = P_1^*$ donne le risque minimum min. La tangente à min est

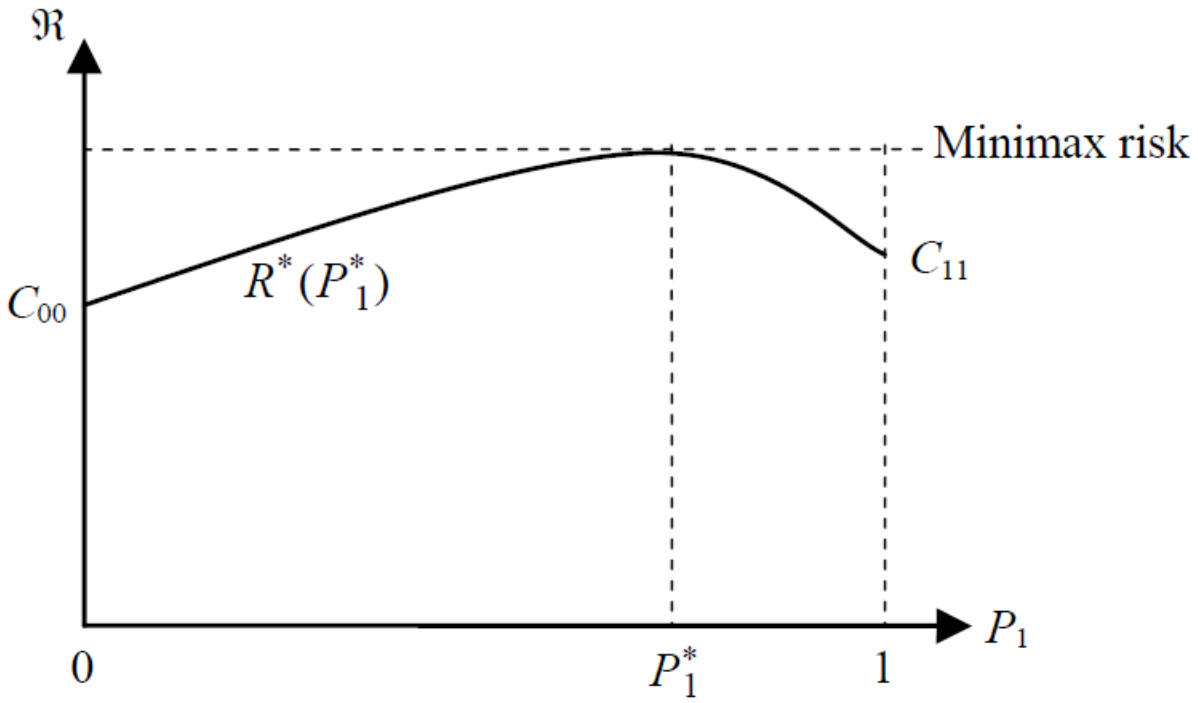


Figure II.15 : Risque en fonction de P1.

Horizontale, et $\mathfrak{R}^*$ ($P1$) à $P1 = P1^*$ représente le coût maximum. Observe cecila courbe de Bayes doit être concave vers le bas. Ainsi, le coût moyen ne sera pas dépassé $R^*(P_1^*)$. Prendre la dérivée de par rapport à P_1 et la mettre égaleà zéro, on obtient l'équation minimax à être

$$(C_{11} - C_{00}) + (C_{01} - C_{11})P_M - (C_{10} - C_{00})P_F + 0 \tag{II.83}$$

Si le coût d'une décision correcte est nul (C00 = C11 = 0), alors l'équation minimax pour $P_1 = P_1$ se réduit à

$$C_{01}P_M = C_{10}P_F \tag{II.84}$$

De plus, si le coût d'une mauvaise décision est d'un (C01 = C10 = 1), alors la probabilité de fausse alarme est égale à la probabilité d'oubli. C'est-à-dire,

$$\mathrm{P_F} = \mathrm{P_M} \tag{II.85}$$

Et le coût minimax est

$$R = P_F(1 - P_1) + P_1P_M = P_0P_F + P_1P_M \tag{II.86}$$

II.2.3 CRITÈRE DE NEYMAN-PEARSON

Dans le début du chapitre, nous avons vu que pour le critère de Bayes nous exigeons connaissance des probabilités a priori et des affectations de coûts pour chaque décision. Ensuite, nous avons étudié le critère minimax, qui est utile dans les situations où la connaissance des probabilités a priori n'est pas possible. Dans beaucoup d'autres situations physiques, telles que la détection radar, il est très difficile d'attribuer des coûts et probabilités a priori. Pour surmonter cette difficulté, nous utilisons les probabilités conditionnelles de fausse alarme, P_F et détection P_D. Le test de Neyman-Pearson exige que P_F soit fixé à une certaine valeur tandis que P_D est maximisé.

Puisque $P_M = 1 - P_D$, maximiser PD équivaut à minimiser P_M.

Pour minimiser PM (maximiser P_D) sous la contrainte que $P_F = \alpha$,nous utilisons le calcul des extrema et formons la fonction objective J pour être

$$J = P_M + \lambda(P_F - \alpha) \tag{II.87}$$

Où ($\lambda \geq 0$) est le multiplicateur de Lagrange. Notons qu'étant donné l'observation dans l'espace Z, il existe de nombreuses régions de décision Z_1 pour lesquelles $P_F = \alpha$. La question est de déterminer les régions de décision pour lesquelles la P est minimale. En conséquence, nous réécrirons la fonction objective J en fonction de la région de décision pour obtenir

$$J = \int_{Z_0} f_Y|_{H_i}(y|H_1)dy + \lambda[(\int_{Z_1} f_Y|_{H_0}(y|H_0)dy - a] \tag{II.88}$$

L'utilisation de (II.1), (II.88) peut être réécrite comme

$$J = \int_{Z_0} f_Y|_{H_i}(y|H_1)dy + \lambda[(\int_{Z_1} f_Y|_{H_0}(y|H_0)dy - a] \tag{II.89}$$

$$= \lambda(1 - a) + \int_{Z_0} [\,f_Y|_{H_i}(y|H_1)dy - \lambda\, f_{\ Y}|_{H_0}(y|H_0)]dy$$

Par conséquent, J est minimisé lorsque les valeurs pour lesquelles $f_{\ Y}|_{H_1}(y|H_1)\ f_{\ Y}|_{H_0}(y|H_0)$ sont affecté à la région de décision Z_1. La règle de décision est donc

$$\Lambda(y)\frac{f_{\ Y|H_i}(y|H_1)}{f_{\ Y|H_0}(y|H_0)} \begin{matrix} H_1 \\ > \\ < \\ H_0 \end{matrix} \lambda \tag{II.90}$$

Si nous définissons la densité conditionnelle d'étant donné que H_0 est vrai car $f_{\Lambda|_{H_0}}(\lambda|H_0)$, alors $P_F = \alpha$ peut être réécrit comme

$$P_F = \int_{Z_1} f_{Y|H_i}(y|H_1)dy = \int_{\lambda}^{\infty} f_{\Lambda_Y}|_{H0}[\lambda(y)|H_0]dy \quad \text{(II.91)}$$

II.2.4 TEST D'HYPOTHÈSE COMPOSITE

Dans le problème simple de test d'hypothèses considéré précédemment, les paramètres caractérisant une hypothèse étaient tous connus. Dans de nombreuses situations, les paramètres caractérisant une hypothèse peut ne pas être connue. Dans ce cas, l'hypothèse est appelée hypothèse composite.

II.2.5 **DÉTECTION SÉQUENTIELLE**

Dans les sections précédentes, nous avons considéré la théorie des tests d'hypothèses, telle que le nombre d'observations sur lequel le test était basé était fixé. Dans de nombreux situations pratiques, les observations peuvent être effectuées de manière séquentielle afin que le test est effectué après chaque observation. Chaque fois qu'une observation est effectuée, l'un des les trois décisions possibles sont prises :

1. Décidez H_1
2. Décidez H_0
3. Pas assez d'informations pour décider en faveur de H_1 ou H_0.

Si les décisions (1) ou (2) sont prises, la procédure de test d'hypothèse s'arrête. Sinon, une observation supplémentaire est effectuée et le test est à nouveau effectué. Ce processus est pour suivi jusqu'à ce qu'une décision soit prise en faveur de H_1 ou H_0 . Notez que le nombre d'observation K n'est pas fixe, mais est une variable aléatoire.

Le test à effectuer après chaque observation consiste à calculer le rapport vraisemblance et le comparer à deux seuils, η_0 et η_1 . Un tel test qui fait l'un des trois décisions possibles mentionnées ci-dessus après le renvoi de la $kième$ observation comme test de rapport de vraisemblance séquentiel.

Soit $Y_k, k = 1, 2, \ldots, K$, représenter le kième échantillon d'observation du vecteur Y_K défini comme

$$Y_K^T = [Y_1 \;\; Y_2 \;\; \ldots Y_k] \quad \text{(II.92)}$$

Le rapport de vraisemblance basé sur les K premières observations est

$$\Lambda(y_K) = \frac{f_Y|_{H_1}(y|H_1)}{f_Y|_{H_1}(y|H_0)} \quad \text{(II.93)}$$

Pour calculer le rapport de vraisemblance de (II.93), nous devons connaître la fonction densité conjointe de ces K observations. Pour simplifier, nous supposons que les observations sont distribuées de manière identique et sont pris indépendamment. Le rapport de vraisemblance peut être écrit comme un produit de K rapports de vraisemblance pour obtenir

$$\Lambda\,(\mathrm{y_K}) = \frac{\mathrm{f_{Y|H_1}(y|H_1)}}{\mathrm{f_{Y|H_1}(y|H_0)}} = \prod_{\mathrm{k=1}}^{\mathrm{K}} \frac{\mathrm{f_{Y|H_1}(y|H_1)}}{\mathrm{f_{Y|H_1}(y|H_0)}} \tag{II.94}$$

Le but est de déterminer η_0 et η_1 en termes de P_F, la probabilité de fausse alarme, et P_M, la probabilité d'oubli. Nous fixons

$$\mathrm{P_F} = \mathrm{a} \tag{II.95}$$

Et

$$\mathrm{P_M} = \beta \tag{II.96}$$

Et effectuez le test suivant. Si

$$\Lambda\,(\mathrm{y_K}) \leq \eta_1 \tag{II.97}$$

Nous décidons en faveur de H_1. Si

$$\Lambda\,(\mathrm{y_K}) \leq \eta_0 \tag{II.98}$$

Nous décidons en faveur de H_0. Sinon, si

$$\eta_0 < \quad \Lambda\,(y_K) < \eta_1 \tag{II.99}$$

Nous prenons une observation supplémentaire et effectuons un autre test. La probabilité de détection, P_D, en termes d'intégrale sur l'espace d'observation est

$$\mathrm{P_D} = \mathrm{P(decide\ H_1\ |H_1\ true)} = \int_{\mathrm{Z_1}} \mathrm{f}_{\ \mathrm{Y_K|H_1}}(\mathrm{y}_{\ \mathrm{K}}\,|\mathrm{H_1})\mathrm{dy}_{\ \mathrm{k}} \tag{II.100}$$

En utilisant (II.93), P_D peut être écrit comme

$$\mathrm{P_D} = \int_{\mathrm{Z_1}} \Lambda\,(\mathrm{y_K})\,\mathrm{f_Y|_{H_0}}(\mathrm{y_K}|\mathrm{H_0})\mathrm{dy}_{\ \mathrm{K}} \tag{II.101}$$

La décision en faveur de H_1 signifie que $\Lambda(\,y\,K\,) \geq \eta_1$. Par conséquent, en substituant (II.97) à (II.101), on obtient l'inégalité

$$\mathrm{P_D} \geq \eta_1 \mathrm{f_Y|_{H_0}}(\mathrm{y_K}|\mathrm{H_0})\mathrm{dy}_{\ \mathrm{K}} \tag{II.102}$$

Notons que l'intégrale

$$\int_{Z_1} f_Y|_{H_0}(y_K|H_0)dy_{\ K} = P_F = a \tag{II.103}$$

Et puisque $P_D = 1 - PM = 1 - \beta$, (II.103) se réduit à

$$1 - \beta \geq \eta_0 \alpha \tag{II.104}$$

Ou, le seuil η_1 est

$$\eta_1 \leq \frac{1-\beta}{\alpha} \tag{II.105}$$

De même, on peut montrer que le seuil η_0 est

$$\eta_0 \geq \frac{\beta}{1-\alpha} \tag{II.106}$$

II.3 Les Paramètres d'Estimation

II.3.1 ESTIMATION MAXIMALE DE PROBABILITE

Comme mentionné dans la fonction précédente, la procédure couramment utilisée pour estimer les paramètres non aléatoires est l'estimation du maximum de vraisemblance (ML). Soient $Y1, Y2, \ldots, YK$ être K observations de la variable aléatoire Y, avec des valeurs d'échantillon$1, y2, \ldots, yk$. Ces variables aléatoires sont indépendantes et identiquement distribuées. Soit $f_Y|_H\ (y \mid \theta)$ la fonction de densité conditionnelle de la variable Y. Notez que la fonction de densité de Y dépend du paramètre θ, qui doit être estimé. La fonction de vraisemblance, L(θ), est

$$L(\theta) = f_{Y_1,\ldots Y_K|\Theta}(y_1, y_2, \cdots, y_K|\theta) = f_Y|_\Theta(y|\theta) = \prod_{k=1}^{K} f_Y|\Theta\ .\ (y\ _K\ |\theta) \tag{II.107}$$

La valeur $\widehat{\theta}$ qui maximise la fonction de vraisemblance est appelée le maximum estimateur de vraisemblance de θ. Afin de maximiser la fonction de vraisemblance, la norme des techniques de calcul peuvent être utilisées. Parce que la fonction logarithmique ln x est une fonction croissante monotone de x, maximisant $L(\theta)$ équivaut à maximiser $L(\theta)$. Ainsi, on peut montrer qu'une condition nécessaire mais non suffisante pour obtenir l'estimation ML est de résoudre l'équation de vraisemblance.

$$\frac{\partial}{\partial\theta} \ln f_Y|\Theta\ (y|\Theta) = 0 \tag{II.108}$$

II.3.2 TEST GÉNÉRALISÉ DU RATIO DE PROBABILITÉ

Nous avons résolu le problème de test d'hypothèse où l'alternative l'hypothèse était composite. Le paramètre m sous l'hypothèse H_1 était inconnu, bien que l'on sache que m était soit positif, soit négatif. Quand j'étais positif seulement (négatif seulement), un test UMP existait et la règle de décision était

$$y \underset{H_0}{\overset{H_1}{\gtrless}} \frac{\sigma^2}{m} \ln \eta + \frac{m}{2} = y_1 \tag{II.109}$$

Pour m positif, et

$$\frac{f_Y|_{H_1}(y|H_1)}{f_Y|_{H_0}(y|H_0)} \underset{H_1}{\overset{H_0}{\gtrless}} \frac{P_0(C_{10}-C_{00})}{P_1(C_{01}-C_{11})} \tag{II.110}$$

Pour m négatif. Étant donné que le test conçu pour un m positif n'était pas le même que le test conçu pour m négatif, nous avons conclu qu'un test UMP n'existait pas pour toutes les valeurs possibles de m; c'est-à-dire positif et négatif. Cela nécessite que différents tests être utilisé. C'est-à-dire qu'utiliser les données requises pour estimer, comme si l'hypothèse H_1 était vraie. Ensuite, on utiliser ces estimations dans le test du rapport de vraisemblance comme s'il s'agissait des valeurs correctes. Là il existe de nombreuses façons d'estimer θ, comme nous le montrerons dans cette section. Si les estimations utilisées sont les estimations du maximum de vraisemblance, alors le résultat est appelé le test du rapport de vraisemblance généralisé et est donné par

$$\Lambda_g(y) = \frac{\theta_1^{max} f_{Y|\Theta}(y|\theta_1)}{\theta_0^{max} f_{Y|\Theta}(y|\theta_1)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta \tag{II.111}$$

$\theta_1$et θ_0 sont les paramètres inconnus à estimer sous les hypothèses H_1et H_0, respectivement.

II.3.3 **QUELQUES CRITÈRES POUR DE BON ESTIMATEURS**

Puisque l'estimateur θˆ est une variable aléatoire et peut prendre plus d'une valeur, certaines caractéristiques d'une « bonne » estimation doivent être déterminées.

Estimation sans biais Nous disons que $\widehat{\theta}$ est un estimateur sans biais pour θ si

$$E[\widehat{\theta}] = \theta \; for \; all \; \theta \tag{II.112}$$

Le biais de l'estimateur devient

$$E[\widehat{\theta}] = \theta + b(\theta) \tag{II.113}$$

1/ Si $b(\theta)$ ne dépend pas de$[b(\theta) = b]$, on dit que l'estimateur $\widehat{\theta}$ a un biais connu. Autrement dit,

$(\widehat{\theta} - b)$ est une estimation sans biais.

2/ Lorsque $(\theta) \neq b$, une estimation sans biais ne peut pas être obtenue, puisque θ est inconnu. Dans ce cas, on dit que l'estimateur a un biais inconnu.

Le fait que l'estimateur soit sans biais, ce qui signifie que la valeur moyenne de l'estimation est proche de la vraie valeur, ne garantit pas nécessairement que l'estimateur est « bon ». Ceci est facilement visible par la fonction de densité conditionnelle de l'estimateur illustré à la figure II.9. Nous observons que même si l'estimation est des erreurs importantes et non biaisées sont susceptibles de se produire, puisque la variance de l'estimation est grande. Cependant, si la

variance est faible, la variabilité de l'estimateur autour de sa valeur attendue est également faible. Par conséquent, la variabilité de l'estimateur est proche à la vraie valeur, puisque l'estimation est sans biais, ce qui est une caractéristique souhaitée. D'où, on dit que la deuxième mesure de la qualité de l'estimation est d'avoir une petite variance.

La variance minimale $\widehat{\theta}$ non biaisée est une variance minimale et non biaisée (MVU) estimation de θ si, pour toutes les estimations θ′ telles que $E[\theta'] = \theta$, on a $var[\theta] \leq \text{var}[\widehat{\theta}]$ pour tous.θ' C'est-à-dire que $\widehat{\theta}$ a la plus petite variance parmi toutes les estimations sans biais de θ.

L'estimation cohérente $\hat{\theta}$ est une estimation cohérente du paramètre, basée sur Kéchantillons observés, si

$$\lim_{K\to\infty} P(|\hat{\theta} - \boldsymbol{\theta}| > \varepsilon) = 0 \quad \text{for all } \varepsilon > 0 \tag{II.114}$$

Où $P(\cdot)$ désigne la probabilité.

L'application de la définition ci-dessus pour vérifier la cohérence d'une estimation n'est pas Facile. Le théorème suivant est utilisé à la place.

Théorème. Soit $\hat{\theta}$un estimateur sans biais de θ basé sur K échantillons observés. Si

$$\lim_{K\to\infty} E[\hat{\theta}] = \boldsymbol{\theta} \tag{II.115}$$

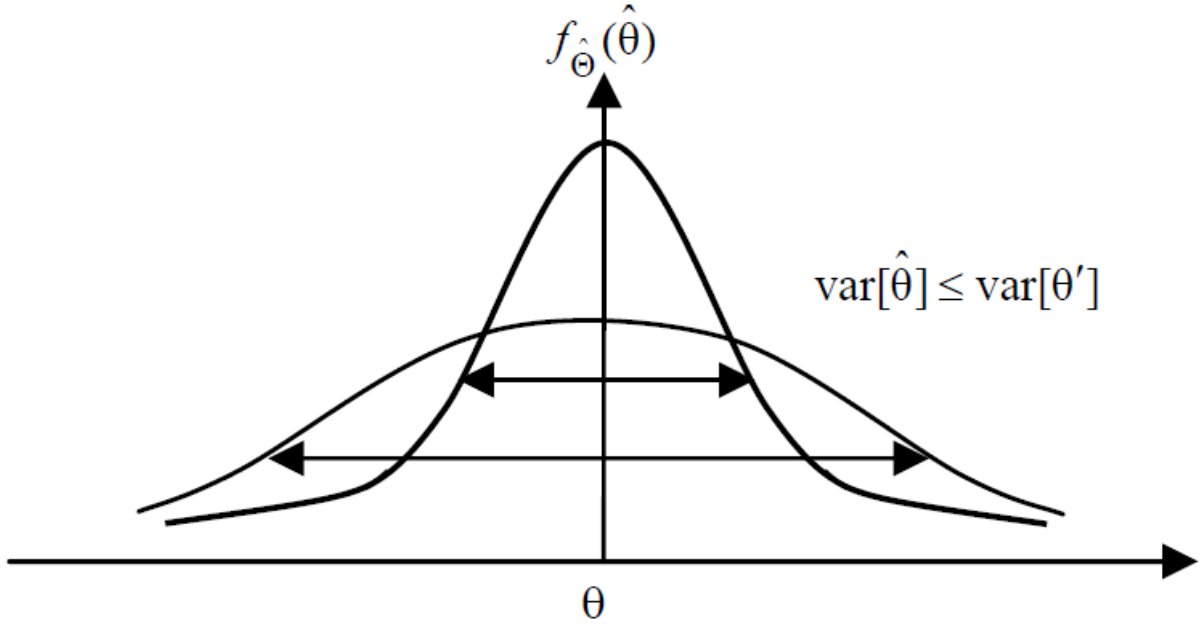


Figure II.16 : Fonction de densité de l'estimateur sans biais (θ).

$$\lim_{K\to\infty} \text{var}\, E[\hat{\theta}] = \boldsymbol{\theta} \tag{II.116}$$

Alors $\widehat{\theta}$ est un estimateur consistant de θ.

II.3.4 **ESTIMATION DE BAYES :**

Dans l'estimation de Bayes, nous attribuons un coût $C(\theta, \widehat{\theta})$ à toutes les paires $(\theta, \widehat{\theta})$ Le coût est unfonction valeur réelle non négative des deux variables aléatoires θ et $\widehat{\theta}$. Comme dans la Détection de Bayes, la fonction de risque est définie comme étant la valeur moyenne du coût c'est :

$$R = E[C(\theta, \widehat{\theta}) \tag{II.117}$$

Le but est de minimiser la fonction de risque afin d'obtenir l'estimation optimale. Dans de nombreux problèmes, seule l'erreur $\tilde{\theta}$ entre l'estimation et la vraie valeur est intéressante ; C'est :

$$\tilde{\theta} = \theta - \widehat{\theta} \tag{II.118}$$

Par conséquent, nous ne considérerons que les coûts qui sont fonction de l'erreur. Trois les cas seront étudiés, et leurs croquis correspondants sont montrés dans la Figure II.10.

1. Erreur au carré

$$c(\theta, \hat{\theta}) = (\theta, \hat{\theta})^2 \tag{II.119}$$

2. Valeur absolue de l'erreur

$$c(\theta, \hat{\theta}) = |\theta, \hat{\theta}| \tag{II.120}$$

3. Fonction de coût uniforme

$$c(\theta, \hat{\theta}) = \begin{cases} 1, |\theta, \hat{\theta}| \geq \frac{\Delta}{2} \\ 0, |\theta, \hat{\theta}| \geq \frac{\Delta}{2} \end{cases} \tag{II.121}$$

Le paramètre inconnu est supposé être une variable aléatoire continue avec fonction de densité $f_\theta(\theta)$ La fonction de risque peut alors être exprimée sous la forme

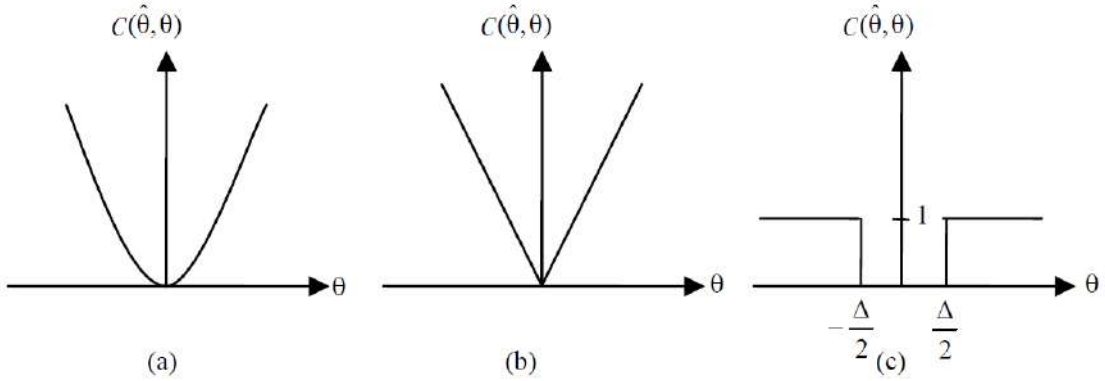


Figure II.17 : fonctions de couts: (a) squared error, (b) absolute value of error, and (c) uniform.

$$R = E[C(\theta - \hat{\theta}) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} c(\theta - \hat{\theta})^2 f_{\Theta}|_Y(\theta|y) d\theta dy \quad \text{(II.122)}$$

Notez que nous prenons la moyenne des coûts sur toutes les valeurs possibles de θ et Y, où Y est le vecteur $[Y_1 Y_2 \ ... \ Y_k]$ Nous trouvons maintenant l'estimateur pour les trois coûts fonctions considérées.

II.3.4.1 Estimation de l'erreur quadratique moyenne minimale :

L'estimateur qui minimise la fonction de risque pour le coût donné en (II.119) est appelée estimation de la moyenne quadratique minimale (MMSE). Le risque correspondant fonction est notée R_{ms}On a

$$R_{ms} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\theta - \hat{\theta})^2 f_{\Theta}|_Y(\theta|y) d\theta = \int_{-\infty}^{\infty} d\theta \int_{-\infty}^{\infty} (\theta - \hat{\theta})^2 f_{\Theta}|_Y(\theta|y) d\theta dy \quad \text{(II.123)}$$

$$f_y(y|x) = \frac{f_{xy}(x,y)}{f_x(x)} \quad \text{(II.124)}$$

En utilisant (II.124), la fonction de risque peut être réécrite comme

$$R_{ms} = \int_{-\infty}^{\infty} dy f_Y(y) \left[\int_{-\infty}^{\infty} (\theta - \hat{\theta})^2 f_{\Theta}|_Y(\theta|y) d\theta \right] \quad \text{(II.125)}$$

Puisque la fonction de densité $f_y(y)$est non négative, minimiser $\Re_{ms}$ est équivalent à minimiser l'expression entre parenthèses de l'équation ci-dessus. Par conséquent, en prenant la dérivée par rapport à $\hat{\theta}$ et en la fixant égale à zéro, on a

$$\frac{d}{d\hat{\theta}} \int_{-\infty}^{\infty} (\theta - \hat{\theta})^2 f_{\Theta}|_Y(\theta|y) d\theta = 0 \quad \text{(II.126)}$$

$$\frac{d}{d_x} \int_{a(x)}^{b(x)} F(u,x) du = \int_{a(x)}^{b(x)} \frac{\partial F}{\partial x} du + F[b(x),x] \frac{db(x)}{dx} - F[a(x),x] \frac{da(x)}{dx} \quad \text{(II.127)}$$

En utilisant la règle de Leibniz donnée en (II.127), on obtient

$$\hat{\theta}_{ms} = \int_{-\infty}^{\infty} f_{\Theta}|_Y(\theta|y) d\theta = E[(\theta|y)] \quad \text{(II.128)}$$

Autrement dit, l'estimation quadratique moyenne minimale $\hat{\theta}_{ms}$représente la moyenne conditionnelle de θ étant donné Y. On peut facilement montrer que la dérivée seconde par rapport à ms θ est définie positivement, ce qui correspond à un unique minimum de $\Re_{ms}$, et est donnée par

$$R_{ms} = \int_{-\infty}^{\infty} f_Y(y) \int_{-\infty}^{\infty} (\theta - \hat{\theta}_{ms})^2 f_{\Theta}|_Y(\theta|y) d\theta \quad \text{(II.129)}$$

$$= \int_{-\infty}^{\infty} \mathrm{dy} f_{\mathrm{Y}}(\mathrm{y}) \int_{-\infty}^{\infty} \{\theta - \mathrm{E}[\theta|\mathrm{y}]\}^2 \, \mathrm{f}_{\Theta}|_{\mathrm{Y}}(\theta|\mathrm{y}) \mathrm{d}\theta$$

La variance conditionnelle de θ étant donné Y est

$$var[\theta|y] = \int_{-\infty}^{\infty} \{\theta - E[\theta|y]\}^2 f_{\Theta}|_{y} \, d\theta \tag{II.130}$$

Par conséquent, $\Re_{ms}$, est juste la variance conditionnelle de θ étant donné Y, moyennée sur tout valeurs possibles de Y. Cette procédure d'estimation utilisant le critère d'erreur quadratique est parfois appelée variance minimale (MV) de l'estimation de l'erreur.

II.3.4**.2 Valeur absolue moyenne minimale de l'estimation de l'erreur :**

Dans ce cas, la fonction de coût est donnée par (II.120), et le risque est

$$\mathrm{R}_{\mathrm{abs}} = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |\theta - \widehat{\theta}| \mathrm{f}_{\Theta}|_{\mathrm{Y}}(\theta|\mathrm{y}) \mathrm{d}\theta = \int_{-\infty}^{\infty} \mathrm{f}_{\mathrm{Y}}(\mathrm{y}) \left[\int_{-\infty}^{\infty} |\theta - \widehat{\theta}| \mathrm{f}_{\Theta}|_{\mathrm{Y}}(\theta|\mathrm{y}) \mathrm{d}\theta \right] \mathrm{dy} \tag{II.131}$$

Le risque peut être minimisé en minimisant l'intégrale entre parenthèses, qui est donnée par

$$\int_{-\infty}^{\widehat{\theta}} \left(\widehat{\theta} - \theta\right) \mathrm{f}_{\Theta}|_{\mathrm{Y}}(\theta|\mathrm{y}) \mathrm{d}\theta + \int_{\widehat{\theta}}^{\infty} \left(\widehat{\theta} - \theta\right) \mathrm{f}_{\Theta}|_{\mathrm{Y}}(\theta|\mathrm{y}) \mathrm{d}\theta \tag{II.132}$$

En différenciant (II.130) par rapport à $\widehat{\theta}$, et en fixant le résultat égal à zéro, nous obtenir

$$\int_{-}^{\widehat{\theta}_{\mathrm{abs}}} \mathrm{f}_{\Theta}|_{\mathrm{Y}}(\theta|\mathrm{y}) \mathrm{d}\theta = \int_{\widehat{\theta}_{\mathrm{abs}}}^{\infty} \mathrm{f}_{\Theta}|_{\mathrm{Y}}(\theta|\mathrm{y}) \mathrm{d}\theta \tag{II.133}$$

Autrement dit, l'estimation $\widehat{\theta}_{abs}$ n'est que la médiane de la fonction de densité conditionnelle $f_{\Theta}|_{y}(\theta|y)$. Cette estimation est également connue sous le nom de valeur absolue moyenne minimale d'erreur (MAVE) estimation, et donc $\widehat{\theta}_{abs} = \widehat{\theta}_{mave}$.

II.3.4.3 **Estimation maximale a posteriori :**

Pour la fonction de coût uniforme donnée par (II.121), le risque de Bayes devient

$$R_{unf} = \int_{-\infty}^{\infty} f_{\ Y}(y) \left[\int_{-\infty}^{\widehat{\theta} + \frac{\Delta}{2}} \partial f_{\Theta}|_{Y}(\theta|y) d\theta + \int_{\widehat{\theta} + \frac{\Delta}{2}}^{\infty} \partial f_{\Theta}|_{Y}(\theta|y) d\theta \right] dy \tag{II.134}$$

$$R_{unf} = \int_{-\infty}^{\infty} f_{Y}(y) \left[1 - \int_{\widehat{\theta} - \frac{\Delta}{2}}^{\widehat{\theta} + \frac{\Delta}{2}} \partial f_{\Theta}|_{Y}(\theta|y) d\theta \right] dy$$

Où

$$\int_{\widehat{\theta} - \frac{\Delta}{2}}^{\widehat{\theta} + \frac{\Delta}{2}} \partial \mathrm{f}_{\Theta}|_{\mathrm{Y}}(\theta|\mathrm{y}) \mathrm{d}\theta = \mathrm{P}\left[\widehat{\theta} - \frac{\Delta}{2} \leq \Theta \leq \widehat{\theta} + \frac{\Delta}{2} | \mathrm{y} \right] \tag{II.135}$$

$P[\cdot]$ Désigne la probabilité. Par conséquent, le risque $\Re_{unf}$ est minimisé en maximisant (II.133). Notons qu'en maximisant (II.133) (en minimisant $\Re_{unf}$, on cherche l'estimation $\widehat{\theta}$,

qui minimise $f_{\Theta}|_{Y}(\theta \mid y)$. C'est ce qu'on appelle le maximum a estimation a posteriori (MAP), $\widehat{\theta}$ qui est définie comme

$$\frac{\partial f_{\Theta}|_{Y}(\theta|y)}{\partial \theta}|_{\theta=\hat{\theta}_{map}} = 0 \tag{II.136}$$

En utilisant le logarithme, qui est une fonction monotone croissante, (II.134) devient

$$\frac{\partial \ln f_{Y}|_{\Theta}(y|\theta)}{\partial \theta} = 0 \tag{II.137}$$

L'équation (II.135) est appelée l'équation MAP. C'est une nécessité mais pas suffisante condition, puisque $f\Theta|Y\ (\theta \mid y)$ peut avoir plusieurs maximaux locaux. Utiliser la règle de Bayes

$$f_{Y}|_{\Theta}(y|\theta) = \frac{\partial \ln f_{Y}|_{\Theta}(y|\theta)}{f_{Y}(y)} \tag{II.138}$$

Et le fait que

$$\ln f_{Y}|_{\Theta}(y|\theta) = \ln f_{Y}|_{\Theta}(y|\theta) + \ln f_{\Theta}(\theta) - \ln f_{Y}(y) \tag{II.139}$$

Alors l'équation MAP peut être réécrite comme

$$\frac{\partial \ln f_{Y}|_{\Theta}(y|\theta)}{\partial \theta} = \frac{\partial \ln f_{Y}|_{\Theta}(y|\theta)}{\partial \theta} + \frac{\partial \ln f_{Y}|_{\Theta}}{\partial \theta} = 0 \tag{II.140}$$

Nous supposons toujours qu'est suffisamment petit, de sorte que la carte d'estimation est donnée par l'équation MAP. C'est-à-dire que la fonction de coût illustrée à la figure II.10 peut être défini comme

$$C(\widehat{\theta} - \theta) = 1 - \delta(\theta - \widehat{\theta}) \tag{II.141}$$

II.3.5 **INÉGALITÉ CRAMER-RAO :**

A partir de l'équation MAP de (II.138), si nous fixons la fonction de densité de à zéro, pour tout, on obtient l'équation de vraisemblance de (II.108). C'est-à-dire que l'estimation ML peut être considéré comme un cas particulier de l'estimation MAP. Dans ce cas, pour vérifier si l'estimation est « bonne », nous devons calculer sa variance de biais et d'erreur et déterminer sa consistance. Il peut être très difficile d'obtenir une expression pour l'écart d'erreur. Dans ce cas, la « bonté » de l'estimateur est étudiée en termes de borne inférieure de la variance d'erreur. Cette borne est connue sous le nom de borne Cramer-Rao. La borne de Cramer-Rao d'un paramètre constant est donnée par le suit théorème qui suit.

Théorème. Soit le vecteur $Y = [Y_1, Y_2, \dots, Y_k]^T$représentent K observations, et $\widehat{\theta}$ soit l'estimateur sans biais de θ. Puis

$$var[(\widehat{\theta}-\theta)|\theta] \geq \frac{1}{E\left\{\left[\frac{\partial \ln f_{Y}|_{\Theta}(y|\theta)}{\partial\theta}\right]^{2}\right\}} \tag{II.142}$$

Où

$$\mathrm{E}\left\{\left[\frac{\partial \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta}\right]^{2} = -\mathrm{E}\left[\frac{\partial^{2} \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta}\right]\right\} \tag{II.143}$$

Preuve. Pour un estimateur sans biais $\widehat{\theta}$, nous avons

$$\mathrm{E} = \left[\left(\widehat{\theta}-\theta\right)\right] = \theta \tag{II.144}$$

Par conséquent,

$$\mathrm{E}\left[\left(\widehat{\theta}-\theta\right)^{2}\middle|\theta\right] = \int_{-\infty}^{\infty}\left(\widehat{\theta}-\theta\right)\mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)\mathrm{dy} = 0 \tag{II.145}$$

En différenciant (II.143) par rapport à θ, on obtient

$$\int_{-\infty}^{\infty}\left(\widehat{\theta}-\theta\right)^{2}\frac{\partial\ \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta}\mathrm{dy} = \int_{-\infty}^{\infty}\mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)\mathrm{dy} \tag{II.146}$$

La seconde intégrale est égale à un. Utilisant le fait que

$$\frac{\partial \ln \mathrm{g(x)}}{\partial \mathrm{x}} = \frac{1}{\mathrm{g(x)}}\frac{\partial \mathrm{g(x)}}{\partial \mathrm{x}} \tag{II.147}$$

Où g(x) est une fonction de x, on peut exprimer $\partial f_{Y}|_{\Theta}\ (y\mid\theta)\ /\partial\theta$ comme

$$\frac{\partial \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta} = \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)\frac{\partial \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta} \tag{II.148}$$

En substituant (II.146) dans (II.144), on obtient

$$\int_{-\infty}^{\infty}\left(\widehat{\theta}-\theta\right)^{2}\mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)\frac{\partial \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta}\mathrm{dy} = 1 \tag{II.149}$$

L'inégalité de Schwarz indique que

$$\left[\int_{-\infty}^{\infty}\mathrm{x}^{2}(\mathrm{t})\mathrm{dt}\right]\left[\int_{-\infty}^{\infty}\mathrm{y}^{2}(\mathrm{t})\mathrm{dt}\right] \geq \left[\int_{-\infty}^{\infty}\mathrm{x(t)y(t)}\ (\mathrm{t})\mathrm{dt}\right] \tag{II.150}$$

Où x(t) et y(t) sont deux fonctions de t. L'égalité est vraie si et seulement si $y(t)\ =\ c_{x}(t)$, avec (c) une constante. Réécriture (II.146) afin d'utiliser le Schwarz inégalité, on a

$$\int_{-\infty}^{\infty}\left[\frac{\partial \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta}\sqrt{\mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}\right]\left[\left(\widehat{\theta}-\theta\right)^{2}\sqrt{\mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}\,\mathrm{dy}\right] = 1 \tag{II.151}$$

Où

$$\left[\int_{-\infty}^{\infty}\left(\widehat{\theta}-\theta\right)^{2}\sqrt{\mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}\ \mathrm{dy}\right]\left\{\int_{-\infty}^{\infty}\left[\frac{\partial \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta}\right]^{2}\mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)\mathrm{dy}\right\} \geq 1 \tag{II.152}$$

La première intégrale entre parenthèses est en fait $var\left[\left(\hat{\theta}-\ \theta\right)\middle|\theta\right]$. D'où le l'inégalité devient

$$\mathrm{var}[(\widehat{\theta}-\theta)|\theta] \geq \frac{1}{\mathrm{E}\left\{\left[\frac{\partial \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta}\right]^2\right\}} \tag{II.153}$$

Ce qui prouve (II.140).

Nous prouvons maintenant (II.141), qui dit que la borne de Cramer-Rao peut être exprimée sous une forme différente. Nous savons que

$$\int_{-\infty}^{\infty} \mathrm{f}_{\mathrm{Y}}|_{\theta}(\mathrm{y}|\theta) = 1 \tag{II.154}$$

Différencier les deux côtés de l'équation par rapport à θ aboutit à

$$\int_{-\infty}^{\infty} \frac{\partial \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta} = 0 \tag{II.155}$$

En réécrivant (II.153) et en utilisant (II.143), on a

$$\int_{-\infty}^{\infty} \frac{\partial \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta} \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)\mathrm{dy} = 0 \tag{II.156}$$

En se différenciant à nouveau par rapport à θ, on obtient

$$\int_{-\infty}^{\infty} \frac{\partial^2 \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta^2} \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)\mathrm{dy} + \int_{-\infty}^{\infty} \frac{\partial \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta} \frac{\partial\, \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta} = 0 \tag{II.157}$$

Substituer (II.154) pour le deuxième terme de la deuxième intégrale de (II.155), et réarrangement des termes rendements

$$\mathrm{E}\left[\frac{\partial^2 \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta^2}\right] = -\mathrm{E}\left\{\left[\frac{\partial \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta}\right]^2\right\} \tag{II.158}$$

Ce qui est le même que (II.141), et la preuve du théorème est complète.

Une observation importante concernant (II.152) est que l'égalité est vraie si et seulement si

$$\frac{\partial \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta} = \mathrm{c}(\theta)\left[\hat{\theta} - \theta\right] \tag{II.159}$$

Tout estimateur sans biais qui satisfait l'égalité dans l'inégalité de Cramer-Rao de (II.140) est dit un estimateur efficace.

Si existe un estimateur efficace, on peut facilement montrer qu'il est égal au ML estimation. L'équation ML est donnée par

$$\frac{\partial \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta}|_{|_{\theta=\hat{\theta}_{\mathrm{ml}}}} = 0 \tag{II.160}$$

En utilisant (II.157), à condition qu'il existe une estimation efficace, nous avons

$$\frac{\partial \ln \mathrm{f}_{\mathrm{Y}}|_{\Theta}(\mathrm{y}|\theta)}{\partial\theta} = \mathrm{c}(\theta)\left[\hat{\theta} - \theta\right]|_{\theta=\hat{\theta}_{\mathrm{ml}}} \tag{II.161}$$

Qui est égal à zéro lorsque $\hat{\theta} = \hat{\theta}_{ml}$.

II.3.6 MEILLEUR ESTIMATEUR LINÉAIRE NON BIAISÉ

Dans de nombreux problèmes pratiques, il peut être impossible de déterminer le MMSE estimateurs d'un paramètre aléatoire ou inconnu, même s'il existe. Par exemple, nous ne connaissons pas la fonction de densité de probabilité des données, mais nous connaissons les moments de premier et de second ordre. Dans ce cas, les méthodes développées en l'estimation des paramètres et la détermination de la borne inférieure de Cramer-Rao ne peuvent être appliqué. Cependant, nous aimerions toujours obtenir une valeur raisonnable (sous-optimale) ou « meilleur » estimateur, dans le sens où il est sans biais et a une variance minimale, généralement appelé estimateur MVU. Pour ce faire, nous limitons l'estimateur à un linéaire fonction des données, et ainsi il devient possible d'obtenir une expression explicite pour le meilleur estimateur linéaire sans biais (BLEU).

Nous donnons d'abord l'estimation du carré moyen minimum linéaire à un paramètre à présenter les concepts fondamentaux, puis les généraliser à plusieurs paramètres.

II.3.6.1 Estimation de la moyenne quadratique linéaire à un paramètre

L'estimation linéaire au carré minimum d'un paramètre aléatoire est donnée par

$$\hat{\theta}_{\text{lms}} = \text{aY} + \text{b} \tag{II.162}$$

La fonction de risque correspondante est

$$\begin{aligned} \text{R}_{\text{lms}} = \text{E}\big[\text{C}\big(\theta, \hat{\theta}\big)\big] &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \big(\theta, \hat{\theta}\big)^2 \ \text{f}_{\ \Theta,\text{Y}}(\theta, \text{y}) \text{d}\theta \text{dy} \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\theta - \text{ay} - \text{b})^2 \ \text{f}_{\ \Theta,\text{Y}}(\theta, \text{y}) \text{d}\theta \text{dy} \end{aligned} \tag{II.163}$$

Nous observons que minimiser le risque implique de trouver les constantes a et b, de sorte que $\mathfrak{R}_{ims}$ soit le minimum. Par conséquent, en prenant les dérivées de $\mathfrak{R}_{ims}$ par rapport à a et b et en les mettant égaux à zéro, nous avons

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\theta - \text{ay} - \text{b}) \text{y} \ \text{f}_{\ \Theta,\text{Y}}(\theta, \text{y}) \text{d}\theta \text{dy} = 0 \tag{II.164}$$

Et

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\theta - \text{ay} - \text{b}) \text{y} \ \text{f}_{\ \Theta,\text{Y}}(\theta, \text{y}) \text{d}\theta \text{dy} = 0 \tag{II.165}$$

En utilisant $E[g(x)] = \int_{-3}^{-1} x\frac{1}{8}dx + \int_{-1}^{1} x\frac{1}{4}dx + \int_{1}^{3} x\frac{1}{8}dx = 0$ et $E[g(x,y)] = \iint_{-\infty}^{\infty} g(x,y) f_{xy}(x,y)dx\,dy$, (II.164) et (II.165) peuvent être réécrits comme

$$\text{aE}[\text{Y}^2] + \text{bE}[\text{Y}] = \text{E}[\theta\ \text{Y}] \tag{II.166}$$

Et

$$aE[Y]+b= E[\theta] \tag{II.167}$$

On a deux équations à deux inconnues. En résolvant pour a et b, on obtient

$$a = \frac{E[\theta\, Y] - E[\theta]E[Y]}{E[Y^2] - E^2(Y]} \tag{II.168}$$

Et

$$b = E[\theta] - E[Y]\frac{E[\theta\, Y] - E[\theta]E[Y]}{E[Y^2] - E^2(Y]} \tag{II.169}$$

Sachant que le coefficient de corrélation $\rho_{\theta Y}$ est donné par

$$p_{\theta y} = \frac{E[(\hat{\theta} - m_{\hat{\theta}})}{\sigma_\theta \sigma_y} \tag{II.170}$$

Avec $m_\theta = E[\theta]$, $m_y = E[y]$, $\sigma_\theta = \sqrt{E[(\theta - m_\theta)^2]}$, $\sigma_y = \sqrt{E[\left(y - m_y\right)^2]}$, après

$$a = \Theta_{\theta y}\frac{\sigma_\theta}{\sigma_y} \tag{II.171}$$

Et

$$b = m_\theta - \Theta_{\theta_y} m_y \frac{\sigma_\theta}{\sigma_y} \tag{II.172}$$

La fonction de coût optimale peut être obtenue pour être

$$R_{lms} = \sigma_\theta^2\left(1 - p_{\theta y}^2\right) \tag{II.173}$$

On peut montrer que si la fonction de densité jointe $f_{Y,\Theta}\,(y, \theta)$ est gaussienne, alors la moyenne conditionnelle $E[\theta \mid y]$ est linéaire dans les données d'observation, et donc le l'estimation quadratique moyenne minimale est linéaire. De plus, on suppose généralement pour commodité que le paramètre et l'observation Y ont des moyennes nulles. Dans ce cas $\hat{\theta}_{ims}$ est sans biais, et est donné par

$$\theta_{blue} = C_\theta C_{yy}^{-1} y \tag{II.174}$$

Où $C_{\theta y} = E[\theta Y] et\ C_{yy}^{-1} = \frac{1}{E[Y^2]}$. On peut maintenant généraliser le résultat de (II.174) pour l'estimation de paramètres multiples.

II.3.6.**2 Vecteur aléatoire**

Si maintenant θ est un paramètre vectoriel aléatoire et Y sont supposés être zéro, alors on peut montrer que le BLEU qui minimise l'erreur quadratique moyenne (écart minimum) est donné par

$$\hat{\theta}_{blue} = C_{\theta Y} C_{YY}^{-1} Y \tag{II.175}$$

Et l'erreur quadratique moyenne est

$$\mathrm{E}\left[(\theta-\hat{\theta}_{\mathrm{blue}})(\theta-\hat{\theta}_{\mathrm{blue}}\ )^{\mathrm{T}}\right]=\mathrm{C}_{\theta\theta}-\mathrm{C}_{\theta\mathrm{Y}}\mathrm{C}_{\mathrm{YY}}^{-1}\mathrm{C}_{\mathrm{Y}\theta} \tag{II.176}$$

C_{yy} Est la matrice de covariance du vecteur d'observation Y. C_{yy}^{-1} est son inverse, et $C_{\theta y}$ est la matrice de covariance croisée entre Y et θ . Notez que la moyenne et la covariance des données est inconnue, et les moyennes de Y et θ sont supposés être zéro, et donc l'estimateur de la moyenne quadratique linéaire est sans biais.

Preuve. Nous donnons maintenant une dérivation du résultat donné en (II.175). Puisque est restreint à être un estimateur linéaire pour Y, c'est-à-dire une fonction linéaire des données, alors $\hat{\theta}$ peut être écrit comme

$$\hat{\theta}=\mathrm{DY} \tag{II.177}$$

Le problème est de sélectionner la matrice D de telle sorte que le carré moyen donné par (II.176) soit minimisé. L'équation (II.176) est appelée la perte d'erreur quadratique à valeur matricielle une fonction. En remplaçant (II.177) dans (II.176), on a

$$\mathrm{E}\left[(\theta-\hat{\theta})(\theta-\hat{\theta})^{\mathrm{T}}\right]=\mathrm{E}[(\theta-\mathrm{DY})(\theta-\mathrm{DY})^{\mathrm{T}}] \tag{II.178}$$

Utilisant le fait que

$$\mathrm{E}\left[(\theta-\hat{\theta})(\theta-\hat{\theta})^{\mathrm{T}}\right]=\mathrm{tr}\,\mathrm{E}[(\theta-\mathrm{DY})(\theta-\mathrm{DY})^{\mathrm{T}}] \tag{II.179}$$

Alors, (II.189) devient

$$\begin{aligned}\mathrm{E}\left[(\theta-\hat{\theta})(\theta-\hat{\theta})^{\mathrm{T}}\right]&=\mathrm{E}[(\theta-\mathrm{DY})(\theta-\mathrm{DY})^{\mathrm{T}}]\\&=\mathrm{tr}[\mathrm{C}_{\theta\theta}-\mathrm{C}_{\theta\mathrm{Y}}\ \mathrm{D}^{\mathrm{T}}-\mathrm{DC}\ _{\mathrm{Y}\theta}+\mathrm{DC}_{\mathrm{YY}}\mathrm{D}^{\mathrm{T}}]\end{aligned} \tag{II.180}$$

Noter que

$$\begin{aligned}&(\mathrm{D}-\mathrm{C}_{\theta\mathrm{Y}}\mathrm{C}_{\mathrm{YY}}^{-1})\mathrm{C}_{\mathrm{YY}}(\mathrm{D}-\mathrm{C}_{\theta\mathrm{Y}}\mathrm{C}_{\mathrm{YY}}^{-1})^{\mathrm{T}}\\&=\mathrm{DC}_{\mathrm{YY}}\mathrm{D}^{\mathrm{T}}-\mathrm{DC}_{\mathrm{YY}}(\mathrm{C}_{\theta\mathrm{Y}}\mathrm{C}_{\mathrm{YY}}^{-1})^{\mathrm{T}}-\mathrm{C}_{\theta\mathrm{Y}}\mathrm{D}^{\mathrm{T}}+\mathrm{C}_{\theta\mathrm{Y}}\mathrm{C}_{\mathrm{YY}}^{-1}\mathrm{C}_{\theta\mathrm{Y}}\\&=\mathrm{DC}_{\mathrm{YY}}\mathrm{D}^{\mathrm{T}}-\mathrm{DC}_{\mathrm{Y}\theta}-\mathrm{C}_{\theta\mathrm{Y}}\mathrm{D}^{\mathrm{T}}+\mathrm{C}_{\theta\mathrm{Y}}\mathrm{C}_{\mathrm{YY}}^{-1}\mathrm{C}_{\theta\mathrm{Y}}\end{aligned} \tag{II.181}$$

En utilisant (II.179), nous pouvons écrire

$$\begin{aligned}\mathrm{E}[(\theta-\hat{\theta})(\theta-\hat{\theta})^{\mathrm{T}}]=\mathrm{tr}[&\mathrm{C}_{\mathrm{YY}}+(\mathrm{D}-\mathrm{C}_{\theta\mathrm{Y}}\mathrm{C}_{\mathrm{YY}}^{-1})\mathrm{C}_{\mathrm{YY}}(\mathrm{D}-\mathrm{C}_{\theta\mathrm{Y}}\mathrm{C}_{\mathrm{YY}}^{-1})^{-1}-\\&\mathrm{C}_{\theta\mathrm{Y}}\mathrm{C}_{\mathrm{YY}}^{-1}\mathrm{C}_{\theta\mathrm{Y}}\end{aligned} \tag{II.182}$$

On observe que la matrice de gain D n'apparaît que dans le second terme de droite côté de (II.182). Ainsi, chaque élément diagonal de la matrice $E[(\theta\ -\ \hat{\theta})\ (\theta\ -\ \hat{\theta})^{T}\ ]$ est minimisé lorsque D est donné par

$$\mathrm{D}=\mathrm{C}_{\theta\mathrm{Y}}\mathrm{C}_{\mathrm{YY}}^{-1} \tag{II.183}$$

En remplaçant (II.181) dans (II.175), on a

$$\hat{\theta} = \hat{\theta}_{\text{blue}} = C_{\theta Y} C_{YY}^{-1} Y \tag{II.184}$$

Et la preuve est complète.

Notons que si Y et θ ne sont pas des moyennes nulles, telles que $E[Y] = m_y$ et $E[\theta] = m\theta$, ensuite

$$\hat{\theta}_{\text{lms}} = AY + b \tag{II.185}$$

Où la matrice A et le vecteur b sont donnés par

$$A = \{E[YY^T] - E[Y^T]\}^{-1}\{E[\theta Y^T] - E[\theta]E[Y^T]\} = C_{YY}^{-1} C_{\theta Y} \tag{II.186}$$

Et

$$b = E[\theta] - AE[Y] \tag{II.187}$$

Par substitution directe, on obtient

$$\hat{\theta}_{blue} = m_\theta + C_{\theta Y} C_{YY}^{-1} (Y - m_Y) \tag{II.188}$$

Le BLEU donné en (II.182) a plusieurs propriétés intéressantes :

$$E\left[\hat{\theta}_{\text{blue}} Y^T\right] = C_{\theta Y} \tag{II.189}$$

$$E\left[\hat{\theta}_{\text{blue}} \hat{\theta}_{\text{blue}}^T\right] = C_{\theta Y}\, C_Y^{-1} C_{Y\theta} = C_{\hat{\theta}_{\text{blue}} \hat{\theta}_{\text{blue}}} \tag{II.190}$$

$$E\left[(E\left[\hat{\theta}_{\text{blue}} \hat{\theta}_{\text{blue}}\right])(\hat{\theta}_{\text{blue}} \hat{\theta}_{\text{blue}}^T)\right] = C_{\theta\theta} - C_{\hat{\theta}_{\text{blue}} \hat{\theta}_{\text{blue}}} \tag{II.191}$$

$$E\left[\left(\theta - \hat{\theta}_{\text{blue}}\right) Y^T\right] = 0 \tag{II.192}$$

$$E\left[\left(\theta - \hat{\theta}_{\text{blue}}\right) \hat{\theta}_{\text{blue}}^T\right] = 0 \tag{II.193}$$

Nous observons que la propriété (II.190) signifie que l'erreur dans l'estimation est orthogonale aux données Y, tandis que la propriété (II.191) signifie que l'erreur dans l'estimation est orthogonale à l'estimateur Blue$\widehat{\theta}$.

II.3.6.3 **BLEU dans le bruit blanc gaussien :**

Considérons le problème général de l'estimation d'un vecteur aléatoire avec N paramètres (notés vecteurs à N dimensions), à estimer à partir de K observations (noté vecteur Y de dimension K), dans le bruit blanc gaussien. Les paramètres θ et les mesures Y sont supposées être liées par le modèle dit linéaire

$$Y = H\,\theta + N \tag{II.194}$$

H est une matrice de mappage connue K × N, Y est le vecteur aléatoire observé K × 1 , est un vecteur aléatoire N ×1 à estimer, et N est un vecteur K ×1 représentant les erreurs dans la mesure (bruit). En supposant que et N ont des moyennes nulles, alors Y a zéro moyenne. La matrice de covariance de Y est

$$C_{YY} = [(H\,\theta + N)(H\,\theta + N)^T] = HC_{\theta\theta} + C_{N\theta}H^T + C_{NN} \quad \text{(II.195)}$$

Tandis que la matrice de covariance croisée de Y et est

$$C_{Y\theta} = HC_{\theta\theta} + C_{N\theta} \quad \text{(II.196)}$$

En substituant (II.195) et (II.196) dans (II.184), nous obtenons l'estimation BLEUE de θ à être

$$\hat{\theta}_{blue} = [C_{\theta\theta}H^T + C_{\theta N}][HC_{\theta\theta}H^T + HC_{\theta N} + C_{N\theta}H^T + C_{NN}]^{-1}Y \quad \text{(II.197)}$$

Avec matrice de covariance d'erreur

$$C_{\tilde{\theta}\tilde{\theta}} = C_{\theta\theta} - (C_{\theta\theta}H^T + C_{\theta N})(HC_{\theta\theta}H^T + HC_{\theta N} + C_{N\theta}H^T + C_{NN})^{-1}(HC_{\theta\theta} + C_{N\theta}) \quad \text{(II.198)}$$

Lorsque θ et N ne sont pas corrélés, ce qui est le cas habituel, $C_{\theta N} = 0$, et le BLEU de se réduit à

$$\hat{\theta} = C_{\theta\theta}H^T(HC_{\theta\theta}H^T + C_{NN})^{-1}Y \quad \text{(II.199)}$$

Tandis que la matrice d'erreur devient

$$C_{\tilde{\theta}\tilde{\theta}} = C_{\theta\theta} - C_{\theta\theta}H^T(HC_{\theta\theta}H^T + C_{NN})^{-1}HC_{\theta\theta} \quad \text{(II.200)}$$

En utilisant le lemme d'inversion de matrice, nous avons

$$\hat{\theta}_{blue} = C_{\tilde{\theta}\tilde{\theta}}H^T C_{NN}^{-1}Y \quad \text{(II.201)}$$

Où

$$C_{\tilde{\theta}\tilde{\theta}} = \left(C_{\theta\theta}^{-1} + H^T C_{NN}^{-1}H\right)^{-1} \quad \text{(II.202)}$$

Si aucune information a priori sur θ n'est disponible, et donc si -1 $C_{\theta\theta}$ est supposé nul, le BLEU d'est donné par

$$\hat{\theta} = (H^T C_{NN}^{-1}H)^{-1}H^T C_{NN}^{-1}Y \quad \text{(II.203)}$$

Notez que dans ces résultats, nous avons seulement supposé qu'est un paramètre aléatoire. Considérons maintenant le problème de l'estimation du vecteur inconnu, mais qui est contraint à être une fonction linéaire des données (mesures). L'estimateur comme fonction linéaire des données

Dans ce cas, nous demandons

$$\hat{\theta} = \sum_{k=1}^{K} a_{iK}Y_K + b_K, \quad i = 1,2,\ldots,M \quad \text{(II.204)}$$

Ou, sous forme matricielle

$$\hat{\theta} = AY + b \tag{II.205}$$

Où A est une matrice $M * K$ et Y et b sont des vecteurs $K * 1$. Pour que $\hat{\theta}$ soit impartial, nous devons avoir

$$E[\hat{\theta}|\theta] = 0 \tag{II.206}$$

D'où,

$$E[AY + b|\theta] = AY[Y|\theta] + b = AE[H\,\theta + N] + b = AH\,\theta + b = 0 \tag{II.207}$$

Seulement si

$$AH = I \tag{II.208}$$

Et

$$b = 0 \tag{II.209}$$

L'estimation BLEUE est alors donnée par

$$\hat{\theta} = (H^T\,C_{NN}^{-1}H)^{-1}H^T C_{NN}^{-1}Y \tag{II.210}$$

Par conséquent, avec le bruit gaussien dans le modèle linéaire, nous pouvons énoncer le résultat suivant donné par le théorème de Gauss-Markov. Si les données sont de la forme du modèle linéaire général

$$Y = H\,\theta + N \tag{II.211}$$

Où H est une matrice $K \times M$ connue, θ est un vecteur $M \times 1$ de paramètres à estimer, et N est un vecteur de bruit $K \times 1$ avec une moyenne nulle et une matrice de covarianceC_{NN} , alors le BLEU de θ qui minimise l'erreur quadratique moyenne est

$$\hat{\theta} = (H^T C_{NN}^{-1}H)^{-1}H^T C_{NN}^{-1}Y \tag{II.212}$$

Avec matrice de covariance d'erreur

$$C_{\hat{\theta}\hat{\theta}} = E\left[(\theta - \hat{\theta}_{blue})(\theta - \hat{\theta}_{blue})^T\middle|\theta\right] = (H^T C_{NN}^{-1}H)^{-1} \tag{II.213}$$

La variance minimale de $\hat{\theta}_k$ est alors

$$var[\hat{\theta}_k] = [(H^T C_{NN}^{-1}H)^{-1}]_{kk} \tag{II.214}$$

II.3.7 **ESTIMATION DES MOINDRES CARRÉS**

En étudiant l'estimation des paramètres dans les sections précédentes, nos critères étaient de trouver un « bon » estimateur sans biais et ayant une variance minimale. A l'estimation des moindre carré, le critère est uniquement de minimiser la différence au carré entre les données (signal plus bruit) et les données de signal supposées.

Supposons que nous voulions estimer M paramètres, dénotant le M-dimensionnel vecteur, à partir des K mesures, désignant le vecteur K-dimensionnel Y avec K M. La relation entre les paramètres et les données observées Y est donnée par le modèle linéaire

$$Y = H\theta + N \tag{II.215}$$

Où H est une matrice connue $(K \times M)$ et N est le vecteur d'erreur inconnu $(K \times 1)$qui se produit dans la mesure de. L'estimateur des moindres carrés (LSE) de θ choisit les valeurs qui font $X = H\,\theta$ plus proche des données observées Y. Par conséquent, nous minimisons

$$\begin{aligned} J(\theta) = \sum_{k=1}^{K} (Y_k - X_k)^2 &= (Y - H\,\theta)^T (Y - H\,\theta) \\ &= YY^T - Y^T H\,\theta - \theta^T H^T Y + \theta^T H^T\,\theta \\ &= YY^T - Y^T H\,\theta - \theta^T H^T Y + \theta^T H^T \theta \end{aligned} \tag{II.216}$$

Notez que $Y^T H$ est un scalaire. Prendre la dérivée partielle du premier ordre du fonction coût $J(\theta)$ par rapport à θ (c'est-à-dire le gradient) et en la fixant égale à zéro, nous obtenons l'ensemble des équations linéaires

$$\frac{\partial J(\theta)}{\partial \theta} = -2H^T Y + 2H^T H\,\theta = 0 \tag{II.217}$$

Et le LSE est

$$\hat{\theta}_{ls} = (H^T H)^1 H^T Y \tag{II.218}$$

Notez que la dérivée partielle du second ordre est

$$\frac{\partial^2 J(\theta)}{\partial \theta^2} = H^T H \tag{II.219}$$

Cette matrice est définie positive tant que H est supposé être de rang complet pour garantir l'inversion de H^T H. Ainsi, la solution (II.216) est unique et minimise J (θ). Les équations

$$H^T H\,\theta = H^T Y \tag{II.220}$$

À résoudre pour $\widehat{\theta} = \widehat{\theta}_{ls}$sont appelés équations normales.

Nous observons que l'erreur dans l'estimateur ls $\widehat{\theta}$ est une fonction linéaire de l'erreur de mesure N, car

$$\begin{aligned} \tilde{\theta}_{ls} \triangleq \theta - \tilde{\theta}_{ls} = \theta - (H^T H)^{-1} H^T Y &= \theta - (H^T H)^{-1} H^T [H\,\theta + N] \\ &= \theta - (H^T H)^{-1} H^T H\theta - (H^T H)^{-1} H^T N \\ &= -(H^T H)^{-1} H^T \end{aligned} \tag{II.221}$$

Le minimum des moindres carrés J_{min} peut être montré, après une opération matricielle,

$$\begin{aligned} J_{min} = J(\hat{\theta}_{ls}) = (Y - H\,\hat{\theta})^T (Y - H\hat{\theta}) &= YY^T - Y^T H (H^T H)^{-1} H^T Y \\ &= Y^T (Y - H\hat{\theta}) \end{aligned} \tag{II.222}$$

Généralisation du problème des moindres carrés

La fonction de coût des moindres carrés peut être généralisée en introduisant un $K \times K$ matrice positive de pondération définie W pour donner

$$\mathrm{J}(\hat{\theta}) = (\mathrm{Y} - \mathrm{H}\,\theta)^{\mathrm{T}}\mathrm{W}(\mathrm{Y} - \mathrm{H}\theta) \tag{II.223}$$

Les éléments de la pondération peuvent être choisis pour mettre l'accent sur des valeurs spécifiques des données les plus fiables pour l'estimation $\widehat{\theta}$ .On peut montrer que la forme générale de l'estimateur des moindres carrés est

$$\hat{\theta} = (\mathrm{H}^{\mathrm{T}}\mathrm{WH})^{-1}\mathrm{H}^{\mathrm{T}}\mathrm{WY} \tag{II.224}$$

Tandis que son erreur minimale des moindres carrés est

$$\mathrm{J}_{\min} = \mathrm{Y}^{\mathrm{T}}[\mathrm{W} - \mathrm{WH}(\mathrm{H}^{\mathrm{T}}\mathrm{WH})^{1}\mathrm{H}^{\mathrm{T}}\mathrm{W}]\mathrm{Y} \tag{II.225}$$

La matrice de covariance d'erreur devient

$$C_{NN} = (H^T WH)^{-1} H^T WRWH(H^T\, WH)^{-1} \tag{II.226}$$

Où $\mathrm{R_{NN}}$ est une matrice de covariance définie positive connue donnée par

$$R_{NN} = E[NN^T] \tag{II.227}$$

Puisque $E[N] = 0\ (c'est - \text{à} - dire, R_{NN} = C_{NN}\ )$.

Si les erreurs de mesure N ne sont pas corrélées et ont une variance 2 identique, alors $R = \sigma^2\, I$; et si $W = \sigma^2\, I$, alors (II.222) se réduit à (II.216). C'est un la mise à l'échelle constante n'a aucun effet sur l'estimation.

On peut aussi montrer que l'estimateur des moindres carrés et le minimum linéaire estimateur quadratique moyen sont identiques lorsque la matrice de pondération W est choisie comme

$$\mathrm{W} = \mathrm{R}^{-1} \tag{II.228}$$

C'est-à-dire l'inverse de la matrice de covariance du bruit de mesure.

II.4 **Conclusions :**

Nous avons développé les concepts de base du test d'hypothèses. En Première étape nous avons étudié le critère de Bayes, qui suppose la connaissance de la probabilité à priori de chaque hypothèse et l'affectation des coûts de chaque décision possible. Le coût moyen, connu sous le nom de fonction de risque, a été réduit au minimum pour obtenir la règle de décision optimale. Le critère minimax, qui minimise le coût moyen d'une probabilité a priori sélectionnée, P1, a été étudié dans section II.2.2. Le critère minimax s'applique aux situations où les probabilités à priori ne sont pas connues, même si des affectations des coûts réalistes aux divers décisions sont possibles. Dans les cas où il n'est pas possible l'affectation réaliste des coûts et les probabilités a priori ne sont pas connues, nous avons considéré le critère Neyman-

Pearson. Dans le critère de Neyman-Pearson, la probabilité de détection est maximisée, tandis que la probabilité de fausse alarme est fixée à une valeur désignée. La caractéristique de fonctionnement du récepteur, qui est un tracé de la probabilité de détection par rapport à la probabilité d'une fausse alarme, était utile dans l'analyse du rendement des détecteurs selon l'approche de Neyman-Pearson.

Nous avons développé le concept d'estimation des paramètres. Nous avons utilisé l'estimation du maximum de vraisemblance pour estimer les paramètres non aléatoires. Nous avons d'abord obtenu la fonction de vraisemblance en fonction des paramètres à estimer. Ensuite, nous avons maximisé la fonction de vraisemblance pour obtenir l'estimateur résultant de la résolution de l'équation de vraisemblance. Nous avons présenté le test du rapport de vraisemblance généralisé dans la section II.3.2. Dans le test du rapport de vraisemblance généralisé, nous avons utilisé l'estimation du maximum de vraisemblance du paramètre inconnu dans l'hypothèse composite comme sa vraie valeur, puis nous avons effectué le test du rapport de vraisemblance. Il s'agissait d'une alternative au cas où les tests UMP n'exister. Des critères de mesure de l'estimateur, tels que le biais et la cohérence, ont été présentés pour déterminer la qualité de l'estimateur.

Lorsque le paramètre à estimer était une variable aléatoire, nous avons utilisé l'estimation de Bayes. Dans l'estimation de Bayes, nous avons minimisé le risque, qui est fonction de l'erreur entre l'estimation et la vraie valeur. Trois cas ont été considérés ;l'erreur au carré, l'erreur en valeur absolue et la fonction de coût uniforme. Il a été montré que l'erreur quadratique moyenne minimale représente la moyenne conditionnelle du paramètre (associé à la variable aléatoire d'observation) à estimer. Le risque minimum résultant était la variance conditionnelle. Dans le cas de l'erreur en valeur absolue, l'estimation s'est avérée être la médiane de la fonction de densité conditionnelle du paramètre à estimer, compte tenu de la variable aléatoire d'observation.

Pour le coût uniforme de Bayes, l'estimateur était en fait la solution de l'équation MAP. En comparant l'estimation ML et l'estimation MAP, il a été observé que l'estimation ML était un cas particulier de l'estimation MAP et est obtenue en mettant à zéro la fonction de densité du paramètre à estimer dans l'équation MAP. Afin de mesurer la « qualité » de l'estimateur, la borne de Cramer-Rao a été donnée comme une autre façon de mesurer la variance d'erreur, car une expression de la variance d'erreur était difficile à obtenir. Les résultats ci-dessus ont été généralisés à l'estimation de paramètres multiples.

Ensuite, nous avons présenté une estimation quadratique moyenne linéaire pour les situations où il aurait pu être difficile de trouver le MMSE, même s'il existait. Nous avons défini

le BLEU dans le sens où la valeur quadratique moyenne est minimisée. Nous avons vérifié que pour une fonction de densité gaussienne conjointe de l'observation et du paramètre à estimer, l’estimateur quadratique moyen linéaire est le MMSE optimal. Une introduction à l’estimation des moindres carrés a été présentée. Nous avons remarqué que l'estimation des moindres carrés n'était pas basée sur les critères de l’estimateur de variance sans biais et minimum, mais plutôt sur la minimisation de la différence au carré entre les données et les données de signal supposées.

CHAPITRE III :

Résultat et simulation

III.1 : Introduction :

Avec les avancées significatives dans les domaines de la mise en réseau, communications sans fil, micro fabrication et microprocesseurs, la mise en réseau de capteurs sans fil est devenue un domaine de recherche en plein essor. Un réseau de capteurs sans fil typique(WSN) est composé d'un grand nombre de petits capteurs, de faible puissance, qui sont déployés dans l'environnement pour recueillir des observations. Chaque nœud de capteur prétraite et extrait des informations des observations brutes. Chaque capteur a également la capacité de communiquer avec d'autres capteurs ou un nœud central (centre de fusion) via des canaux sans fil. Habituellement, il existe un centre de fusion qui traite conjointement les données de capteurs locaux et donne une évaluation précise. Parce que les capteurs sont connectés via un réseau sans fil, les WSN peuvent être déployés rapidement et de manière souple. Cette flexibilité, ainsi que l'amélioration la couverture de surveillance et les performances de détection, rendent les WSN particulièrement attrayant pour un certain nombre d'applications, telles que surveillance du champ de bataille, sécurité, surveillance du trafic et environnement.

De nombreux aspects des WSNs ont été étudiés récemment, tels que les protocoles de routage efficaces [22], la compression de données distribuée et la transmission, et le traitement collaboratif du signal [23]. Nous sommes particulièrement intéressés par la tâche de traitement de l'information au centre de fusion, et plus particulièrement, dans la tache de décision. Pour un système de détection distribué, l'idée est de considérer la communication et la fusion des décisions comme entités indépendantes et de définir ces deux fonctions séparément. Il existe aussi de nombreux travaux consacrés au problème de la détection distribuée avec des ressources système limitées [24] – [25]. Plus précisément, ce sont proposé des solutions pour une allocation optimale des bits (ou sélection du capteur) étant donné une contrainte sur la quantité totale de communications. Bien que bon nombre des résultats obtenus soient assez éclairant, l'hypothèse de la communication sans perte, Augmentation de la puissance et/ou l'utilisation de codes de correction d'erreur puissante n'est pas toujours possible car la plupart des WSN sont limités en ressources (énergie et bande passante). De plus, dans un environnement hostile, la puissance du signal transmis doit être maintenue au minimum pour atteindre une faible probabilité d'interception/détection (LPI/LPD). Par conséquent, il peut être nécessaire dans de nombreuses situations pour tolérer la perte pendant la transmission de données dans une certaine mesure. Pour surmonter cette perte, il est hautement souhaitable d'intégrer la communication et la fusion de décision pour atteindre un niveau acceptable de performances du système sans dépenser des

ressources du système supplémentaires. Ceci motive l'étude de la fusion de décisions locales corrompu pendant le processus de transmission en raison de l'évanouissement du canal/du bruitdépréciation.

III.2 Formulation du problème :

Le modèle du système qu'on envisage à étudier est représenté sur la figure III.1, un certain nombre de capteurs est implanté pour détecter un phénomène donné, ces capteurs représentent la première couche du système. Les informations collectées par ces derniers sont envoyées à travers un canal de Rayleigh qui représente la deuxième couche du système. et finalement ces informations sont collectées et traitées dans un centre de fusion pour donner la décision optimale, cette dernière étape est la troisième couche du système.

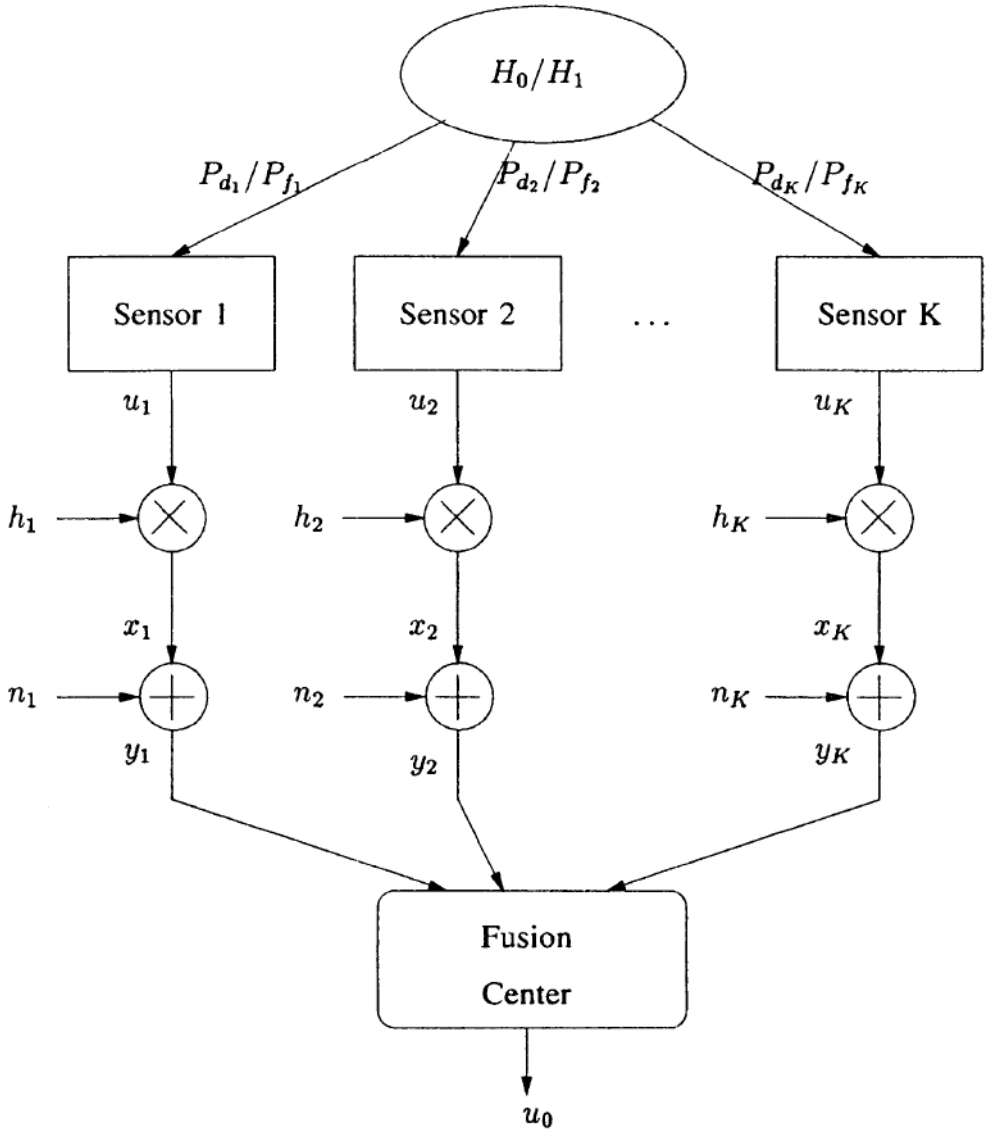


Figure III.18 : Modèle de fusion parallèle on présence des chaînes « fading » et « noisy » entre les capteurs locaux et le centre de fusion uk.

III.2.1 : Model du système à trois couches :

Le modèle à trois couches pour un système de détection distribué dans la présence de canaux d'évanouissement est illustré à la figure III.1. On a deux hypothèses, H_1 et H_0 sous test, chaque capteur obtient ses propres observations, les traite et prend une décision locale. Ces décisions locales sont transmises via un canal à évanouissements vers un centre de fusion. Dans un système de détection distribuée traditionnelle, toutes les décisions locales sont supposées être récupérées parfaitement au centre de fusion. Une chaîne couche doit être incorporée dans notre modèle pour permettre le développementde règles de fusion de décision correcte.

III.2.1.1 Couche capteur local :

C'est toutes les K observations collectés sous une hypothèse donnée. Dans ce travail, nous supposons que les observations sont indépendantes les unes les autres. Après avoir reçu

son observation, chaque capteur prend une décision ferme (binaire) : $u_k = 1$ est envoyé si H_1 est décidé, et $u_k = -1$ est envoyé dans l'autre cas c'est-à-dire H_0 décidé, où $k = 1, \dots, K$ Les performances de détection de chaque capteur local peut être caractérisé par sa probabilité correspondante de fausse alarme et de détection, désignées par P_{f_k} et P_{d_k} respectivement, pour le K^{ieme} capteur :

$$P_{d_K} = P(u_K = 1\ H_1) \quad \text{(III.1)}$$

$$P_{f_K} = P(u_K = 1\ H_1) \quad \text{(III.2)}$$

En général, ces paires (P_{f_k}, P_{d_k}) ne sont pas identiques et elles sont fonctions du SNR ainsi que des seuils des capteurs locaux.

III.2.1.2 Couche canal « Fading »

Les décisions au niveau des capteurs locaux, notées par u_k pour $k = 1, \dots, K$ sont transmises sur des canaux parallèles supposés subir des évanouissements indépendants. Nous supposons des canaux d'évanouissement plats entre les capteurs locaux et le centre de fusion. Cette hypothèse est raisonnable car la plupart des WSN fonctionne à courte portée (d'où un petit écart de retard) et à un taux faible (longue durée de symbole) en raison des limitations de puissance et d'énergie. Chaque décision locale u_k est transmet via un fading canal et la sortie du canal (ou entrée du centre fusion) pour le capteur est :

$$y_K = x_k + n_k = h_k u_k + n_k \quad \text{(III.3)}$$

Où h_k est le coefficient d'évanouissement (fading) du canal et n_k est le bruit gaussien de moyenne nulle et de variance σ^2.

III.2.1.3 Couche centre de fusion :

Sur la base des données reçues y_k pour tous *K*, le centre de fusion décide quelle est l'hypothèse la plus susceptible d'être vraie. Cela se fait en construisant et en évaluant une statistique de fusion en utilisant les observations y_k ainsi que certains paramètres du système, si disponible.

III.2.2 Aperçu sur les règles de fusion existantes :

Pour faciliter nos comparaisons plus tard, nous donnons ici un bref aperçudes règles de fusion proposées dans [26].

III.2.2.1 Règle de décision optimale basée sur le rapport de vraisemblance :

En supposant la connaissance de l'état instantané du canal concernant le canal et les indices de performance des capteurs locaux, c'est-à-dire les valeurs de P_{f_k} et P_{d_k}, la règle de fusion optimale basée sur le rapport de vraisemblance (LR) a été dérivée dans [26], est donnée par :

$$\Lambda(y) = log\left[\frac{f(y|H_1)}{f(y|H_0)}\right] \quad \text{(III.4)}$$

$$= \sum_{k=1}^{K} log\left[\frac{P_{d_k} e - \frac{(y_k - h_k)^2}{2\sigma^2} + \left(1 - P_{d_k}\right) e - \frac{(y_k - h_k)^2}{2\sigma^2}}{P_{f_k} e - \frac{(y_k - h_k)^2}{2\sigma^2} + \left(1 - P_{f_k}\right) e - \frac{(y_k - h_k)^2}{2\sigma^2}}\right]$$

Où $Y = [y_1, \dots, y_k]^T$ est un vecteur contenant des données reçues de tous les K capteurs, et σ^2 est la variance du Bruit blanc additif gaussien pour tous les canaux. Cette règle de fusion requiert à la fois les indices de performance des capteurs locaux et l'état du canal instantané (CSI).

III.2.2.2 Règle de fusion Chair-Varshney :

La règle de décision (statistique) de Chair-Varshney [26] est démontrée pour une approximation SNR élevé.

$$\Lambda_1 = \sum_{sign(y_k)=1} \log\frac{P_{dk}}{P_{fk}} + \sum_{sign(y_k)=-1} log\frac{1 - P_{d_k}}{1 - P_{f_k}} \quad \text{(III.5)}$$

$\Lambda_1$ne nécessite aucune connaissance concernant le canal mais nécessite la connaissance de P_{d_k} et P_{f_k} pour tous les capteurs. Cette approche, cependant, subit une perte en performance significative à moyen et faible SNR.

III.2.2.3 Règle de fusion MRC :

Il a été montré dans [26] que pour petites valeurs du SNR du canal, Λ en (III.4) se réduit à

$$\widehat{\Lambda_2} = \sum_{k=1}^{K}\left(P_{d_k} - P_{f_k}\right)h_k y_k. \quad \text{(III.6)}$$

De plus, si les capteurs locaux sont identiques, c'est-à-dire $P_{d_k} = P_d$ et $P_{f_k} = P_f$ pour tous k_s, puis Λ réduit encore à une forme analogue à une statistique MRC [27]

$$\Lambda_2 = 1/K\sum_{k=1}^{K}\left(P_{d_k} - P_{f_k}\right)h_k yk\,. \quad \text{(III.7)}$$

$\Lambda_2$en (III.7) ne requiert pas la connaissance de P_d et P_f fourni $P_d - P_f > 0$ La connaissance de gain du canal est cependant obligatoire.

III.2.2.4 Règle de fusion EGC :

Motivé par le fait que $\Lambda_2$ressemble à une statistique MRC pour la combinaison de la diversité, une troisième alternativesous la forme d'un EGC a été proposé, ce qui nécessite un minimum quantité d'informations :

$$\Lambda_3 = 1/K\sum_{k=1}^{K} y_k \quad \text{(III.8)}$$

Chose intéressante, cette alternative simple surpasse les deux Λ_1 et Λ_2 pour une large gamme de SNR en termes de détection performances [26].

III.3 Rapport de vraisemblance basé sur les statistiques du canal :

La règle de fusion optimale basée sur le rapport de vraisemblance (LR) nécessite la connaissance du CSI instantané. Cependant, pour un WSN avec des ressources très limitées (énergie et bande passante), il est prohibitif de dépenser des ressources pour estimer le canal chaque fois qu'un capteur local envoie sa décision au centre de fusion. Ainsi nous devons établit de nouvelles règles de fusion qui ne nécessitent pas de CSI instantané. Dans de nombreux Scénarios WSN, les statistiques du canal fading (aléatoire)et le bruit gaussien additif peut être estimé à l'avance, et utilisé comme information préalable. Notre objectif est de développer une nouvelle règle de fusion basée sur LR avec uniquement les informations préalables concernantles statistiques du canal au lieu du CSI instantané.

En supposant un canal d'évanouissement de Rayleigh avec une puissance unitaire (c'est-à-dire, $E[h_k^2] = 1$), le p_{d_f} de h_k est :

$$\mathrm{f(h_k)} = 2\mathrm{hk}\, e^{-\mathrm{h}_\mathrm{k}^2}, \quad \mathrm{h_k} \geq 0 \tag{III.9}$$

De (III.9) et en utilisant le fait que$x_k = u_k h_k$, c'est facile d'obtenir

$$f(x_k\, u_k) = 2u_k x_k e^{-x_k^2} U(u_x x_{\,k}) \tag{III.10}$$

Où $U(\cdot)$ est une fonction échelon définie comme suit :

$$\mathrm{U(x)} = \begin{cases} 1 & \mathrm{x} \geq 0 \\ 0 & \mathrm{x} < 0 \end{cases} \tag{III.11}$$

De plus, nous avons le résultat suivant :

$$\mathrm{f(y_k|u_K)} = \frac{\int \mathrm{p(u_k) f(x_k|u_k) f(y_k|x_k, u_k) dx_k}}{\mathrm{p(u_k)}} \tag{III.12}$$

$$= \int f(x_k u_k) f(y_k x_k) d_{x_k}$$

Où l'identité $f(y_k|x_k, u_k) = f(y_k|x_k)$, qui est obtenu à cause du fait que u_k, x_k et y_k forment une chaîne de Markov. Combiné au fait que n_k est une variable aléatoire gaussienne à moyenne nulle avec variance σ^2 .

Lemme 1 :

La fonction densité de probabilité (pdf) d'une observation y_kd'un capteur k n donnant une décision u_k est :

$$\mathrm{f(y_k|u_k)} = \frac{2\sigma}{\sqrt{2\pi}(1+2\sigma^2)} e^{\frac{\mathrm{y_k^2}}{2\sigma^2}} * \left[1 + \mathrm{u_K}\sqrt{2\pi}\mathrm{a y_k} e^{\frac{(\mathrm{a y_k})^2}{2}} \mathrm{Q}(1 - \mathrm{a u_k y_k})\right] \tag{III.13}$$

Où $a = \frac{1}{\sigma\sqrt{1+2\sigma^2}}$, et $Q(\cdot)$ est la fonction complémentaire de la distribution gaussienne standard, c'est-à-dire,

$$\mathrm{Q(x)} = \int_{\mathrm{x}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{\mathrm{x}^2}{2}\mathrm{dt}}. \tag{III.14}$$

Une fois que nous avons $f(y_k|u_k)$, il est facile d'obtenir $f(y_k|H_1)$, $f(y_k|H_0)$ et donc le rapport de vraisemblance est établi, comme indiqué dans le théorème qui suit.

Théorème 1 : Le rapport de vraisemblance logarithmique basé sur la connaissance des statistiques de canal et des indices de performance de détection locale est

$$\Lambda_4 = \log\left[\frac{\mathrm{f(y\ H_1)}}{\mathrm{f\ (y\ H_0)}}\right] = \sum\nolimits_{\mathrm{k=1}}^{\mathrm{K}} \log\left\{\frac{1+\left[\mathrm{P_{d_k}} - \mathrm{Q(ay_k)}\right]\sqrt{2\pi}\mathrm{ay_k}e^{\frac{(\mathrm{ay_k})^2}{2}}}{1+\left[\mathrm{P_{f_k}} - \mathrm{Q(ay_k)}\right]\sqrt{2\pi}\mathrm{ay_k}e^{\frac{(\mathrm{ay_k})^2}{2}}}\right\} \tag{III.15}$$

Preuve : Sous l'hypothèse H_1, on a

$$\mathrm{f(\ y_k\ H_1)} = \frac{\sum_{\mathrm{u_k}}[\mathrm{p(H_1)p(u_k|H_1)f(y_k u_k)}]}{\mathrm{p(H_1)}} \tag{III.16}$$

$$= \sum_{\mathrm{u_k}} [\mathrm{p(u_k\ H_1)p(y_k|u_k)}]$$

$$= \mathrm{P_{dk}f\ (y_k|u_k = 1)} + \left(1 - \mathrm{P_{d_k}}\right)\mathrm{f(y_k u_k = -1)}$$

$$= \frac{2\sigma}{\sqrt{2\pi}(1+2\sigma^2)} e^{-\frac{\mathrm{y_k^2}}{2\sigma^2}} * \left\{1+\left[\mathrm{P_{d_k}} - \mathrm{Q(ayk)}\right]\sqrt{2\pi}\mathrm{ay_k}e^{\frac{(\mathrm{ay_k})^2}{2}}\right\}$$

Où (III.1), lemme 1, et l'identité $Q(ay_k) = 1 - Q(-ay_k)$ ont étés utilisées.

En suivant la même procédure et en remplaçant $P_{d\,k}$ avec $P_{f\,k}$ on obtient $f(y_k|H_0)$ comme suit :

$$\mathrm{f(y_k H_{0)}} = \frac{2\sigma}{\sqrt{2\pi}(1+2\sigma^2)} e^{-\frac{\mathrm{y_k^2}}{2\sigma^2}} * \left\{1+\left[\mathrm{P_{f_k}} - \mathrm{Q(ay_k)}\right]\sqrt{2\pi}\mathrm{ay_k}e^{\frac{(\mathrm{ay_k})^2}{2}}\right\} \tag{III.17}$$

De toute évidence, le rapport de vraisemblance est

$$\frac{\mathrm{f(y\ H_1)}}{\mathrm{f(y\ H_0)}} = \prod\nolimits_{\mathrm{k=1}}^{\mathrm{K}} \frac{\mathrm{f(y\ H_1)}}{\mathrm{f(y\ H_0)}} \tag{III.18}$$

$$= \prod\nolimits_{\mathrm{k=1}}^{\mathrm{K}} \frac{1+\left[\mathrm{P_{d_k}} - \mathrm{Q(ay_k)}\right]\sqrt{2\pi}\mathrm{ay_k}e^{\frac{(\mathrm{ay_k})^2}{2}}}{1+\left[\mathrm{P_{f_k}} - \mathrm{Q(ay_k)}\right]\sqrt{2\pi}\mathrm{ay_k}e^{\frac{(\mathrm{ay_k})^2}{2}}}$$

Où les hypothèses d'indépendance conditionnelle des décisions locales et l'indépendance des différents canaux d'évanouissement ont étés utilisés. En prenant le logarithme de l'équation ci-dessus des deux côtés,on obtient le rapport de vraisemblance log Λ_4 donné en (III.15).

Comme on peut le voir, cette statistique de fusion nécessite la connaissance des statistiques de canal et les indices de performances de détection des capteurs locaux. Pour distinguer cette nouvelle règle de fusion de la règle optimale règle de fusion basée sur LR qui

nécessite un CSI instantané, nous l'appelons le test du rapport de vraisemblance basé sur les statistiques de canal (LRT-CS).

Ensuite, nous déterminons les approximations de la statistique de fusion LRT-CS pour SNR élevé et faible. Ces approximations sont dérivées et énoncés dans les deux propositions suivantes.

Proposition 1 : Comme la variance du bruit du canal $\sigma^2 \to 0$ c'est-à-dire SNR $\to \infty$, $\Lambda_4$défini en (III.15) se réduit à Λ_1 défini en (III.5), c'est-à-dire,

$$\lim_{\sigma_{\to 0}^2} \Lambda_4 = \Lambda_1 \tag{III.19}$$

Étant donné que $a \to \infty$ (hypothèse de SNR élevé), on a, pour $k \epsilon S_0 Q(ay_k) \to Q(-\infty) = -1$ et $\left| ay_k e^{\frac{(ay_k)^2}{2}} \right| \gg 1$ tandis que pour $k \epsilon S_1, Q(ay_k) \to Q(\infty) = 0$, et $\left| ay_k e^{\frac{(ay_k)^2}{2}} \right| \gg 1$. D'où, l'hypothèse $\sigma^2 \to 0$ peut être utilisé pour obtenir

$$\lim_{\sigma_{\to 4}^2} = \sum_{k \in S_0} \log\left[\frac{1 - P_{d_k}}{1 - P_{f_k}}\right] + \sum_{k \in S_1} \log\left[\frac{P_{d_k}}{P_{f_k}}\right] = \Lambda_1 \tag{III.20}$$

Qui est exactement la statistique de fusion de décision Chair-Varshney.

Proposition 2 : Comme $\sigma^2 \to \infty$, Λ_4 dans (III.15) se réduit à

$$\widehat{\Lambda_3} = \sum_{K=1}^{K} (P_{d_k} - P_{f_k}) y_k \tag{III.21}$$

De plus, si les capteurs locaux sont identiques, c'est-à-dire, $P_{d_k} = P_d$ et $P_{f_k} = P_f$ pour tout k et $P_d > P_f$, après Λ_4 réduit encore à une forme analogue à une statistique EGC.

$$\Lambda_3 = 1/k \sum_{k=1}^{K} y_k \tag{III.22}$$

Preuve : Pour un faible SNR, c'est-à-dire $\sigma^2 \to \infty$, $a = \frac{1}{\sigma\sqrt{1+2\sigma^2}} \to 0$. L'expansion en série de Taylor de $Q(x)$ est

$$Q(x) = \frac{1}{2} - \frac{1}{\sqrt{2\pi}}\left(x - \frac{x^3}{2} + \frac{x^5}{2.4} - \frac{x^7}{2.4.8} + \cdots\right). \tag{III.23}$$

Par conséquent, $\lim_{a \to 0} Q(ay_k) \approx \frac{1}{2} - \frac{ay_k}{\sqrt{2\pi}}$. En plus $e^{\frac{(ay_k)^2}{2}}$ peut-être approximé par la série de Taylor du premier ordre, c'est-à-dire $e^{\frac{(ay_k)^2}{2}} \approx 1 + \frac{a^2 y_k^2}{2}$. D'où

$$\lim_{a \to 0}\left\{1 + \left[P_{d_k} - Q(ay_k)\right]\sqrt{2\pi} a y_k e^{\frac{(ay_k)^2}{2}}\right\} \tag{III.24}$$

$$= \lim_{a \to 0}\left\{1 + \left[P_{d_k} - \left(\frac{1}{2} - \frac{ay_k}{\sqrt{2\pi}}\right)\right]\right\}\sqrt{2\pi} a y_k * \left\{1 + \frac{a^2 y_k^2}{2}\right\}$$

$$= \lim_{a \to 0} \left\{ 1 + \left(P_{d_k} - \frac{1}{2} \right) \sqrt{2\pi} a y_k \right\}$$

Où les termes d'ordre supérieur ont été ignorés. De même

$$\lim_{a \to 0} \left\{ 1 + \left[P_{d_k} - Q(a y_k) \right] \sqrt{2\pi} a y_k e^{\frac{(a y_k)^2}{2}} \right\} \tag{III.25}$$

$$= \lim_{a \to 0} \left\{ 1 + \left(P_{f_k} - \frac{1}{2} \right) \right\} \sqrt{2\pi} a y_k$$

Avec les résultats ci-dessus, il est facile de voir que

$$\lim_{a \to 0} \mathbf{\Lambda_4} = \lim_{a \to 0} \sum_{k=1}^{K} \log \left\{ \frac{1 + (P_{d_k} - \frac{1}{2}) \sqrt{2\pi} a y_k}{1 + (P_{d_k} - \frac{1}{2}) \sqrt{2\pi} a y_k} \right\} \tag{III.26}$$

$$= \lim_{a \to 0} \sum_{k=1}^{K} \left[(P_{d_k} - \frac{1}{2}) \sqrt{2\pi} a y_k \right] - \lim_{a \to 0} \sum_{k=1}^{K} \left[(P_{d_k} - \frac{1}{2}) \sqrt{2\pi} a y_k \right]$$

$$= \lim_{a \to 0} \sum_{k=1}^{K} (P_{d_k} - P_{f_k}) \sqrt{2\pi} a y_k$$

Où nous avons utilisé le fait $\lim_{x \to 0} \log(1 + x) = \lim_{x \to 0} x$. Cette statistique équivaut à $\widehat{\Lambda_3}$. De plus, si $P_{d_k} = P_d$ et $P_{f_k} = P_f$ pour tous k, la statistique de fusion peut être encore réduite à comme $\Lambda_3$en (III.8).

III.4 ANALYSE DE PERFORMANCE :

Dans le développement de la règle de fusion LRT-CS, nous avons déjà obtenu la pdf de $f(y_k|H_1)$ et $f(y_k|H_0)$. Avec ces pdf disponibles, il est possible de dériver la distribution sous forme fermée de la statistique Chair–Varshney, qui est très utile pour analyse de performance. Définir$K_0 = |S_0|$, ou $S_0 = \{k : y_k < 0\}$. C'est-à-dire K_0 est la cardinalité de S_0, et $K_1 = |S_1|$, où $S_1 = \{k : y_k \geq 0\}$. Ainsi, $K_1 + K_0 = K$. Avec ces définitions et en supposant que tous les capteurs sont identiques, c'est-à-dire $P_{d_k} = P_d$ et $P_{f_k} = P_f$ pour tout k (III.5) devient

$$\Lambda_1 = \log \left(\frac{P_d}{p_f} \right) K_1 + \log \left(\frac{1 - P_d}{1 - P_f} \right) (K - K_1) \tag{III.27}$$

$$= \log \left[\frac{P_d (1 - P_f)}{P_f (1 - P_d)} \right] K_1 + \log \left(\frac{1 - P_d}{1 - P_f} \right) K$$

Qui est une fonction affine de K_1 quand $\frac{P_d(1-P_f)}{P_f(1-P_d)} > 1$ (ou $P_d > P_f$), la statistique Λ_1 est équivalente à K_1. Lorsque tous les capteurs sont identiques, tous y_k sont indépendants et identiquement distribué (i.i.d.). Par conséquent,K_1 est une distribution binomiale, où la probabilité de succès est définie comme

$$p = P(y_k \geq 0) \tag{III.28}$$

Nous désignons P_1 et p_0 comme les probabilités de succès sous H_1et H_0, respectivement. Les solutions fermées sont fournies dansle lemme suivant.

Lemme 2 : Les probabilités d'une observation non négative $y_k$sous hypothèses H_1 et H_0 sont, respectivement

$$\mathrm{p_1 = p(y_k \geq 0|H_1) = \frac{1}{2} + \frac{P_d - \frac{1}{2}}{\sqrt{1 + 2\sigma^2}}} \tag{III.29}$$

Et

$$\mathrm{p_0 = p(y_k \geq 0|H_0) = \frac{1}{2} + \frac{P_d - \frac{1}{2}}{\sqrt{1 + 2\sigma^2}}} \tag{III.30}$$

Les performances de détection au niveau du système peuvent être évaluées avec précision en utilisant la distribution binomiale, à savoir

$$\mathrm{P_{d_0} = \sum_{i=K_T}^{K} \binom{K}{i} p_1^i (1 - p_1)^k - i} \tag{III.31}$$

Où K_T est le seuil. Pour le même K_T, le correspondant P_{f_0} est

$$\mathrm{Pf_0 = \sum_{i=K_T}^{K} \binom{K}{i} p_1^i (1 - p_1)^K - i} \tag{III.32}$$

III.5. Simulation et résultats

Afin de comparer la règle de décision établie, une simulation, en utilisant Matlab est réalisée pour le scenario suivant : un rapport signal sur bruit SNR de 5 décibels pour un canal de Rayleigh, une probabilité de fausse alarme P_f =0.05 et une probabilité de détection fixée à P_d=0.5, tandis que le nombre de capteurs est fixé à 8 capteurs.

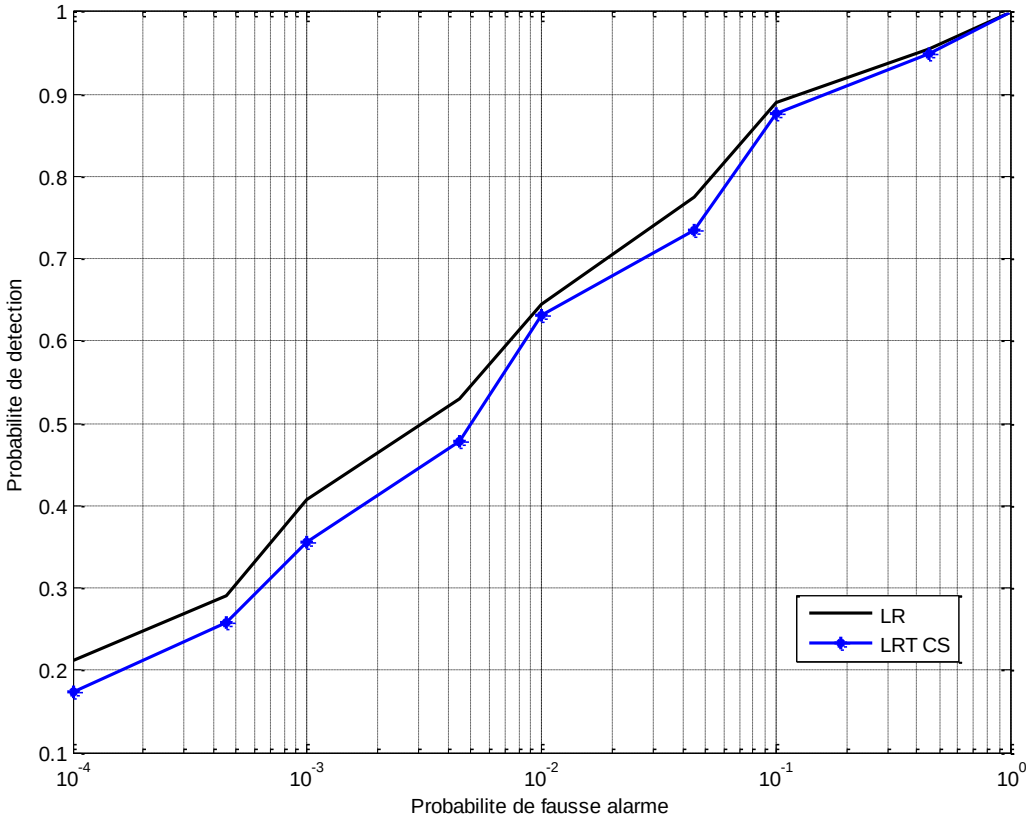


Figure III.19 : Courbes ROC (Receiver Operating Characteristic) pour règles de fusion LR et LRT-CS.

La figure III.2 nous montre que la règle de fusion basée sur le rapport de vraisemblance (LR) fournit des performances de détection plus puissantes que la règle LRT-CS, mais cela nécessite la connaissance des statistiques instantanés du canal (CSI).

Pour mieux comprendre les performances deux règles LR et LRT-CS, on a évalué la probabilité de détection en fonction du rapport signal sur bruit (SNR), comme le montre la figure suivante :

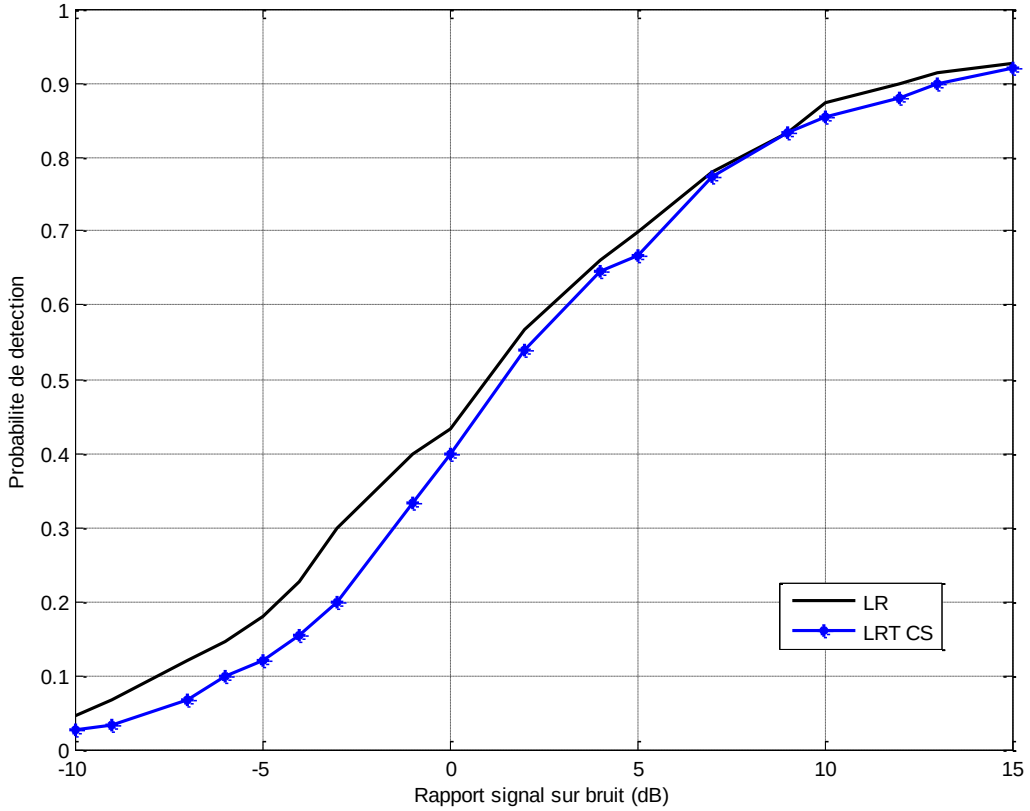


Figure III.20 : Niveau de probabilité de détection en fonction du SNR du canal de Rayleigh

La règle de fusion base sur le LR est plus performante que celle basée sur LRT-CS.

III.6 **Conclusion :**

La simulation du scenario proposé précédemment nous donne deux courbes importantes pour évaluer les performances et établir un comparatif entre les deux règles LR et LRT-CS en termes de performances.

La règle LRT-CS basée sur le rapport de vraisemblance est la plus performante, mais elle nécessite la connaissance du CSI. Cependant avoir les informations du canal (CSI) est tes couteux.

On note que la règle LT basée sur le rapport de vraisemblance sans avoir recours au CSI présente une légère dégradation par rapport à la règle LRT-CS, ce qui nous permet de dire que la règle LT est plus performante que LRT-CS vu qu'elle nécessite pas la connaissance du CSI.

Conclusion générale :

Dans la première partie de cette mémoire nous donnons un aperçu les réseaux des capteurs sans fil leur utilisation leur domaine d'application et l'intérêt des WSN dans notre vie quotidienne.

Les réseaux de capteurs sans fil sont une technologie récente. Les progrès de miniaturisation et d'allongement de la durée de vie des batteries, annoncent un futur prometteur à cette technologie. De plus, le développement de nouveaux capteurs permettra d'étendre les domaines d'applications déjà nombreux

Les réseaux de capteurs sans fil vont sans doute dans les années à venir constituer un développement technologique majeur apportant des solutions aux différents problèmes dans plusieurs domaines d'applications liés à la sécurité, la santé, l'agronomie, etc.

En deuxième partie dans ce travail, on a traité le problème de détection et d'estimation, un développement des différents méthodes d'estimation et de détection est réalisé, tel que le critère de Bayes, le critère minimax, NEYMAN-PEARSON et autres.

En dernière partie on a établi une formulation du problème et on a traité ce model de trois couches physiques (capteurs, canal de Rayleigh, centre de fusion) un rapport de vraisemblance est obtenu.

Une étude comparative est réalisée dans le but de discuter les performances des deux règles (LR) et (LRT-CS).

WEBOGRAPHIE

[1] *https://moodle.utc.fr/file.php/498/SupportWeb/co/Module_RCSF_1.html*

[2] https://fr.wikipedia.org/wiki/Capteur.

[3] http://fr.wikipedia.org/wiki/OFDM.

[4] www.afjv.fr/press0512/051219_standards_sans_fil_wifi.htm.

REFERENCES ET BIBLIOGRAPHIQUES

[5] El zein.G. guguen. P 0 (2004). Les téchnique multi-antennes pour les réseaux sans fil. Lavoisier. ISBN : 2-7462-0883-0.

[6] Collin L., "Optimisation de systèmes multi-antennes basée sur la distance minimale",thèse *de doctorat*, 17 décembre 2002,

[7] Plicanic V., "Antenna Diversity Studies and Evaluation" *Master of science Thesis*,Lund University Sweden May 2004, http://www.es.lth.se/teorel/Publications/TEAT-5000-series/TEAT-50 64.pdf.

[8] R. Vaughan and J. B. Andersen "Channel propagation and antennas for mobile communications" *The IEE Electromagnetic waves* 67eries 50, 2003.

[9] Y. Sankarasubramaniam I.F. Akyildiz, W. Su and E. Cayirci. A survey on sensor networks. IEEE Communications Magazine, 40 :102–114, Aug. 2002.

[10] W. B. Heinzelman. Application-specific protocol architectures for wireless networks. PhD thesis, Massachusetts Institute of Technology, May 2000.

[11] Chee-Yee Chong and S P Kumar. Sensor networks: evolution, opportunities, and challenges. Proceedings of the IEEE, 91 :1247â $\breve{A}$ ¸S1256, Aug. 2003.

[12] C.L. Ham D.A. Knapp Z.M. Koenig S.J. Luke T.B. Gosnell, J.M. Hall. Gamma-ray identification of nuclear weapon materials. Technical report, Lawrence Livermore National Lab, Livermore,CA (USA), 1997.

[13] Aline Baggio. Wireless sensor networks in precision agriculture. In Real-World Wireless Sensor Networks, Sweden, 2005.

[14] McCarthy P. Guru S. M. Peng W. Hugo D. McCulloch, J. and Terhorst. Wireless sensor network deployment for water use efficiency in irrigation. In Real-World Wireless Sensor Networks, Scotland, 2008.

[15] Jacques Panchard, Seshagiri Rao, T.V. Prabhakar, H.S. Jamadagni, and JeanPierre Hubaux. Commonsense net : A wireless sensor network for resource-poor agriculture in the semiarid areas of developing countries.

The MIT Press Journal publication, 4(1) :51â $\check{A}S$ ¸S67, 2008.

[16] Soto F. Suardiaz J. Sanchez P. Iborra A. Lopez Riquelme, J. A. and J. A Vera. Wireless sensor networks for precision horticulture in southern spain. Computers and Electronics in Agriculture, 2009.

[17] Projet agro-sens. https ://www.hds.utc.fr/heudiasyc/catalogue-131/projets-133/projets-ro/article/agrosens/ ?lang=fr. Responsable : Yacine Challal, 2009-2013.

[18] Philippe Hunel Harry Gros-Desormeaux Nathalie Dessart, Hac‘ene Fouchal and Nicolas Vidot. Distributed decision for medical alerts using wireless sensors. IEE World Of Wireless, Mobile and Multimedia Networks (WOWMOM), 2009.

[19] I. Bachelet S.M. Douglas and G.M. Church. A logic-gated nanorobot for targeted transport of molecular payloads. Science journal, 335 :831–834, 2012.

[20] Technnology review: Tracking a shopper’s habits. http ://www.technologyreview.com/computing/21161/. Michael Fitzgerald, August 2008.

[21] D.C. Petriu D. Makrakis E.M. Petriu, N.D. Georganas and V.Z. Groza. Sensor-based information appliances. IEEE Instrumentation Measurement Magazine, , December 2000.

[22] C. Intanagonwiwat, R. Govindan, D. Estrin, J. Heidemann, and F. Silva, “Directed diffusion for wireless sensor networking,” *IEEE/ACM Trans. Netw.*, vol. 11, no. I, pp. 2–16, Feb. 2003.

*[23] IEEE Signal Processing Mag. (*Special Issue on Collaborative Signal and Information Processing in Microsensor Networks*)*, vol. 19, no. 2, Mar. 2002.

[24] C. Rago, P. K. Willett, and Y. Bar-Shalom, “Censoring sensors: A low-communication-rate scheme for distributed detection,” *IEEE Trans. Aerosp. Electron. Syst.*, vol. 32, no. 2, pp. 554–568, Apr. 1996.

[25] J. Chamberland and V. V.Veeravalli, “Decentralized detection in sensor networks,” *IEEE Trans. Signal Process.*, vol. 51, no. 2, pp. 407–416, Feb. 2003.

[26] B. Chen, R. Jiang, T. Kasetkasem, and P. K. Varshney, “Channel aware decision fusion in wireless sensor networks,” *IEEE Trans. Signal Process.*, vol. 52, no. 12, pp. 3454–3458, Dec. 2004.

[27] G. L. Stüber, *Principles of Mobile Communication*, 2nd ed. Boston, MA : Kluwer, 2001.

[28] A. HAMOU and S. HARKATI, « La Gestion des Données dans les Réseaux de Capteurs sans Fil, ». 2013.

[29] S. Abdessamad and M. Wafa, " Développement d'une application orientée surveillance pour les réseaux de capteurs sous Contiki**,**" vol. 2015, pp. 2014–2015, 2015.

[30] F.BERKANI and L.HAMMOUI, " Le routage hiérarchique sous contrainte d'énergie dans les réseaux de capteurs sans fil," Mémoire de Master, U. A. M. De Béjaia, Algeria.