HAL
open science

Détection et localisation de particules de très hautes énergies en acoustique sous-marine.

Nicolas Juennard

To cite this version:

Nicolas Juennard. Détection et localisation de particules de très hautes énergies en acoustique sous-marine.. Traitement du signal et de l'image [eess.SP]. Université du Sud Toulon Var, 2007. Français. NNT : . tel-00295026v2

HAL Id: tel-00295026

https://theses.hal.science/tel-00295026v2

Submitted on 12 Feb 2008

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Université du Sud Toulon Var

THÈSE

Présentée par

Nicolas JUENNARD

pour obtenir

le grade de DOCTEUR DE L'UNIVERSITÉ DU SUD TOULON VAR
Mention TRÉS HONORABLE

Équipe d'accueil : Laboratoire L2MP - Equipe Signaux et Systèmes
École Doctorale : USTV

Titre de la thèse :

Détection et localisation de particules de très hautes énergies en acoustique sous-marine.

soutenue le 10 décembre 2007 devant le jury composé de :

M.	Régis LENGELLE	Professeur des Universités (UTT)	Rapporteur.
M.	Michel GUGLIELMI	Ingénieur de recherche CNRS (IRCCYN)	Rapporteur.
M.	Jean-Jacques AUBERT	Professeur des Universités (UDM)	Examinateur.
M.	Vincent BERTIN	Chargé de recherche CNRS (CPPM)	Examinateur.
M.	Claude JAUFFRET	Professeur des Universités (USTV)	Directeur de thèse.
M.	Bernard XERRI	Maître de Conférences (USTV)	Co-directeur de thèse.

Remerciements

Je tiens tout d'abord à remercier Claude Jauffret pour m'avoir proposé ce sujet de thèse à la fin de mon année de DEA. Son encadrement sans failles pendant toutes ces années de thèse et nos nombreuses discussions m'ont aidé à sans cesse me remettre en question, qualité fondamentale lorsque l'on fait de la recherche. Je lui dois de m'avoir accueilli dans l'équipe Signaux et Systèmes où j'ai pu travailler sur cette thèse et sur bien d'autres sujets passionnants. L'ambiance toujours chaleureuse de l'équipe GESSY m'a permis de mener ces travaux dans des conditions idéales.

Je tiens à remercier tout particulièrement Bernard Xerri pour son intuition et ses nombreux conseils. La qualité de son encadrement m'a particulièrement aidé pour le choix des orientations de ma recherche. Nos nombreuses discussions m'ont permis de mieux comprendre les mystères du filtrage adapté et sont à l'origine de bon nombre des résultats obtenus dans les travaux décrits dans cet exposé.

Je remercie également Claire Noël, Directrice de l'entreprise *Semantic TS* d'avoir bien voulu être le partenaire socio-économique de cette thèse.

Je remercie bien sûr également les membres de l'équipe GESSY, Jean Barrère, Bruno Borloz, Jean-François Cavassillas et Gilles Chabriel pour leur soutien et les nombreuses discussions que nous avons pu avoir durant ces quelques années.

De nombreuses discussions avec Valentin Niess du CPPM m'ont permis de mieux comprendre et appréhender les mécanismes physiques à l'origine des phénomènes étudiés dans ces travaux. Je lui suis reconnaissant d'avoir adapté le langage des physiciens à celui des traiteurs du signal. Je tiens également à le remercier d'avoir assisté à ma soutenance.

Je suis reconnaissant à Régis Lengelle et Michel Guglielmi d'avoir accepté d'être rapporteurs de ma thèse. Leurs commentaires et leurs questions m'ont permis de clarifier ma rédaction et m'ont donné de nouvelles pistes de réflexion.

Je remercie également Jean-Jacques Aubert et Vincent Bertin, d'avoir fait l'honneur d'assister à ma soutenance et d'avoir fait parti de mon jury de thèse.

Merci enfin à toute ma famille et mes proches de m'avoir supporté et aidé.

Table des matières

Remerciements **-1**

Table des matières **0**

Introduction **5**

1 Le signal acoustique et le bruit marin Protocole de simulations **9**

1.1 Le signal acoustique 10

1.1.1 Modélisation issue de l'approche thermodynamique 10

1.1.1.1 Cas général 10

1.1.1.2 Etude en champ lointain 12

1.1.2 Modélisation effectuée au CPPM - Modèle de Niess 14

1.1.2.1 Contexte des simulations du CPPM 14

1.1.2.2 Signal acoustique produit par une cascade hadronique et sa propagation 15

1.1.3 Détectabilité de la cascade d'énergie par un hydrophone 24

1.1.3.1 Indice de détectabilité 25

1.1.3.2 Densité de probabilité de la distance R d'une cascade détectable 29

1.1.3.3 Fonction de répartition de la distance R d'une cascade détectable 31

1.2 Le bruit marin 34

1.2.1 Modélisation réelle 34

1.2.2 Modélisation synthétique 37

1.3 Protocole des simulations appliquées à la détection 38

1.3.1 Paramètres de la cascade d'énergie et des simulations 39

1.3.2 Estimation des matrices de variance covariance 40

1.3.3 Génération du signal acoustique déterministe 40

1.3.3.1 Modèle d'Askaryian 40

1.3.3.2 Modèle de Niess 41

1.3.4 Génération du signal Antares aléatoire 42

1.3.5 Mise en forme du bruit de mer 43

1.3.5.1 Bruit de mer issu d'enregistrements 43

1.3.5.2 Bruit de mer synthétique 44

1.3.6 Notion de rapport signal-à-bruit 44

1.3.7 Les signaux de test 46

1.3.8 Test de détection 47

1.3.9 Calcul des valeurs de P_d, P_{fa} et affichage des courbes COR 47

1.3.10 Précision des courbes COR 47
1.4 Conclusion du chapitre 48

2 Détection par rapport de vraisemblance 50
2.1 Introduction 50
2.2 Résumé 50
2.3 Approximation polynômiale du signal acoustique 51
2.3.1 Quel ordre pour le polynôme ? 52
2.4 Cas d'un bruit gaussien centré de variance connue et constante 53
2.4.1 Quelle valeur pour le seuil ? 54
2.4.2 Calcul et tracé des courbes COR théoriques 55
2.4.2.1 Calcul du seuil à partir de la valeur de P_{fa} 55
2.4.2.2 Calcul de P_d 56
2.4.2.3 Résumé : 57
2.5 Cas d'un bruit gaussien centré de variance inconnue et constante 59
2.5.1 Calcul et tracé des courbes COR théoriques 61
2.5.1.1 Résumé : 64
2.5.1.2 Vérification par Monte-Carlo 64
2.6 Cas d'un bruit gaussien de matrice de variance covariance estimée 66
2.7 Détection en situation réelle 66
2.7.1 Cas du signal acoustique d'Askaryian 67
2.7.1.1 Première distance de détection : $d = 100m$ 67
2.7.1.2 Seconde distance de détection : $d = 200$ m 68
2.7.1.3 Interprétations des résultats 69
2.7.2 Cas du signal acoustique de Niess 71
2.8 Conclusion 74

3 Détection par le filtrage adapté et ses extensions 76
3.1 Introduction 76
3.2 Résumé 77
3.3 Le signal acoustique, un signal aléatoire ? 78
3.4 Détection par analyse linéaire discriminante [16] 80
3.4.1 Rappel de la théorie dans le cas de la détection 80
3.4.2 Application au cas Antares 83
3.5 Le filtrage adapté stochastique FAS [3] 86
3.5.1 Théorie 86
3.5.1.1 Fonctionnelle associée 88
3.5.2 Application au cas Antares 88
3.6 Le filtrage adapté stochastique étendu FASE [3] 90
3.6.1 Théorie 90
3.6.1.1 Fonctionnelle utilisée 91
3.6.1.2 Rapport de vraisemblance dans le cas gaussien 91
3.6.2 Application au cas Antares 92
3.6.2.1 Fonctionnelle utilisée 93
3.6.3 Résultats expérimentaux 94
3.6.3.1 Cas du bruit gaussien 95
3.6.3.2 Cas du bruit réel 95
3.7 Le filtrage adapté stochastique sous contrainte FASC [3] 98
3.7.1 Théorie 98

3.7.1.1 Algorithme . . . 100
3.7.1.2 Fonctionnelle associée . . . 101
3.7.2 Application au cas Antares . . . 102
3.7.3 Résultats expérimentaux . . . 103
3.7.3.1 Recherche de la valeur de p optimale . . . 103
3.8 Adaptation du FASC au cas multibruits : le FASCM . . . 104
3.8.1 Théorie . . . 105
3.8.1.1 Sous-espace de dimension 1 . . . 106
3.8.1.2 Sous-espace de dimension p . . . 106
3.8.1.3 Fonctionnelle associée . . . 107
3.8.2 Adaptation au cas Antares . . . 108
3.8.2.1 Signal d'intérêt non centré . . . 109
3.8.3 Résultats expérimentaux . . . 110
3.9 Bilan des méthodes proposées : comparaison des résultats expérimentaux . . 111
3.9.1 Cas mono-bruit . . . 112
3.9.2 Cas multi-bruits . . . 115
3.10 Extensions et perspectives . . . 119
3.11 Conclusion . . . 121

4 Structure d'antenne et estimation de la trajectoire 123
4.1 Introduction . . . 123
4.2 Résumé . . . 123
4.3 Estimation des temps de retard . . . 124
4.3.1 Estimation des dates d'arrivée du front d'onde sur un hydrophone . 124
4.3.1.1 Résultats expérimentaux . . . 126
4.3.2 Estimation de la différence du temps d'arrivée sur un couple d'hydrophone . . . 128
4.3.2.1 Description des deux méthodes d'estimation du TDOA . . 128
4.3.2.2 Les contraintes liées aux deux méthodes . . . 129
4.3.2.3 Choix de la méthode d'estimation du TDOA . . . 131
4.4 Structure d'antenne : proposition . . . 133
4.4.1 Contraintes physiques . . . 133
4.5 Estimation de la trajectoire du neutrino . . . 135
4.5.1 Mise en équation du TDOA . . . 135
4.5.2 Mise en équation de l'estimateur . . . 138
4.6 algorithme . . . 139
4.6.1 Structure . . . 139
4.6.2 Calcul du pas de descente . . . 139
4.6.3 Singularités éventuelles de $Hess_M(\theta)$ . . . 140
4.6.4 Traitement des angles . . . 140
4.6.5 Test d'acceptation . . . 140
4.7 Simulations et résultats . . . 141
4.7.1 Initialisation et valeurs numériques . . . 143
4.7.2 $A^{(0)}_{vrai}$ à l'intérieur de la structure de capteurs
Zone **A** . . . 145
4.7.2.1 Tous les hydrophones détectent - $P_d = 1$ . . . 145
4.7.2.2 9 couples détectent - $P_d = 0.75$ . . . 147
4.7.2.3 5 couples détectent - $P_d = 0.416$ . . . 149

4.7.3 $A^{(0)}_{vrai}$ à l'extérieur de la structure de capteurs Zone **B** . . . 151
4.7.3.1 Tous les hydrophones détectent - $P_d = 1$ . . . 151
4.7.3.2 9 couples détectent - $P_d = 0.75$ . . . 152
4.7.3.3 5 couples détectent - $P_d = 0.416$ . . . 155
4.7.4 Résumé et interprétations . . . 157
4.8 Conclusion . . . 158

5 Simulation globale 160
5.1 Introduction . . . 160
5.2 Contexte physique de la simulation . . . 160
5.2.1 La structure d'hydrophones . . . 160
5.2.2 La cascade d'énergie . . . 161
5.3 Protocole et résultats expérimentaux . . . 162
5.3.1 Validation des éléments du système . . . 164
5.3.1.1 Modules de détection . . . 164
5.3.1.2 Module d'estimation du TDOA . . . 165
5.3.1.3 Module d'estimation de la trajectoire . . . 166
5.3.2 Simulation complète . . . 168
5.3.2.1 Cas n°1 : RSB = -45 dB . . . 168
5.3.2.2 Cas n°2 : RSB = -55 dB . . . 171
5.4 Conclusion et extensions . . . 173

A Etude en champ lointain du signal d'Askaryian 178

B A propos de l'échantillonnage 181

C Variances et écarts-types du détecteur classique dans le cas d'un bruit gaussien centré de variance connue et constante 182

D Calcul de la loi de $\frac{1}{\sigma}(\mathbf{M}_d^T\mathbf{M}_d)^{-1/2}\mathbf{M}_d^T X$ dans le cas de l'hypothèse H_0 183

E Dépendance de $X^T\Pi_1 X$ et $X^T X$ 185

F Indépendance de $X^T(\mathbf{I}_N - \Pi_1)X$ et $X^T\Pi_1 X$ 186

G Evolution des performances de détection en fonction de la distance 187
G.1 Atténuation du bruit en fonction de la distance pour des performances de détection constantes . . . 188
G.2 Evolution du RSB en fonction de la distance pour des performances de détection constantes . . . 189
G.3 Vérification des résultats . . . 190

H Calcul du rapport de vraisemblance dans le cas Antares 192

I Détail des calculs de l'estimateur de la trajectoire du neutrino 195
I.1 Calcul de $\nabla_\theta M(\theta)$ . . . 196
I.1.1 Calcul de $\nabla_\theta S(\theta)$ . . . 196
I.2 Calcul de $Hess_M(\theta)$ . . . 197
I.3 Régions de confiance . . . 197
I.3.1 Calcul de la matrice d'information de Fisher . . . 198

I.3.2 Région de confiance pour $d = 2$: ellipse de confiance 198
I.3.3 Borne de Cramer-Rao 199

J Protocole et simulations de la détection appliquée au cas réel - Cas monobruit 200
J.1 Paramètres des simulations 201
J.2 Génération du signal Antares 202
J.3 Génération du bruit de mer 203
J.4 Notion de rapport signal à bruit 204
J.5 Calcul des matrices de covariance 204
J.6 Les signaux de test 204
J.7 Test de détection 205
J.8 Calcul des P_d et P_{fa} et affichage des courbes COR 205

K Calcul de l'intercorrélation de deux signaux filtrés par le FASE 206
K.1 Filtrage par le FASE 206
K.2 Calcul de l'intercorrélation 207

Glossaire 209

Bibliographie 212

Introduction

Pourquoi les neutrinos ?

La plupart de nos connaissances sur l'Univers proviennent de l'observation des photons. En astronomie, ceux-ci possèdent de nombreux avantages : ils sont en effet produits en grand nombre, sont électriquement neutres, stables, faciles à détecter et leurs spectres contiennent une multitude d'informations sur les propriétés chimiques et physiques de leurs sources. Un inconvénient apparaît néanmoins lorsque l'on veut étudier certaines parties de la galaxie ou certaines sources d'énergie de type astrophysique : celles-ci sont complètement opaques aux photons et l'étude directe est impossible. Par exemple, les photons qui nous arrivent du soleil proviennent de sa photosphère, et non du coeur d'hydrogène en fusion de l'étoile.

Pour observer de tels phénomènes, les scientifiques ont besoin d'un objet physique qui soit électriquement neutre, dont la trajectoire ne soit pas affectée par les champs magnétiques et qui possèdent la propriété physique d'interaction faible (leur permettant ainsi de pénétrer dans des régions inaccessibles aux photons). Ces objets, décrits par les scientifiques depuis les années 30, existent et se nomment les *neutrinos*. D'après certains modèles, ceux-ci pourraient même porter 23 % de l'énergie de l'univers. En étudiant ces particules, il devient donc possible d'obtenir des informations uniques sur certains phénomènes astrophysiques de très hautes énergies.

Certaines sources astrophysiques émettent naturellement des neutrinos : la fusion de l'hydrogène, les réactions dans le coeur des supernovas, etc. De part leur propriété d'interaction faible, ces particules très difficilement observables directement, les scientifiques sont donc partis à la recherche de niches naturelles susceptibles d'être exploitées pour leur observation. Il apparaît que l'eau de mer est une matière assez dense pour que les neutrinos manifestent leur présence en interagissant avec certaines molécules. Ce sont ces observations qui ont motivé le développement du télescope *Antares* (**A**stronomy with a **N**eutrino **T**elescope and **A**byss **RES**earch).

Le téléscope Antares

Le projet international **Antares** a débuté en 1996. Son objectif est la construction d'un télescope à neutrinos situé dans un environnement marin à fortes profondeurs. Après de nombreuses études sur les propriétés requises de l'eau de mer [22] [1], le site du télescope a été choisi en mer Méditerranée à 40km au sud de Toulon, en France.

Interaction des neutrinos avec la matière

Le neutrino a beau être la particule la plus abondante de l'univers (il y a environ un milliard de fois plus de neutrinos que d'atomes d'hydrogène, qui est pourtant l'élément le plus répandu), l'attraper est pourtant loin d'être une chose aisée. Comme il n'a pas de charge électrique, il ne se lie pas avec d'autres particules chargées. Il n'est pas non plus sensible à l'interaction nucléaire forte, qui assemble les nucléons au sein du noyau atomique. Le neutrino agit si faiblement avec la matière qu'elle lui est pratiquement transparente : il pourrait traverser pendant mille années-lumière de matière solide avant d'être arrêté.

Néanmoins, lorsque la matière est suffisamment dense, le neutrino peut heurter un électron. Le résultat de cette collision est la génération d'un flash de lumière associé à une onde ultra-sonore[1]. Un télescope à neutrinos est constitué d'un réseau tridimensionnel de capteurs. Jusqu'à présent, la recherche s'est focalisée sur la détection de l'onde lumineuse issue de l'interaction neutrino/électron (utilisations de photo-capteurs).
Plusieurs télescopes ont été développés :

- ⋆ Le premier détecteur est celui du lac Baïkal, constitué de 8 lignes supportant au total 192 photo-capteurs. Ce télescope a permis d'établir la faisabilité de l'expérience mais est limité en profondeur à cause de la nature du site.

- ⋆ Le réseau *Amanda*, situé au pôle sud, est basé sur un milieu constitué de glace et non plus d'eau de mer, permettant ainsi de gagner en profondeur.

- ⋆ Le détecteur *Super Kamiokande*, est une gigantesque cuve d'eau entourée de tubes "à lumière Cherenkov" et est située au Japon. Les premiers résultats de cette expérience ont montré que le flux de neutrinos cosmique est plus faible que celui attendu par les scientifiques.

[1] En toute rigueur, devraient être considérés plusieurs types de réactions suivant le type de neutrino rencontré (neutrino-électron, neutrino-muon ou neutrinon-tau), mais toutes ces réactions produisant une onde lumineuse et ultra-sonore, nous n'en donnerons pas le détail.

Le télescope *Antares* est complémentaire du projet *Amanda*. En effet, un système basé sur un milieu marin est plus flexible, même si le bruit de fond est plus important. Ce système s'articule autour de plusieurs lignes immergées à une profondeur de 2350m. Le volume couvert par les lignes sera à terme de l'ordre du kilomètre cube. Sur ces lignes seront disposés des détecteurs de photons (photo-détecteurs), ainsi que des hydrophones capables de détecter le signal acoustique issu de la réaction neutrino/électron.

L'objectif principal de cette thèse est de proposer plusieurs approches, classiques et originales, pour appréhender le problème de la détection d'un signal acoustique généré par une cascade d'énergie après propagation dans un milieu marin. L'exposé se focalisera sur les méthodes et protocoles mis en jeu et essayera autant que possible d'en proposer des extensions pour de futures études. Plusieurs modèles de signaux et de bruits seront étudiées et les résultats illustrés par des simulations à partir de données réelles et/ou simulées.
Le second objectif de cette thèse est d'étudier le problème d'estimation de la position de la cascade d'énergie et de la trajectoire du neutrino correspondant grâce à un réseau d'hydrophones. L'exposé se voudra là encore assez général et sera articulé autour de la description d'une méthode d'estimation classique, tout en proposant d'autres orientations pour la résolution de ce problème.

Organisation du manuscrit de thèse

C'est autour de ce signal acoustique, appelé dans la suite signal Antares, que s'articule le sujet de cette thèse. Le premier chapitre est consacré à l'étude théorique et expérimentale de celui-ci. Ce chapitre revêt un aspect chronologique. En effet, les informations sur le signal acoustique nous sont fournies par les physiciens et les différents modèles étudiés nous ont été fournis tout au long de cette thèse. Le premier modèle étudié est un modèle théorique basé sur les travaux du physicien russe G.A. Askariyan [8], tandis que le second est basé sur une approche expérimentale basé sur les travaux de V. Niess [23] au CPPM de Marseille. Ces modèles ont aboutit a un modèle analytique du signal acoustique dont certains paramètres sont des variables aléatoires. Le signal Antares sera donc représenté par un signal aléatoire.

Basée sur les données issues des hydrophones (réelles ou simulées), l'étude d'un système de détection du signal acoustique noyé dans le bruit ambiant (constitué du bruit de mer et éventuellement de divers bruits biologiques/humains), fera l'objet des chapitres 2 et 3.

Le chapitre 2 traite d'une méthode dite "classique" de détection. Celle-ci est basée sur le calcul d'un rapport de vraisemblance et donc sur la connaissance *a priori* des densités

de probabilité des variables aléatoires mises en jeu. Plusieurs situations seront étudiées, du cas simple de la variance connue et constante au cas plus général de variance inconnue et variable. Le cas dit "réel", pour lequel les matrices de variance/covariance des processus mis en jeu sont estimées expérimentalement sera également traité.

Le chapitre 3 décrit également d'autres méthodes de détections. Celles-ci sont basées sur une approche totalement stochastique des signaux mis en jeu, par le biais de leurs moments d'ordre 1 et 2 (i.e. par le biais de leurs espérances et de leurs matrices de variance covariance). Les méthodes étudiées sont basées sur la notion de filtre adapté stochastique. Plusieurs extensions seront décrites et leurs performances évaluées.
Le cas dit multi-bruits, dans lequel le bruit de mer ambiant n'est plus le seul à perturber le signal acoustique est également abordé. En effet, l'écosystème marin ou l'activité humaine (bateaux, sonars, etc.) génèrent des signaux susceptibles de perturber le détecteur. Pour s'affranchir de ces nouvelles contraintes, on étudiera une méthode de détection basée sur une extension du filtrage adapté stochastique et qui tiendra compte de ces nouveaux bruits, par le biais de leurs matrices de variance covariance.
Les performances de chacune des méthodes étudiées précédemment seront comparées dans un cas dit "semi-réel", dans lequel le bruit de mer est un bruit réel issu de campagne de mesure effectuées en Méditerranée et le signal acoustique est généré synthétiquement à partir des travaux effectués au CPPM.

Dans le dernier chapitre de cette thèse, nous nous intéresseront à l'estimation de la trajectoire d'un neutrino dont on aura détecté la présence. Nous supposerons pour cela disposer d'un détecteur capable de réagir au passage d'une de ces particules. Après une étude d'observabilité, nous en déduirons une première architecture simple de capteurs nous permettant d'estimer la trajectoire de la particule. Cette estimation est basée sur la minimisation en puissance d'un critère basé sur les différences de temps d'arrivée sur les hydrophones de la structure.

Chapitre 1

Le signal acoustique et le bruit marin Protocole de simulations

Le but de ce chapitre est de se familiariser avec le signal acoustique issu de la collision d'un neutrino et d'un électron en milieu marin. La physique et la modélisation de ce phénomène ont donné lieu à de nombreuses études et publications [8] [14] [20].

La génération d'une telle onde acoustique met en jeu plusieurs mécanismes physiques. Au plus bas de l'échelle, l'interaction du neutrino avec les atomes du milieu est décrite par la physique des interactions faibles. Le neutrino possédant une masse très faible, pour que le signal sonore soit détectable sur des distances de l'ordre du kilomètre, il faut que son énergie soit très importante, de l'ordre de 10^{18}eV. Néanmoins, ces énergies n'étant pas accessibles aux chercheurs sur des systèmes actuels tels que les grand accélérateurs de particules (énergie de l'ordre de 10^{12}eV), l'interaction du neutrino avec les molécules de l'eau de mer a été extrapolée à partir des résultats connus aux plus faibles énergies. Ce comportement conduit au développement d'une gerbe (ou cascade) de particules secondaires. Le volume d'eau contenant cette cascade est ionisé, et il en résulte un échauffement local avec génération de l'onde acoustique par un phénomène de dilatation-compression.

Dans la suite de ce chapitre, deux modélisations du signal acoustique seront présentées. Le premier modèle est basé sur les travaux du physicien russe Gurgen Ashotovich Askariyan du P.N. Lebedev Physical Institute [9] qui montre que l'onde acoustique est correctement décrite à l'échelle mésoscopique en considérant une approche thermodynamique. Les études et simulations effectuées par Valentin Niess au CPPM de Marseille [23] aboutiront au second modèle. Il est à noter que ces simulations informatiques lourdes sont désormais possibles grâce aux outils informatiques actuels.

1.1 Le signal acoustique

1.1.1 Modélisation issue de l'approche thermodynamique

1.1.1.1 Cas général

Les calculs suivant sont basés sur les travaux du physicien russe G.A. Askaryian [9] datant de 1979. Les notations utilisées dans ce chapitres sont regroupées sur la figure 1.1. L'origine du repère est choisie au centre de la cascade d'énergie. Le point de l'espace r_c représente la position d'un éventuel hydrophone qui captera l'onde sonore générée par cette déposition d'énergie.

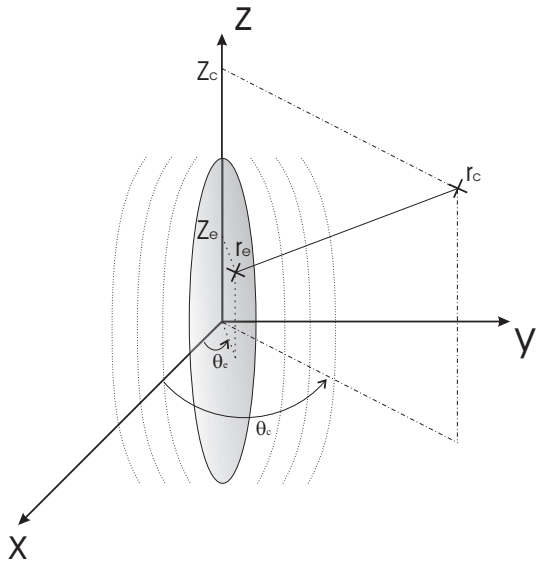

FIG. 1.1 – Radiation acoustique de la cascade d'énergie (coordonnées cylindriques)

Si l'on considère une approche thermodynamique, l'équation décrivant la propagation d'une onde acoustique générée par un tel phénomène dans un milieu aquatique est :

$$\Delta P - \frac{1}{C_s^2}\frac{\partial^2 P}{\partial t^2} = -\frac{\alpha}{C_p}\frac{\partial^2 q(r,t)}{\partial t^2} \tag{1.1}$$

$P(r_c, t)$ représente la valeur de la pression acoustique et $q(r,t)$ la densité d'énergie. Ces quantités sont évaluées à un point r_c et à un instant t. C_s est la célérité du son dans l'eau, α le coefficient d'expansion thermique, et C_p la capacité thermique du milieu.

La solution de cette équation est donnée par l'intégrale de Kirchhoff :

$$P(r_c, t) = \frac{\alpha}{4\pi C_p}\int_\Omega \frac{1}{|r_e - r_c|}\frac{\partial^2}{\partial t^2} q(r_e, t - \frac{|r_e - r_c|}{C_s})\ \mathrm{d}V \tag{1.2}$$

où Ω représente l'intégralité de l'espace ($\Omega = \mathbb{R}^3$ dans le cas de coordonnées cartésiennes).

L'expansion de la cascade d'énergie se fait dans un temps de l'ordre de $\tau_h \approx 10^{-8}s$. Ce temps est beaucoup plus court que la durée effective de l'onde acoustique ($\tau_s \approx 10^{-5}s$). On peut donc physiquement faire l'hypothèse suivante sur la déposition de l'énergie q :

$$\frac{\partial q(r,t)}{\partial t} = q(r)\delta(t) \tag{1.3}$$

Simplifions l'équation 1.2 en y injectant l'équation précédente et en posant $\rho = |r_e - r_c|$:

$$P(r_c,t) = \frac{\alpha}{4\pi C_p}\int_\Omega \frac{1}{\rho}\, q(r_e)\, \frac{\partial}{\partial t}\delta\Big(t - \frac{\rho}{C_s}\Big)\, \mathrm{d}V \tag{1.4}$$

$$= \frac{\alpha}{4\pi C_p}\frac{\partial}{\partial t}\int_\Omega \frac{1}{\rho}\, q(r_e)\, \delta\Big(t - \frac{\rho}{C_s}\Big)\, \mathrm{d}V \tag{1.5}$$

Simplifions encore cette expression en plaçant l'origine du repère au point r_c. L'expression de la pression acoustique devient donc :

$$P(r_c,t) = \frac{\alpha}{4\pi C_p}\frac{\partial}{\partial t}\int_{\mathbb{R}^3} \frac{1}{\sqrt{x_e^2+y_e^2+z_e^2}}\, q(r_e)\, \delta\left(t - \frac{\sqrt{x_e^2+y_e^2+z_e^2}}{C_s}\right)\, \mathrm{d}x_e\, \mathrm{d}y_e\, \mathrm{d}z_e \tag{1.6}$$

Au vu de cette expression, il paraît intéressant de passer à un système de coordonnées sphériques, décrit sur la figure 1.2.

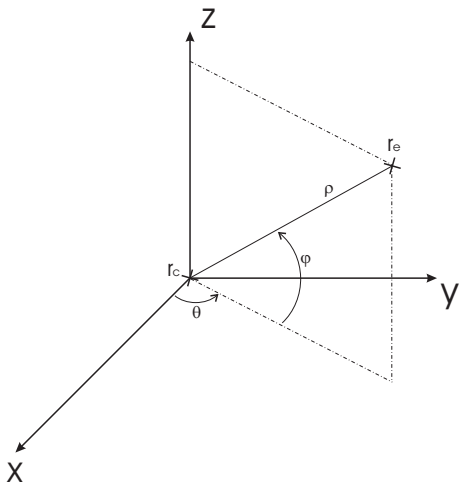

FIG. 1.2 – Passage aux coordonnées sphériques

Pour cela, nous avons besoin de calculer le déterminant du jacobien $J_{c\to s}$ du changement de repère cartésien $\longrightarrow$ cylindrique. Le jacobien $J_{s\to c} = J_{c\to s}^{-1}$ de changement de repère inverse cylindrique $\longrightarrow$ cartésien est facile à calculer. Il suffit ensuite d'utiliser le résultat $\det(J_{s\to c}^{-1}) = \frac{1}{\det(J_{s\to c})}$.

Nous avons :

$$J_{s\to c} = \begin{pmatrix} cos(\varphi)\ cos(\theta) & -\rho\ cos(\varphi)\ sin(\theta) & -\rho\ sin(\varphi)\ cos(\theta) \\ cos(\varphi)\ sin(\theta) & \rho\ cos(\varphi)\ cos(\theta) & -\rho\ sin(\varphi)\ sin(\theta) \\ sin(\varphi) & 0 & \rho\ cos(\varphi) \end{pmatrix} \tag{1.7}$$

Son déterminant est

$$det(J_{s\to c}) \quad = \quad \rho^2\ cos(\varphi) \tag{1.8}$$

En posant

$$\overline{r}_e = \begin{cases} \rho cos(\varphi) cos(\theta) \\ \rho cos(\varphi) sin(\theta) \\ \rho sin(\varphi) \end{cases} \tag{1.9}$$

et en utilisant le résultat précédent, on en déduit l'expression de la pression acoustique en coordonnées sphériques :

$$P(r_c, t) = \frac{\alpha}{4\pi C_p} \frac{\partial}{\partial t} \int_0^{2\pi} \int_0^{\pi} \rho\ q(\overline{r}_e)\ cos(\varphi)\ \mathrm{d}\theta\ \mathrm{d}\varphi \tag{1.10}$$

Si l'on pose $\rho = C_s t$, ce qui implique $dR = C_s dt$, on peut simplifier l'équation 1.10 et on obtient :

$$P(r_c, t) = \frac{\alpha}{4\pi C_p}\ C_s\ \frac{\partial}{\partial \rho} \int_0^{2\pi} \int_0^{\pi} q(\overline{r}_e)\ \rho\ cos(\varphi)\ \mathrm{d}\theta\ \mathrm{d}\varphi \tag{1.11}$$

Le lieu de l'intégration est la surface d'une sphère centrée au point r_c (grâce au 1er changement de variable) et de diamètre $\rho = C_s t$.
Nous avons obtenu ici l'équation du champ de pression acoustique reçu sur un capteur positionné en r_c à l'instant t.

1.1.1.2 Etude en champ lointain

L'expression du champ de pression acoustique ne tient pour l'instant pas compte de l'atténuation due au milieu marin, ni des caractéristiques des transducteurs d'écoute. Nous allons maintenant considérer que l'on utilise des transducteurs ayant une bande passante suffisamment élevée et que la distance du détecteur à la cascade est $\rho \ll L(f)$, où $L(f)$ est la longueur d'absorption du son à la fréquence f.

Plaçons-nous maintenant dans le cas du champ lointain, i.e. le signal acoustique est capté par un hydrophone à une grande distance de la cascade d'énergie. Dans ce cas de figure, le signal acoustique subit un filtrage fréquentiel caractéristique du milieu marin. Pour introduire la réponse impulsionnelle $\chi(w)$ de ce filtre, il est pratique de travailler sur

la transformée de Fourier de l'équation 1.2. Dans le domaine fréquentiel, l'expression de la pression acoustique est :

$$P_w = iw\frac{\alpha}{4\pi C_p}\chi(w)\int_\Omega \frac{1}{|r_e - r_c|}\, q(r_e)\; e^{i\frac{w}{C_s}|r_e - r_c|}\; \mathrm{d}V \tag{1.12}$$

Dans ce cas précis, il est intéressant d'utiliser un système de coordonnées cylindriques (cf. figure 1.3), en effet, la multiplicité de cette intégrale sera réduite car la distribution d'énergie $q(r_e)$ dans la cascade est à symétrie axiale. Nous supposerons également que cette distribution peut s'écrire :

$$q(Z_e, r_e) = \frac{1}{2\pi}\sum_{k=1}^{N} A_k(Z_e)e^{-r_e/\lambda_k(Z_e)} \tag{1.13}$$

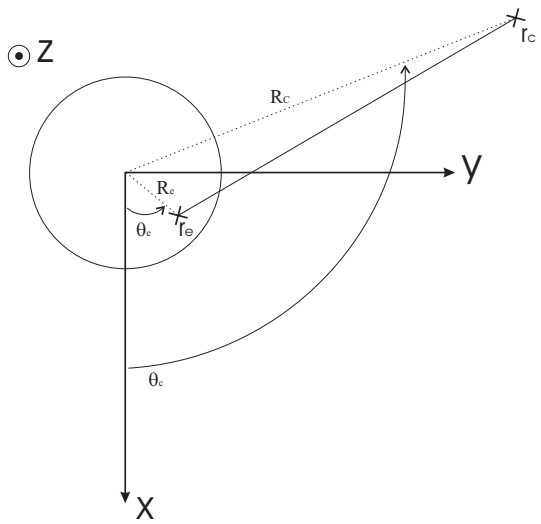

FIG. 1.3 – Radiation acoustique de la gerbe (vue de dessus et coordonnées cylindriques)

Le calcul mené en annexe A nous indique l'expression du champ de pression $s_{\text{Askaryian}}$:

$$\boxed{s_{\text{Askaryian}}(r_c, t) = i\frac{\alpha}{4\pi^2 C_p}\int_{-\infty}^{\infty}\int_0^{\infty} w\frac{\chi(w)}{r_c\sqrt{\gamma}}\sum_{k=0}^{N}\frac{A_k(Z_e)\lambda_k^2(Z_e)}{(1+\beta_k^2)^{3/2}}\; e^{iw(T_0 - t)}\; \mathrm{d}Z_e\; \mathrm{d}w} \tag{1.14}$$

avec $\gamma = 1 + \frac{(Z_c - Z_e)^2}{r_c^2}$, $T_0 = \frac{r_c\sqrt{\gamma}}{C_s}$ et $\beta_k^2 = \left(\frac{w}{C_s}\right)^2 \frac{\lambda_k^2}{\gamma}$.

Nous avons ainsi calculé l'expression de la pression acoustique reçue par le capteur situé au point r_c à l'instant t. Pour une expression de $\chi(w)$ donnée, il est donc possible de coder cette relation afin de calculer numériquement l'allure du signal sonore reçu au niveau d'un capteur. La figure 1.4 montre un exemple de l'allure et du spectre du signal Antares

reçu par un capteur situé à 100m de la cascade d'énergie pour un neutrino d'énergie 10^{18}eV.

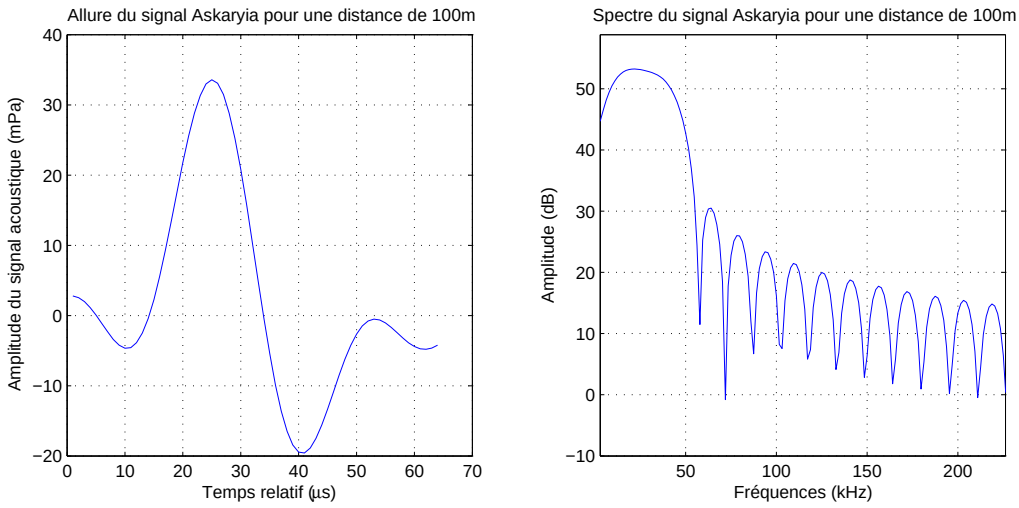

FIG. 1.4 – Allure et spectre fréquentiel du modèle thermo-dynamique d'Askaryian pour un capteur situé à une distance de 100m.

L'algorithme fourni par Askaryian limite volontairement la bande passante à 50kHz. Ne disposant pas d'autre modèle de référence à ce stade de l'étude, nous avons décidé de garder cette bande passante pour toutes les simulations utilisant ce modèle.

Remarque : Ce modèle du signal Antares était le seul disponible au début de ce travail de thèse. Il est basé sur un grand nombre d'hypothèses et d'approximations, mais a permis d'avoir une première idée de l'allure du signal à détecter.

1.1.2 Modélisation effectuée au CPPM - Modèle de Niess

Les travaux menés par Valentin Niess au CPPM de Marseille de 2002 à 2005 ont abouti à un nouveau modèle du signal Antares basé sur des simulations physiques et sur un modèle de propagation plus précis et adapté au site Antares. Pour ce modèle, nous admettons que la physique des particules élémentaires mise en jeu par les neutrinos considérés de très hautes énergies (de l'ordre de 10^{20} eV) reste bien décrite dans le cadre du Modèle Standard.

1.1.2.1 Contexte des simulations du CPPM

Ce modèle est basé sur des simulations utilisant les codes GEANT4. Ceux-ci permettent de simuler un grand nombre de processus de particules libres, non rattachées à un atome, avec la matière environnante. Ils sont donc particulièrement bien adaptés à la simulation de cascades d'énergie. Ces simulations ont permis d'aboutir à la simulation de deux types

de cascades :

- Les cascades hadroniques compactes, de longueur effective environ 10m.

- Les cascades dites électromagnétiques, pouvant s'étendre sur des centaines de mètres pour une énergie de 10^{20}eV.

Les occurrences relatives de ces deux types d'évènements sont délicates à déterminer *a priori*, car elles dépendent notamment de la composition en saveur des neutrinos (neutrino-électron, neutrino-tau ou neutrino-muon). Par ailleurs, dans le cas de très hautes énergies, apparait un effet dit LPM (formalisé par Migdal [2]) affectant la partie électromagnétique des cascades, et introduisant une élongation du dépôt d'énergie et un comportement fortement aléatoire. Aussi, est-il proposé de considérer ces deux types d'évènements comme des signaux physiques distincts. Dans tous les travaux de cette thèse, seules les cascades hadroniques compactes ont été prises en compte.

1.1.2.2 Signal acoustique produit par une cascade hadronique et sa propagation

Dans l'étude qui suit, le signal acoustique est pris à une distance de 100m de la cascade d'énergie. C'est la distance la plus faible disponible parmi toutes les simulations fournies par le CPPM.

Dépendance en énergie

Nous allons étudier ici la relation entre l'énergie de la cascade et l'amplitude des signaux acoustiques correspondant, pour différentes altitudes z le long de la cascade. Les trois énergies disponibles dans les fichiers de simulations CPPM sont considérées : 1, 10 et 100 EeV . On trace sur la figure 1.5 pour chaque valeur de l'énergie, les valeurs des amplitudes normalisées à 1 des signaux pour différentes positions le long de la cascade.

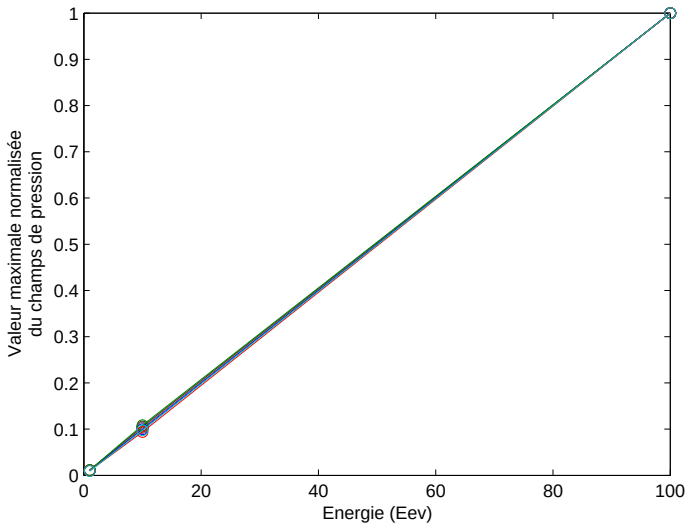

FIG. 1.5 – Dépendance en énergie de l'amplitude des signaux pris à différentes positions le long de la cascade

La courbe obtenue est quasiment une droite, on peut donc conclure à une dépendance linéaire entre énergie de la cascade et amplitude du signal acoustique. L'énergie se traduira donc un simple coefficient dans le modèle que nous sommes en train de construire.

Atténuation radiale

Lors de sa propagation, le signal subit deux types d'atténuation. La première, appelée atténuation radiale correspond à une atténuation en puissance du signal (ou en amplitude) et est directement proportionnelle à la répartition de l'énergie du signal sur la surface du front d'onde. La seconde dépend de la nature du milieu traversé et modifie spectre et amplitude du signal.

Pour une distance d, l'atténuation radiale est notée $g_L(d)$ et ne dépend que de la distance d. Celle-ci est représentée sur la figure 1.6.

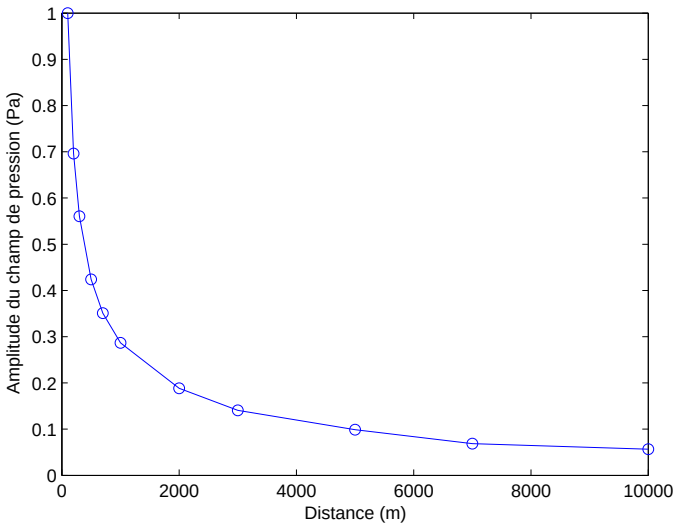

FIG. 1.6 – Atténuation radiale de l'amplitude du signal acoustique

Un modèle de cette atténuation a été étudié par J.C. Learned [14] et conduit à la relation suivante :

$$\frac{1}{g_L(d)} = d^{\frac{1}{2}}.\left(1+\left(\frac{d}{d_1}\right)^n\right)^{\frac{1}{2n}}.\left(1+\left(\frac{d}{d_2}\right)^n\right)^{\frac{1}{n}} \tag{1.15}$$

L'amplitude décroît comme la racine de la distance pour des petites distances de la cascade (inférieures à une distance d_1 fixée, ce qui correspond au champ proche). Pour de plus grandes distances (supérieures à d_1 et inférieures à une distance d_2 fixée), elle décroît plutôt en $1/d$ à cause de la perte des hautes fréquences et atteint une valeur asymptotique en $frm[o]--/d^2$ en champ lointain (distances supérieure à d_2).

Pour les intervalles de distance qui nous intéressent (de l'ordre du km), les paramètres optimaux ont été évalués empiriquement à $d_1 = 4000$m, $d_2 = 10000$m et $n = 1$. Pour ces valeurs, on superpose sur la figure 1.7 la courbe déduite des simulations et celle issue du modèle de Learned (l'atténuation, exprimée en dB, est calculée à partir du signal à 100m).

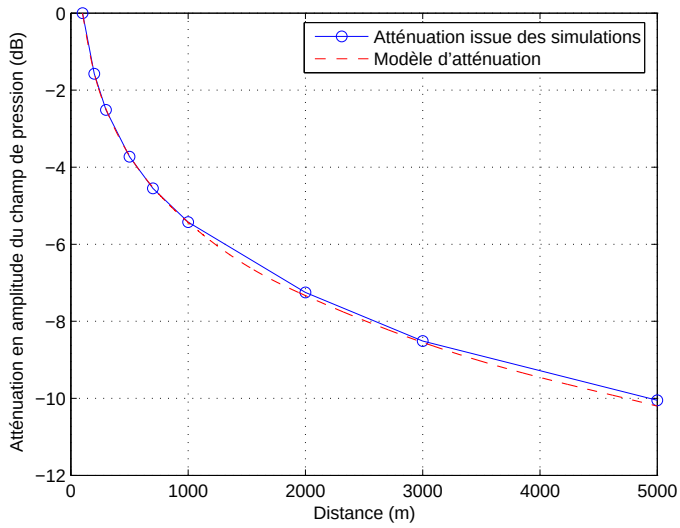

FIG. 1.7 – Validation du modèle de Learned

Les courbes sont superposées, et le modèle est donc validé pour cet intervalle de distances. Néanmoins, dans le cas qui nous intéresse, une distance maximale de 10000m est bien trop importante. En effet, dans le cadre du télescope de 1km[3], la distance maximale serait plutôt de l'ordre de 500 ou 1000m. Pour ces dimensions, l'atténuation radiale est représentée sur la figure 1.8.

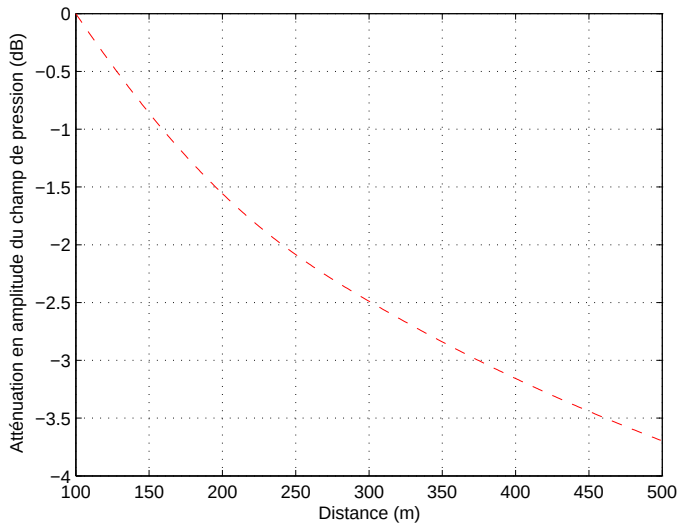

FIG. 1.8 – Validation du modèle de Learned des dimensions de l'ordre du km

Dans cet intervalle, le terme prépondérant pour des distances comprises entre d_1 et d_2 (et *a posteriori* pour des distances supérieures à d_2) peut être négligé, et on peut se

ramener à une formulation plus simple du coefficient d'atténuation radial :

$$\boxed{\frac{1}{g_L(d)} = \sqrt{d\left(1+\frac{d}{d_1}\right)}} \tag{1.16}$$

avec $d_1 = 1000$m.

Atténuation due au milieu marin

Le signal acoustique subit également une atténuation fréquentielle due à la nature du milieu marin lors de sa propagation. Celle-ci est caractérisée par un filtre de fonction de transfert H_d dont l'expression est de la forme :

$$H_d(f) = 10^{-\frac{\alpha(f)d}{20}} \tag{1.17}$$

La synthèse des travaux effectués par François et Garrison [7] et par Liebermann [18] conduisent au coefficient d'absorption complexe suivant :

$$\alpha(f) = A_1P_1f_1\frac{j\omega}{\omega_1+j\omega} + A_2P_2f_2\frac{j\omega}{\omega_2+j\omega} + A_3p_3(\frac{\omega}{2\pi})^2 \tag{1.18}$$

avec pour dépendance des paramètres :

$$\begin{cases} P_1 &= 1 & \text{[Pa]} \\ P_2 &= 1-1.37.10^{-4}P+6.2.10^{-9}P^2 & \text{[Pa]} \\ P_3 &= 1-3.83.10^{-4}P+4.9.10^{-10}P^2 & \text{[Pa]} \end{cases}$$

$$\begin{cases} A_1 &= 8.86.10^{0.78pH-5}/C_s \\ A_2 &= 21.44S(1+0.0025.T)/C_s \\ A_3 &= 4.937.10^{-4}-2.59.10^{-5}T+9.11.10^{-7}T^2-1.5.10^{-8}T^3 \end{cases}$$

$$\begin{cases} f_1 &= 2.8(S/35)^{0.5}.10^{4-1245/(T+273.15)} & \text{[kHz]} \\ f_2 &= 8.17.10^{8-1990/(T+273.15)}/(1+0.0018(S-35)) & \text{[kHz]} \end{cases}$$

Le premier terme correspond à la contribution de l'absorption de l'acide borique, le second à la contribution du sulfate de magnésium et le dernier à la contribution de l'eau pure. P_1, P_2 et P_3 représentent des pressions, A_1, A_2 et A_3 sont des coefficients sans unités et f_1 et f_2 sont les fréquences de relaxation associées à l'acide borique et au sulfate de magnésium.

Dans les expressions précédentes, la température T est exprimée en degrés Celsius, la pression P en déci-bar, la salinité S en PSU et la vitesse du son C_s en m/s. Avec ces unités, le coefficient d'atténuation est exprimé en dB/km. Pour la mer Méditerranée, les

valeurs suivantes sont utilisées :

- ⋆ une température de fond de 13.2^{o}C.
- ⋆ une salinité de 38.5 PSU.
- ⋆ une vitesse du son de 1520 m/s.
- ⋆ un pH de 8.2.

Dans la suite, nous noterons :

$$h_d(t) = \mathrm{TF}^{-1}\left[H_d(f)\right] \tag{1.19}$$

où TF représente l'opération de transformée de Fourier.

Signal acoustique reçu par un hydrophone

Nous pouvons maintenant construire un modèle pour le signal acoustique. Nous avons déjà vu que l'amplitude de celui-ci est directement proportionnelle à l'énergie E de la cascade. Notre modèle peut donc se mettre sous la forme :

$$\boxed{s_{\mathrm{Niess}}(t,d,E) = k.E.g_L(d).\mathrm{Re}\left(h_d * s_0\right)(t)} \tag{1.20}$$

avec :

- ⋆ k est un coefficient nécessaire à l'homogénéité de la formule, déterminé numériquement à partir des simulations disponibles.
- ⋆ E est l'énergie de la cascade.
- ⋆ $g_L(d)$ est l'atténuation radiale du signal pour une distance d.
- ⋆ $\mathrm{Re}(x)$ représente la partie réelle du complexe x.
- ⋆ h_d est le filtre temportel caractérisant l'atténuation du milieu marin pour une distance d.
- ⋆ $s_0(t)$ représente le signal acoustique pris à la distance minimale disponible dans les simulations, notée $d_{\min}$.

Il faut remarquer que cette modélisation n'est valable que pour des distances $d > d_{\min}$. Les simulations disponibles pour cette étude ont fixé cette valeur à :

$$d_{\min} = 100\mathrm{m}$$

La figure 1.9 représente l'allure de ce modèle de signal acoustique pour des distance à la cascade de 100 et 500m, une énergie $E = 10^{19}$eV et une longueur transverse $L_{eff} = 10$m. On peut remarquer qu'après une propagation de 500m, le signal a subit une forte atténuation en amplitude, ainsi qu'un filtrage fréquentiel du type passe bas.

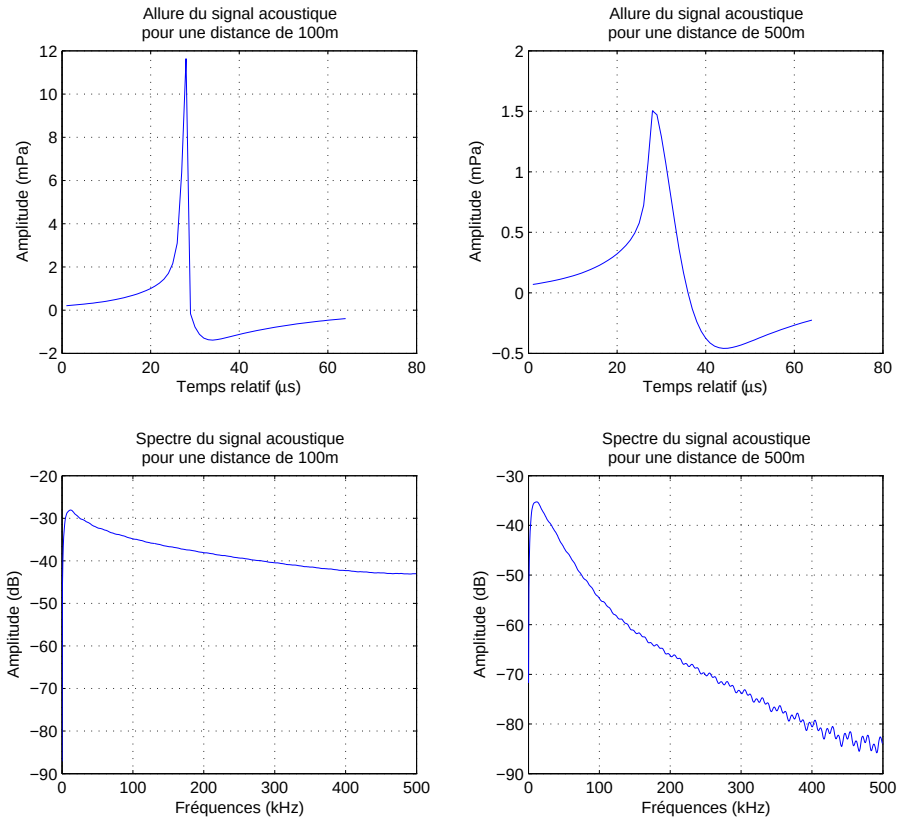

FIG. 1.9 – Allure du signal acoustique pour plusieurs distances à la cascade - Modèle de Niess

Allure du front d'onde

Il est important de s'intéresser maintenant à l'allure du front d'onde portant les signaux acoustiques. On représente sur la figure 1.10 les temps de retard des signaux acoustiques pour différentes altitudes le long de l'axe de la cascade. On représente également pour les mêmes altitudes les amplitudes du champ de pression des mêmes signaux acoustiques.

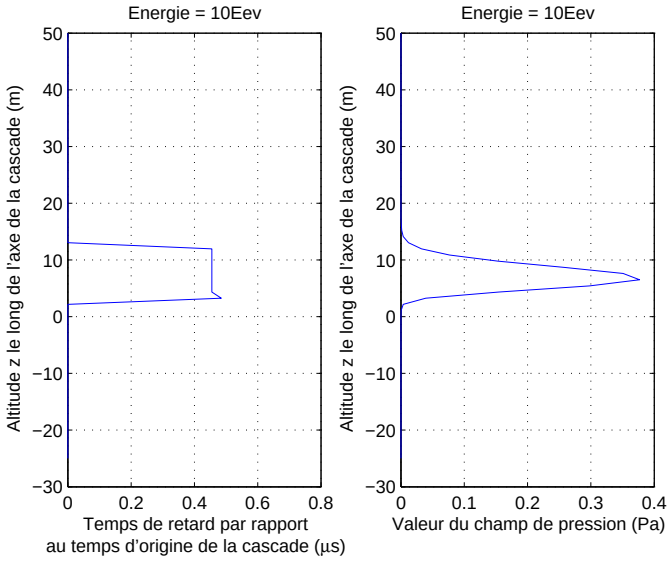

FIG. 1.10 – Allure du front d'onde des signaux acoustiques générés par la cascade

Il apparaît que le front d'onde est quasiment plan, tous les signaux acoustiques ont donc la même phase et se déplacent dans le milieu marin à la même vitesse. Néanmoins, l'amplitude de ces même signaux pour différentes valeurs de z n'est pas constante.
L'idée maintenant est de modéliser cette atténuation en altitude par un modèle simple basé sur une gaussienne. Considérons la fonction :

$$g_z(z) = e^{-\frac{1}{2}\left(\frac{z-\mu}{\sigma}\right)^2} \tag{1.21}$$

Nous prendrons pour μ la valeur expérimentale de z laquelle l'amplitude du signal acoustique est maximale. La valeur de σ a été empiriquement estimée à 2. On superpose sur la figure 1.11 le coefficient d'atténuation en altitude z déduite des simulations CPPM et la courbe de g_z pour les paramètres précédents.

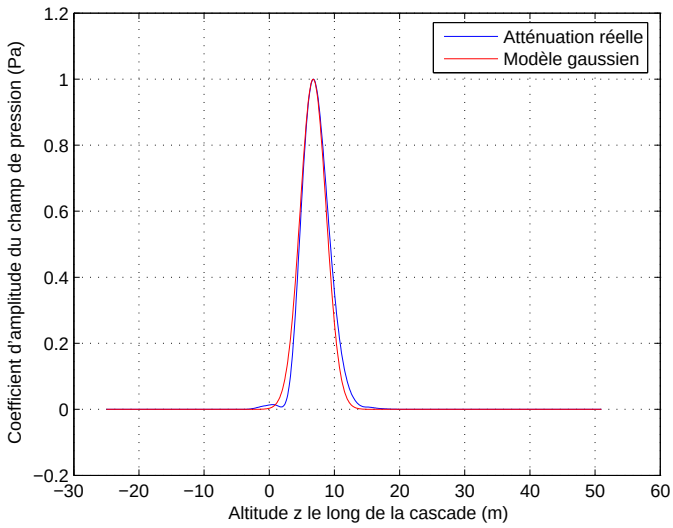

FIG. 1.11 – Atténuation de l'amplitude du signal acoustique suivant l'altitude z

Ces courbes sont très proches, et la fonction d'atténuation $g_z(z)$ peut donc être utilisée dans cette modélisation.

Néanmoins, dans une première approche, et afin de diminuer le nombre de variables, nous considérerons un profil en amplitude constant quelle que soit la position z le long de la cascade. La valeur de cette amplitude sera celle correspondant au maximum de la figure 1.10. Elle correspond à une altitude :

$$\boxed{z_{\text{opt}} = 6.52\text{m}}$$

Atténuation du front d'onde

L'atténuation en puissance que subit le signal en fonction de la distance parcourue d est due à la contribution de deux termes du modèle précédent : le coefficient d'atténuation radiale, noté $g_L(d)$, et l'atténuation due au filtrage fréquentiel $H_d(f)$. Il est intéressant à ce stade de l'étude de tracer en fonction de la distance de propagation d l'atténuation en puissance (exprimée en dB) subie par le signal. Cette courbe est représentée sur la figure 1.12.

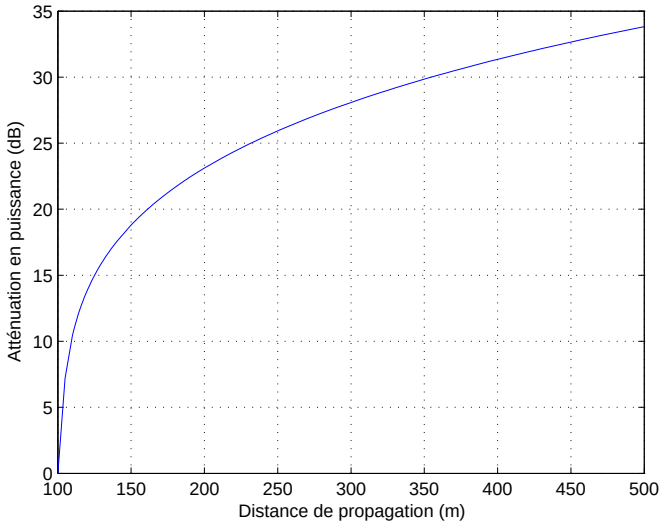

FIG. 1.12 – Atténuation en puissance subie par le signal en fonction de la distance de propagation.

1.1.3 Détectabilité de la cascade d'énergie par un hydrophone

Nous appelons ici évènement l'apparition d'une cascade d'énergie générant un signal acoustique. Intéressons nous maintenant à la répartition statistique de ces évènements dans l'espace. Nous avons déjà vu que la déposition d'énergie de la cascade se fait le long d'un segment de longueur L_{eff}. D'après la figure 1.10, le front d'onde se propage dans une direction quasi orthogonale à la cascade. Dans la suite de cette étude, nous allons faire l'hypothèse que cette propagation est parfaitement orthogonale à l'axe de la cascade. Considérons dans l'espace une cascade d'énergie caractérisée par :

- ⋆ L_{eff} : Longueur efficace de la cascade.
- ⋆ $\vec{r}$: Position du centre C de la cascade.
- ⋆ $\vec{u}$: Vecteur unitaire d'orientation de la cascade.
- ⋆ d_{max} : Extension maximale du front d'onde.

Considérons également un hydrophone localisé au point $\vec{r_H}$, ces informations sont regroupées sur la figure 1.13.

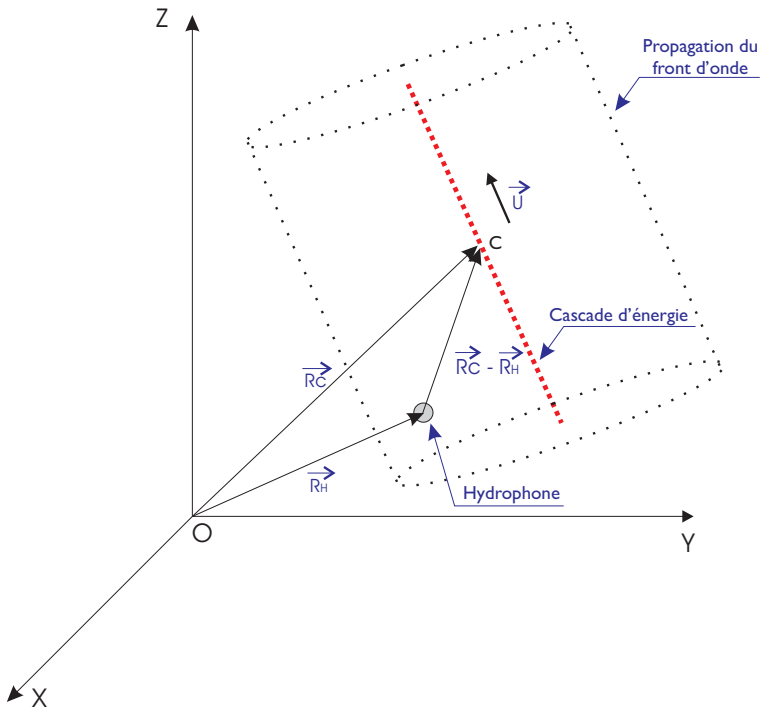

FIG. 1.13 – Représentation dans l'espace de la propagation du signal acoustique issu d'une cascade d'énergie. L'hydrophone est localisé à la position $\vec{r_H}$.

1.1.3.1 Indice de détectabilité

Introduisons l'indice de détectabilité [1] Δ. Celui-ci vaut 1 si le front d'onde rencontre l'hydrophone au cours de sa propagation, 0 sinon. Son expression peut donc s'écrire :

$$\Delta\left(\vec{u},\vec{r},\vec{r_H}\right) = \begin{cases} 1 \text{ si } |\vec{r}-\vec{r_H}|^2 - (\vec{r}.\vec{u})^2 \leq d_{max}^2 \text{ et } \left|\left(\vec{r}-\vec{r_H}\right).\vec{u}\right| \leq \frac{L_{eff}}{2} \\ 0 \text{ sinon} \end{cases} \tag{1.22}$$

Elle dépend de l'orientation $\vec{u}$ de la cascade, de la position $\vec{r}$ du centre de la cascade et $\vec{r_H}$ de l'hydrophone dans l'espace. Si l'on change l'origine du repère au point défini par le capteur, l'expression devient indépendante de $\vec{r_H}$ et s'écrit :

$$\Delta\left(\vec{u},\vec{r}\right) = \begin{cases} 1 \text{ si } |\vec{r}|^2 - (\vec{r}.\vec{u})^2 \leq d_{max}^2 \text{ et } |\vec{r}.\vec{u}| \leq \frac{L_{eff}}{2} \\ 0 \text{ sinon} \end{cases} \tag{1.23}$$

avec les notations de la figure 1.14.

[1] Il est important de remarquer que l'assertion "*un signal est détectable*" n'implique pas l'assertion "*ce signal sera détecté*".

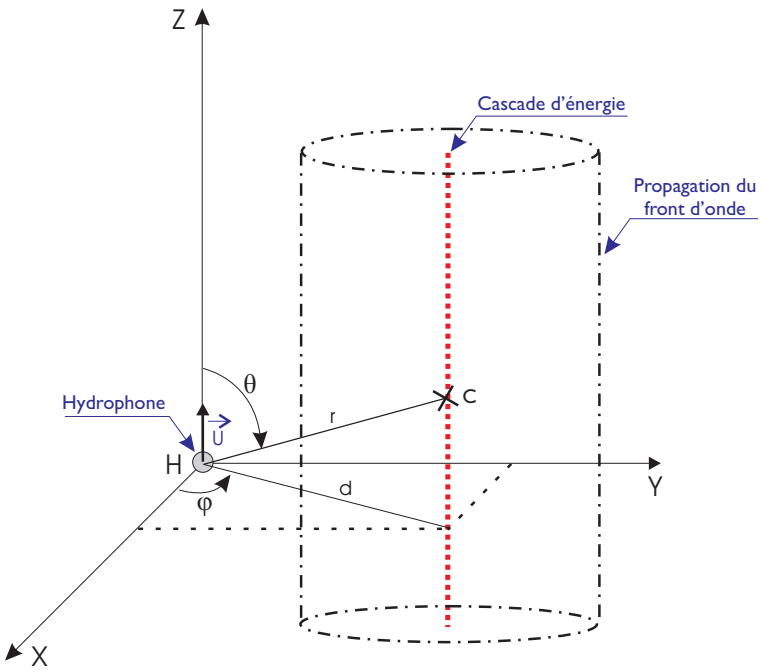

FIG. 1.14 – Représentation de la cascade d'énergie dans le repère du capteur.

Note : L'orientation du repère n'a pas d'importance, nous avons donc orienté l'axe O_z suivant le vecteur $\vec{u}$. Dans ce repère, on peut écrire :

$$\vec{u}\left|\begin{array}{l}0\\0\\1\end{array}\right. \quad \text{et} \quad \vec{r}\left|\begin{array}{l}x\\y\\z\end{array}\right. \tag{1.24}$$

La fonction Δ se réécrit donc :

$$\Delta\left(\vec{u},\vec{r}\right) = \Delta_{\vec{u}}(x,y,z) = \begin{cases} 1 \text{ si } x^2+y^2+z^2-z^2 \leq d_{max}^2 \text{ et } z \leq \frac{L_{eff}}{2} \\ 0 \text{ sinon} \end{cases} \tag{1.25}$$

Passons maintenant en coordonnées sphériques :

$$\begin{cases} x &= r\sin\theta\cos\varphi \\ y &= r\sin\theta\sin\varphi \\ z &= r\cos\theta \end{cases} \tag{1.26}$$

on obtient :

$$\Delta(r,\theta,\varphi) = \begin{cases} 1 \text{ si } & \left.\begin{array}{l} r^2\sin^2\theta \leq d_{max}^2 \\ 1-\cos^2\theta \leq \frac{d_{max}^2}{r^2} \end{array}\right\} \cos^2\theta \geq 1 - \frac{d_{max}^2}{r^2} \text{ et } \cos\theta \leq \frac{L_{eff}}{2r} \\ \\ 0 \text{ sinon} \end{cases} \tag{1.27}$$

Il est intéressant de discuter des cas suivants :

$\underline{r < d_{max}}$ **:** La première condition de 1.27 est toujours vérifiée.

$\underline{d_{max} < r < d_{max} + \frac{L_{eff}}{2}}$ **:** Pour certaines distances r supérieures à la distance maximale de détection d_{max}, il existe des orientations de cascades pour lesquelles la détection est possible. Ceci est dû au choix de localiser celle-ci par son milieu.

Néanmoins, il est raisonnable de considérer les hypothèses suivantes :

$$L_{eff} \quad << \quad d_{max} \tag{1.28}$$

$$d \quad < \quad d_{max} \tag{1.29}$$

En effet, pour les cascades hadroniques, la longueur efficace ($\approx$ 10m) est beaucoup plus petite que les distances de propagation considérées ($\approx$ 500 à 1000m). Dans ces conditions, on peut également écrire $d \approx r$.

Hypothèses importantes :

Avec les hypothèses suivantes :

$$L_{eff} \quad << \quad d_{max}$$

$$d \quad < \quad d_{max}$$

on peut écrire :

$$r \approx d$$

et si on note D la variable aléatoire associée à la distance d et R la variable aléatoire associée à la distance r, alors la densité de probabilité de D est approximativement la même que celle de R :

$$p_D(u) \approx p_R(u) \ \forall \ u$$

L'intérêt de cette approximation est que la détermination de la densité de probabilité de la variable aléatoire d est plus facile. Celle-ci sera calculée dans la suite de cette section.

Avec ces hypothèses, la première condition est donc toujours vérifiée, et la détectabilité se ramène donc à la seconde condition :

$$\cos\theta \leq \frac{L_{eff}}{2r} \tag{1.30}$$

Pour simplifier l'expression de Δ, plusieurs cas sont considérés pour r :

Cas n°1 : $r \leq \frac{L_{eff}}{2}$

On peut écrire : $\frac{L_{eff}}{2r} \geq 1$. Or, on sait que $\forall\theta$, $\cos\theta \leq 1$.
La condition 1.30 est donc toujours vérifiée.

Cas n°2 : $\frac{L_{eff}}{2} \leq r \leq d_{max}$

Ici, nous avons $\frac{L_{eff}}{2r} \geq 1$, or comme précédemment, $\forall\theta$, $\cos\theta \leq 1$. La condition 1.30 n'est donc pas toujours vérifiée et doit donc être testée.

Cas n°3 : $r > d_{max}$ Nous venons de faire l'hypothèse $d \approx r$, or nous avons défini la distance d_{max} comme la distance au delà de laquelle la détection est impossible. L'indice Δ prendra donc toujours pour valeur 0 dans ce cas.

Résumé :

Hypothèses :

$$Leff << d_{max} \quad \text{et} \quad d < d_{max}$$

$$\Rightarrow \; r \approx d$$

L'indice de détectabilité est :

$$\Delta(d,\theta,\varphi) = \begin{cases} 1 & \text{si} \quad d \leq \frac{L_{eff}}{2} \\ 1 & \text{si} \quad \frac{L_{eff}}{2} < d \leq d_{max} \text{ et } |\cos\theta| \leq \frac{L_{eff}}{2d} \\ 0 & \text{sinon} \end{cases} \tag{1.31}$$

Probabilité de détectabilité une cascade située à une distance r

Intéressons nous maintenant à la probabilité de détectabilité d'une cascade sachant qu'elle est située à une distance r. Celle-ci est déduite de l'indice de détectabilité Δ :

$$P\left(\Delta = 1 | R = r\right) \tag{1.32}$$

R représente la variable aléatoire associée à la distance r à la cascade.
Le calcul de cette densité est évident, et on obtient :

$$\begin{aligned} P\left(\Delta = 1 | R = r\right) &= 1 & \text{si } r \leq \frac{L_{eff}}{2} \\ &= P\left(|\cos\theta| \leq \frac{L_{eff}}{2r}\right) & \text{si } \frac{L_{eff}}{2} < r \leq d_{max} \\ &= 0 & \text{si} \end{aligned}$$

Résumé :

Hypothèses :

$$Leff << d_{max} \quad \text{et} \quad d < d_{max}$$

$$\Rightarrow \; r \approx d$$

La probabilité de détecter une cascade située à une distance d est :

$$P\left(\Delta = 1 | R = d\right) = \begin{cases} 1 & \text{si} \quad d \leq \frac{L_{eff}}{2} \\ \frac{L_{eff}}{2\ d} & \text{si} \quad \frac{L_{eff}}{2} < d \leq d_{max} \\ 0 & \text{si} \quad d > d_{max} \end{cases} \tag{1.33}$$

1.1.3.2 Densité de probabilité de la distance R d'une cascade détectable

Cherchons maintenant à calculer la densité de probabilité que la cascade soit à une distance r sachant qu'on peut la détecter. La loi de Bayes nous dit que :

$$p_R(r|\Delta = 1) \;=\; p\left(R = r\right) \frac{P\left(\Delta = 1 | R = r\right)}{P\left(\Delta = 1\right)} \tag{1.34}$$

Calcul de $p(R = r)$

Revenons à la définition d'une probabilité uniforme.

Soit X une variable aléatoire uniforme, alors :

$$X : \Omega \to I \subset \mathbb{R}$$

X est une variable aléatoire uniforme sur le domaine A si et seulement si

$$P\{X \in A\} = \frac{|A|}{|I|} \tag{1.35}$$

où $|.|$ représente la surface (volume) du domaine considéré.

La fonction de répartition de R est :

$$F_R(r) = P\{R \leq r\} = \frac{\frac{4}{3}\pi r^3}{\frac{4}{3}\pi d_{max}^3} = \frac{r^3}{d_{max}^3} \tag{1.36}$$

La densité de probabilité de la variable aléatoire R est donc :

$$p(\text{dist} = r) = \frac{\partial F_R(r)}{\partial r} = \frac{3r^2}{d_{max}^3} \tag{1.37}$$

Calcul de $P(\Delta = 1)$
On peut ici écrire :

$$\begin{aligned}
P(\Delta = 1) &= \int_0^{d_{max}} P(\Delta = 1|R = r)\ p(R = r)\ \mathrm{d}r && (1.38)\\
&= \frac{3}{d_{max}^3}\left[\int_0^{\frac{L_{eff}}{2}} r^2\ \mathrm{d}r + \frac{L_{eff}}{2}\int_{\frac{L_{eff}}{2}}^{d_{max}} r\ \mathrm{d}r\right] && (1.39)\\
&= \frac{3}{d_{max}^3}\left[\frac{L_{eff}^3}{24} + \frac{L_{eff}}{4}\left[d_{max}^2 - \frac{L_{eff}^2}{4}\right]\right] && (1.40)\\
&= \frac{3}{4}\frac{L_{eff}}{d_{max}} - \frac{1}{16}\frac{L_{eff}^3}{d_{max}^3} && (1.41)\\
&\triangleq K_c && (1.42)
\end{aligned}$$

Cette constante est notée K_c, et il est facile de vérifier que $K_c < 1$ avec l'hypothèse 1.28.

Résumé :

Hypothèses :

$$Leff << d_{max} \quad \text{et} \quad d < d_{max}$$

$$\Rightarrow \; r \approx d$$

La densité de probabilité que la cascade soit située à une distance r sachant qu'elle est détectable est :

$$p_R(d|\Delta = 1) = \begin{cases} 0 & \text{si} \quad d \leq 0 \\ \frac{3d^2}{d_{max}^3 Kc} & \text{si} \quad 0 < d \leq \frac{L_{eff}}{2} \\ \frac{3\ L_{eff}\ d}{2d_{max}^3 Kc} & \text{si} \quad \frac{L_{eff}}{2} < d \leq d_{max} \\ 0 & \text{si} \quad d > d_{max} \end{cases} \tag{1.43}$$

1.1.3.3 Fonction de répartition de la distance R d'une cascade détectable

Intéressons nous maintenant à la variable aléatoire R représentant la distance entre l'axe de la cascade supposée détectable et un hydrophone. Dans certaines parties de cette étude, le signal Antares sera considéré aléatoire. La seule variable aléatoire intervenant dans l'équation 1.20 est la distance de la cascade à l'hydrophone notée R. Il est donc nécessaire de déterminer la fonction de répartition de cette variable aléatoire afin d'être capable de programmer un générateur aléatoire de distance à la cascade.
La fonction de répartition de la variable aléatoire R est :

$$F_R(r) = \int_0^r p_R(u|\Delta = 1)\ \mathrm{d}u \tag{1.44}$$

Cas n°1 : $r \leq \frac{L_{eff}}{2}$

$$F_R(r) = \int_0^r \frac{3u^2}{d_{max}^3 Kc}\ \mathrm{d}u = \frac{r^3}{d_{max}^3 Kc} \tag{1.45}$$

$$= \frac{8\ r^3}{L_{eff}^3 + 6\ L_{eff}\left[d_{max}^2 - \frac{L_{eff}^2}{4}\right]} \tag{1.46}$$

Cas n°2 : $\frac{L_{eff}}{2} \leq r \leq d_{max}$

$$F_R(r) = \int_0^{\frac{L_{eff}}{2}} \frac{3u^2}{d_{max}^3 Kc}\,\mathrm{d}u + \int_{\frac{L_{eff}}{2}}^{r} \frac{3L_{eff}u}{2d_{max}^3 Kc}\,\mathrm{d}u = \frac{r^3}{d_{max}^3 Kc} \tag{1.47}$$

$$= \frac{L_{eff}^2 + 6\left[r^2 - \left(\frac{L_{eff}}{2}\right)^2\right]}{L_{eff}^2 + 6\left[d_{max}^2 - \left(\frac{L_{eff}}{2}\right)^2\right]} \tag{1.48}$$

Cas n°3 : $r > d_{max}$ Le résultat est direct :

$$F_R(r) = 1 \tag{1.49}$$

Résumé :

Hypothèses :

$$Leff << d_{max} \quad \text{et} \quad d < d_{max}$$

$$\Rightarrow \; r \approx d$$

La fonction de répartition de la variable aléatoire D associée à la distance cascade/hydrophone est :

$$F_D(d) = \begin{cases} 0 & \text{si} \quad d \leq 0 \\ \frac{8\ d^3}{L_{eff}^3 + 6\ L_{eff}\left[d_{max}^2 - \frac{L_{eff}^2}{4}\right]} & \text{si} \quad 0 < d \leq \frac{L_{eff}}{2} \\ \frac{L_{eff}^2 + 6\left[d^2 - \left(\frac{L_{eff}}{2}\right)^2\right]}{L_{eff}^2 + 6\left[d_{max}^2 - \left(\frac{L_{eff}}{2}\right)^2\right]} & \text{si} \quad \frac{L_{eff}}{2} < d \leq d_{max} \\ 1 & \text{si} \quad d > d_{max} \end{cases} \tag{1.50}$$

Génération des réalisations de la variable aléatoire D

Pour générer des réalisations de la variable aléatoire D, on utilise le résultat classique suivant :

Théorème : si U est une variable aléatoire uniforme sur l'intervalle $[0,1]$, si $F_D(d)$ est la fonction de répartition d'une variable aléatoire D, alors $F_D^{-1}(U)$ est une variable aléatoire suivant la loi de D.

Pour appliquer ce théorème, calculons la fonction quantile $F_D^{-1}(u)$ de D.

Cas n°1 : $d \leq 0$

La fonction $F_D(d)$ est croissante, donc $d \leq 0 \Rightarrow F_D(d) \leq F_D(0)$. Posons $u = F_d(d)$, il vient donc $u \leq 0$. Sur cet intervalle, $u = F_D(d) = 0$, il vient donc $F_D^{-1}(u) = 0$.

Cas n°2 : $0 \leq d \leq \frac{L_{eff}}{2}$

Dans ce cas, $0 \leq u \leq F_D\left(\frac{L_{eff}}{2}\right)$. Nous avons :

$$F_D\left(\frac{L_{eff}}{2}\right) = \frac{L_{eff}^2}{L_{eff}^2 + 6\ \left[d_{max}^2 - \frac{L_{eff}^2}{4}\right]} \tag{1.51}$$

Et dans cet intervalle, nous avons :

$$F_D^{-1}(u) = \frac{1}{2}\left[u\ \left(L_{eff}^3 + 6\ L_{eff}\left[d_{max}^2 - \frac{L_{eff}^2}{4}\right]\right)\right]^{1/3} \tag{1.52}$$

Cas n°3 : $\frac{L_{eff}}{2} \leq d \leq d_{max}$

Ici, $F_D(d_{max}) = 1$, et nous avons :

$$\frac{L_{eff}^2}{L_{eff}^2 + 6\ \left[d_{max}^2 - \frac{L_{eff}^2}{4}\right]} \leq u \leq 1 \tag{1.53}$$

La valeur de la fonction est :

$$F_D^{-1}(u) = \left[\frac{L_{eff}^2}{12} + \frac{u}{6}\left(L_{eff}^2 + 6\left[d_{max}^2 - \frac{L_{eff}^2}{4}\right]\right)\right]^{1/2} \tag{1.54}$$

Cas n°4 : $d > d_{max}$ Le résultat est direct :

$$F_D^{-1}(u) = 1 \tag{1.55}$$

Résumé :

La fonction quantile de la variable aléatoire D associée à la distance cascade/hydrophone est :

$$F_D^{-1}(u) = \begin{cases} 0 & \text{si} \quad u \leq 0 \\ \frac{1}{2}\left[u\ \left(L_{eff}^3 + 6\ L_{eff}\left[d_{max}^2 - \frac{L_{eff}^2}{4}\right]\right)\right]^{1/3} & \text{si} \quad 0 < u \leq \frac{L_{eff}^2}{L_{eff}^2 + 6\left[d_{max}^2 - \frac{L_{eff}^2}{4}\right]} \\ \left[\frac{L_{eff}^2}{12} + \frac{u}{6}\left(L_{eff}^2 + 6\left[d_{max}^2 - \frac{L_{eff}^2}{4}\right]\right)\right]^{1/2} & \text{si} \quad \frac{L_{eff}^2}{L_{eff}^2 + 6\left[d_{max}^2 - \frac{L_{eff}^2}{4}\right]} < u \leq 1 \\ 1 & \text{si} \quad u > 1 \end{cases} \tag{1.56}$$

1.2 Le bruit marin

Le signal acoustique précédent se propageant dans le milieu marin pollué par un bruit de fond toujours présent, deux approches sont considérées : dans la première, des enregistrements de bruit réel seront superposés au signal acoustique synthétique, dans la seconde, le bruit marin sera généré artificiellement.

1.2.1 Modélisation réelle

Dans cette approche, le bruit marin considéré est issu de campagnes de mesures effectuées au large de l'île du Levant, en mer Méditerranée au cours de l'année 2001 par des membres de l'équipe GESSY du laboratoire SIS.
Les mesures dont on dispose ont été effectuées pour des conditions de mer normale à agitée, avec un vent de force 15 à 30 noeuds. La fréquence d'échantillonnage d'origine est de 500 kHz. La fréquence de travail que nous avons décidé d'utiliser est de 1 MHz, il a donc été nécessaire de ré-échantillonner ce signal.

Il est maintenant possible de tracer l'histogramme des mesures de ce bruit. Nous y avons superposé l'allure d'une densité de probabilité gaussienne de même écart type. Ces résultats sont affichés sur la figure 1.15.

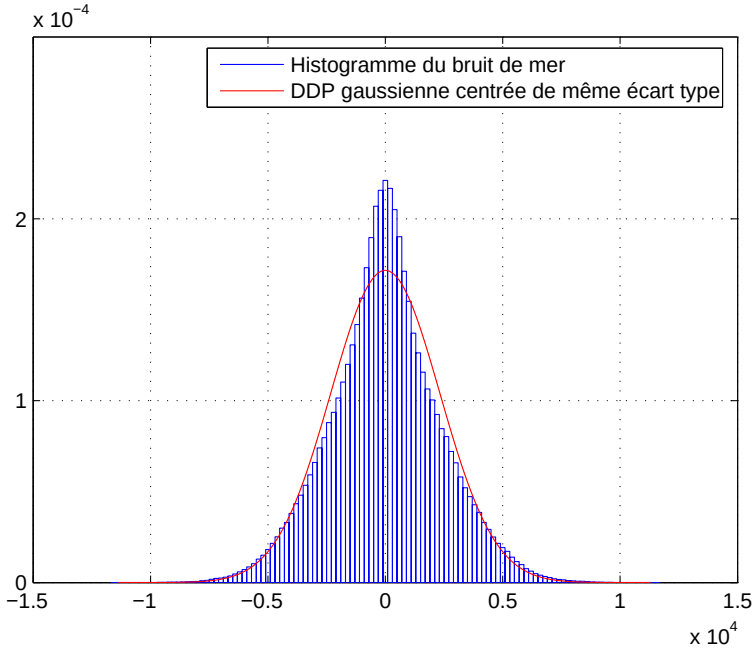

FIG. 1.15 – Histogramme du bruit de mer réel et densité de probabilité gaussienne.

Un test de gaussianité basé sur la *méthode des moments* de Van Der Waerden [5] a été appliqué à ce bruit. Ce test repose sur l'évaluation de deux coefficients H_1 et H_2 dits de "*symétrie*" et "*d'aplatissement*" à partir des moments d'ordre 2, 3 et 4 de la variable aléatoire à tester. On montre que, correctement normalisés, ces coefficients suivent une loi gaussienne de moyenne nulle et d'écart type égal à 1.
Pour les bruits de mer utilisés dans nos simulations, les valeurs de ces coefficients sont :

$$\begin{aligned} H_1 &= 0.6782 \\ H_2 &= 1.1021 \end{aligned}$$

Ces valeurs sont compatibles avec l'hypothèse de variable aléatoire gaussienne centrée, et de variance 1. On peut également remarquer que :

$$H_1^2 + H_2^2 \sim \chi_2^2$$

Or pour une variable aléatoire suivant une loi du χ_ν^2, la probabilité que cette variable soit supérieure à une certaine valeur est notée $P = prob(\chi_\nu^2 > x)$. Pour une valeur de $P = 1\%$, on a :

ν \ P	0.01
1	6.635
2	9.210
3	11.345
4	13.277
5	15.086
6	16.812
7	18.475
8	20.090

TAB. 1.1 – Valeurs de x pour la loi du χ^2_ν

Ici, le degrés de liberté est de 1, et nous avons :

$$H_1^2 + H_2^2 = 1.6748$$

Cette valeur est compatible avec le tableau précédent. La variable aléatoire $H_1^2 + H_2^2$ semble donc suivre une loi du χ^2_2, ce qui confirme le résultat précédent.

On affiche sur la figure 1.16 le spectre moyen du bruit de mer utilisé pour nos simulations. On trace également en pointillé les limites de la zone à l'intérieur de laquelle vont se trouver 90% des spectres des différentes réalisations considérées.

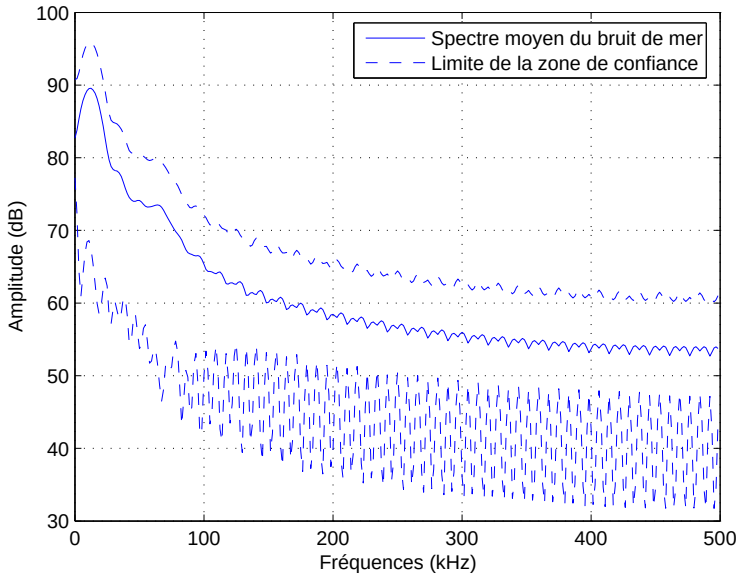

FIG. 1.16 – Spectre moyen du bruit de mer. Les pointillés représentent les limites de la zone de confiance à 90%.

1.2.2 Modélisation synthétique

Le bruit de mer ici est maintenant synthétisé. Afin que celui-ci garde des caractéristiques proches du bruit "réel" précédent, nous avons décidé que ce bruit synthétique (noté $b_m^{(s)}$) suivrait la loi suivante :

$$b_m^{(s)} \sim G\left(0, \mathbf{B}\right) \tag{1.57}$$

où $\mathbf{B}$ est la matrice de variance covariance du bruit de mer réel.

La marche à suivre pour générer un tel bruit disposant de la matrice $\mathbf{B}$ est simple : notons $\mathbf{V}$ la matrice contenant les vecteurs propres de $\mathbf{B}$ et $\mathbf{D}$ la matrice diagonale contenant les valeurs propres correspondantes. Si X est un vecteur gaussien de moyenne nulle et de matrice de variance covariance l'identité : $X \sim G\left(0, \mathbf{I}_N\right)$, alors le vecteur aléatoire $Y = \mathbf{V}\,\sqrt{\mathbf{D}}\,X$ suit la loi :

$$Y \sim G\left(0, \mathbf{B}\right) \tag{1.58}$$

Seule la connaissance de la matrice de variance covariance du bruit réel est nécessaire pour générer ce bruit de mer synthétique.

1.3 Protocole des simulations appliquées à la détection

Nous allons maintenant décrire les protocoles de simulation que nous allons utiliser pour comparer les performances de toutes les méthodes de détection étudiées ultérieurement. Les méthodes de détection étudiées dans les chapitres suivants sont basées sur deux hypothèses différentes concernant le signal acoustique :

- ⋆ Dans la première, le signal Antares est considéré comme déterministe, et correspond donc à une distance d fixée. Deux modèles sont alors utilisés pour en générer une réalisation : celui d'Askaryian et celui de Niess.

- ⋆ Dans la seconde, le signal Antares est modélisé par un signal aléatoire. La variable aléatoire D correspondant à la distance d séparant la cascade d'énergie à l'hydrophone suit la loi de probabilité définie par l'équation 1.43. Générer un grand nombre de distances d suivant cette loi permet donc d'obtenir autant de réalisations du signal acoustique. Dans cette hypothèse, les simulations seront générées uniquement à partir du modèle de Niess.

Comme nous l'avons vu précédemment, le bruit marin peut également être de deux sortes : réel ou simulé. L'un ou l'autre peut être utilisé indifféremment dans les protocoles de simulations décrits ci-dessous.

Les protocoles de simulations associés aux deux hypothèses précédentes sur le signal Antares sont les suivants :

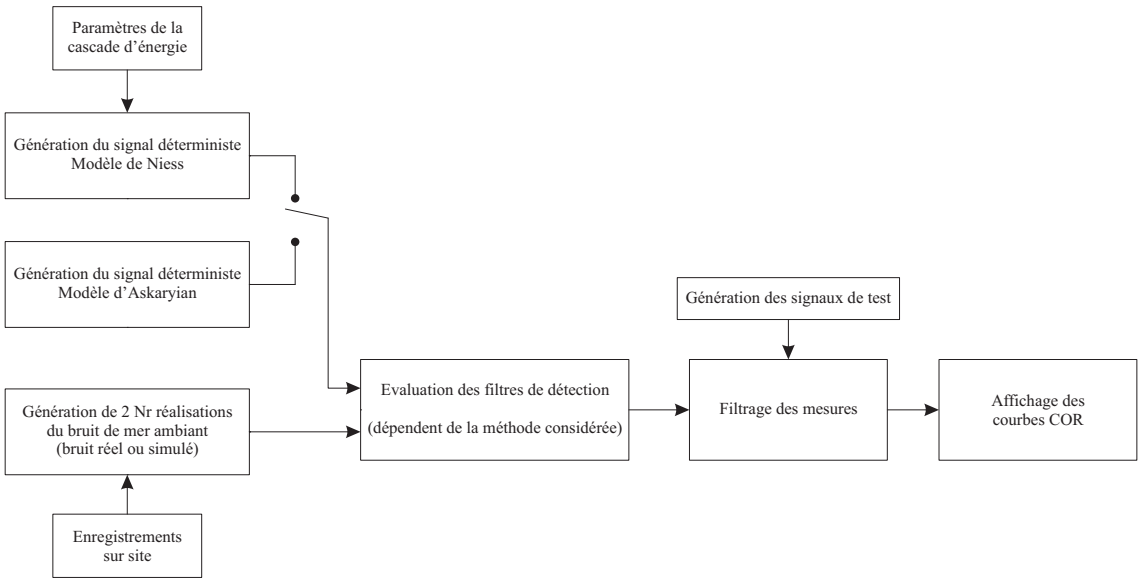

FIG. 1.17 – Schéma bloc du protocole de simulation appliqué à la détection d'un signal Antares déterministe.

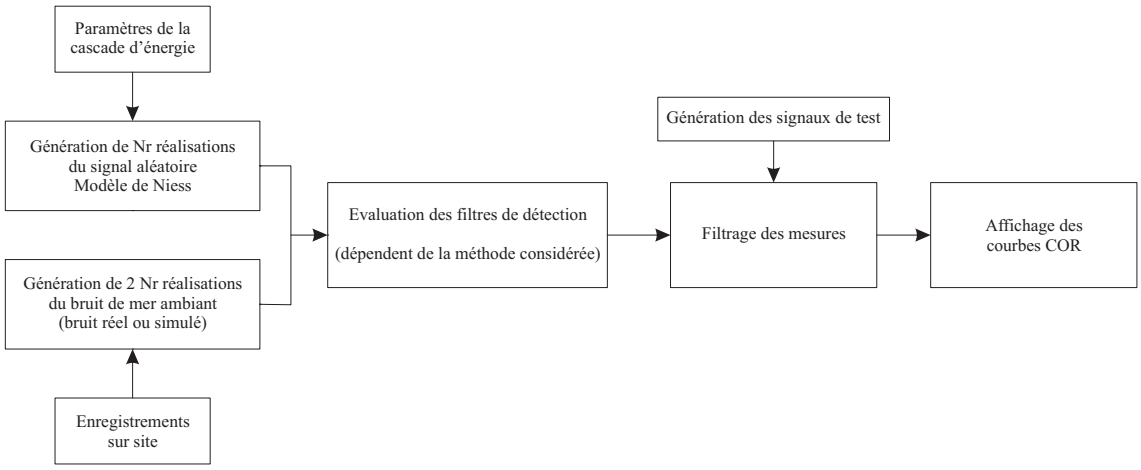

FIG. 1.18 – Schéma bloc du protocole de simulation appliqué à la détection d'un signal Antares aléatoire.

Chacun des blocs précédents sera décrit précisément dans les sections suivantes.

1.3.1 Paramètres de la cascade d'énergie et des simulations

Les valeurs de paramètres suivantes ont été utilisées pour les simulations :

⋆ **Paramètres généraux des simulations**

- Fréquence d'échantillonnage : $F_e = 1$ MHz.
- Nombre de réalisations : $N_r = 1000$.

- Longueur d'une réalisation : $N = 64$.

⋆ **Paramètres de la cascade d'énergie**
- Energie de la cascade : $E = 10^{19}$ eV.
- Longueur efficace : $L_{eff} = 10$ m.
- Distance minimale de détection : $D_{min} = 100$ m.
- Distance maximale de détection : $D_{max} = 500$ m.

1.3.2 Estimation des matrices de variance covariance

Certaines méthodes décrites dans les sections suivantes nécessitent l'utilisation de matrices de variance covariance des signaux mis en jeu. A chaque fois, celles-ci sont estimées par la méthode suivante :

Si s est un signal aléatoire échantillonné de dimension N, la matrice de variance/covariance $\mathbf{M}_s$ de s est estimée par l'estimateur sans biais suivant :

$$\mathbf{M}_s = \mathsf{E}\left(ss^T\right) \approx \frac{1}{N_r - 1} \sum_{i=1}^{N_r} \left(s_i - \mathsf{E}\left(s_i\right)\right)\left(s_i - \mathsf{E}\left(s_i\right)\right)^T$$

où N_r est le nombre de réalisations disponibles de s et s_i est la réalisation n°i de s.

1.3.3 Génération du signal acoustique déterministe

Le signal acoustique déterministe est caractérisé par une distance cascade / hydrophone fixée. Cette distance est notée d_0 dans cette section.

1.3.3.1 Modèle d'Askaryian

Le modèle d'Askaryian du signal acoustique est défini par l'équation 1.14. Cette modélisation dépend de deux paramètres physiques : l'énergie initiale de la cascade E et la distance de la cascade à l'hydrophone d. Une intégration numérique de l'équation 1.14 pour une distance d_0 mène à une réalisation du signal acoustique.

Filtrage dans la bande passante utile

Dans la pratique, tous les signaux reçus sur les hydrophones seront filtrés dans une bande de fréquences dite *bande utile*. La largeur de cette bande est déduite de la nature du signal que l'on cherche à détecter.

Dans le cas du modèle d'Askaryian, la plus grande des bandes passantes est obtenue pour le signal ayant parcouru la distance la plus faible possible. Lors de l'étude du modèle

d'Askaryian, nous avions remarqué que la bande passante des signaux était volontairement limitée à 50kHz (figure 1.19). La limite basse peut être estimée à 3 kHz.

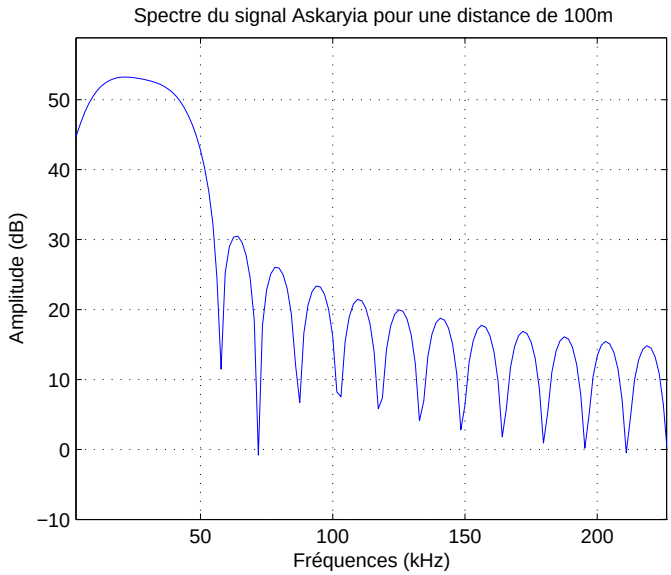

FIG. 1.19 – Spectre du signal acoustique pris à une distance de 100m - Modèle d'Askaryian.

La bande passante utile lors de l'utilisation du modèle d'Askaryian est donc :

$$\mathrm{BP}_{Askaryian} = [3\mathrm{kHz} \ - \ 50\mathrm{kHz}]$$

Tout signal reçu au niveau de l'hydrophone sera filtré dans cette bande.

1.3.3.2 Modèle de Niess

Pour le modèle de Niess, l'expression du signal acoustique déterministe pour une énergie E et pris à une distance d_0 est donné par l'équation 1.20 :

$$s_{\mathrm{Niess}}(t, d_0, E) = k.E.g_L(d_0).\mathrm{Re}\left(h_{d_0} * s_0\right)(t)$$

Filtrage dans la bande passante utile

Pour les mêmes raisons que celles évoquées dans le paragraphe précédent, il faut définir la bande passante utile lors de l'utilisation du modèle de Niess. On rappelle sur la figure 1.20 le spectre fréquentiel obtenu pour un signal ayant parcouru une distance de 100m.

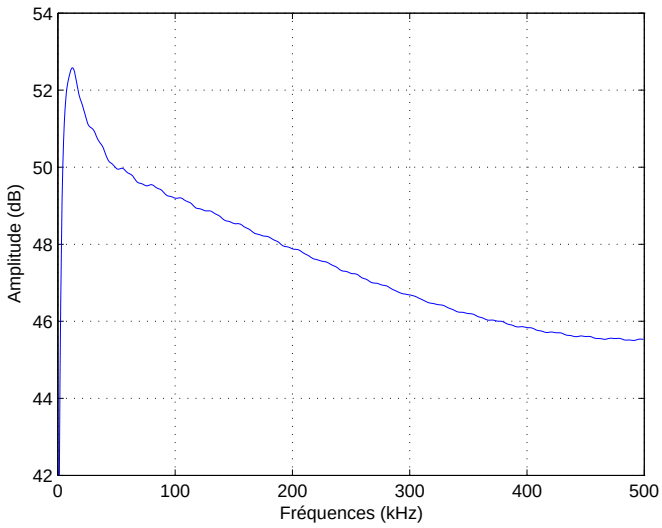

FIG. 1.20 – Spectre du signal acoustique pris à une distance de 100m - Modèle de Niess.

La bande passante à -3 dB est estimée à :

$$\mathrm{BP}_{Niess} = [7\mathrm{kHz} \ - \ 150\mathrm{kHz}]$$

On peut déjà remarquer qu'un signal issu du modèle CPPM possède une occupation spectrale plus importante que celle obtenue pour un signal issu du modèle d'Askaryian.

Tout signal reçu au niveau de l'hydrophone sera filtré dans cette bande.

Le filtre passe bas est de type FIR à ondulation constante (*equiripple*), et d'ordre 40. Le filtre passe haut est également un filtre FIR de type *least-square* et d'ordre 30.

1.3.4 Génération du signal Antares aléatoire

Dans l'hypothèse du signal acoustique aléatoire, nous avons décidé de n'utiliser que le modèle de Niess (équation 1.20). Celui-ci fait en effet appel à un modèle de propagation et d'atténuation due au milieu marin plus précis et actuel que celui utilisé dans le modèle d'Askaryian.

Pour pouvoir considérer le signal acoustique comme aléatoire, l'idée est de considérer ce signal comme une réalisation d'un signal aléatoire défini par l'équation :

$$S_{\mathrm{Niess}}(t, D, E) = k.E.g_L(D).\mathrm{Re}\left(h_D * s_0\right)(t)$$

où D est une variable aléatoire correspondant à la distance cascade / hydrophone et dont la densité de probabilité a été calculée plus tôt dans ce chapitre. Le reste de l'expression étant déterministe, il suffit donc, pour générer différentes réalisations du signal, de générer un lot de réalisation de D et d'en déduire les réalisations de S_A correspondantes. Chacune de ces réalisations est ensuite filtrée dans la bande passante utile décrite dans le paragraphe 1.3.3.2.

Les étapes sont résumées ci-dessous :

Génération des réalisations aléatoire du signal Antares
Modèle de Niess

- ⋆ Tirage de N_r réalisations de la variable aléatoire D représentant la distance cascade/hydrophone.
- ⋆ Evaluation des N_r signaux acoustiques correspondants.
- ⋆ Rééchantillonnage des signaux à la fréquence de travail.
- ⋆ Filtrage de chaque signal dans la bande passante utile 7 kHz - 150 kHz.
- ⋆ Récupération de la partie utile du signal sur N points.
- ⋆ Rangement de ces réalisations dans une matrice de taille $[N_r \times N]$.

1.3.5 Mise en forme du bruit de mer

Nous avons vu précédemment que le bruit de mer pouvait être de deux types : réel ou synthétique.

1.3.5.1 Bruit de mer issu d'enregistrements

Les étapes de génération du bruit de mer réel sont :

Le bruit de mer réel

- ⋆ Récupération des enregistrements issus de campagnes de mesures.
- ⋆ Rééchantillonnage à la fréquence de travail.
- ⋆ Filtrage dans la bande passante utile.
- ⋆ Découpage en 2 N_r réalisations de N points.
- ⋆ Rangement de ces réalisations dans une matrice de taille $[2\ N_r \times N]$.

La matrice de variance covariance déduite de ces réalisations est notée $\mathbf{B}_{\text{réel}}$.

1.3.5.2 Bruit de mer synthétique

Les étapes de génération du bruit de mer synthétique sont :

Générations d'un bruit de mer synthétique gaussien de matrice de variance covariance $\mathbf{B}_{\text{réel}}$

- ⋆ Génération de $2N_r$ vecteurs gaussiens $X_i \sim G(0, \mathbf{I}_N)$ de taille N.
- ⋆ Evaluation des valeurs propres (rangées dans la matrice diagonale $\mathbf{D}$) et des vecteurs propres (rangés dans la matrice $\mathbf{V}$) de la matrice de variance covariance $\mathbf{B}_{\text{réel}}$ déduite du bruit de mer réel précédent.
- ⋆ Projection des vecteurs X_i sur les vecteurs propres précédents : $Y_i = \mathbf{V}\ \sqrt{\mathbf{D}}\ X_i$.
- ⋆ Filtrage des vecteurs Y_i dans la bande passante utile.
- ⋆ Rangement de ces réalisations dans une matrice de taille $[2\ N_r \times N]$.

1.3.6 Notion de rapport signal-à-bruit

La notion de rapport signal-à-bruit est essentielle dans le problème de détection qui nous intéresse. Le rapport signal-à-bruit que nous utiliserons est le rapport entre la puissance

du signal P_s et la puissance du bruit P_b :

$$RSB = \frac{P_s}{P_b}$$

Son expression en décibel est :

$$RSB_{dB} = 10\log_{10}(RSB)$$

Dans la suite, nous serons amenés à travailler avec des signaux déterministes et des signaux aléatoires. Deux cas seront à considérer pour les définitions des puissances utilisées :

- ⋆ Le signal acoustique est déterministe et le bruit de mer est aléatoire.
  Rappelant que N est la longueur du signal s, la puissance de celui-ci est :

  $$P_s = \frac{s^T\ s}{N}$$

  et celle du bruit b de longueur N est :

  $$P_b = \mathsf{E}\left(\frac{b^T\ b}{N}\right)$$

  Elle peut être estimée en pratique par :

  $$\widehat{P_b} \approx \frac{1}{N_r}\sum_{i=0}^{N_r}\frac{b_i^T\ b_i}{N}$$

  Le rapport signal-à-bruit en pratique est donc estimé par :

  $$\widehat{RSB} = \frac{P_s}{\widehat{P_b}} \tag{1.59}$$

- ⋆ Le signal acoustique et le bruit de mer sont aléatoires.
  Les puissances du signal et du bruit sont :

  $$\begin{aligned} P_s &= \mathsf{E}\left(\frac{s^T\ s}{N}\right) \\ P_b &= \mathsf{E}\left(\frac{b^T\ b}{N}\right) \end{aligned}$$

Ces puissances sont estimées en pratique par :

$$\widehat{P_s} = \frac{1}{N_r}\sum_{i=0}^{N_r}\frac{s_i^T\ s_i}{N}$$
$$\widehat{P_b} = \frac{1}{N_r}\sum_{i=0}^{N_r}\frac{b_i^T\ b_i}{N}$$

Le rapport signal-à-bruit en pratique est donc estimé par :

$$\widehat{RSB} = \frac{\widehat{P_s}}{\widehat{P_b}}$$

1.3.7 Les signaux de test

Pour estimer les courbes COR représentatives de l'efficacité d'un détecteur, on utilise des signaux de test prédéfinis. Ces signaux doivent couvrir toutes les hypothèses que l'on peut rencontrer. Dans le cas d'un signal perturbé par un bruit, il y a deux hypothèses :

- ⋆ La mesure est constituée de la superposition additive du signal acoustique et du bruit de mer.
- ⋆ La mesure est constituée uniquement de bruit.

Le signal de test utilisé est donc composé d'un ensemble de mesures appartenant à chacune de ces deux hypothèses. Un exemple d'un tel signal est affiché sur la figure 1.21

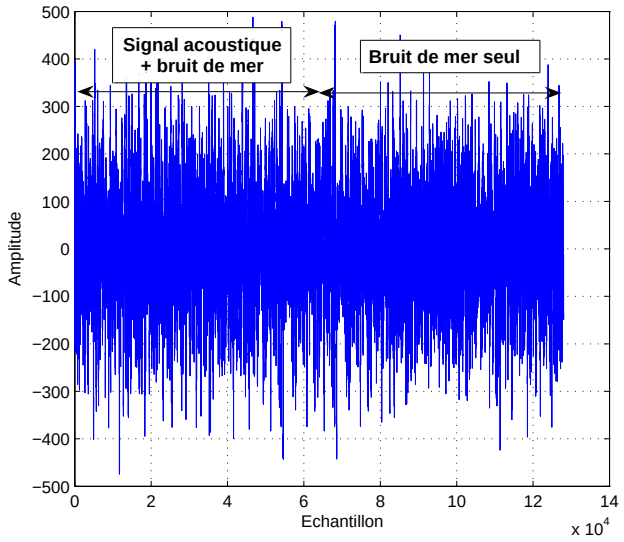

FIG. 1.21 – Signal de test pour un signal acoustique perturbé par un bruit de mer.

1.3.8 Test de détection

Le test de détection dépend de la méthode considérée. Pour chacune d'elle, nous avons défini une fonctionnelle de test. Le résultat de cette fonctionnelle appliquée au signal de test qui est comparée à un seuil prédéfini (dépendant de la valeur de P_{fa}). Le résultat de cette comparaison nous donne une information sur la présence du signal acoustique.

1.3.9 Calcul des valeurs de P_d, P_{fa} et affichage des courbes COR

Pour un seuil donné, connaissant la structure du signal de test, on estime les valeurs du couple $\{P_d, P_{fa}\}$ en comptant le nombre de fois où la valeur de la fonctionnelle dépasse la valeur du seuil.

Les performances du détecteur sont visualisées en affichant toutes les valeurs de P_d en fonction des valeurs de P_{fa} correspondantes. Ces courbes sont appelées courbes COR.

1.3.10 Précision des courbes COR

La précision des courbes COR peut être quantifiée par le calcul de la variance correspondante. Pour une P_{fa} fixée, on montre que la variance correspondant à une valeur p de P_d suit une loi de Bernoulli :

$$\sigma^2_{P_d} = \frac{p(1-p)}{N}$$

où N est le nombre de points utilisé pour estimer cette valeur de P_d. Connaissant cette formule, il est possible d'estimer l'intervalle de confiance à 95 % de la courbe COR. Pour une valeur de P_d donnée, cet intervalle est $[P_d - 2\sigma_{P_d}, P_d + 2\sigma_{P_d}]$. Ce résultat est illustré sur la figure 1.22.

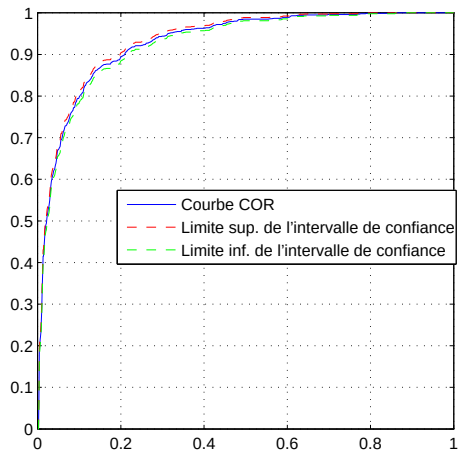

FIG. 1.22 – Intervalle de confiance à 95% pour les courbes COR.

1.4 Conclusion du chapitre

Seule disponible au début de cette thèse, la modélisation thermo-acoustique de G.A. Askariyan, développée au Lebedev Physical Institute en 1979 a été la première étudiée. Elle présente l'intérêt d'aboutir à l'équation de l'onde acoustique à un instant t et une distance d de la cascade d'énergie, mais le modèle de propagation mis en jeu est assez simple et n'est pas forcément adapté au milieu marin méditerranéen.

Les travaux effectués par Valentin Niess au CPPM de 2002 à 2005 ont permis d'aboutir à un second modèle du signal acoustique. L'étude de propagation est basée sur un modèle adapté à la mer Méditerranée et les simulations de cascades d'énergie ont pu être menées grâce aux moyens informatiques actuels. Les méthodes de détection que nous allons étudier dans les chapitres suivants se basent sur ces modèles de signaux et de bruits.

Dans un premier temps, le signal est considéré comme déterministe et est calculé à partir du modèle d'Askaryian et du modèle de Niess.

Dans un second temps, pour les méthodes de détection basées sur les notions de filtrage adapté stochastique, le signal acoustique est considéré comme aléatoire et les réalisations de celui-ci sont évaluées à partir du modèle de Niess.

Dans chacun des cas précédent, le bruit perturbateur sera de deux types : réel ou synthétique.

Lors de l'étude de ce protocole d'analyse, l'accent a été mis davantage sur les méthodes utilisées plutôt que sur les signaux considérés. *Ce protocole se veut ainsi pérenne et pourra être conservé et utilisé avec d'autres types de signaux (acoustiques ou bruits).*

Les différents cas traités pour la détection dans cette thése sont résumés ci-dessous :

⋆ Le signal acoustique est déterministe :

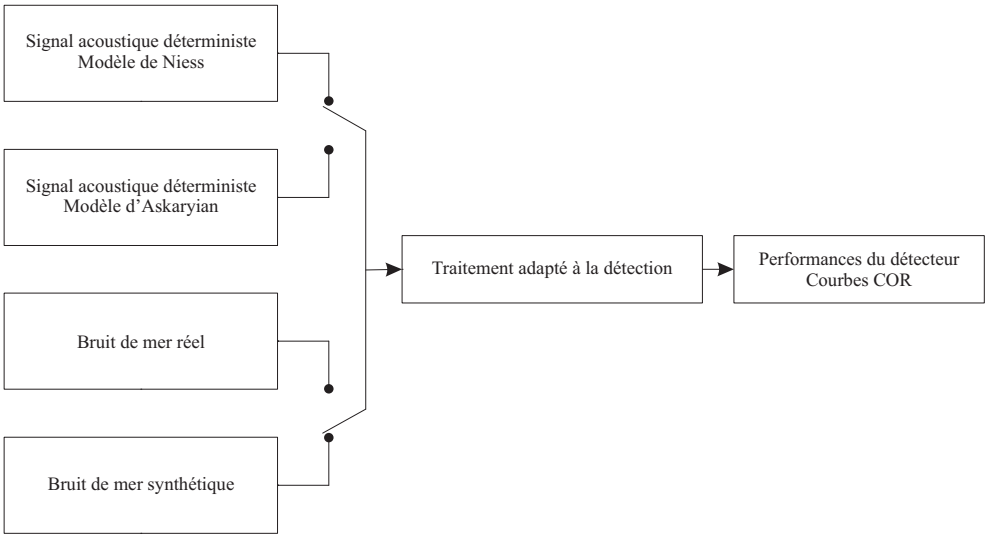

FIG. 1.23 – Résumé des cas traités pour un signal acoustique déterministe.

⋆ Le signal acoustique est aléatoire :

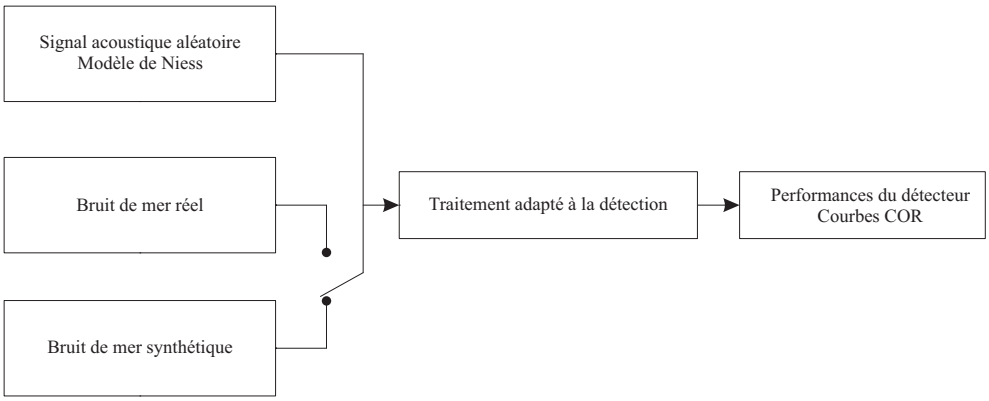

FIG. 1.24 – Résumé des cas traités pour un signal acoustique aléatoire.

Chapitre 2

Détection par rapport de vraisemblance

2.1 Introduction

Nous venons de définir plusieurs modélisations du signal acoustique généré par une cascade d'énergie. La première étape de ce travail de thèse est d'étudier un système capable de détecter la présence de ce signal acoustique dans le bruit du milieu. Dans ce chapitre, les hypothèses de travail seront simples : le signal acoustique est considéré comme déterministe et est perturbé par un unique bruit (le bruit de mer ambiant). Dans ce cadre, nous allons décrire et évaluer les performances de la méthode de détection classique basée un rapport de vraisemblance sous hypothèses gaussiennes.

2.2 Résumé

Le modèle thermo-dynamique proposé par Askaryian a été chronologiquement le premier à être à notre disposition. Avant d'étudier d'autres méthodes de détection, il est d'abord nécessaire d'étudier des méthodes dites classiques. Dans un premier temps, le signal acoustique est approximé par un polynôme déterministe d'ordre fixé *a priori*. Cette approximation nous permettra de mettre en place un test de détection basé sur le calcul du rapport de vraisemblance sous hypothèses gaussiennes.

Ce test sera calculé dans plusieurs configurations : bruit gaussien de variance connue et constante, bruit gaussien de variance inconnue et constance, bruit gaussien de variance inconnue et variable, puis enfin dans un cas proche de la réalité en prenant comme bruit ambiant un bruit de mer réel issu de campagnes de mesures.
Fournis par l'approche classique, les résultats issus de ce dernier cas serviront de référence aux méthodes de détection originales présentées dans le chapitre 3.

2.3 Approximation polynômiale du signal acoustique

L'idée est d'essayer d'écrire le signal acoustique échantillonné sous la forme :

$$S_A(\theta) = M \ . \ \theta \tag{2.1}$$

M est une matrice fixe et θ un vecteur dont les composantes caractérisent le signal S_A. Si le signal acoustique est pris dans sa version échantillonnée sur N points, nous avons à disposition le vecteur :

$$S_A = \begin{pmatrix} S_{A1} \\ S_{A2} \\ \vdots \\ S_{Ak} \\ \vdots \\ S_{A_N} \end{pmatrix} \tag{2.2}$$

Supposons que l'on puisse approximer ce signal par un polynôme d'ordre $\delta - 1$ fixé, l'approximation à l'instant $k\Delta t$ de S_{Ak} donne :

$$S_{Ak} = \sum_{j=0}^{\delta-1} a_j \ (k\Delta t)^j \tag{2.3}$$

Si on pose :

$$\mathbf{M}_\delta^{(k)}(\Delta t) = \left[(k\Delta t)^0 \ (k\Delta t)^1 \ (k\Delta t)^2 \ \ldots \ (k\Delta t)^{\delta-1} \right] \tag{2.4}$$

$$\theta = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{\delta-1} \end{pmatrix} \tag{2.5}$$

et que l'on note Δt le pas d'échantillonnage, on peut écrire :

$$S_{Ak}(\theta) = \mathbf{M}_\delta^{(k)}(\Delta t) \ . \ \theta \tag{2.6}$$

Notons maintenant

$$\mathbf{M}_\delta(\Delta t) = \begin{pmatrix} \mathbf{M}_\delta^{(1)}(\Delta t) \\ \mathbf{M}_\delta^{(2)}(\Delta t) \\ \vdots \\ \mathbf{M}_\delta^{(N)}(\Delta t) \end{pmatrix} \tag{2.7}$$

il vient :

$$S_A(\theta) = \mathbf{M}_\delta(\Delta t) \ . \ \theta \tag{2.8}$$

Il est donc possible d'écrire le signal acoustique échantillonné S_A sous la forme d'un produit de deux matrices. La matrice $\mathbf{M}_\delta(\Delta t)$ étant fixée, il reste maintenant à estimer les coefficients du polynôme d'approximation par l'estimateur des moindres carrés linéaire. Dans le cas sans bruit, nous avons :

$$\theta = \left[\mathbf{M}_\delta(\Delta t)^T \ \mathbf{M}_\delta(\Delta t)\right]^{-1} . \ \mathbf{M}_\delta(\Delta t)^T . \ S_A \tag{2.9}$$

2.3.1 Quel ordre pour le polynôme ?

Intéressons nous maintenant à l'ordre du polynôme d'approximation. On trace pour cela sur la figure 2.1 l'écart maximal relatif entre l'allure du signal acoustique réel et son approximation polynômiale en fonction du degré de celui-ci. Le signal réel considéré est basé sur le modèle thermo-dynamique d'Askariyan pour une distance de 200m.

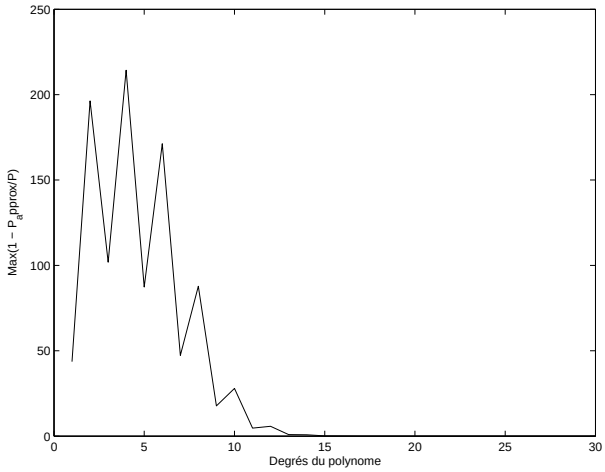

FIG. 2.1 – Degré du polynôme en fonction de l'écart maximal entre la le signal acoustique réel et son approximation polynômiale.

Un degré de 15 semble être une valeur adéquate, donc une valeur de $\delta = 16$. Pour un tel degré, on représente sur la figure 2.2 la superposition du signal acoustique et de son approximation.

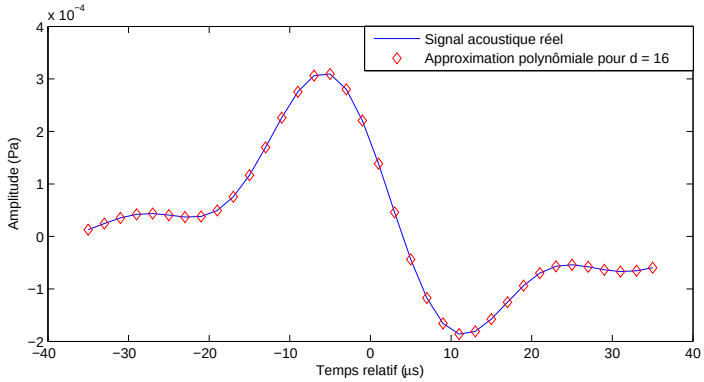

FIG. 2.2 – Comparaison du signal acoustique issu du modèle d'Askaryian et de son approximation - Distance de la cascade à l'hydrophone = 200m.

2.4 Cas d'un bruit gaussien centré de variance connue et constante

Le cas le plus simple pour le bruit ε est de considérer un bruit gaussien centré de variance σ^2 connue et constante. Si nous disposons de N mesures, nous noterons :

$$\varepsilon \sim G\left(0, \sigma^2 \ \mathbf{I}_n\right) \tag{2.10}$$

On suppose disposer également de N mesures issues d'un hydrophone regroupées dans le vecteur $x = (x_1, x_2, \ldots, x_N)^T$. Ce vecteur est une réalisation du vecteur aléatoire

$$X = \begin{pmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{pmatrix} \tag{2.11}$$

Nous nous retrouvons dans le cas classique de détection avec deux hypothèses :

$$\begin{aligned} H_0 &: X = \varepsilon \\ H_1 &: X = S_A + \varepsilon \end{aligned}$$

Dans l'hypothèse H_1, la mesure peut se réécrire $X = \mathbf{M}_\delta(\Delta t)\ \theta + \varepsilon$.

Dans les calculs suivants, θ sera remplacé par son estimateur des moindre carrés $\hat{\theta}$:

$$\widehat{\theta} = \left[\mathbf{M}_\delta(\Delta t)^T \ \mathbf{M}_\delta(\Delta t)\right]^{-1} . \ \mathbf{M}_\delta(\Delta t)^T . \ X \tag{2.12}$$

Le lemme de Neyman Pearson nous conduit à évaluer le rapport de vraisemblance généralisé :

$$\Lambda_G(x) = \frac{p(x|H_1)}{p(x|H_0)} \tag{2.13}$$

L'utilisation du logarithme népérien pour simplifier le calcul du test mène à :

$$\begin{aligned}
\ln\left[\Lambda_G(x)\right] &= -\frac{1}{2}\left(x - \mathbf{M}_\delta(\Delta t)\hat{\theta}\right)^T \frac{1}{\sigma^2}\left(x - \mathbf{M}_\delta(\Delta t)\hat{\theta}\right) + \frac{1}{2}\frac{x^T x}{\sigma^2} && (2.14)\\
& && (2.15)\\
&= \frac{1}{\sigma^2}x^T\mathbf{M}_\delta(\Delta t)\hat{\theta} - \frac{1}{2\sigma^2}\hat{\theta}^T\mathbf{M}_\delta(\Delta t)^T\mathbf{M}_\delta(\Delta t)\hat{\theta} && (2.16)
\end{aligned}$$

Si on remplace $\hat{\theta}$ par son expression, on obtient :

$$\ln\left[\Lambda_G(x)\right] = \frac{1}{2\sigma^2}x^T\mathbf{M}_\delta(\Delta t)\left[\mathbf{M}_\delta(\Delta t)^T\mathbf{M}_\delta(\Delta t)\right]^{-1}\mathbf{M}_\delta(\Delta t)^T x$$

Le test revient donc à calculer la valeur numérique suivante et à la comparer à un seuil

$$T_1(x) = \frac{1}{\sigma^2}x^T\mathbf{M}_\delta(\Delta t)\left[\mathbf{M}_\delta(\Delta t)^T\mathbf{M}_\delta(\Delta t)\right]^{-1}\mathbf{M}_\delta(\Delta t)^T x \underset{d_0}{\overset{d_1}{\gtrless}} seuil \tag{2.17}$$

Pour un ensemble de N mesures regroupées dans le vecteur de mesure x et pour un seuil donné, il prendre la décision d_1 si la valeur de $T_1(x)$ est supérieure à ce seuil, la décision d_0 le cas échéant.

Il est intéressant à ce stade de remarquer que la matrice $\mathbf{M}_\delta(\Delta t)\left[\mathbf{M}_\delta(\Delta t)^T\mathbf{M}_\delta(\Delta t)\right]^{-1}\mathbf{M}_\delta(\Delta t)^T$ ne dépend pas du pas d'échantillonnage choisi Δt. La démonstration en est faite dans l'annexe B. Le test peut donc être réécrit :

$$T_1(x) = \frac{1}{\sigma^2}x^T\mathbf{M}_\delta\left(\mathbf{M}_\delta^T\mathbf{M}_\delta\right)^{-1}\mathbf{M}_\delta^T x \underset{d_0}{\overset{d_1}{\gtrless}} seuil \tag{2.18}$$

2.4.1 Quelle valeur pour le seuil ?

La valeur du seuil à fixer est directement proportionnelle à la probabilité de fausse-alarme (probabilité de détecter le signal dans une mesure alors que celui-ci n'est pas présent, notée P_{fa}). Dans la pratique, on se fixe une valeur de P_{fa} qui caractérise les performances de notre détecteur, et on en déduit la valeur du seuil correspondante. Cette valeur de seuil nous permet de calculer la probabilité de détection associée, notée P_d. Le calcul du seuil à partir de la valeur de la P_{fa} nécessite la connaissance de la loi de la variable aléatoire $T_1(X)$ dont $T_1(x)$ est réalisation.

2.4.2 Calcul et tracé des courbes COR théoriques

2.4.2.1 Calcul du seuil à partir de la valeur de P_{fa}

Tout le problème est maintenant de déterminer la valeur du seuil correspondant à une valeur de P_{fa} donnée. Sous l'hypothèse H_0, la loi de la mesure X est :

$$X \sim G(0, \sigma^2 \mathbf{I}_N) \tag{2.19}$$

Il faut maintenant déterminer la loi de la variable aléatoire $T_1(X) = \frac{1}{\sigma^2} X^T \mathbf{M}_\delta \left(\mathbf{M}_\delta^T \mathbf{M}_\delta\right)^{-1} \mathbf{M}_\delta^T X$. On montre (annexe D) que :

$$\frac{1}{\sigma} \left(\mathbf{M}_\delta^T \mathbf{M}_\delta\right)^{-1/2} \mathbf{M}_\delta^T X \sim G(0, \mathbf{I}_N) \tag{2.20}$$

Ainsi,

$$\left\| \frac{1}{\sigma} \left(\mathbf{M}_\delta^T \mathbf{M}_\delta\right)^{-1/2} \mathbf{M}_\delta^T X \right\|^2 \sim \chi_d^2$$

où χ_d^2 représente l'allure de la densité de probabilité d'un Chi-deux centré à d degrés de liberté. Il faut maintenant remarquer que :

$$\begin{aligned} \left\| \frac{1}{\sigma} \left(\mathbf{M}_\delta^T \mathbf{M}_\delta\right)^{-1} \mathbf{M}_\delta^T X \right\|^2 &= \frac{1}{\sigma^2} X^T \mathbf{M}_\delta \left(\mathbf{M}_\delta^T \mathbf{M}_\delta\right)^{-1/2} \left(\mathbf{M}_\delta^T \mathbf{M}_\delta\right)^{-1/2} \mathbf{M}_\delta^T X & (2.21) \\ &= \frac{1}{\sigma^2} X^T \mathbf{M}_\delta \left(\mathbf{M}_\delta^T \mathbf{M}_\delta\right)^{-1} \mathbf{M}_\delta^T X & (2.22) \end{aligned}$$

On peut donc affirmer le résultat suivant :

$$\frac{1}{\sigma^2} X^T \mathbf{M}_\delta \left(\mathbf{M}_\delta^T \mathbf{M}_\delta\right)^{-1} \mathbf{M}_\delta^T X \sim \chi_\delta^2$$

Disposant d'une valeur de P_{fa}, il est maintenant possible de calculer numériquement la valeur du seuil correspondant (fig. 2.3).

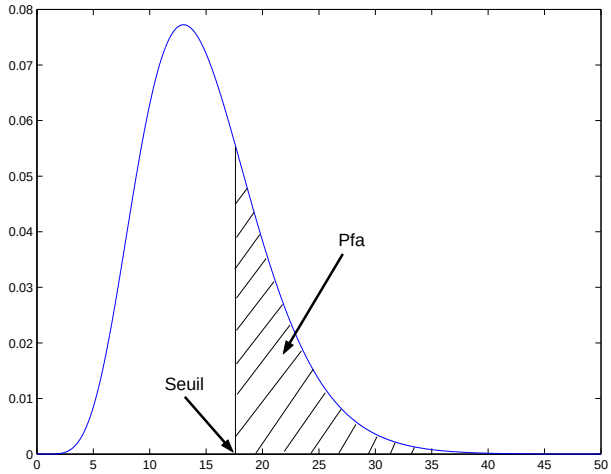

FIG. 2.3 – Calcul du seuil à partir d'une valeur de P_{fa} donnée (loi du χ^2_{15})

2.4.2.2 Calcul de P_d

La valeur du seuil évaluée, il est possible de calculer la valeur de P_d correspondante. Pour cela, il faut trouver la loi de $\frac{1}{\sigma^2}X^T\mathbf{M}_\delta\left(\mathbf{M}_\delta^T\mathbf{M}_\delta\right)^{-1}\mathbf{M}_\delta^T X$ en se plaçant sur l'hypothèse H_1. Nous avons donc :

$$X \sim G(\mathbf{M}_\delta\theta, \sigma^2\mathbf{I}_N) \tag{2.23}$$

On en déduit

$$\mathbf{M}_\delta^T X \sim G(\mathbf{M}_\delta^T\mathbf{M}_\delta\theta, \sigma^2\mathbf{M}_\delta^T\mathbf{M}_\delta) \tag{2.24}$$

En utilisant le même raisonnement que précédemment, il vient :

$$\frac{1}{\sigma}\left(\mathbf{M}_\delta^T\mathbf{M}_\delta\right)^{-1/2}\mathbf{M}_\delta^T X \sim G\left(\frac{1}{\sigma}\left(\mathbf{M}_\delta^T\mathbf{M}_\delta\right)^{-1/2}\mathbf{M}_\delta^T\mathbf{M}_\delta\theta, \mathbf{I}_N\right) \tag{2.25}$$

On remarque encore que :

$$\left\|\frac{1}{\sigma}\left(\mathbf{M}_\delta^T\mathbf{M}_\delta\right)^{-1}\mathbf{M}_\delta^T X\right\|^2 = \frac{1}{\sigma^2}X^T\mathbf{M}_\delta\left(\mathbf{M}_\delta^T\mathbf{M}_\delta\right)^{-1/2}\left(\mathbf{M}_\delta^T\mathbf{M}_\delta\right)^{-1/2}\mathbf{M}_\delta^T X \tag{2.26}$$

$$\tag{2.27}$$

$$= \frac{1}{\sigma^2}X^T\mathbf{M}_\delta\left(\mathbf{M}_\delta^T\mathbf{M}_\delta\right)^{-1}\mathbf{M}_\delta^T X \tag{2.28}$$

et on peut ainsi en déduire :

$$\frac{1}{\sigma^2}X^T\mathbf{M}_\delta\left(\mathbf{M}_\delta^T\mathbf{M}_\delta\right)^{-1}\mathbf{M}_\delta^T X \sim \chi^2_\delta(\lambda) \tag{2.29}$$

où λ est le décentrement de la fonction Chi-deux avec

$$\begin{aligned}
\lambda &= \frac{1}{\sigma^2}\left\|\left(\mathbf{M}_\delta^T\mathbf{M}_\delta\right)^{1/2}\theta\right\|^2 && (2.30)\\
&= \frac{1}{\sigma^2}\theta^T\left(\mathbf{M}_\delta^T\mathbf{M}_\delta\right)^{1/2}\left(\mathbf{M}_\delta^T\mathbf{M}_\delta\right)^{1/2}\theta && (2.31)\\
&= \frac{1}{\sigma^2}\theta^T(\mathbf{M}_\delta^T\mathbf{M}_\delta)\theta && (2.32)\\
&= \frac{1}{\sigma^2}\left\|\mathbf{M}_\delta\theta\right\|^2 && (2.33)
\end{aligned}$$

Connaissant la valeur du seuil, on calcule la valeur de la P_d par la formule suivante :

$$P_d = \int_{seuil}^{\infty} p_{\chi_\delta^2(\lambda)}(u)\ du \tag{2.34}$$

On peut remarquer que le terme $\frac{\|\mathbf{M}_\delta\theta\|^2}{N\sigma^2}$ correspond au rapport signal-à-bruit (noté RSB). La valeur du décentrement λ influe donc directement sur celui-ci.

2.4.2.3 Résumé :

Le test retenu est :

$$T_1(x) = \frac{1}{\sigma^2}x^T\mathbf{M}_\delta\left(\mathbf{M}_\delta^T\mathbf{M}_\delta\right)^{-1}\mathbf{M}_\delta^T x \underset{d_0}{\overset{d_1}{\gtrless}} \text{seuil} \tag{2.35}$$

Tracé des courbes COR

Il est maintenant possible de tracer les courbes COR (Caractéristiques Opérationnelles du Récepteur) du test. Ces courbes permettent d'étudier les variations de la spécificité et de la sensibilité d'un test pour différentes valeurs du seuil de discrimination. Le terme de courbe COR peut être envisagé comme une "courbe de caractéristiques d'efficacité". Chaque courbe correspond à une valeur de RSB distincte (obtenue en faisant varier la valeur de λ). La valeur d, proportionnelle au degré du polynôme, est fixée à 15.
Le tableau C donné dans l'annexe C nous donne pour chaque valeur de λ les valeurs de variance et d'écart-type correspondants. L'amplitude maximale du signal S_A sera noté S_A^{max}.

Les courbes COR théoriques sont présentées sur la figure 2.4.

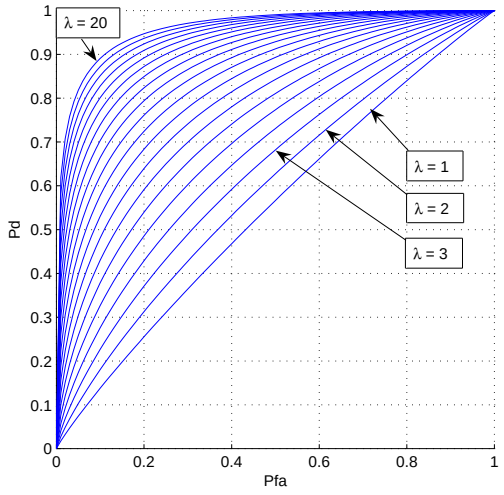

FIG. 2.4 – Courbes COR théoriques pour λ variant de 1 à 20

Vérification par Monte-Carlo

On peut maintenant vérifier la validité de ces courbes en utilisant la méthode de Monte-Carlo. On se fixe pour cela des valeurs de λ et de P_{fa}, et on calcule la valeur de P_d correspondante. Ces valeurs sont ensuite reportées sur les courbes COR théoriques calculées précédemment.

Les résultats sont reportés sur la figure 2.5. On remarque que les valeurs issues des simulations empiriques se positionnent sur les courbes théoriques, le modèle est donc validé.

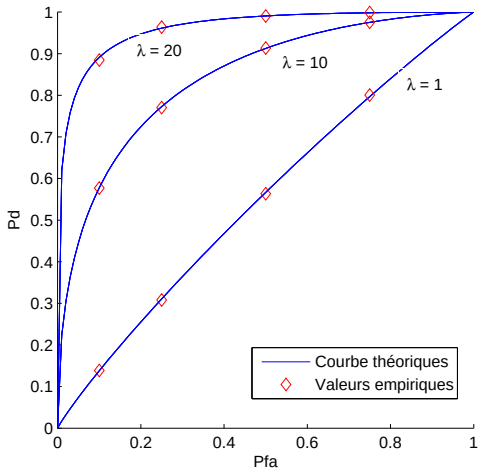

FIG. 2.5 – Courbes COR théoriques et mesures empiriques - λ connu

2.5 Cas d'un bruit gaussien centré de variance inconnue et constante

Reprenons le cas précédent en considérant que la variance du bruit σ^2 est maintenant inconnue.

Sous l'hypothèse H_0, l'inconnu est

$$\theta_0 = \sigma_0^2 \tag{2.36}$$

et dans l'hypothèse H_1, l'inconnu est

$$\theta_1 = \begin{pmatrix} \theta \\ \sigma_1^2 \end{pmatrix} \tag{2.37}$$

Dans les cas des ces deux hypothèses, et notant x une réalisation de la variable aléatoire X correspondant à la mesure, il est possible d'estimer ces inconnues :

- ⋆ Sous H_0 : $\hat{\sigma}_0^2 = \frac{1}{N}\|x\|^2$
- ⋆ Sous H_1 : $\begin{cases} \hat{\theta} & = & \left(\mathbf{M}_\delta^T \mathbf{M}_\delta\right)^{-1} \mathbf{M}_\delta^T x \\ \hat{\sigma}_1^2 & = & \frac{1}{N-\delta}\left\|x - \mathbf{M}_\delta \hat{\theta}\right\|^2 \end{cases}$

Il faut maintenant étudier le rapport de vraisemblance $\Lambda_G(x) = \frac{p(x|H_1)}{p(x|H_0)}$.
On peut écrire :

$$p(x|H_1) = \frac{1}{(2\pi\hat{\sigma}_1)^{N/2}}\, e^{\left(-\frac{1}{2}\frac{\|x-\mathbf{M}_\delta\hat{\theta}\|^2}{\hat{\sigma}_1^2}\right)} = \frac{C_1}{\hat{\sigma}_1^{N/2}} \tag{2.38}$$

et

$$p(x|H_0) = \frac{1}{(2\pi\hat{\sigma}_0)^{N/2}}\, e^{\left(-\frac{1}{2}\frac{\|x-\mathbf{M}_\delta\hat{\theta}\|^2}{\hat{\sigma}_0^2}\right)} = \frac{C_0}{\hat{\sigma}_0^{N/2}} \tag{2.39}$$

Ainsi

$$\Lambda_G(x) = \frac{C_1}{C_0}\frac{\hat{\sigma}_0^{N/2}}{\hat{\sigma}_1^{N/2}} \tag{2.40}$$

Le test revient à utiliser la règle de décision suivante :

$$\Lambda_G(x) \underset{d_0}{\overset{d_1}{\gtrless}} t \tag{2.41}$$

où t est le seuil de détection. On peut écrire :

$$\frac{\hat{\sigma}_0^{N/2}}{\hat{\sigma}_1^{N/2}} \underset{d_0}{\overset{d_1}{\gtrless}} t_1 \tag{2.42}$$

$$\frac{\hat{\sigma}_0^2}{\hat{\sigma}_1^2} \underset{d_0}{\overset{d_1}{\gtrless}} t_2 \tag{2.43}$$

$$\tag{2.44}$$

$$\frac{\|x\|^2}{\left\|x - \mathbf{M}_\delta \hat{\theta}\right\|^2} \underset{d_0}{\overset{d_1}{\gtrless}} t_2 \tag{2.45}$$

$$\frac{\left\|x - \mathbf{M}_\delta \hat{\theta}\right\|^2}{\|x\|^2} \underset{d_1}{\overset{d_0}{\gtrless}} t_3 \tag{2.46}$$

Or on sait que $\mathbf{M}_\delta \hat{\theta} = \mathbf{M}_\delta \left(\mathbf{M}_\delta^T \mathbf{M}_\delta\right)^{-1} \mathbf{M}_\delta^T x$.
Posons maintenant

$$\Pi_1 \triangleq \mathbf{M}_\delta \left(\mathbf{M}_\delta^T \mathbf{M}_\delta\right)^{-1} \mathbf{M}_\delta^T \tag{2.47}$$

On remarque que $\Pi_1^T = \Pi_1$ et que $\Pi_1^2 = \Pi_1$. Il s'agit donc de la matrice d'un projecteur. On peut donc écrire

$$\left\|x - \mathbf{M}_\delta \hat{\theta}\right\|^2 = \|x - \Pi_1 x\|^2 \tag{2.48}$$

$$= \|(\mathbf{I}_N - \Pi_1)x\|^2 \tag{2.49}$$

En parallèle, notons

$$\Pi_2 \triangleq \mathbf{I}_N - \Pi_1 \tag{2.50}$$

Nous avons encore $\Pi_2^T = \Pi_2$ et $\Pi_2^2 = \Pi_2$. Il s'agit donc encore de la matrice d'un projecteur. Il vient

$$\left\|x - \mathbf{M}_\delta \hat{\theta}\right\|^2 = \|\Pi_2 x\|^2 \tag{2.51}$$

$$= x^T \Pi_2^T \Pi_2 x \tag{2.52}$$

$$= x^T \Pi_2 x \tag{2.53}$$

Le test 2.46 peut maintenant se réécrire de la manière suivante :

$$\frac{x^T \Pi_2 x}{x^T x} \underset{d_0}{\overset{d_1}{\gtrless}} t_3 \tag{2.54}$$

$$1 - \frac{x^T \Pi_1 x}{x^T x} \underset{d_0}{\overset{d_1}{\gtrless}} t_3 \tag{2.55}$$

$$\frac{x^T \Pi_1 x}{x^T x} \underset{d_0}{\overset{d_1}{\gtrless}} t_4 \tag{2.56}$$

ie

$$\frac{x^T \mathbf{M}_\delta (\mathbf{M}_\delta^T \mathbf{M}_\delta)^{-1} \mathbf{M}_\delta^T x}{x^T x} \underset{d_0}{\overset{d_1}{\gtrless}} t_4 \tag{2.57}$$

2.5.1 Calcul et tracé des courbes COR théoriques

Pour choisir correctement le seuil t_4 en fonction d'une valeur de P_{fa} donnée, il est nécessaire de déterminer la loi du test $\frac{X^T \mathbf{M}_\delta (\mathbf{M}_\delta^T \mathbf{M}_\delta)^{-1} \mathbf{M}_\delta^T X}{X^T X}$.

En utilisant les résultats de la section précédente, nous connaissons les lois des variables aléatoires correspondant au numérateur et dénominateur précédents.

$$\left\{ \begin{array}{rcl} X^T \mathbf{M}_\delta (\mathbf{M}_\delta^T \mathbf{M}_\delta)^{-1} \mathbf{M}_\delta^T X & \sim & \chi_\delta^2 \\ X^T X & \sim & \chi_\delta^2 \end{array} \right. \tag{2.58}$$

Ces variables aléatoires n'étant pas indépendantes (démonstration annexe E), il n'est donc pas possible d'en déduire directement la loi du test. Repartons de l'équation 2.56 et essayons de nous ramener à un rapport de variables aléatoires indépendantes.

$$\begin{aligned}
\frac{x^T \Pi_1 x}{x^T x} &\underset{d_0}{\overset{d_1}{\gtrless}} t_4 \\
\frac{x^T x}{x^T \Pi_1 x} &\underset{d_1}{\overset{d_0}{\gtrless}} t_5 \\
\frac{x^T (\mathbf{I}_N - \Pi_1 + \Pi_1) x}{x^T \Pi_1 x} &\underset{d_1}{\overset{d_0}{\gtrless}} t_5 \\
\frac{x^T (\mathbf{I}_N - \Pi_1) x}{x^T \Pi_1 x} + \frac{x^T \Pi_1 x}{x^T \Pi_1 x} &\underset{d_1}{\overset{d_0}{\gtrless}} t_5 \\
\frac{x^T (\mathbf{I}_N - \Pi_1) x}{x^T \Pi_1 x} + 1 &\underset{d_1}{\overset{d_0}{\gtrless}} t_5 \\
\frac{x^T (\mathbf{I}_N - \Pi_1) x}{x^T \Pi_1 x} &\underset{d_1}{\overset{d_0}{\gtrless}} t_6 \\
\frac{x^T \Pi_1 x}{x^T (\mathbf{I}_N - \Pi_1) x} + 1 &\underset{d_0}{\overset{d_1}{\gtrless}} t_7 \\
\frac{x^T \Pi_1 x}{x^T (\mathbf{I}_N - \Pi_1) x} &\underset{d_0}{\overset{d_1}{\gtrless}} t_8
\end{aligned} \tag{2.59}$$

On démontre (annexe F) que numérateur et dénominateur sont des variables aléatoires indépendantes et qui suivent une loi du χ^2_δ.
Nous noterons pour la suite :

$$T(X) \triangleq \frac{X^T \Pi_1 X}{X^T(\mathbf{I}_N - \Pi_1)X} \tag{2.60}$$

$$N(X) \triangleq X^T \Pi_1 X = \|\Pi_1 X\|^2 \tag{2.61}$$

$$D(X) \triangleq X^T(\mathbf{I}_N - \Pi_1)X = \|(\mathbf{I}_N - \Pi_1)X\|^2 \tag{2.62}$$

Loi de $N(X)$

On se retrouve ici dans le cas précédent, et nous pouvons affirmer que :

- Sous H_0, $\frac{1}{\sigma^2}N(X) \sim \chi^2_\delta$
- Sous H_1, $\frac{1}{\sigma^2}N(X) \sim \chi^2_\delta\left(\lambda = \frac{\|\mathbf{M}_\delta \theta\|^2}{\sigma^2}\right)$

Loi de $D(X)$

$\Pi_1 X$ est un projecteur sur un sous-espace vectoriel (noté SEV) de dimension δ.
$\Pi_2 = \mathbf{I}_N - \Pi_1$ est également un projecteur sur un SEV de dimension $N - \delta$.

- Sous H_0,
$$\frac{1}{\sigma^2}\|(\mathbf{I}_N - \Pi_1)X\|^2 \sim \chi^2_{N-\delta}$$

- Sous H_1,
$$\frac{1}{\sigma^2}\|(\mathbf{I}_N - \Pi_1)X\|^2 \sim \chi^2_{N-\delta}(\lambda)$$

Il reste à calculer le décentrement λ.
Pour cela, on remarque que, quel que soit le vecteur X,

$$X = (\mathbf{I}_N - \Pi_1)X + \Pi_1 X$$

Ainsi :

$$\|X\|^2 = \|(\mathbf{I}_N - \Pi_1)X\|^2 + \|\Pi_1 X\|^2 \tag{2.63}$$

$$E\left\{\|X\|^2\right\} = E\left\{\|(\mathbf{I}_N - \Pi_1)X\|^2\right\} + E\left\{\|\Pi_1 X\|^2\right\} \tag{2.64}$$

$$E\left\{\frac{\|X\|^2}{\sigma^2}\right\} = E\left\{\frac{\|(\mathbf{I}_N - \Pi_1)X\|^2}{\sigma^2}\right\} + E\left\{\frac{\|\Pi_1 X\|^2}{\sigma^2}\right\} \tag{2.65}$$

On peut donc écrire :

$$N + \frac{\|\mathbf{M}_\delta \theta\|^2}{\sigma^2} = (N - \delta) + \lambda + \delta + \frac{\|\mathbf{M}_\delta \theta\|^2}{\sigma^2} \tag{2.66}$$

et ainsi

$$\lambda = 0$$

<u>Conclusion :</u> sous H_0 et sous H_1, $\frac{1}{\sigma^2}\|(\mathbf{I}_N - \Pi_1)X\|^2$ suit la loi du χ^2_{N-d} centrée.

Loi de $T(X)$

Commençons par remarquer que

$$T(X) = \frac{N(x)/\sigma^2}{D(x)/\sigma^2}$$

- Sous H_0,
$$\frac{N(X)}{\sigma^2} \sim \chi^2_\delta$$
- Sous H_1,
$$\frac{D(X)}{\sigma^2} \sim \chi^2_{N-\delta}$$

Ces variables aléatoires étant indépendantes, on peut conclure quand à la loi suivie par $T(X)$:

- Sous H_0,
$$T(X) \sim \beta_2\left(\frac{\delta}{2}, \frac{N-\delta}{2}, 0\right)$$
- Sous H_1,
$$T(X) \sim \beta_2\left(\frac{\delta}{2}, \frac{N-\delta}{2}, \gamma\right)$$

où le décentrement est $\gamma = \frac{\|\mathbf{M}_\delta \theta\|^2}{\sigma^2}$.

Il est également possible d'utiliser les lois de Fisher-Snédécor :

$$\begin{aligned}\frac{\frac{1}{N}\,N(X)}{\frac{1}{N-\delta}\,D(X)} = \frac{N-\delta}{\delta}\,T(X) &\sim F(\delta, N-\delta) \quad \text{sous } H_0 \\ &\sim F(d, N-\delta, \gamma) \quad \text{sous } H_1\end{aligned}$$

où γ est le décentrement de la fonction F dont la fonction de répartition est calculée numériquement.

2.5.1.1 Résumé :

Le test retenu lorsque λ est inconnu est :

$$T_2(x) = \frac{N-\delta}{\delta} \frac{x^T \Pi_1 x}{x^T(\mathbf{I}_N - \Pi_1)x} \underset{d_0}{\overset{d_1}{\gtrless}} \text{seuil} \tag{2.67}$$

avec

$$\Pi_1 \triangleq \mathbf{M}_\delta(\mathbf{M}_\delta^T \mathbf{M}_\delta)^{-1}\mathbf{M}_\delta^T \tag{2.68}$$

.

Les courbes COR théoriques pour différentes valeurs de γ sont présentées sur la figure 2.6.

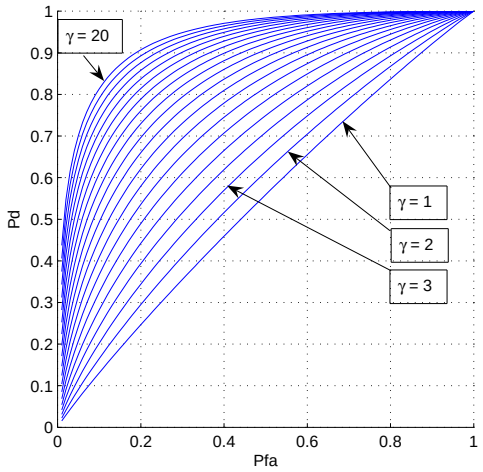

FIG. 2.6 – Courbes COR pour γ variant de 1 à 20

2.5.1.2 Vérification par Monte-Carlo

Le protocole de simulation est le même que celui de la section précédente. Là encore, les résultats issus des simulations de Monte-Carlo sont identiques aux résultats théoriques. Ceux-ci sont présentés sur la figure 2.7.

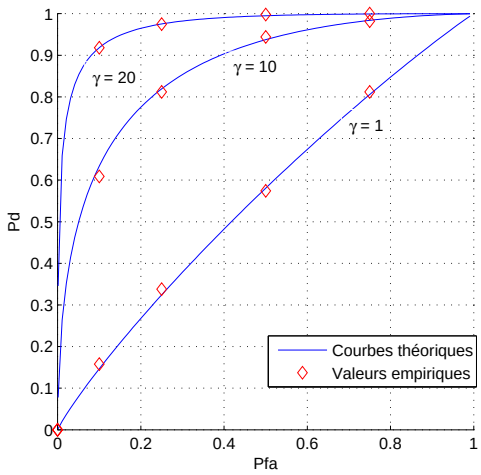

FIG. 2.7 – Courbes COR pour γ variant de 1 à 20 - λ inconnu

2.5.1.2.1 Cas particulier : la variance n'est pas constante Il peut être intéressant de tester la robustesse du test précédent à une variation de la variance. La variance σ^2 suit ici une loi uniforme centrée autour de la valeur σ_0^2 précédente et donc les valeurs extrêmes sont 1.5 σ_0^2 et 0.5 σ_0^2. Les résultats sont présentés sur la figure 2.8.

Même si l'on observe un tassement des courbes COR, les résultats issus des simulations sont proche des résultats théoriques malgré le changement d'une des hypothèses de départ. Le test est donc robuste à des variations de la variance autour de sa valeur moyenne.

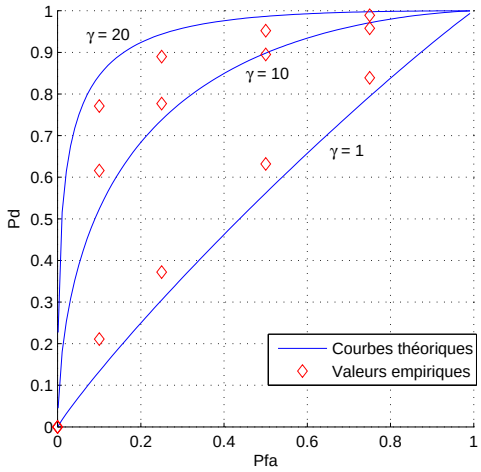

FIG. 2.8 – Courbes COR théoriques et mesures empiriques avec une variance inconnue et variable

2.6 Cas d'un bruit gaussien de matrice de variance covariance estimée

Plaçons-nous maintenant dans un cas plus général. Le bruit de mesure ε est toujours supposé gaussien mais n'est plus indépendant. Sa matrice de covariance sera notée $\mathbf{B}$ et supposée connue et estimée par ailleurs à partir de mesures de bruits de mer réels. La méthode utilisée pour calculer pratiquement ce genre de matrice est décrite dans le chapitre 1.

Nous sommes encore confrontés au problème de détection classique :

$$\begin{aligned} H_0 &: X = \varepsilon \\ H_1 &: X = \mathbf{M}_\delta \theta + \varepsilon \end{aligned}$$

Le calcul du rapport de vraisemblance (ou plutôt de son logarithme) mène au résultat suivant :

$$Ln\left[\frac{p(x|H_1)}{p(x|H_0)}\right] \propto x^T\mathbf{B}^{-1}x - (x - \mathbf{M}_\delta\theta)^T\,\mathbf{B}^{-1}\,(x - \mathbf{M}_\delta\theta) \tag{2.69}$$

$$= x^T\mathbf{B}^{-1}\mathbf{M}_\delta\theta + \theta^T\mathbf{M}_\delta^T\mathbf{B}^{-1}x - \theta^T\mathbf{M}_\delta^T\mathbf{B}^{-1}\mathbf{M}_\delta\theta \tag{2.70}$$

θ étant inconnu, il convient de le remplacer dans l'expression précédente par son estimé du maximum de vraisemblance. On aboutit ainsi à l'expression du nouveau test de détection :

$$T_3(x) = x^T\mathbf{B}^{-1}\mathbf{M}_\delta\left(\mathbf{M}_\delta^T\mathbf{B}^{-1}\mathbf{M}_\delta\right)^{-1}\mathbf{M}_\delta^T\mathbf{B}^{-1}x \underset{d_0}{\overset{d_1}{\gtrless}} \text{seuil} \tag{2.71}$$

Résumé :

Le test retenu lorsque le bruit est caractérisé par sa matrice de variance/covariance $\mathbf{B}$ est :

$$T_3(x) = x^T\mathbf{B}^{-1}\mathbf{M}_\delta\left(\mathbf{M}_\delta^T\mathbf{B}^{-1}\mathbf{M}_\delta\right)^{-1}\mathbf{M}_\delta^T\mathbf{B}^{-1}x \underset{d_0}{\overset{d_1}{\gtrless}} \text{seuil} \tag{2.72}$$

Cas particulier : Si la matrice de covariance du bruit peut s'écrire $\mathbf{B} = \sigma^2\mathbf{I}_N$, on se ramène au premier cas étudié et on retombe bien sur le test 2.35.

2.7 Détection en situation réelle

En accord avec les différents protocoles de simulations décrits dans le premier chapitre, nous allons tester le détecteur précédent dans les deux configurations suivantes :

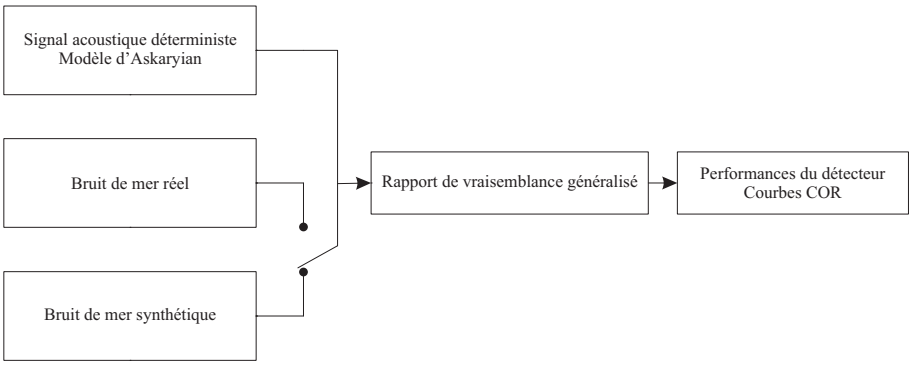

FIG. 2.9 – Configuration 1 : signal acoustique d'Askaryian perturbé par deux types de bruits.

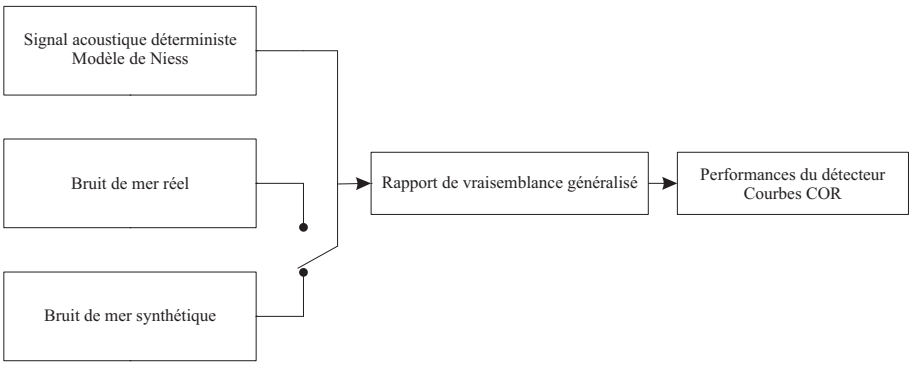

FIG. 2.10 – Configuration 2 : signal acoustique de Niess perturbé par deux types de bruits.

Les performances dans ces deux cas seront présentées sous forme de courbes COR.

2.7.1 Cas du signal acoustique d'Askaryian

Comme décrit sur la figure 2.9, le signal acoustique est considéré comme déterministe et est calculé à partir du modèle d'Askaryian décrit dans le chapitre précédent. Deux distances vont être considérées. La comparaison des courbes COR correspondant à chacune de ces distances nous permettra de déduire des résultats sur l'influence de la propagation du signal et sur la validité du modèle d'approximation dans chacun des cas. On rappelle (c.f. premier chapitre) que la distance maximale de détection a été fixée à 500m.

2.7.1.1 Première distance de détection : $d = 100m$

Le signal acoustique est calculé pour une distance de 100m. Deux cas sont testés : dans le premier, le bruit est gaussien, dans le second, le bruit est réel. La matrice de covariance **B** du bruit de mer est la même dans les deux cas.

Les performances du détecteur sont présentés sur la figure 2.11.

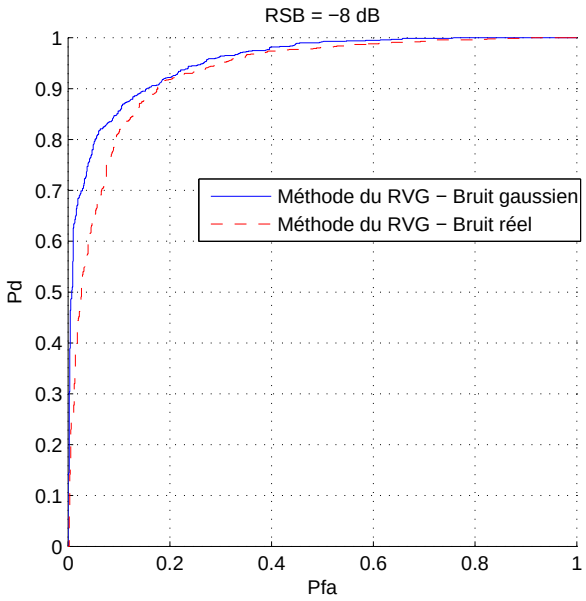

FIG. 2.11 – Comparaison des courbes COR obtenues pour un signal acoustique déterministe issu du modèle d'Askaryian pour une distance de 100m.

Ces courbes correspondent à un RSB de -8 dB.
En accord avec la théorie, la meilleure courbe COR est obtenue avec le bruit gaussien. Nous avions remarqué dans le premier chapitre que la densité de probabilité du bruit réel que nous avions à disposition était proche d'une gaussienne. Ce résultat est confirmé par la courbe COR correspondante, proche de celle obtenue pour le cas gaussien.

2.7.1.2 Seconde distance de détection : $d = 200$ m

La même expérience est menée pour un signal acoustique maintenant calculé à une distance de 200m pour une même puissance de bruit. Les courbes COR vont donc nous renseigner sur l'influence de la propagation sur les performances du détecteur.
Les résultats sont présentés sur la figure 2.12.

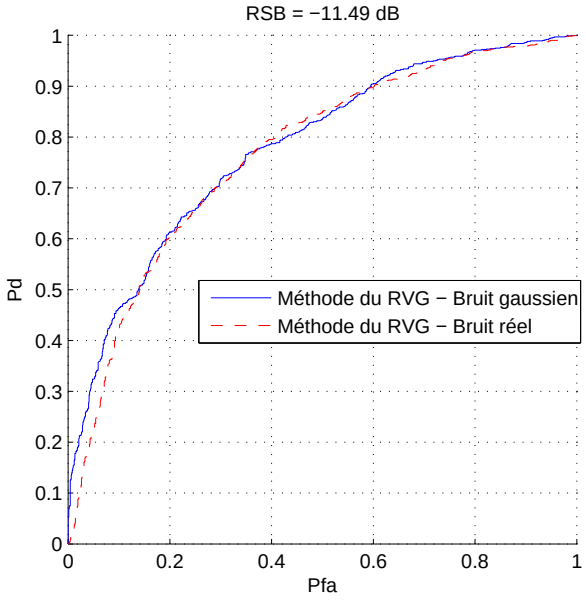

FIG. 2.12 – Comparaison des courbes COR obtenues pour un signal acoustique déterministe issu du modèle d'Askaryian pour une distance de 200m.

Pour une distance de 200m (donc une propagation de 100m), le rapport signal à bruit est maintenant d'environ -11.5 dB. Les courbes COR se sont tassées, illustrant ainsi une perte de performances du détecteur.

Concernant le changement du type de bruit, on retrouve le même résultat que précédemment. La courbe correspondant au bruit gaussien est toujours équivalente ou au dessus de celle correspondant au bruit réel.

2.7.1.3 Interprétations des résultats

Les signaux du modèle d'Askaryian respectivement aux distances 100 et 200m sont représentés sur la figure 2.13

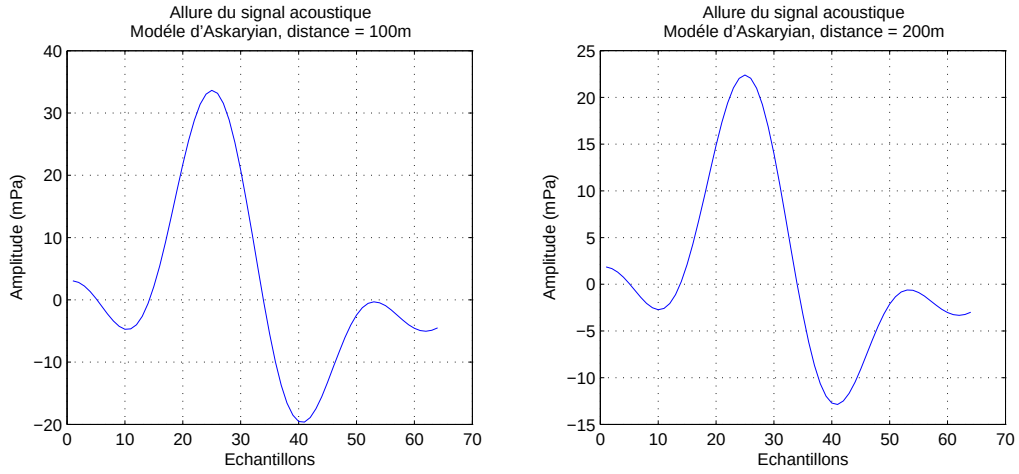

FIG. 2.13 – Comparaison de l'allure du signal acoustique généré à 100m et à 200m de l'hydrophone - Modèle d'Askaryian

A une différence d'amplitude prés, les signaux semblent très proches. Ceci est confirmé par les spectres fréquentiels présentés sur la figure 2.14.

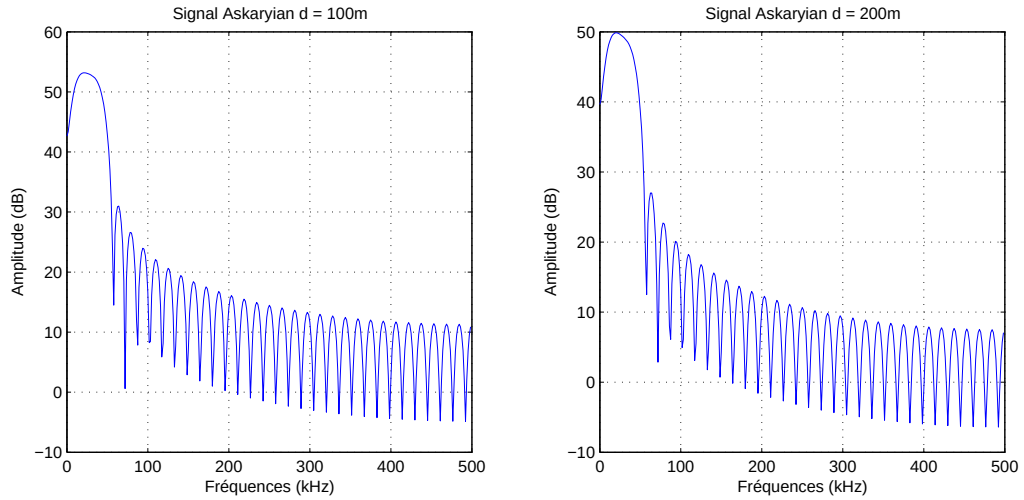

FIG. 2.14 – Comparaison des spectres fréquentiels des signaux acoustiques générés à 100m et 200m de l'hydrophone - Modèle d'Askaryian

Au vu de ces résultats, il est raisonnable de penser que les pertes de performances du détecteur ne sont dues qu'à la perte de puissance qu'a subit le signal lors de sa propagation et non aux possibles pertes dues aux approximations polynômiales des signaux. En effet, les formes d'ondes des signaux sont très proches, et l'approximation doit donc fonctionner de la même manière dans les deux cas.

Pour confirmer ces résultats, nous avons mesuré le rapport de puissance α entre le signal

à 100m et le signal à 200m :

$$\alpha = 0.447$$

qui correspond à

$$\alpha_{\text{dB}} = -3.49 \text{ dB}$$

Si l'on augmente de 3.49 dB la puissance du signal à 200m, et que les pertes de performances sont uniquement dues à la propagation et non à une mauvaise approximation polynômiale de ce signal, on doit retomber sur une courbe COR équivalent à celle correspondant au signal à 100m. Cette courbe COR est tracée sur la figure 2.15

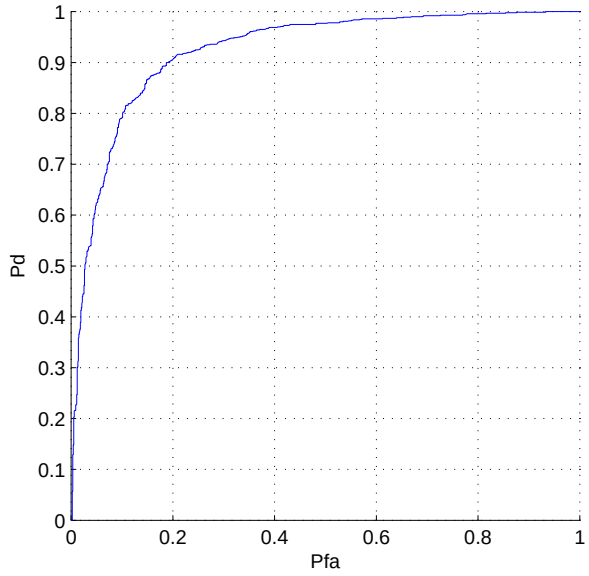

FIG. 2.15 – Courbe COR obtenue avec le signal généré à 200m et dont la puissance a été rehaussée de 3.49 dB. Bruit de mer réel et modèle d'Askaryian.

Cette courbe COR est quasiment identique à celle tirée de la figure 2.11. On peut donc en conclure que les pertes de performances ne sont dues qu'à la propagation dans le milieu marin.

2.7.2 Cas du signal acoustique de Niess

Dans le cas du signal de Niess et au vu des résultats précédents, nous allons plutôt ici étudier l'évolution des performances du détecteur en fonction de la distance d'apparition de la cascade. Le détail des calculs et la description du protocole de mesure de ces performances sont décris dans l'annexe G.

On représente sur les figures 2.16 et 2.17 pour chaque distance, la valeur du coefficient

d'atténuation α (en dB) à appliquer au bruit de mer pour obtenir une courbe COR identique à la courbe de référence choisie.

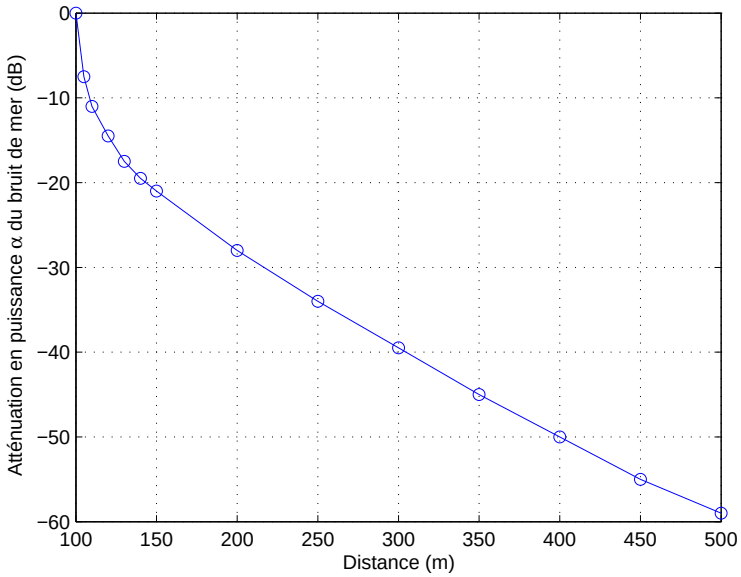

FIG. 2.16 – Evolution du coefficient α en fonction de la distance pour des performances correspondant à la courbe COR de référence 1 - Méthode du RVG.

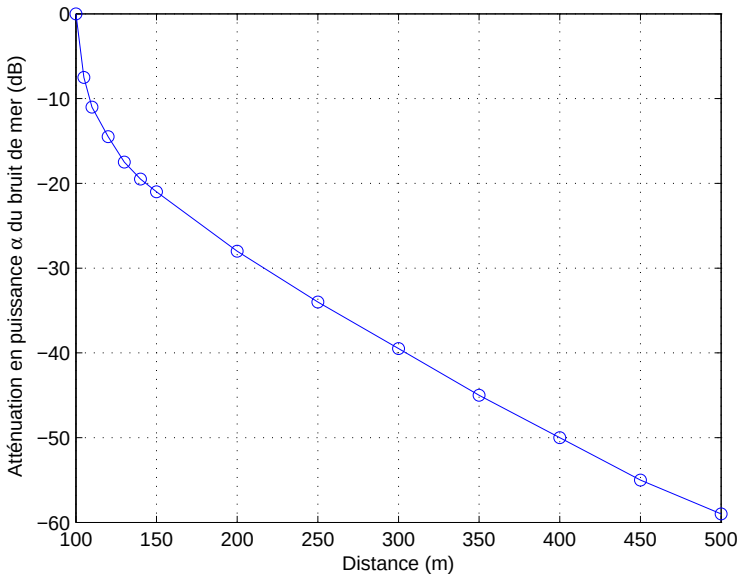

FIG. 2.17 – Evolution du coefficient α en fonction de la distance pour des performances correspondant à la courbe COR de référence 2 - Méthode du RVG.

On remarque que ces courbes sont indépendantes de la courbe COR de référence choi-

sie, et ne dépendent donc que de la physique du problème.

On représente également sur les figures 2.18 et 2.19 en fonction de la distance les différentes valeurs du RSB au niveau de l'hydrophone calculées après compensation du bruit par le coefficient α pour les deux courbes COR de référence.

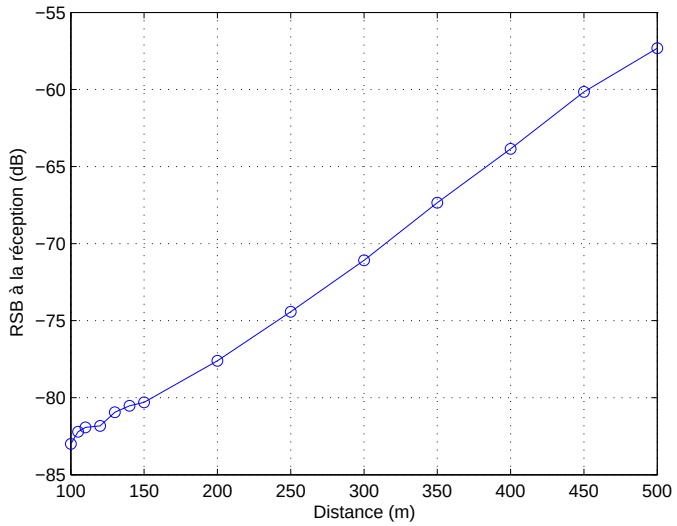

FIG. 2.18 – Evolution du RSB au niveau de l'hydrophone en fonction de la distance pour des performances correspondant à la courbe COR de référence 1 - Méthode du RVG.

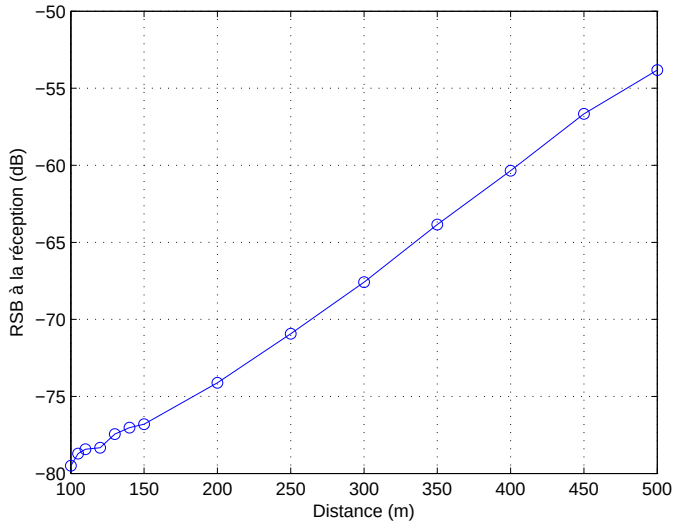

FIG. 2.19 – Evolution du RSB au niveau de l'hydrophone en fonction de la distance pour des performances correspondant à la courbe COR de référence 2 - Méthode du RVG.

A une valeur constante prés, les deux courbes correspondant aux deux courbes COR de références présentent une évolution identique. Le choix d'une courbe COR n'influence donc que le rapport signal à bruit initial (RSB_{d_0} où $d_0 = 100$m) et non l'évolution du RSB en fonction de la distance. le RSB à l'origine ne dépendant donc que de la méthode de détection étudiée, la comparaison de ce type de courbe correspondant à plusieurs méthodes nous permettra donc d'en déduire le meilleur des détecteurs.

2.8 Conclusion

La méthode de détection décrite dans ce chapitre est basée sur le calcul d'un rapport de vraisemblance et sur la connaissance *a priori* des densités de probabilité des processus mis en jeu.

Plusieurs cas ont été étudiés :

- ⋆ le bruit est gaussien, centré et de variance connue et constante.
- ⋆ le bruit est gaussien, centré et de variance inconnue et constante.
- ⋆ le bruit est gaussien, centré et de matrice de variance covariance estimée expérimentalement à partir de données réelles.
- ⋆ le bruit est réel, issu de campagnes de mesures effectuées en Méditerrannée.

Les deux derniers cas sont les plus intéressants, car proches de la réalité. Pour chacun d'entre eux, le signal acoustique est généré grâce à l'un des deux modèles à notre disposition, et est ensuite approximé par un polynôme de degrés fixé *a priori*.

Pour chacun des signaux acoustiques, on constate que les performances du détecteur (illustrées par les courbes COR) se dégradent légèrement lorsqu'on passe d'un bruit gaussien à un bruit réel, ce qui est conforme à la théorie. Malgré des rapports signal-à-bruit beaucoup plus défavorables, le signal issu du modèle de Niess semble être mieux détecté que le signal issu du modèle d'Askaryian. De plus, le modèle de propagation utilisé dans le modèle d'Askaryian semble être un peu frustre comparé à celui utilisé dans le modèle de Niess qui semble plus proche de la réalité. C'est à la lumière de ces résultats que nous décidons d'utiliser uniquement le modèle de Niess dans les chapitres suivants.

Un des inconvénients de cette méthode de détection est que le signal acoustique est défini comme déterministe. En pratique, cela revient à ne chercher à détecter que les signaux situés à une certaine distance. On a ainsi montré que si l'on cherche à détecter un signal généré à 100m de l'hydrophone, on aura beaucoup de mal à détecter un signal généré à 150m et plus. Ceci est d'autant plus vrai avec le modèle de Niess car l'atténuation en

distance de celui-ci est bien plus importante pour des petits écarts de distance que celle imposée par le modèle d'Askaryian.
Une solution à ce problème serait de faire travailler en parallèle plusieurs détecteurs, chacun basé sur un signal correspondant à une distance précise. Si théoriquement, cette solution peut fonctionner, la mise en pratique risque d'être difficile et de demander de trop importants moyens. Nous verrons dans le chapitre suivant qu'une autre solution est de ne plus considérer le signal acoustique comme déterministe mais comme aléatoire.

La modélisation polynômiale du signal d'intérêt présente également plusieurs inconvénients : les coefficients du polynôme sont en effet les seules informations que l'on utilise pour caractériser le signal acoustique. Or, il existe dans le milieu marin d'autres signaux possédants des caractéristiques proches de celles du signal acoustique, par exemple le bruit des crevettes claqueuses ou certains bruits issus de l'activité marine humaine (bateau, sous-marins, ...). Ces signaux seraient donc interprétés par le détecteur comme des signaux d'intérêt et aboutiraient à une augmentation du nombre de fausse alarme.

Pour essayer de s'affranchir de tous ces inconvénients, le signal acoustique sera modélisé par un signal aléatoire dont chaque réalisation correspondra à une distance cascade/hydrophone différente. Plusieurs méthodes de détection basées sur la notion de filtrage adapté seront étudiées et leurs performances comparées avec celles du détecteur étudié dans ce chapitre.

Chapitre 3

Détection par le filtrage adapté et ses extensions

3.1 Introduction

Jusqu'ici, nous avons considéré que le signal acoustique n'était perturbé que par le bruit ambiant généré par le milieu marin. Mais il ne faut pas oublier que la mer méditerranée (comme toute autre mer ou océan) possède un écosystème complet générant à lui seul de nombreux bruits parasites qu'on ne peut négliger. A ceux-ci viennent également s'ajouter ceux issus de l'activité humaine (bateaux, sous-marins, sonars, ...). Notre signal acoustique n'est donc plus seulement perturbé par le bruit marin ambiant. Un bon système de détection se devra de prendre en compte ces nouvelles sources sonores et de s'en affranchir au mieux.

La méthode de détection étudiée dans le chapitre précédent possède certains inconvénients : les densités de probabilité des processus aléatoires mis en jeu sont supposées gaussiennes et le signal acoustique est considéré comme déterministe. Dans la réalité, les densités de probabilité sont la plupart du temps inconnues et un signal déterministe ne l'est jamais vraiment, celui-ci présentant toujours une certaine variabilité autour d'une valeur moyenne.

Ainsi, au bruit de mer ambiant, dont la densité de probabilité est proche d'une gaussienne, viennent s'ajouter d'autres sources de bruits. La densité de probabilité de bruit global issu de cette superposition n'a aucune raison d'être gaussienne.

Une approche plus réaliste du signal acoustique est de considérer celui-ci non plus comme un signal déterministe mais comme un signal aléatoire dont chaque réalisation cor-

respond à une distance cascade d'énergie / hydrophone différente. Nous avons décidé que toute cette approche serait confrontée aux signaux issus du modèle de Niess.

L'objectif de ce chapitre est de mettre en oeuvre des techniques de détection originales pour essayer de palier ces inconvénients, et par conséquent d'obtenir de meilleures performances que celles du détecteur précédent. Les détecteurs sont basés sur une approche totalement stochastique des processus mis en jeu (signal acoustique et différents bruits) et reposent sur des notions d'optimisation de rapport signal-à-bruit.

3.2 Résumé

Dans le chapitre précédent, nous avions considéré que le signal acoustique était déterministe (mais d'équation inconnue). Dans le souci de se rapprocher de la réalité, nous allons le modéliser par un signal aléatoire. Dans tout ce chapitre, le modèle de signal utilisé est le plus récent, celui de Niess. Deux cas de figure seront étudiés dans ce chapitre :

⋆ Dans un premier temps, nous nous placerons dans le cas simple où le signal acoustique n'est perturbé que par le bruit de mer ambiant. Avant de se focaliser sur les méthodes basées sur le filtrage adapté, nous allons étudier un détecteur basé sur la notion d'analyse linéaire discriminante (ALD) [16]. Cette théorie est basée sur la connaissance *a priori* des statistiques d'ordre un et deux des signaux mis en jeu et consiste à trouver un ensemble de fonctions dites discriminantes permettant de ranger un signal donné dans une classe parmi plusieurs décrites *a priori*.

Nous étudierons ensuite un détecteur basé sur la méthode du filtrage adapté stochastique (FAS) [3] [11]. Cette méthode consiste trouver un sous-espace vectoriel (SEV) de dimension 1 dans lequel un critère basé sur la puissance des signaux mis en jeu est maximisé. Ce critère est ici le RSB.

Cette notion de filtrage adapté sera ensuite étendu à une projection sur un SEV de dimension $p > 1$ connue *a priori*. Deux méthodes seront étudiées : celle du filtrage adapté stochastique étendu (FASE) [3] et celle du filtrage adapté stochastique sous contrainte (FASC) [3]. La première est une extension directe et naturelle du FAS tandis que la seconde est basée sur la recherche d'un SEV de dimension p dans lequel le RSB est maximal. Nous allons nous rendre compte que le modèle stochastique du signal acoustique n'est pas de moyenne nulle. Or, les théories auxquelles nous faisons référence se basent toujours sur l'hypothèse d'espérance nulle pour les signaux mis en jeu, nous allons donc étudier l'impact de cette nouvelle hypothèse et essayer de les étendre au cas de signaux de moyennes non

nulles. Les performances de chacune des méthodes seront comparées sous forme de courbes COR.

⋆ Dans un second temps, deux autres signaux perturbateurs seront considérés. Ces signaux occupent la même bande passante que le signal acoustique d'intérêt. Le nouveau détecteur ne devra donc pas générer de fausses alarmes en présence de ces nouveaux signaux perturbateurs. Pour cela, on choisira un nouveau critère à maximiser, basé sur les puissances de tous les signaux mis en jeu.

Les résultats issus de cette méthode seront discutés et confrontés à ceux issus des méthodes précédentes (en considérant comme bruit la moyenne de tous les processus perturbateurs).

Sauf mention contraire, dans toute la suite de ce chapitre, nous nous placerons dans le cas d'hypothèses suivant :

Si on note x la mesure, s le signal d'intérêt (le signal acoustique ici) et b le bruit de mer, les hypothèses quand à la détection sont :

$$H_0 \quad : \quad x = b \tag{3.1}$$

$$H_1 \quad : \quad x = s + b \tag{3.2}$$

Le choix de H_0 ou H_1, dépendant de la valeur d'une fonctionnelle dépendant de la méthode étudiée, nous permet de prendre une décision sur la présence du signal s dans la mesure.

3.3 Le signal acoustique, un signal aléatoire ?

Les nouvelles méthodes de détection décrites dans ce chapitre sont basées sur une approche stochastique de tous les signaux mis en jeu. Le bruit de mer étant par définition aléatoire, intéressons nous au signal acoustique. Nous avons vu dans le premier chapitre que l'expression du signal acoustique est :

$$s_A(t, d, E) = k.E.g_L(d).\mathrm{Re}\left(h_d * s_0\right)(t)$$

Ce signal est il aléatoire ?

La réponse est oui si l'on considère que la distance d est une réalisation d'une variable aléatoire D représentant la distance entre l'hydrophone et la cascade d'énergie. A chaque réalisation de d correspond donc une réalisation du signal acoustique. Nous noterons S_A le

signal aléatoire dont s_A est une réalisation.

En pratique, un tirage de la variable aléatoire D suivant la méthode décrite dans le chapitre 1 permet de construire la réalisation de S_A correspondante. Nous avons tracé sur la figure 3.1 les formes des réalisations du signal acoustique correspondant à différentes distances.

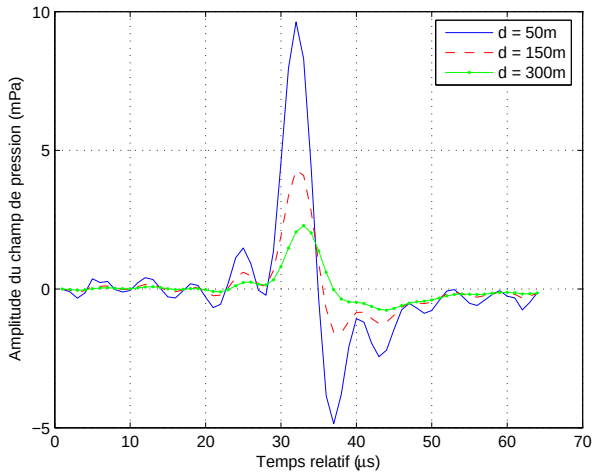

FIG. 3.1 – Allure du signal acoustique pour plusieurs distances après filtrage dans la bande passante utile.

Quelle est la moyenne de ce signal aléatoire ?

Le signal acoustique étant maintenant défini comme un signal aléatoire, il est possible d'en calculer la valeur moyenne. Ce signal moyen est noté $\overline{s_A}$ et représenté sur la figure 3.2. Il faudra donc adapter les méthodes de détection existantes, basées à l'origine sur l'hypothèse de moyenne nulle, au cas d'un signal de valeur moyenne non nulle.

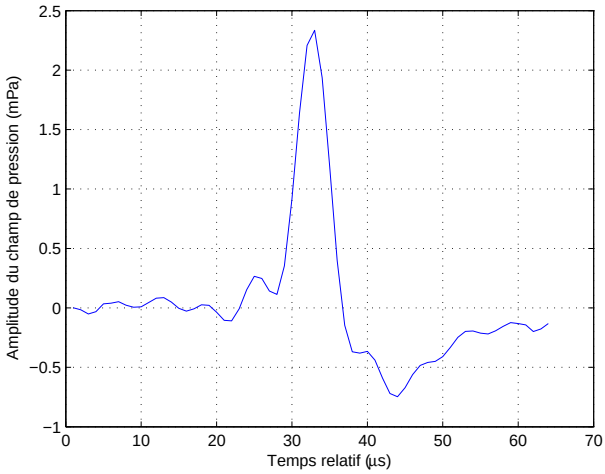

FIG. 3.2 – Allure du signal acoustique moyen pour une distribution de distances dans l'intervalle [0, 500m] après filtrage dans la bande passante utile.

Représentation des performances dans le cas d'un signal aléatoire

Nous venons de voir que le signal acoustique est maintenant modélisé par un signal aléatoire. Dans cette hypothèse, nous avons défini un RSB moyen (équation 1.59), et il est possible de tracer des courbes COR "moyennes" en considérant toutes les réalisations possibles du signal (c'est à dire en considérant toutes les distances entre D_{min} et D_{max}. Ces courbes ne peuvent donc être comparées directement aux courbes COR du chapitre précédent, celles-ci ayant été générées pour des signaux acoustiques déterministes (et donc pour une distance fixée).

3.4 Détection par analyse linéaire discriminante [16]

3.4.1 Rappel de la théorie dans le cas de la détection

Avant de s'intéresser aux techniques de détection basées sur les notions de filtrage adapté, nous allons présenter une méthode basée sur l'analyse linéaire discriminante [16]. Si elle fait intervenir des mathématiques proches de celles utilisées dans les notions de filtrage adapté, la philosophie en est assez différente.

En règle générale, l'analyse linéaire discriminante a pour objet la classification, c'est à dire l'affectation exclusive d'une mesure x à un certain nombre de classes distinctes C_i. Le nombre de classes c est une donnée *a priori* du problème. La définition de ces classes dépend étroitement de la nature de celui-ci. Dans le cas simple de la détection d'un signal noyé dans du bruit, le nombre de classe est $c = 2$ et sont définies de la manière suivante :

Classe	Description
C_1	La mesure est composée du mélange signal + bruit
C_2	La mesure est composée du bruit seul

Une fois les classes définies, il s'agit en général de trouver un ensemble de fonction discriminante que l'on notera $g_i(x)$. La décision de placer la mesure x dans la classe C_i est prise si

$$g_i(x) > g_j(x) \quad \forall j \neq i$$

Dans la suite, nous noterons S le signal d'intérêt de moyenne $\overline{s}$, b le bruit supposé centré. Ces vecteurs sont tous de taille N.

Supposons que l'on dispose de réalisations du signal S et du bruit b. Il est possible, à partir de celle-ci, de générer successivement N_{b1} et N_{b2} vecteurs appartenant respectivement aux classes C_1 et C_2. A partir de ceux-ci, nous allons définir un certain nombre de quantités qui vont nous servir à caractériser chacune des classes. Ces quantités sont énumérées dans le tableau ci-dessous :

Description	Expression
Vecteur moyen d'une classe C_i	$m_i = \frac{1}{N_{bi}} \sum\limits_{x \in C_i} x$
Vecteur moyen total	$m = \frac{1}{N_b} \sum\limits_{x} x = \frac{1}{N_b} \sum\limits_{i=1}^{c} p_i \; m_i$
Matrice de dispersion d'une classe C_i	$\mathbf{S}_i = \sum\limits_{x \in C_i} [x - m_i] \; [x - m_i]^T$
Matrice de dispersion totale intra-classe	$\mathbf{S}_W = \sum\limits_{i=1}^{c} \mathbf{S}_i$
Matrice de dispersion inter-classe	$\mathbf{S}_B = \sum\limits_{i=1}^{c} p_i \, [m_i - m] \; [m_i - m]^T$

p_i est le poids que l'on attribue à la classe C_i (avec $\sum\limits_{i=1}^{c} p_i = 1$). On choisira ici

$$p_1 = p_2 = 0.5$$

On notera également $N_b = N_{b1} + N_{b2}$, c'est le nombre total de vecteur à disposition pour définir les classes.

Dans le cas qui nous intéresse, ces expressions se simplifient et on obtient :

$$m_1 = \overline{s} \tag{3.3}$$

$$m_2 = 0 \tag{3.4}$$

$$m = 0.5\ \overline{s} \tag{3.5}$$

$$\mathbf{S}_B = 0.25\ \overline{s}\ \overline{s}^T \tag{3.6}$$

Il s'agit maintenant de choisir un critère à optimiser aboutissant au choix d'une fonction discriminante. Ce critère doit être fonction des matrices de dispersion, comme par exemple la trace ou le déterminant. Une approche naturelle [3] consiste à opposer deux de ces matrices, par exemple en cherchant à maximiser la trace de la matrice $\mathbf{S}_W^{-1}\mathbf{S}_B$:

$$J_1 = \mathrm{tr}\left(\mathbf{S}_W^{-1}\mathbf{S}_B\right) \tag{3.7}$$

On montre que ce critère est maximal dans le sous-espace vectoriel engendré par les vecteurs propres de la matrice $\mathbf{S}_W^{-1}\mathbf{S}_B$. En effet, les valeurs propres λ_i de cette matrice mesurent, dans la direction des vecteurs propres, la valeur d'un rapport que l'on souhaite le plus grand possible.

Fonction discriminante :
Ici, la matrice $\mathbf{S}_B$ est composée d'un seul vecteur et est donc de rang 1. La matrice $\mathbf{S}_W$ est, quand à elle, de rang plein, ce qui justifie l'existence de la matrice inverse. On peut en déduire le résultat suivant :

$$\mathrm{rang}\left(\mathbf{S}_W^{-1}\mathbf{S}_B\right) = 1 \tag{3.8}$$

La dimension du sous-espace engendré est donc également de 1 (c'est une droite), il n'y aura donc qu'une seule fonction discriminante (ce résultat est trivial, car une seule décision est à prendre pour ranger une mesure dans un ensemble de 2 classes).

Les valeurs propres λ_i de $\mathbf{S}_W^{-1}\mathbf{S}_B$ sont donc comme suit :

$$\lambda_1 \neq 0 \tag{3.9}$$

$$\lambda_i = 0,\ i = 2 \ldots N \tag{3.10}$$

Or la trace d'une matrice vaut la somme de ses valeurs propres :

$$\mathrm{tr}\left(\mathbf{S}_W^{-1}\mathbf{S}_B\right) = \lambda_1 \tag{3.11}$$

On peut également écrire :

$$\begin{aligned} \mathrm{tr}\left(\mathbf{S}_W^{-1}\mathbf{S}_B\right) &= p_1\ p_2 \overline{s}^T \mathbf{S}_W^{-1} \overline{s} && (3.12) \\ &= \frac{1}{4} \overline{s}^T \mathbf{S}_W^{-1} \overline{s} && (3.13) \end{aligned}$$

L'expression du vecteur propre de $\mathbf{S}_W^{-1}\mathbf{S}_B$ associé à λ_1 et engendrant le sous-espace dans lequel J_1 est maximal est :

$$\Phi_1 = \mathbf{S}_W^{-1} \overline{s} \tag{3.14}$$

On vérifie ce résultat aisément :

$$\begin{aligned} \mathbf{S}_W^{-1}\mathbf{S}_B\Phi_1 &= \mathbf{S}_W^{-1}\mathbf{S}_B\mathbf{S}_W^{-1}\overline{s} && (3.15) \\ &= \frac{1}{4}\mathbf{S}_W^{-1}\overline{s}\overline{s}^T\mathbf{S}_W^{-1}\overline{s} && (3.16) \\ &= \lambda_1\mathbf{S}_W^{-1}\overline{s} && (3.17) \\ &= \lambda_1\Phi_1 && (3.18) \end{aligned}$$

Il suffit maintenant de choisir comme fonction discriminante la projection d'un vecteur de mesure x sur ce sous-espace :

$$\boxed{g(x) = \Phi_1^T x = \overline{s}^T \mathbf{S}_W^{-1} x} \tag{3.19}$$

Pour prendre la décision d'appartenance de x à la classe C_1 ou C_2, on compare la valeur de $g(x)$ à un seuil prédéfini (proportionnel à une valeur de P_{fa} que l'on s'impose). Si cette valeur est supérieure au seuil, on range x dans la classe C_1, dans la classe C_2 sinon.

3.4.2 Application au cas Antares

Appliquons ces résultats au cas réel qui nous intéresse : le signal d'intérêt est le signal acoustique S_A et le bruit est le bruit de mer réel b_m.

La courbe COR moyenne est affichée sur la figure 3.3.

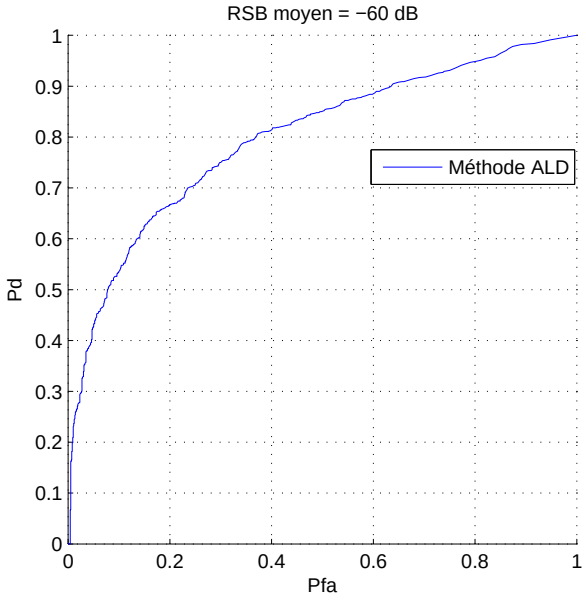

FIG. 3.3 – Courbe COR dans le cas semi-réel associée à la méthode d'analyse linéaire discriminante.

Le rapport signal-à-bruit moyen pour une telle courbe est de -60 dB.

Comme dans le chapitre précédent pour la méthode du RVG, il est intéressant de tracer en fonction de la distance l'évolution des valeurs du coefficient d'atténuation α à appliquer au bruit afin d'obtenir des performances correspondant à une courbe COR de référence choisie *a priori*.

On rappelle que les courbes COR de références et la description de l'étude des performances sont décrites dans l'annexe G. Pour les deux courbes COR de référence choisies, l'évolution du coefficient d'atténuation en puissance α est décrite sur la figure 3.4.

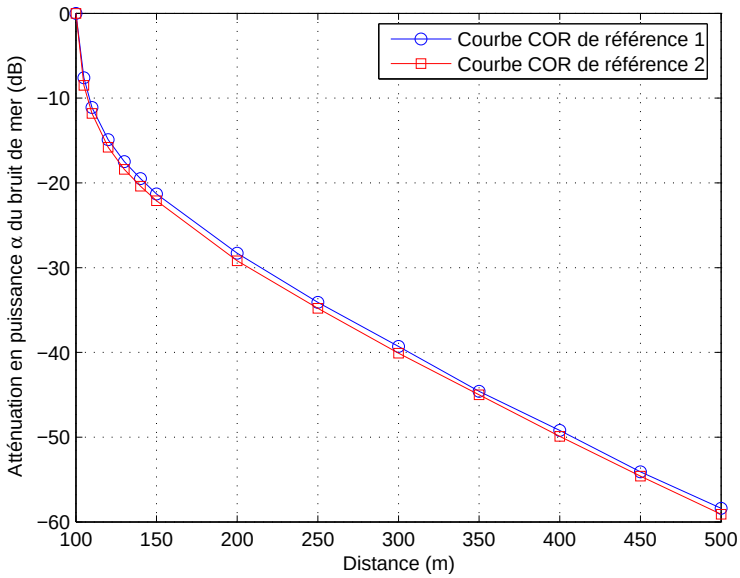

FIG. 3.4 – Evolution du coefficient α en fonction de la distance pour des performances correspondant aux deux courbes COR de référence choisies - Méthode ALD.

On remarque que ces courbes sont quasiment identiques et donc indépendantes de la courbe COR de référence choisie, et ne dépendent donc que de la physique du problème.

On trace maintenant sur la figure 3.5, pour les deux courbes COR de référence choisies, l'évolution du RSB au niveau de l'hydrophone en fonction de la distance et calculé *après compensation en puissance par le coefficient* α.

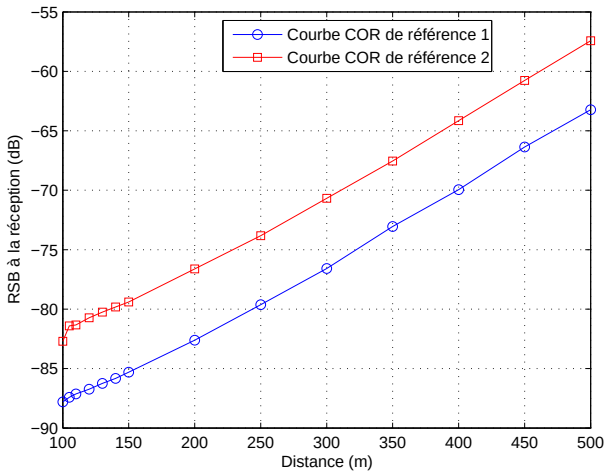

FIG. 3.5 – Evolution du RSB au niveau de l'hydrophone en fonction de la distance pour des performances correspondant aux deux courbes COR de référence choisies - Méthode ALD.

On remarque qu'à une valeur constante prés, les deux courbes précédentes présentent une évolution identique. Le choix d'une courbe COR de référence n'influence donc que le rapport signal à bruit initial (RSB_{d_0} où $d_0 = 100$m) et non l'évolution du RSB en fonction de la distance. Dans le cas présent, l'augmentation des performances du détecteur (par le passage de la courbe COR de référence 1 à la courbe COR de référence 2) implique donc une diminution du RSB d'environ 5 dB.

3.5 Le filtrage adapté stochastique FAS [3]

3.5.1 Théorie

Considérons un signal aléatoire s tel que $\mathsf{E}\left(s\ s^T\right) = \mathbf{A}$ [1] perturbé additivement par un bruit indépendant centré noté b de matrice de variance covariance $\mathbf{B}$. Ces processus sont supposés décorrellés entre eux et les matrices de covariance sont supposées définies positives et de rangs pleins.

Le choix de l'hypothèse H_1 ou H_0 (équations 3.2 et 3.2) sera effectué en comparant la valeur d'un critère (par le biais d'une fonctionnelle dépendant de la mesure) à un seuil prédéfini, dépendant directement de la valeur de P_{fa} souhaitée. Si ce critère est supérieur au seuil, on prendra la décision H_1, H_0 sinon. Celui-ci doit donc prendre des grandes valeurs en présence du mélange signal/bruit et des faibles valeurs en présence du bruit seul. Pour définir les caractéristiques des détecteurs que nous allons étudier, nous avons décidé de nous intéresser à l'écart relatif moyen en puissance (noté ξ_P) entre le signal issu du détecteur pris dans l'hypothèse H_1 et le même signal pris dans l'hypothèse H_0.

L'idée est de trouver un filtre linéaire h qui va maximiser cet écart relatif en sortie du détecteur. Si l'on note z le résultat du filtré de la mesure x par h dans chacune des hypothèses précédentes, il vient :

$$H_0 \ : \ z \ = \ h^T b \tag{3.20}$$

$$H_1 \ : \ z \ = \ h^T (s + b) \tag{3.21}$$

[1] Si s est centré, alors la matrice $\mathbf{A}$ est la matrice de variance covariance de s.

Mettons maintenant cet écart relatif en équation, il vient :

$$\xi_P = \frac{\mathsf{E}\left(z^T z|H_1\right) - \mathsf{E}\left(z^T z|H_0\right)}{\mathsf{E}\left(z^T z|H_0\right)} = \frac{\mathsf{E}\left(h^T(s+b)(s+b)^T h\right) - \mathsf{E}\left(h^T b b^T h\right)}{\mathsf{E}\left(h^T b b^T h\right)} \quad (3.22)$$

$$= \frac{h^T \mathsf{E}\left(ss^T + bb^T + 2sb^T\right) h - h^T \mathsf{E}\left(bb^t\right) h}{h^T \mathsf{E}\left(bb^t\right) h} \quad (3.23)$$

$$= \frac{h^T \mathbf{A} h + h^T \mathbf{B} h - h^T \mathbf{B} h}{h^T \mathbf{B} h} \quad (3.24)$$

$$= \frac{h^T \mathbf{A} h}{h^T \mathbf{B} h} \quad (3.25)$$

Ce quotien de Rayleigh est l'expression du **rapport signal-à-bruit** après projection sur le vecteur h.

La technique du filtrage adapté stochastique [6] (noté FAS) consiste à trouver un sous-espace vectoriel de dimension un dans lequel le RSB est maximum.

Si l'on s'intéresse au gain apporté au RSB initial par le filtrage, il convient de normaliser les matrices en rendant leurs traces unitaires : $\widetilde{\mathbf{A}} = \frac{\mathbf{A}}{\mathrm{tr}(\mathbf{A})}$ et $\widetilde{\mathbf{B}} = \frac{\mathbf{B}}{\mathrm{tr}(\mathbf{A})}$, il vient alors :

$$\rho = \frac{\mathrm{tr}(\mathbf{A})}{\mathrm{tr}(\mathbf{B})} \frac{h^T \widetilde{\mathbf{A}} h}{h^T \widetilde{\mathbf{B}} h} \quad (3.26)$$

$\widetilde{\rho} = \frac{h^T \widetilde{\mathbf{A}} h}{h^T \widetilde{\mathbf{B}} h}$ est alors le gain obtenu sur le RSB.

On considérera dans la suite que les matrices $\mathbf{A}$ et $\mathbf{B}$ sont normalisées par leurs traces respectives.

On montre [3] que le vecteur h qui maximise le RSB est le vecteur solution de l'équation aux valeurs propres généralisées :

$$\mathbf{A} h = \lambda \mathbf{B} h \quad (3.27)$$

associé à la valeur propre la plus grande, qui correspond au gain sur le RSB. On montre que celle-ci est toujours supérieure à 1.

h est donc un vecteur adapté à la classe de signaux définies par la matrice $\mathbf{A}$ et non adapté à la classe de bruits définies par la matrice $\mathbf{B}$.

3.5.1.1 Fonctionnelle associée

Nous pouvons maintenant définir une fonctionnelle dont la valeur sera seuillée pour prendre la décision de détection. Dans le cas du FAS, la fonctionnelle est la puissance de la mesure filtrée : si h est le filtre issu du FAS et x une mesure, elle s'écrit donc :

$$\Lambda_{\text{FAS}}(x) = \left[h^T x\right]^2 \tag{3.28}$$

Résumé FAS

Soient s et b deux vecteurs de dimension N représentant le signal d'intérêt à détecter et le bruit perturbateur, associés respectivement aux matrices de covariances $\mathbf{A}$ et $\mathbf{B}$, alors le filtre linéaire h maximisant le rapport signal à bruit après filtrage $\rho = \frac{h^T \mathbf{A} h}{h^T \mathbf{B} h}$ est le vecteur propre solution de l'équation :

$$\mathbf{B}\ h = \lambda\ \mathbf{A}\ h$$

associé à la valeur propre la plus grande λ_1.

Pour une mesure x, la fonctionnelle associée est :

$$\Lambda_{\text{FAS}}(x) = \left[h^T x\right]^2$$

3.5.2 Application au cas Antares

Là encore, on s'intéresse à l'évolution en fonction de la distance du coefficient d'atténuation en puissance α à appliquer au bruit pour que les performances du détecteur soient identiques à celles décrites par les deux courbes de référence choisies (voir annexe G). Cette évolution est tracée sur la figure 3.6

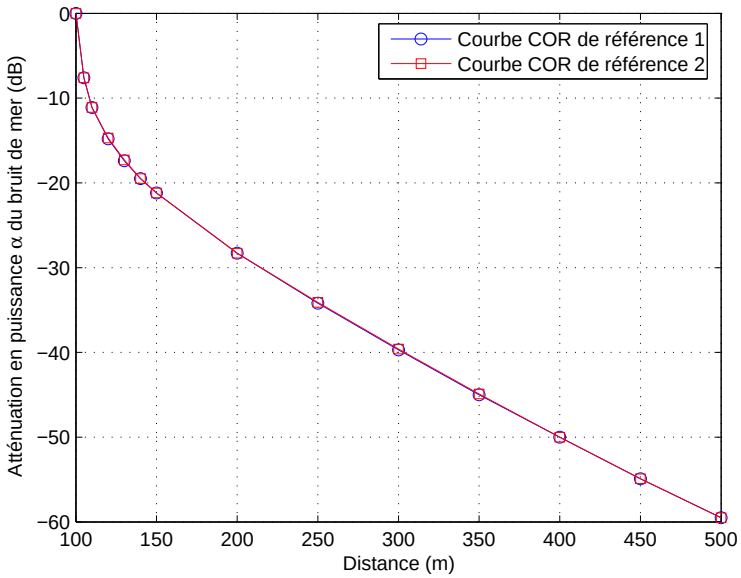

FIG. 3.6 – Evolution du coefficient α en fonction de la distance pour des performances correspondant aux deux courbes COR de référence choisies - Méthode FAS.

On remarque encore que ces courbes sont quasiment identiques et donc indépendantes de la courbe COR de référence choisie, et ne dépendent donc que de la physique du problème.

On trace maintenant sur la figure 3.7, pour les deux courbes COR de référence choisies, l'évolution du RSB au niveau de l'hydrophone en fonction de la distance et calculé *après compensation en puissance par le coefficient* α.

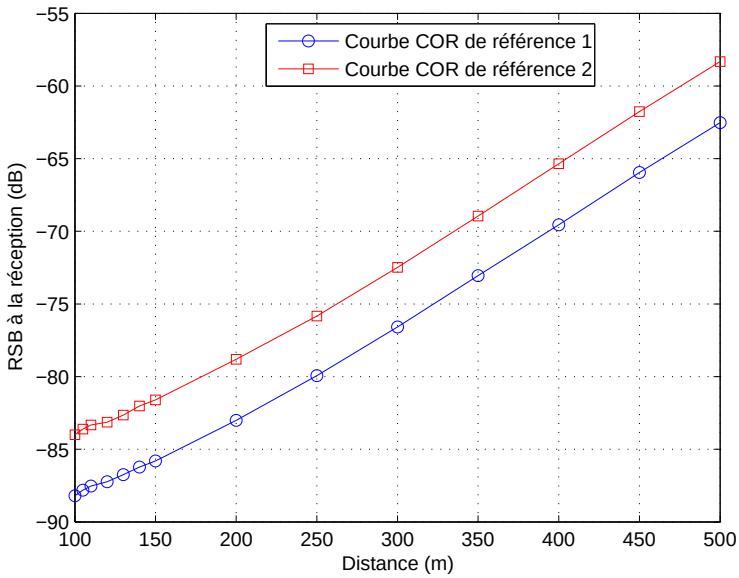

FIG. 3.7 – Evolution du RSB au niveau de l'hydrophone en fonction de la distance pour des performances correspondant aux deux courbes COR de référence choisies - Méthode FAS.

On remarque qu'à une valeur constante prés, les deux courbes précédentes présentent une évolution identique. Le choix d'une courbe COR de référence n'influence donc que le rapport signal à bruit initial (RSB_{d_0} où $d_0 = 100$m) et non l'évolution du RSB en fonction de la distance. Pour un détecteur à base de FAS, l'augmentation des performances du détecteur (par le passage de la courbe COR de référence 1 à la courbe COR de référence 2) implique donc une diminution du RSB d'environ 4.1 dB.

3.6 Le filtrage adapté stochastique étendu FASE [3]

3.6.1 Théorie

Intéressons nous maintenant à la projection du vecteur de mesure de dimension N sur un sous-espace de dimension p, choisie *a priori* et telle que $p < N$. Le filtrage adapté stochastique étendu noté FASE est une extension naturelle du FAS [3]. En effet, il arrive assez fréquemment que les valeurs propres $\lambda_2, \lambda_3, \ldots$ de la matrice $\mathbf{B}^{-1}\mathbf{A}$ (cf. équation 3.27) soient proches de la valeur propre maximale λ_1. Il semble donc naturel de vouloir utiliser l'information stockée par ces valeurs propres par le biais des vecteurs propres associés. C'est de cette constatation qu'est née l'idée de concevoir une fonctionnelle qui combinerait ces vecteurs propres. En toute rigueur, cette fonctionnelle ne conduira pas à un gain maximal sur le RSB puisque celui-ci est atteint pour un SEV de dimension un, mais peut participer à l'amélioration de la détection.

Il est important de remarquer qu'il n'existe pas encore de théorie capable d'évaluer la taille du sous-espace p pour laquelle l'amélioration de la détection est la plus importante. C'est pour cela que cette valeur est estimée expérimentalement dans chaque situation. Quelle fonctionnelle choisir pour le FASE ?

3.6.1.1 Fonctionnelle utilisée

L'idée pour bien choisir une fonctionnelle pour le FASE est de pondérer les projections sur les différentes directions du sous-espace optimal par un coefficient basé sur les valeurs propres correspondantes. Une proposition d'une telle fonctionnelle est donnée par l'équation 3.29.

$$\Lambda_{\text{FASE}}(x) = \sum_{k=1}^{p} \frac{\lambda_k}{1+\lambda_k} \frac{\left(h_k^T\ x\right)^2}{h_k^T \mathbf{B} h_k} \tag{3.29}$$

où p est la dimension choisie *a priori* du sous-espace, avec $p < N$.

Le terme $\frac{\lambda_k}{1+\lambda_k}$ permet de favoriser les vecteurs propres h_k associés aux grandes valeurs propres λ_k.

Le terme $\left(h_k^T\ x\right)^2$ peut être interprété comme la puissance de la projection du vecteur x sur la direction h_k. Chacune de ces puissances est pondérée par un terme positif dépendant des valeurs propres. Ce terme est d'autant plus faible que les valeurs propres sont faibles, seules les valeurs propres de fortes valeurs sont donc favorisées.

Le terme $h_k^T \mathbf{B} h_k$ correspond à une normalisation des filtres par la puissance du bruit.

3.6.1.2 Rapport de vraisemblance dans le cas gaussien

Il est intéressant de remarquer que l'expression de la fonctionnelle précédente est similaire à celle que l'on trouve en calculant le rapport de vraisemblance dans le cas gaussien.

En effet, notons ici s le signal d'intérêt centré et b le bruit centré, associés respectivement aux matrice de variance covariance $\mathbf{A}$ et $\mathbf{B}$. Les densités de probabilité de ces deux processus sont supposées ici gaussiennes. On se retrouve encore confronté au cas classique de détection avec les hypothèses H_0 et H_1.

Le choix de l'une ou l'autre dépendra de la valeur du rapport de vraisemblance que l'on comparera à un seuil. Pour un vecteur x donnée, réalisation de la variable aléatoire X

associée à la mesure, on montre [3] que ce rapport de vraisemblance s'écrit :

$$\Lambda(x) = \sum_{k=1}^{N} \frac{\lambda_k}{1+\lambda_k} \frac{\left(h_k^t\ x\right)^2}{h_k^T \mathbf{B} h_k} \underset{d_0}{\overset{d_1}{\gtrless}} \text{seuil} \tag{3.30}$$

où h_k est le vecteur propre associé à la valeur propre λ_k, solution de l'équation 3.27.

On remarque que l'expression de ce rapport de vraisemblance est similaire à celle de la fonctionnelle décrite précédemment, mais la sommation se fait ici sur les N vecteurs propres, et les densités de probabilité sont supposées gaussiennes.

3.6.2 Application au cas Antares

Intéressons nous maintenant au cas où le signal d'intérêt est notre signal acoustique S_A et le bruit est le bruit de mer ambiant b_m. Le signal acoustique n'est plus centré. Il faut donc maintenant tenir compte de cette valeur moyenne $\overline{s_A}$. Nous allons calculer l'expression du nouveau rapport de vraisemblance avec cette nouvelle hypothèse et en déduire une nouvelle fonctionnelle. Signal acoustique et bruit sont toujours supposés indépendants.

Nous utiliserons les notations suivantes :

- ⋆ $\mathbf{H}$ est la matrice [N×N] contenant les vecteurs propres de l'équation 3.27. Pour plus de clarté dans les calculs qui suivent, on supposera que chaque vecteur h_k est normalisé par la puissance du bruit.

- ⋆ X est le vecteur aléatoire associé à la mesure.

- ⋆ $Z = \mathbf{H}^T X$ est le vecteur aléatoire associé à la mesure filtrée par les filtres contenus dans la matrice $\mathbf{H}$.

- ⋆ b_m est le bruit de mer supposé centré, gaussien et de matrice de variance covariance $\mathbf{B}$.

- ⋆ S_A est le vecteur aléatoire représentant le signal acoustique d'intérêt associé à la matrice de variance covariance $\mathbf{A}$.

- ⋆ $\overline{s_A} = \mathsf{E}(S_A)$ représente le signal acoustique moyen.

Les hypothèses pour la décision sont les mêmes que précédemment :

$$\begin{aligned} H_0 &: X = b_m \\ H_1 &: X = S_A + b_m \end{aligned}$$

Le calcul classique du rapport de vraisemblance mené dans l'annexe H nous donne le résultat suivant :

$$\sum_{k=1}^{N} \frac{\lambda_k}{1+\lambda_k} \left(h_k^T x\right)^2 + 2 \sum_{k=1}^{N} \frac{1}{1+\lambda_k} \left(h_k^T x\right) \left(h_k^T \overline{s_A}\right) \underset{d_0}{\overset{d_1}{\gtrless}} \text{seuil} \tag{3.31}$$

Ce test est optimal dans le cas de signaux gaussiens et pour une sommation sur tous les N éléments.

3.6.2.1 Fonctionnelle utilisée

Nous allons ici déterminer l'expression de la fonctionnelle associée à partir de l'expression du test de rapport de vraisemblance, mais en ne considérant que les vecteurs propres associés aux p plus grandes valeurs propres. La fonctionnelle ainsi créée est :

$$\Lambda_{\text{FASE}}(x) = \sum_{k=1}^{p} \frac{\lambda_k}{1+\lambda_k} \left(h_k^T x\right)^2 + 2 \sum_{k=1}^{p} \frac{1}{1+\lambda_k} \left(h_k^T x\right) \left(h_k^T \overline{s_A}\right) \tag{3.32}$$

où p est la dimension choisie *a priori* du sous-espace, avec $p < N$.

Cette fonctionnelle sera utilisée dans des cas non gaussiens. On remarque également que si la moyenne du signal acoustique est nulle, on retombe sur l'expression 3.29.

Résumé FASE

Soient s et b deux vecteurs de dimension N représentant respectivement le signal d'intérêt d'intérêt et le bruit perturbateur. Leurs sont associés les matrices de variance covariances $\mathbf{A}$ et $\mathbf{B}$, supposées définies positive et de rangs pleins. p est un entier fixé *a priori* avec $p < N$ et représente la dimension du sous-espace de projection.

Alors, la base du sous-espace adapté du FASE est constituée des p vecteurs propres h_k, associés aux p valeurs propres λ_k les plus importantes et solutions de l'équation aux valeurs propres généralisée associée :

$$\mathbf{BH} = \boldsymbol{\lambda}\mathbf{AH}$$

Pour un vecteur de mesure x, on distingue deux cas pour la fonctionnelle utilisée :

⋆ Le signal d'intérêt possède une moyenne nulle :

$$\Lambda_{\text{FASE}}(x) = \sum_{k=1}^{p} \frac{\lambda_k}{1+\lambda_k} \frac{\left(h_k^T\ x\right)^2}{h_k^T \mathbf{B} h_k}$$

⋆ Le signal d'intérêt possède une moyenne non nulle $\overline{s_A}$:

$$\Lambda_{\text{FASE}}(x) = \sum_{k=1}^{p} \frac{\lambda_k}{1+\lambda_k} \frac{\left(h_k^T x\right)^2}{h_k^T \mathbf{B} h_k} + 2 \sum_{k=1}^{p} \frac{1}{1+\lambda_k} \frac{\left(h_k^T x\right)\left(h_k^T \overline{s_A}\right)}{h_k^T \mathbf{B} h_k}$$

3.6.3 Résultats expérimentaux

Nous allons illustrer les résultats théoriques précédents en les appliquant à plusieurs cas. Dans un premier temps, le bruit de mer sera supposé gaussien et de matrice de variance covariance $\mathbf{B}$ et dans un second temps, il sera réel.

La matrice de covariance du signal a été évaluée pour des réalisations du signal correspondants à des distances comprises entre $D_{min} = 100$m et $D_{max} = 500$m. Nous allons ainsi utiliser les courbes COR moyennes pour estimer la dimension optimale du sous-espace de projection.

3.6.3.1 Cas du bruit gaussien

Recherche de la valeur de p optimale

La recherche de la valeur de la taille p du sous-espace se fait de manière expérimentale. Sont tracés sur la courbe 3.8 les courbes COR correspondant à plusieurs valeurs de p.
Le rapport signal à bruit moyen utilisé pour ces simulations est de -60 dB.

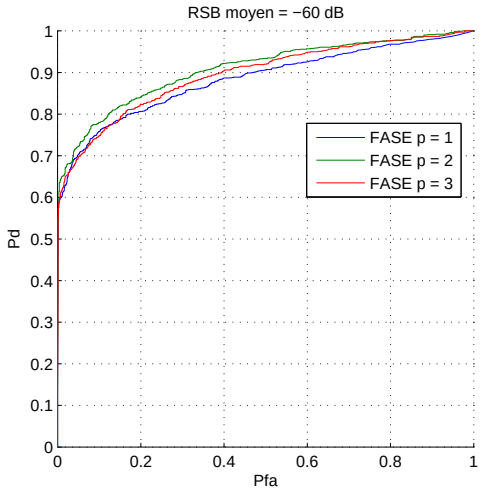

FIG. 3.8 – Courbes COR dans le cas du bruit gaussien pour la méthode du FASE et pour différentes valeurs de p.

A partir de valeurs de p supérieures à 2, les courbes COR ne sont pratiquement plus améliorées voire les performances se dégradent. La valeur de p "optimale" dans le cas du bruit de mer gaussien est choisie à 2.

$$\boxed{p_{\text{opt}}^{(\text{FASE})} = 2 \mid \text{Le bruit est gaussien}} \tag{3.33}$$

3.6.3.2 Cas du bruit réel

Le protocole de simulation est le même, mais le bruit est ici le bruit de mer dit réel comme défini dans le premier chapitre.

Recherche de la valeur de p optimale

La valeur optimale de p est encore évaluée expérimentalement. On trace sur la figure 3.9 les courbes COR correspondantes à différentes valeurs de p.

Le rapport signal à bruit moyen considéré est de -60 dB.

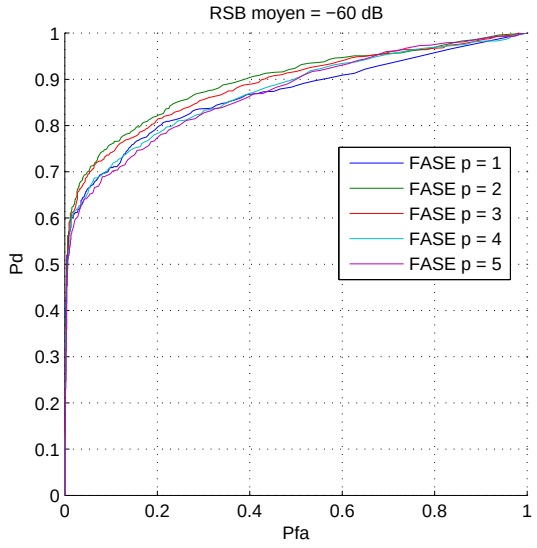

FIG. 3.9 – Courbes COR dans le cas du bruit réel pour la méthode du FASE et pour différentes valeurs de p.

La valeur $p = 2$ donne la meilleure des courbes COR observée. Pour des valeurs supérieures, les performances sont les mêmes voire se dégradent.

$$\boxed{p_{\text{opt}}^{(\text{FASE})} = 2 \mid \text{Le bruit est réel}} \tag{3.34}$$

On retrouve ici des résultats similaires au cas gaussien, ce qui confirme encore l'idée que le bruit de mer réel a un comportement assimilable à celui d'une variable aléatoire gaussienne.

Comme pour les méthodes précédentes, on s'intéresse à l'évolution en fonction de la distance du coefficient d'atténuation en puissance α à appliquer au bruit pour que les performances du détecteur soient identiques à celles décrites par les deux courbes de référence choisies (voir annexe G). Cette évolution est tracée sur la figure 3.10

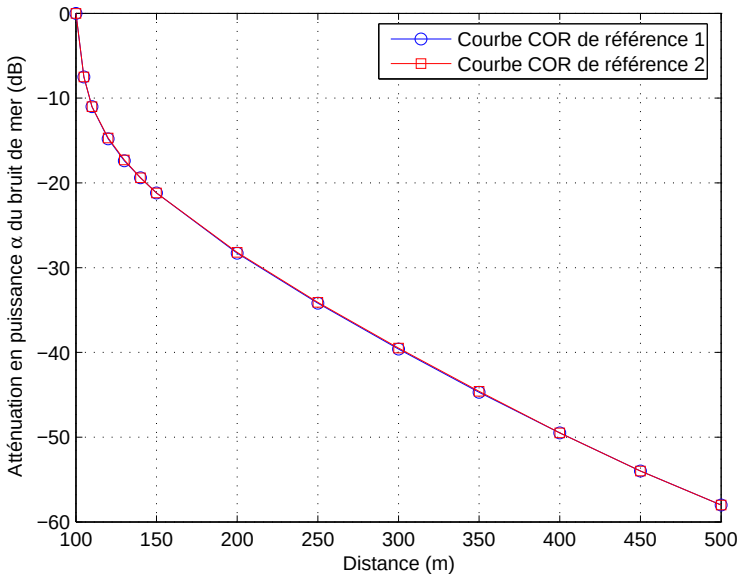

FIG. 3.10 – Evolution du coefficient α en fonction de la distance pour des performances correspondant aux deux courbes COR de référence choisies - Méthode FASE.

On remarque encore que ces courbes sont quasiment identiques et donc indépendantes de la courbe COR de référence choisie, et ne dépendent donc que de la physique du problème.

On trace maintenant sur la figure 3.11, pour les deux courbes COR de référence choisies, l'évolution du RSB au niveau de l'hydrophone en fonction de la distance et calculé *après compensation en puissance par le coefficient* α.

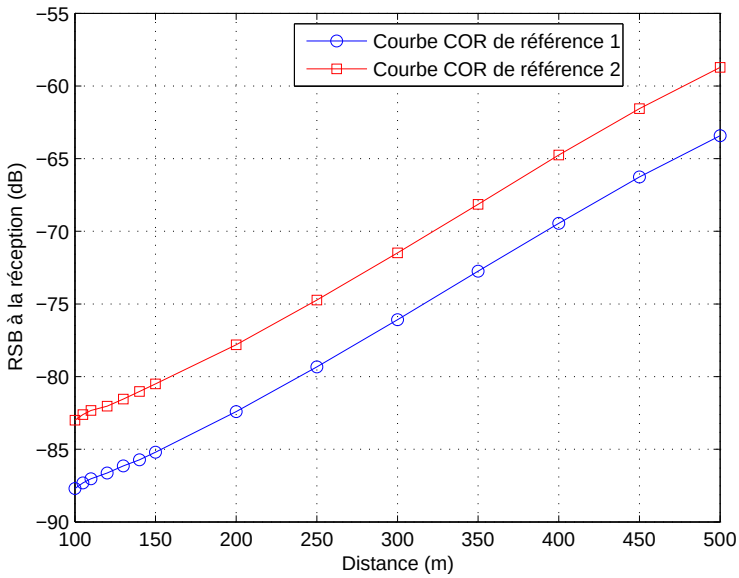

FIG. 3.11 – Evolution du RSB au niveau de l'hydrophone en fonction de la distance pour des performances correspondant aux deux courbes COR de référence choisies - Méthode FASE.

On remarque qu'à une valeur constante prés, les deux courbes précédentes présentent une évolution identique. Le choix d'une courbe COR de référence n'influence donc que le rapport signal à bruit initial (RSB_{d_0} où $d_0 = 100$m) et non l'évolution du RSB en fonction de la distance. Pour un détecteur à base de FASE pour un sous-espace de dimension 2, l'augmentation des performances du détecteur (par le passage de la courbe COR de référence 1 à la courbe COR de référence 2) implique donc une diminution du RSB d'environ 4.7 dB.

3.7 Le filtrage adapté stochastique sous contrainte FASC [3]

3.7.1 Théorie

Replaçons nous maintenant dans le même cadre d'hypothèses de signaux centrés : s et b sont deux signaux aléatoires centrés respectivement de matrice de variance covariance $\mathbf{A}$ et $\mathbf{B}$. La méthode du FAS consistait à maximiser le rapport signal à bruit dans un sous-espace de dimension 1. Cette maximisation faisait intervenir une équation aux valeurs propres généralisée dont le vecteur propre associé à la plus grande valeur propre engendrait ce sous-espace [4].

Une autre extension de cette méthode est appelée Filtrage Adapté Stochastique sous Contrainte (FASC) [3], et consiste à maximiser le RSB dans un sous-espace vectoriel de dimension p. La contrainte se trouve dans le choix *a priori* de p. La notion de projection

dans un SEV de dimension supérieure à 1 a déjà été abordée dans la méthode du FASE, mais ce sous-espace n'était pas optimal en termes de RSB après filtrage.

Le problème est donc de trouver une expression mathématique du RSB après filtrage (ou projection sur un sous-espace de dimension p), et de trouver le SEV qui maximise cette expression, ce qui optimisera la détection.

On montre que la puissance du projeté d'un signal aléatoire centré de matrice de covariance $\mathbf{M}$ sur un sous-espace E_p de dimension p et de base $\mathbf{V}_p = \{\mathrm{v}_1 \dots \mathrm{v}_p\}$ est :

$$\mathsf{E}\left(||y||^2\right) = \sum_{i=1}^{p} \mathrm{v}_i^T \mathbf{M} \mathrm{v}_i = \mathrm{tr}\left(\mathbf{V}_p^T \mathbf{M} \mathbf{V}_p\right)$$

Si on note Π le projecteur utilisé et x le signal de matrice de covariance $\mathbf{M}$, alors on a : $y = \Pi x$.

Pour les mêmes raisons que celles évoquées dans la section précédente, le choix de ce sous-espace est basé sur la maximisation de l'écart relatif moyen en puissance entre une mesure issue de l'hypothèse H_1 projetée sur le SEV et une mesure issue de l'hypothèse H_0 projeté sur le SEV.

L'écart relatif moyen en puissance s'écrit :

$$\xi_P \triangleq \frac{\mathsf{E}\left(||y||^2|H_1\right) - \mathsf{E}\left(||y||^2|H_0\right)}{\mathsf{E}\left(||y||^2|H_0\right)} \tag{3.35}$$

$$= \frac{\mathrm{tr}\left(\mathbf{V}_p^T(\mathbf{A}+\mathbf{B})\mathbf{V}_p\right) - \mathrm{tr}\left(\mathbf{V}_p^T \mathbf{B} \mathbf{V}_p\right)}{\mathrm{tr}\left(\mathbf{V}_p^T \mathbf{B} \mathbf{V}_p\right)} = \frac{\mathrm{tr}\left(\mathbf{V}_p^T \mathbf{A} \mathbf{V}_p\right)}{\mathrm{tr}\left(\mathbf{V}_p^T \mathbf{B} \mathbf{V}_p\right)} \tag{3.36}$$

C'est l'expression du rapport signal-à-bruit dans le sous-espace E_p de base $\mathbf{V}_p = \{\mathrm{v}_1 \dots \mathrm{v}_p\}$.

$$\boxed{RSB_p = \frac{\mathrm{tr}\left(\mathbf{V}_p^T \mathbf{A} \mathbf{V}_p\right)}{\mathrm{tr}\left(\mathbf{V}_p^T \mathbf{B} \mathbf{V}_p\right)} = \frac{\sum_{i=1}^{p} \mathrm{v}_i^T \mathbf{A} \mathrm{v}_i}{\sum_{i=1}^{p} \mathrm{v}_i^T \mathbf{B} \mathrm{v}_i}} \tag{3.37}$$

Si l'on s'intéresse aux matrices de covariance normalisées :

$$\widetilde{\mathbf{A}} \triangleq \frac{\mathbf{A}}{\mathrm{tr}(\mathbf{A})} \tag{3.38}$$

$$\widetilde{\mathbf{B}} \triangleq \frac{\mathbf{B}}{\mathrm{tr}(\mathbf{B})} \tag{3.39}$$

i.e. $\mathrm{tr}(\widetilde{\mathbf{A}}) = 1$ et $\mathrm{tr}(\widetilde{\mathbf{B}}) = 1$, on peut écrire :

$$RSB_p = \frac{\mathrm{tr}(\mathbf{A})}{\mathrm{tr}(\mathbf{B})} \frac{\mathrm{tr}\left(\mathbf{V}_p^T \widetilde{\mathbf{A}} \mathbf{V}_p\right)}{\mathrm{tr}\left(\mathbf{V}_p^T \widetilde{\mathbf{B}} \mathbf{V}_p\right)} = \frac{\mathrm{tr}(\mathbf{A})}{\mathrm{tr}(\mathbf{B})}\ \rho \tag{3.40}$$

Le second terme ρ peut être interprété comme le gain sur le RSB apporté par la projection.

$$\boxed{\rho = \frac{\mathrm{tr}\left(\mathbf{V}_p^T \widetilde{\mathbf{A}} \mathbf{V}_p\right)}{\mathrm{tr}\left(\mathbf{V}_p^T \widetilde{\mathbf{B}} \mathbf{V}_p\right)}} \tag{3.41}$$

Dans la suite, on considérera que les matrices $\mathbf{A}$ et $\mathbf{B}$ sont de trace unité.

L'objectif est donc de trouver le sous-espace E_p^* dit "optimal" dans lequel ρ est maximal.

Recherche de E_p^*

On cherche une base $\mathbf{H} = [h_1 \dots h_p]$ du sous-espace optimal E_p^*. On montre [3] que les vecteurs de cette base sont solutions de l'équation :

$$(\mathbf{A} - \rho\mathbf{B})\, h_i = h_i\ \mu_i\ ,\ \forall\ i = 1 \dots N \tag{3.42}$$

Puisque $\mathbf{A}$ et $\mathbf{B}$ sont des matrices réelles symétriques, ces vecteurs propres sont orthogonaux. Le problème revient à choisir p vecteurs parmi N qui maximisent la valeur de ρ.

$$\rho = \frac{\mathrm{tr}(\mathbf{H}^T\mathbf{A}\mathbf{H})}{\mathrm{tr}(\mathbf{H}^T\mathbf{B}\mathbf{H})} \tag{3.43}$$

On montre également que la somme des valeurs propres correspondantes est nulle :

$$\sum_{i \text{ tq } h_i \in E_p^*}^{p} \mu_i = 0 \tag{3.44}$$

3.7.1.1 Algorithme

Un algorithme itératif a été mis au point [3] pour calculer les vecteurs de base du sous-espace E_p^* :

Algorithme de détermination de E_p^*

⋆ Choix de la valeur initiale de $\rho^{(0)}$, valeur propre la plus grande de la matrice $\mathbf{B}^{-1}\mathbf{A}$.

⋆ Choix d'un scalaire ε déterminant l'arrêt de la boucle.

⋆ Itération sur n jusqu'à convergence

1. Calcul de la matrice $\mathbf{M}_{n+1} = \mathbf{A} - \rho^{(n+1)}\mathbf{B}$.

2. Calcul des valeurs propres et vecteurs propres $\mu_i^{(n+1)}$ et $h_i^{(n+1)}$ de $\mathbf{M}_{n+1}$.

3. Choix d'une partition $I_p^{(n+1)}$ de $\{1 \dots N\}$ satisfaisant $\text{card}\left(I_p^{(n+1)}\right) = p$ dont les vecteurs $\psi_i^{(n+1)}$ maximise le rapport

$$\rho^{(n+1)} = \frac{\sum_{i=1}^{p} h_i^{(n+1)T}\mathbf{A}\ h_i^{(n+1)}}{\sum_{i=1}^{p} h_i^{(n+1)T}\mathbf{B}\ h_i^{(n+1)}}$$

4. Si $|\rho^{(n+1)} - \rho^{(n)}| < \varepsilon$, on sort de la boucle et on prend $\rho^* = \rho^{(n+1)}$ et $I_p^* = I_p^{(n+1)}$.
   Sinon retour en début de boucle.

Fin de la boucle.

⋆ Le SEV optimal E_p^* est engendré par les vecteurs propres $\{h_i^*\}_{i \in I_p^*}$ et le gain sur le RSB est ρ^*.

On démontre que cet algorithme converge vers une unique solution.

3.7.1.2 Fonctionnelle associée

Toute cette méthode est basée sur la maximisation d'un rapport de puissance. Il est donc naturel de choisir une fonctionnelle de test proportionnelle à une puissance. Nous allons ici choisir de tester celle du signal de mesure projeté dans le sous-espace optimal. Nous avons vu précédemment que ce sous-espace optimal était engendré par une base orthonormale, l'expression de la fonctionnelle associée s'en trouve donc simplifiée :

$$\Lambda_{\text{FASC}}(x) = \sum_{i=1}^{p} \left(h_i^T x\right)^2 \tag{3.45}$$

Résumé FASC

Soient s et b deux vecteurs aléatoires de dimension N représentant respectivement le signal d'intérêt et le bruit. On leur associe les matrices de covariances $\mathbf{A}$ et $\mathbf{B}$, supposées définies positive et de rangs pleins.

La méthode du FASC consiste à trouver un sous-espace dit optimal de dimension $p < N$ (définie *a priori*) dans lequel le rapport signal à bruit est maximisé.

Alors, les vecteurs de base h_k du sous-espace optimal sont les p vecteurs propres parmi les N solutions de l'équation :

$$(\mathbf{A} - \rho\mathbf{B})\, h_k = h_k\ \mu_k$$

qui maximisent le gain ρ sur le RSB apporté par la projection sur ce sous-espace avec :

$$\rho = \frac{h_k^T \mathbf{A} h_k}{h_k^T \mathbf{B} h_k}$$

Un algorithme itératif a été développé pour trouver ces vecteurs.

Pour un vecteur de mesure noté x, la fonctionnelle utilisée pour la détection est la puissance du projeté de x sur le sous-espace optimal :

$$\Lambda_{\mathrm{FASC}}(x) = \sum_{i=1}^{p} \left(h_i^T x\right)^2$$

3.7.2 Application au cas Antares

Nous utiliserons ici les mêmes notations que celles utilisées dans la section 3.6.2. Le signal acoustique n'est maintenant plus centré et on peut l'écrire sous la forme :

$$S_A = S_A^c + \overline{s_A} \tag{3.46}$$

S_A^c représente le signal acoustique centré et $\overline{s_A}$ la moyenne de ce même signal. L'idée est toujours de trouver un sous-espace vectoriel de dimension $p < N$, de base $\mathbf{H}_p = \{\mathrm{H}_1 \ldots \mathrm{H}_p\}$, dans lequel l'écart relatif moyen en puissance défini précédemment est maximal.

La puissance du signal acoustique projeté sur un sous-espace de dimension p est :

$$\begin{aligned} P_{S_A}^{(p)} &= \mathrm{tr}\left[\mathsf{E}\left((\mathbf{H}_p\ S_A)(\mathbf{H}_p\ S_A)^T\right)\right] && (3.47)\\ &= \mathrm{tr}\left[\mathsf{E}\left(\mathbf{H}_p^T\ S_A S_A^T \mathbf{H}_p\right)\right] && (3.48)\\ &= \mathrm{tr}\left[\mathbf{H}_p^T \mathsf{E}\left(S_A S_A^T\right) \mathbf{H}_p\right] && (3.49) \end{aligned}$$

On peut développer le terme $\mathsf{E}\left(S_A S_A^T\right)$ pour faire intervenir la matrice de covariance $\mathbf{A}$ du signal acoustique :

$$\mathsf{E}\left(S_A S_A^T\right) = \mathbf{A} + \overline{s_A}\ \overline{s_A}^T \tag{3.50}$$

En posant $\mathbf{A}' = \mathbf{A} + \overline{s_A}\ \overline{s_A}^T$, il vient :

$$P_{S_A}^{(p)} = \mathrm{tr}\left[\mathbf{H}_p^T \left(\mathbf{A} + \overline{s_A}\ \overline{s_A}^T\right) \mathbf{H}_p\right] = \mathrm{tr}\left[\mathbf{H}_p^T \mathbf{A}' \mathbf{H}_p\right] \tag{3.51}$$

L'écart relatif s'écrit donc de la même manière que précédemment en utilisant non plus la matrice de covariance $\mathbf{A}$ du signal mais la matrice $\mathbf{A}'$:

$$\xi_P = \frac{\mathrm{tr}\left[\mathbf{H}_p^T \mathbf{A}' \mathbf{H}_p\right]}{\mathrm{tr}\left[\mathbf{H}_p^T \mathbf{B} \mathbf{H}_p\right]} \tag{3.52}$$

On retombe sur une expression similaire à celle trouvée dans la section précédente, il est donc possible d'utiliser le même algorithme pour trouver le sous-espace dans lequel ce rapport signal-à-bruit est maximal. La moyenne du signal est alors prise en compte.

3.7.3 Résultats expérimentaux

Le protocole de simulation est toujours celui décrit dans le premier chapitre. Le bruit est le bruit de mer réel et le signal acoustique est celui décrit au début de ce chapitre. Les vecteurs de base du sous-espace optimal sont calculés à partir des matrices $\mathbf{A}'$ et $\mathbf{B}$ définies précédemment.

3.7.3.1 Recherche de la valeur de p optimale

Pour trouver la valeur optimale de p, il faut avoir recours à l'expérience.
Le rapport signal-à-bruit moyen considéré est de -60 dB.

La meilleure des courbes COR semble être atteinte pour une valeur de $p = 1$.

$$\boxed{p_{\mathrm{opt}}^{(\mathrm{FASC})} = 1} \tag{3.53}$$

La figure 3.12 nous montre l'écolution des courbes COR pour différentes valeurs de p.

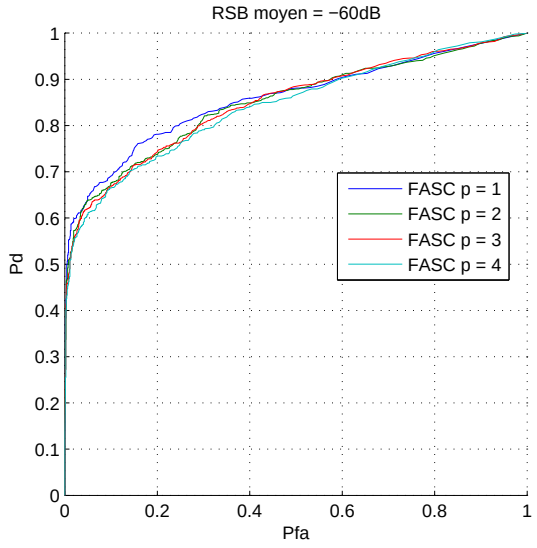

FIG. 3.12 – Courbes COR dans le cas des signaux acoustiques et bruit réels pour différentes valeurs de p.

Notons que lorsque $p = 1$, le FASC est identique au FAS. Les courbes d'évolution du coefficient d'atténuation en puissance α etdu RSB en fonction de la distance parcourue par le front d'onde pour les deux courbes COR de référence choisies sont donc les mêmes que celle du FAS.

Comme pour les méthodes précédentes, on s'intéresse à l'évolution en fonction de la distance du coefficient d'atténuation en puissance α à appliquer au bruit pour que les performances du détecteur soient identiques à celles décrites par les deux courbes de référence choisies (voir annexe G). Cette évolution est tracée sur la figure 3.10

3.8 Adaptation du FASC au cas multibruits : le FASCM

Intéressons nous maintenant au cas où le bruit de mer ambiant n'est plus le seul à venir perturber le signal acoustique. En effet, la mer Méditerranée (comme tout autre mer ou océan) possède un écosystème générant une grande diversité de bruits biologiques. A ceux-ci viennent également s'ajouter ceux issus de l'activité humaine (bateaux, sous-marins, sonars, ...).
Parmi ces nouveaux bruits, certains occupent la même bande passante que le signal acoustique d'intérêt et peuvent donc être détecté à la place de celui-ci et générer une fausse alarme au niveau du détecteur. Lorsqu'ils ne se substituent pas au signal d'intérêt, ces signaux vont dans tous les cas modifier le bruit ambiant.

Afin de s'affranchir au mieux de ces nouvelles contraintes, nous allons adapter une

méthode de classification de signaux dérivée du FASC au problème de détection propre au télescope Antares. Cette méthode sera nommée FASCM.

3.8.1 Théorie

Dans les problèmes précédents, nous supposions que les observations pouvaient suivre deux hypothèses. Dans la première, l'observation était constituée du bruit seul et dans la seconde, de la superposition du signal d'intérêt et du bruit.
Le cas qui nous intéresse maintenant est celui de la multiplicité des bruits que l'on peut rencontrer. Nous utiliserons les notations suivantes :

- ⋆ s le signal d'intérêt centré de matrice de variance covariance $\mathbf{A}$.
- ⋆ b_i le bruit n°i centré de matrice de variance covariance $\mathbf{B}_i$.
- ⋆ N_b est le nombre de bruits perturbateurs.

Chaque signal mis en jeu est représenté par un vecteur échantillonné de dimension N.

Nous ferons l'hypothèse que tous les bruits sont indépendants entre eux (et également indépendant du signal d'intérêt). Les matrices $\mathbf{A}$ et $\mathbf{B}_i$ sont normalisées par leurs trace, nous travaillerons ainsi non pas sur les rapports signal-à-bruit mais sur les gains sur les rapports signal-à-bruit.

Notant x la mesure, il est possible pour chaque bruit b_i de définir le couple d'hypothèses suivant :

$$\begin{aligned} H_{0i} &: x = b_i \\ H_{1i} &: x = s + b_i \end{aligned} \tag{3.54}$$

Le but est toujours d'essayer de détecter le signal d'intérêt s quelque soit le bruit b_i rencontré. L'idée est de trouver un sous-espace vectoriel de dimension $p < N$ qui optimiserait une certaine quantité. Si on note $s^{(p)}$ et $b_i^{(p)}$ respectivement le signal s et le bruit b_i projeté sur ce sous-espace, on peut définir l'écart relatif moyen en puissance entre les mesures issues des deux hypothèses précédentes par :

$$\frac{\mathsf{E}\left((s^{(p)} + b_i^{(p)})^2\right) - \mathsf{E}\left(b_i^{(p)2}\right)}{\mathsf{E}\left(b_i^{(p)2}\right)} = \frac{\mathsf{E}\left(s^{(p)2}\right)}{\mathsf{E}\left(b_i^{(p)2}\right)} \tag{3.55}$$

On retombe sur l'expression du rapport signal-à-bruit dans le sous-espace. Si on note $\mathbf{H}_p = \{H_1 \dots H_p\}$ sa base, celui-ci s'écrit pour le bruit b_i :

$$\rho_i = \frac{\mathrm{tr}\left(\mathbf{H}_p^T \mathbf{A} \mathbf{H}_p\right)}{\mathrm{tr}\left(\mathbf{H}_p^T \mathbf{B}_i \mathbf{H}_p\right)} \tag{3.56}$$

Il faut maintenant trouver un critère à maximiser qui tienne compte de ces différents rapports signal-à-bruit. On montre [3] qu'un critère pertinent est la moyenne géométrique de ces RSB :

$$\boxed{\rho = \sqrt[N_b]{\prod_{i=1}^{N_b} \rho_i} = \sqrt[N_b]{\prod_{i=1}^{N_b} \frac{\mathrm{tr}\left(\mathbf{H}^T \mathbf{A} \mathbf{H}\right)}{\mathrm{tr}\left(\mathbf{H}^T \mathbf{B}_i \mathbf{H}\right)}}} \tag{3.57}$$

3.8.1.1 Sous-espace de dimension 1

Commençons par le cas simple d'un sous-espace de dimension 1. On peut remarquer que maximiser le terme $\sqrt[N_b]{\prod_{i=1}^{N_b} \frac{\mathrm{tr}\left(H_p^T \mathbf{A} H_p\right)}{\mathrm{tr}\left(H_p^T \mathbf{B}_i H_p\right)}}$ revient à maximiser $\prod_{i=1}^{N_b} \frac{\mathrm{tr}\left(H_p^T \mathbf{A} H_p\right)}{\mathrm{tr}\left(H_p^T \mathbf{B}_i H_p\right)}$. On montre que le vecteur h dans la direction duquel ρ est maximal est solution de l'équation aux valeurs propres suivante :

$$\mathbf{A}\ h = \frac{1}{N_b} \sum_{i=1}^{N_b} \rho_i\ \mathbf{B}_i\ h \tag{3.58}$$

Un algorithme a été développé [3] pour trouver ce vecteur. Sa convergence et l'unicité de sa solution ont été démontrées.

3.8.1.2 Sous-espace de dimension p

Considérons maintenant un sous-espace E_p de dimension p. Notons E_p^* le sous-espace dit optimal dans lequel le critére ρ est maximal, $\mathbf{H}^* = \{h_1 \dots h_p\}$ la matrice contenant les vecteurs de base de E_p^*. Ces vecteurs sont supposés orthonormaux. On montre [3] que les vecteurs de cette base sont les p solutions de l'équation aux valeurs propres :

$$\left(\mathbf{A} - \frac{1}{N_b} \sum_{i=1}^{N_b} \rho_i \mathbf{B}_i\right) X = X\mu \tag{3.59}$$

qui maximisent le critère ρ. Un algorithme itératif a été développé :

Algorithme de détermination de E_p^* dans le cas multibruits

⋆ Choix de la valeur initiale de $\rho^{(0)} = \left\{\rho_1^{(0)} \dots \rho_N^{(0)}\right\}$.

⋆ Choix d'un scalaire ε déterminant l'arrêt de la boucle.

⋆ Itération sur n jusqu'à convergence

1. Calcul de la matrice $\mathbf{M}_{n+1}\left(\rho^{(n)}\right) = \mathbf{A} - \frac{1}{N_b}\sum_{i=1}^{N_b} \rho_i^{(n)}\mathbf{B}_i$.

2. Calcul des valeurs propres et vecteurs propres $\mu^{(n+1)}$ et $\Psi^{(n+1)}$ de $\mathbf{M}_{n+1}$ en résolvant l'équation :

$$\mathbf{M}_{n+1}\left(\rho^{(n)}\right)\Psi^{(n+1)} = \Psi^{(n+1)}\mu^{(n+1)}$$

3. On garde parmi ces solutions les p vecteurs $\mathbf{H}^{(n+1)} = \left\{H_1^{(n+1)} \dots H_p^{(n+1)}\right\}$ qui donnent les N_b plus grandes valeurs de :

$$\rho_i^{(n+1)} = \frac{\mathrm{tr}\left(\mathbf{H}^{(n+1)T}\mathbf{A}\mathbf{H}^{(n+1)}\right)}{\mathrm{tr}\left(\mathbf{H}^{(n+1)T}\mathbf{B}_i\mathbf{H}^{(n+1)}\right)}$$

4. Evaluation de $\rho^{(n+1)} = \sqrt[N_b]{\prod_{i=1}^{N_b} \rho_i^{(n+1)}}$.

5. Si $\rho^{(n+1)} - \rho^{(n)} < \varepsilon$, on sort de la boucle et on prend $\rho^* = \rho^{(n+1)}$ et $\mathbf{H}^* = \mathbf{H}^{(n+1)}$.

   Sinon retour en début de boucle.

Fin de la boucle.

⋆ Le SEV optimal E_p^* est engendré par les vecteurs propres contenus dans $\mathbf{H}^*$ et le gain sur le rapport signal-à-bruit est ρ^*.

On démontre que celui-ci converge vers une unique solution.

3.8.1.3 Fonctionnelle associée

Le choix de la fonctionnelle associée à cette méthode de détection est basé sur le même raisonnement que celui décrit dans le paragraphe du FASC. Nous choisissons comme fonc-

tionnelle la valeur de la puissance du signal projeté sur le sous-espace optimal.

$$\Lambda_{\text{CLASSIF}}(x) = \sum_{i=1}^{p} \left(h_i^T x\right)^2 \tag{3.60}$$

3.8.2 Adaptation au cas Antares

Le signal d'intérêt est maintenant le signal acoustique Antares et n'est plus centré. Les notations sont les mêmes que celles utilisées précédemment :

$$S_A = S_A^c + \overline{s_A}$$

S_a^c représente le signal acoustique centré et $\overline{s_A}$ la moyenne de ce même signal.

Le cadre des signaux de l'expérience Antares n'est pas exactement le même que celui décrit par le jeu d'hypothèses 3.54. En effet, le bruit de mer b_m est tout le temps présent, alors que les bruits biologiques ne le sont que par intermittence. Pour deux bruits de type biologique supplémentaires, le signal de test utilisé pour évaluer les performances du détecteur est représenté sur la figure 3.13. Tous les cas de figure que l'on peut rencontrer dans la réalité sont représentés dans ce signal.

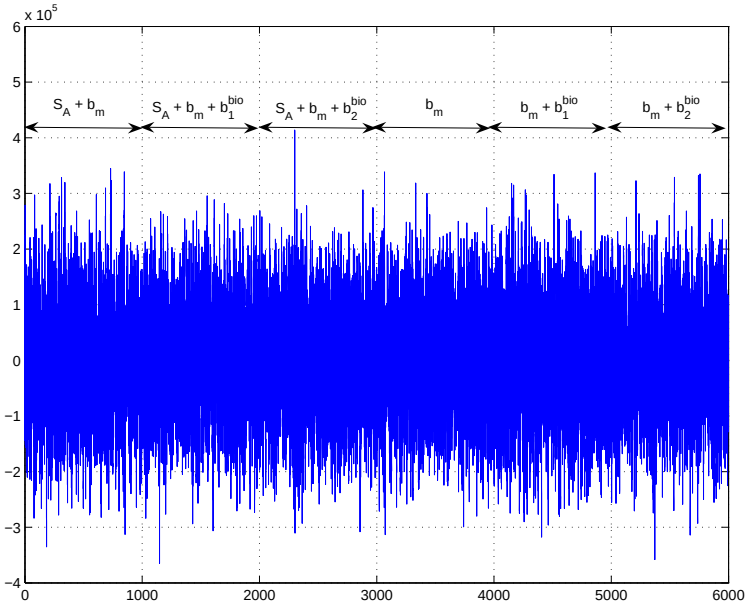

FIG. 3.13 – Allure des signaux de test utilisés pour l'évaluation des performances du détecteur de le cas Antares pour deux bruits biologiques supplémentaires

Il suffit donc d'utiliser les notations suivantes pour se ramener aux hypothèses 3.54 :

$$b_1 = b_m \tag{3.61}$$
$$b_2 = b_m + b_1^{bio} \tag{3.62}$$
$$b_3 = b_m + b_2^{bio} \tag{3.63}$$

Ce résultat est facilement extrapolable à un nombre N_b de bruits biologiques supplémentaires :

$$b_1 = b_m \tag{3.64}$$
$$b_2 = b_m + b_1^{bio} \tag{3.65}$$
$$\vdots \tag{3.66}$$
$$b_{N_b+1} = b_m + b_{N_b}^{bio} \tag{3.67}$$

Néanmoins, on perd l'hypothèse d'indépendance entre les bruits. Ce sont les conditions dans lesquelles la méthode sera utilisée pour détecter le signal acoustique Antares.

3.8.2.1 Signal d'intérêt non centré

Comme nous l'avons rappelé précédemment, le signal d'intérêt n'est maintenant plus centré. Qu'elle est l'influence de cette moyenne sur la méthode du FASCM ? En utilisant les mêmes notations et raisonnements que ceux utilisés dans la section 3.7.2, il vient :

$$\mathbf{A}' = \mathbf{A} + \overline{s_A}\ \overline{s_A}$$

on peut réécrire pour chaque bruit b_i le rapport signal-à-bruit correspondant comme :

$$\rho_i' = \frac{\mathrm{tr}\left(\mathbf{H}_p^T \mathbf{A}' \mathbf{H}_p\right)}{\mathrm{tr}\left(\mathbf{H}_p^T \mathbf{B} \mathbf{H}_p\right)} \tag{3.68}$$

Il faut donc maintenant maximiser le terme :

$$\rho' = \sqrt[N_b]{\prod_{i=1}^{N_b} \rho_i'}$$

L'algorithme utilisé est le même que celui étudié dans la section précédente en utilisant la matrice $\mathbf{A}'$ à la place de la matrice $\mathbf{A}$.

3.8.3 Résultats expérimentaux

Pour illustrer les résultats précédents, nous avons considéré deux autres bruits. On rappelle que ces bruits sont soit d'origine biologique (crevettes, dauphins, etc.), soit d'origine humaine (bateaux, sous-marins, etc.). Nous ne disposons pas, à ce stade de l'étude, d'enregistrements de tels signaux possédant les caractéristiques nécessaires pour nos simulations (fréquence d'échantillonnage, résolution, etc.).

Pour pallier ce problème dans nos simulations, nous avons donc synthétisé deux nouveaux bruits basés sur des signaux biologiques réels. En effet, nous avions à disposition des enregistrements d'animaux marins (fournis par la société *Semantic T.S.*, partenaire de cette thèse). Malheureusement, la bande passante limitée de ces enregistrements (22.05 kHz, dû à la fréquence d'échantillonnage de 44.1 kHz) nous a empêché de les utiliser directement (bande passante utile définie de 7 kHz à 150 kHz). Ces bruits ont donc été transposés en fréquence en utilisant une modulation d'amplitude dont les fréquences porteuses sont respectivement de 60 kHz et 80 kHz. Même si ils ne reflètent plus directement les évènements biologiques sur lesquels ils sont basés, les signaux obtenus opèrent bien dans la bande en fréquence utile.

Le premier bruit biologique est basé sur un enregistrement de *crevettes claqueuses*, le second sur un enregistrement de *langoustes*.

Recherche de la valeur de p optimale

Là encore, la valeur de p optimale est cherchée expérimentalement. A chaque itération de l'algorithme, il est nécessaire de tester chaque sous-espace de dimension p. Le nombre d'itération peut devenir très important. Par exemple, pour une dimension de l'espace $N = 64$, la résolution de l'équation 3.59 donne 64 vecteurs propres. Pour une valeur *a priori* de p, le nombre de sous-espaces générés par une base de 5 vecteurs issus de ces 64 vecteurs propres est donc de C_N^p. Ce nombre devient vite très important (par exemple, pour $p = 5$, il y a déjà près de 8 millions de possibilités), et nous nous retrouvons limités par la puissance de calcul des ordinateurs à disposition. C'est pour cela que nous avons testé dans cette section des valeurs de p variant de 1 à 4.

Nous avons superposé sur la figure 3.14 les courbes COR issues de la méthode du FASCM pour p variant de 1 à 4. Puisque nous sommes en présence de plusieurs bruits, il est possible de définir le même nombre de rapports signal à bruit moyens. Dans une première approche, ces différents RSB sont considérés égaux.

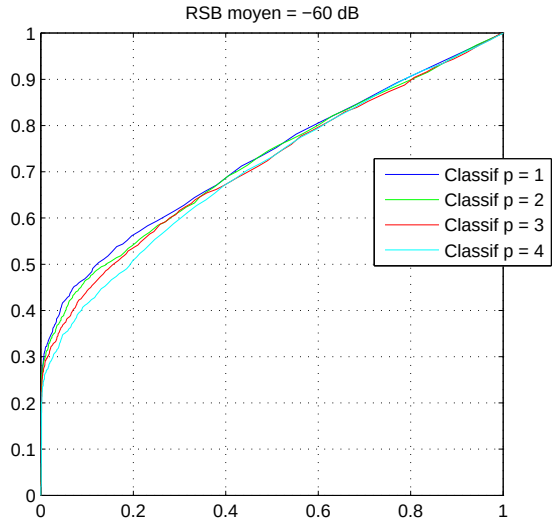

FIG. 3.14 – Courbes COR dans le cas réel pour la méthode du FASCM. Deux bruits biologiques additionnels sont considérés et chaque rapport signal-à-bruit est de -60dB.

La courbe COR correspondant à une de $p = 1$ semble être la meilleure, nous allons donc considérer que :

$$\boxed{p_{\text{opt}}^{(\text{CLASSIF})} = 1} \tag{3.69}$$

en gardant à l'esprit que les valeurs de p supérieures à 5 n'ont pu être testées. Ce résultat n'est donc là que pour valider l'intérêt de la méthode.

Remarque : D'autres signaux complètement synthétiques ont également été générés (i.e. plusieurs signaux aléatoires différemment filtrés dans la bande utile) pour ces simulations, et des résultats équivalents sont obtenus.

3.9 Bilan des méthodes proposées : comparaison des résultats expérimentaux

Comparons maintenant les résultats de chacune des méthodes étudiées précédemment. Dans un premier temps, seul un bruit sera considéré et nous nous placerons ensuite dans le cas multi-bruit étudié avec la méthode du FASCM.

Dans chacun des cas, les performances du détecteur sont caractérisées par deux courbes COR différentes dont les formes sont données dans l'annexe G. On s'intéresse là encore à l'évolution en fonction de la distance parcourue par le front d'onde du coefficient d'atténuation en puissance α à appliquer au bruit pour que les performances du détecteur soient

identiques déduite de la courbe COR de référence choisie.

3.9.1 Cas mono-bruit

Nous nous trouvons ici dans le cadre d'hypothèses classique (équations 3.2 et 3.2) et le seul bruit est le bruit de mer. Sur les figures 3.15 et 3.16 sont superposées les courbes représentant l'atténuation en puissance du bruit en fonction de la distance pour chacune des méthodes étudiées dans les sections précédentes et pour les deux courbes COR de référence.

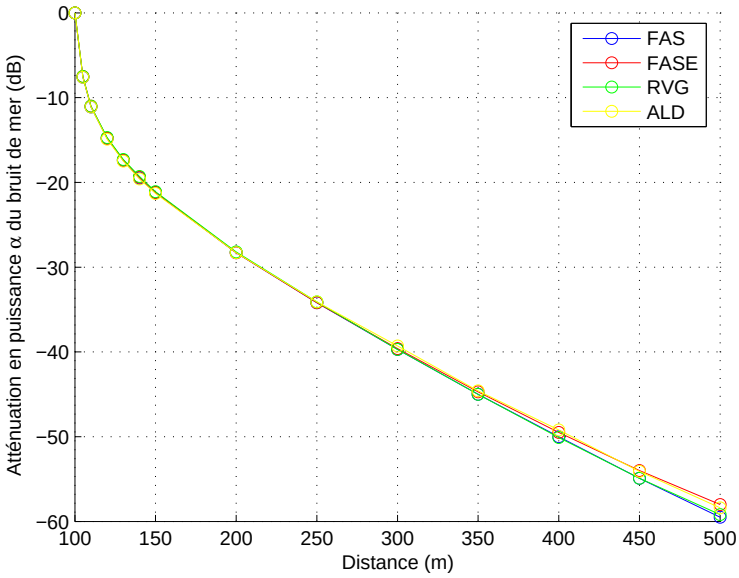

FIG. 3.15 – Evolution du coefficient α en fonction de la distance pour des performances correspondant à la première courbe COR de référence - Comparaison de toutes les méthodes étudiées.

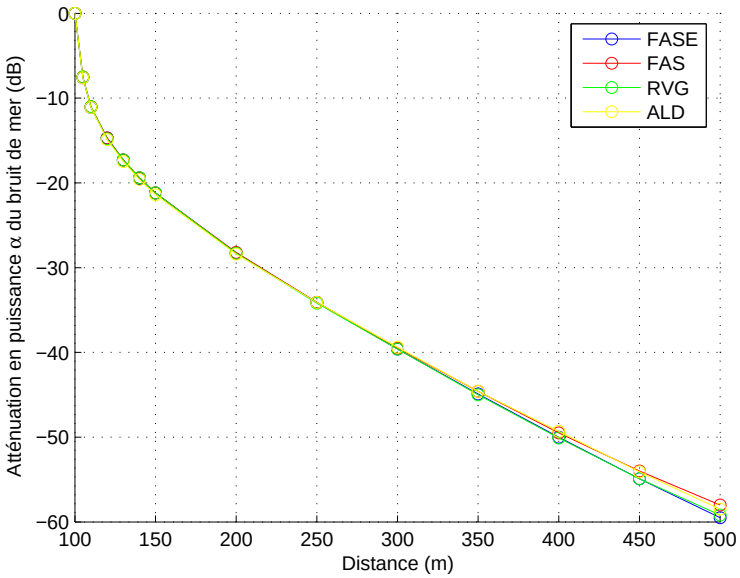

FIG. 3.16 – Evolution du coefficient α en fonction de la distance pour des performances correspondant à la seconde courbe COR de référence - Comparaison de toutes les méthodes étudiées.

Pour les deux courbes COR de référence étudiées, les courbes sont quasiment identiques quelle que soit la méthode de détection considérée. Ce résultat est cohérent car l'atténuation en puissance ne dépend que des phénomènes physiques mis en jeu.

Les figures 3.17 et 3.18 représentent l'évolution en fonction de la distance du RSB au niveau de l'hydrophone calculé après compensation de la puissance du bruit pour chacune des courbes COR de référence.

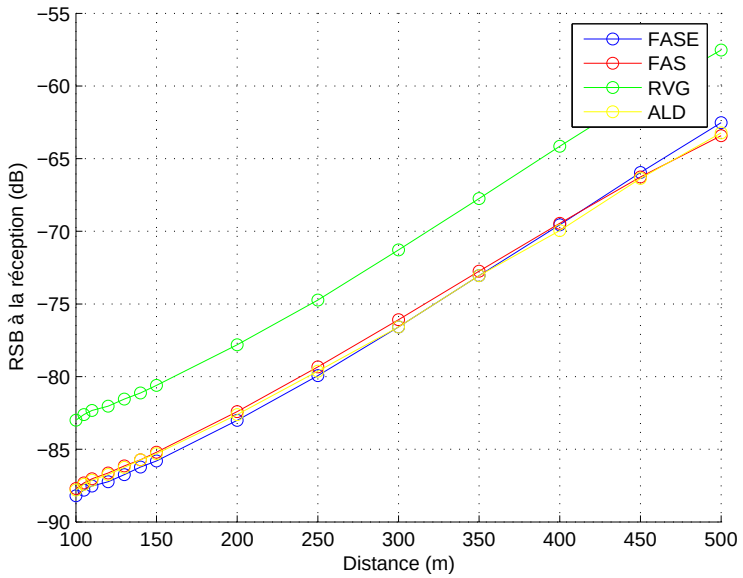

FIG. 3.17 – Evolution du RSB au niveau de l'hydrophone en fonction de la distance pour la première courbe COR de référence - Comparaison de toutes les méthodes étudiées.

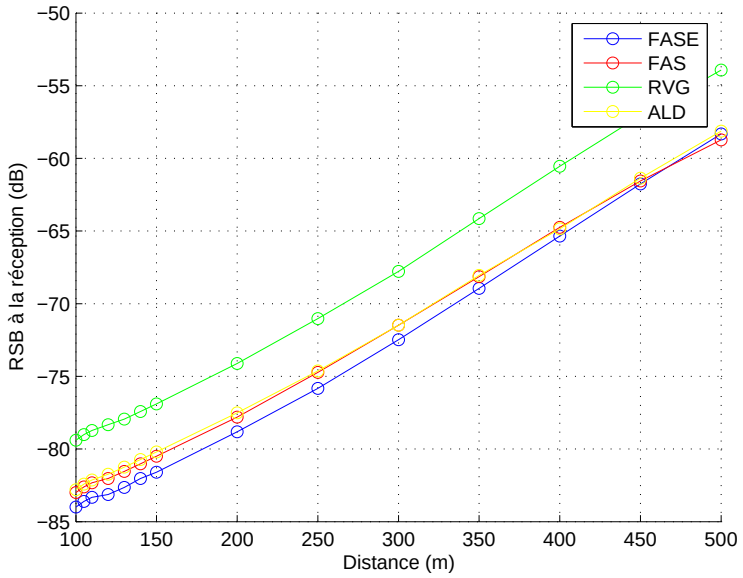

FIG. 3.18 – Evolution du RSB au niveau de l'hydrophone en fonction de la distance pour la seconde courbe COR de référence - Comparaison de toutes les méthodes étudiées.

Interprétations :

Tout d'abord, on remarque que les courbes ont la même allure quelque soit la courbe

COR de référence choisie. Seule change la valeur du RSB à $d = 100$m. En changeant de performance de détection, les courbes sont donc simplement translatées verticalement suivant la valeur du RSB à l'origine.

Dans les deux cas étudiés, la courbe la plus défavorable au vu du RSB est celle issue de la méthode du RVG. Elle se situe en effet au-dessus de toutes les autres, et correspond donc, à mêmes performances, à des valeurs de RSB plus faibles.

Les courbes déduites des méthodes de détections basées sur une approche aléatoire du signal font conclure à de meilleurs performances de détection. Si l'on classe ces méthodes par ordre de performances, on obtient le résultat suivant :

1. Méthode du FASE pour un sous-espace de dimension 2
2. Méthode du FAS / Méthode d'ALD (ex-aequo)

La méthode du FASC n'est pas représenté car nous avons remarqué dans la section précédente que la taille optimale du sous-espace adapté au FASC était 1 et que l'on se ramenait donc à la méthode du FAS.

On peut remarquer que pour des distances de propagation supérieures à 350m, les performances des méthodes FAS/FASE/ALD deviennent équivalentes et pour des distances proches de 500m, la méthode du FAS devient légèrement meilleure.

Remarque : Nous avons vu que les méthodes du FAS et du FASE sont basées sur une approche aléatoire du signal par le biais de sa matrice de covariance. Pour nos simulations, l'intervalle de distance retenu a été de 100 à 500m. La matrice de covariance a donc été estimée à partir de réalisations du signal acoustique générés à des distances tirées aléatoirement dans cet intervalle suivant la loi adéquate (équation 1.43).

3.9.2 Cas multi-bruits

Pour comparer la méthode du FASCM aux méthodes issues du FAS, il est nécessaire de considérer plusieurs bruits. Au bruit de mer toujours présent, nous allons donc rajouter deux bruits d'origine biologique.

Nous disposons à ce stade de réalisations des signaux suivants :

⋆ S_A : signal acoustique d'intérêt.

- ⋆ b_m : bruit de mer ambiant.
- ⋆ b_1^{bio} : bruit biologique supplémentaire 1.
- ⋆ b_2^{bio} : bruit biologique supplémentaire 2.

Pour les méthodes classique étudiées précédemment (RVG, ALD, FAS, FASE et FASC), il est nécessaire de revenir a un seul bruit. Pour cela, nous allons considérer comme signal perturbateur la moyenne des bruits précédents :

$$b_{\text{moyen}} = \frac{1}{3}\left(3\ b_m + b_1^{\text{bio}} + b_2^{\text{bio}}\right)$$

La matrice de covariance $\mathbf{B}$ du bruit sera donc estimée à partir des réalisations de b_{moyen}.

Pour la méthode du FASCM, les bruits considérés sont les suivants :

$$\begin{aligned} b_1 &= b_m \\ b_2 &= b_m + b_1^{\text{bio}} \\ b_3 &= b_m + b_2^{\text{bio}} \end{aligned}$$

A ces trois bruits différents correspondent trois rapports signal-à-bruit distincts : RSB_0, RSB_1 et RSB_2. Nous ferons l'hypothèse que ces trois bruits ont la même puissance et donc que ces trois RSB sont égaux et donc :

$$\text{RSB} = \text{RSB}_0 = \text{RSB}_1 = \text{RSB}_2$$

Dans ces conditions, il est possible de tracer l'évolution du coefficient d'atténuation en puissance en fonction de la distance, à performances de détection constantes.

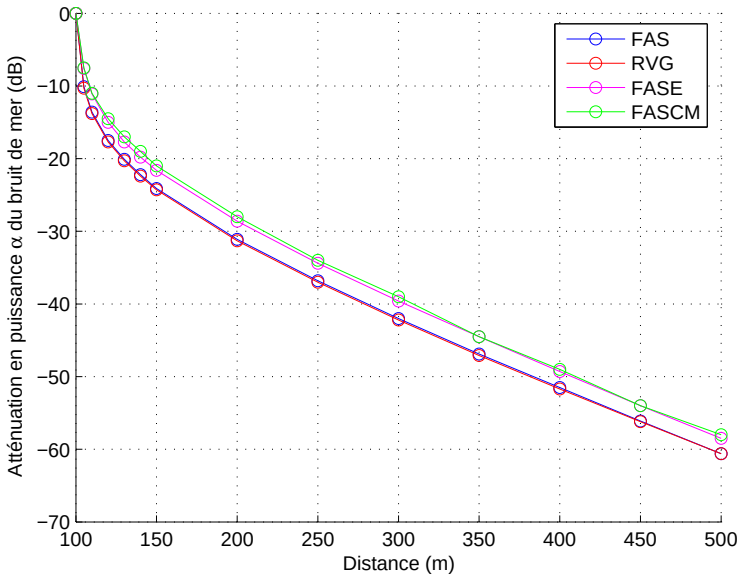

FIG. 3.19 – Evolution du coefficient d'atténuation en puissance en fonction de la distance pour la première courbe COR de référence - Cas de plusieurs bruits perturbateurs.

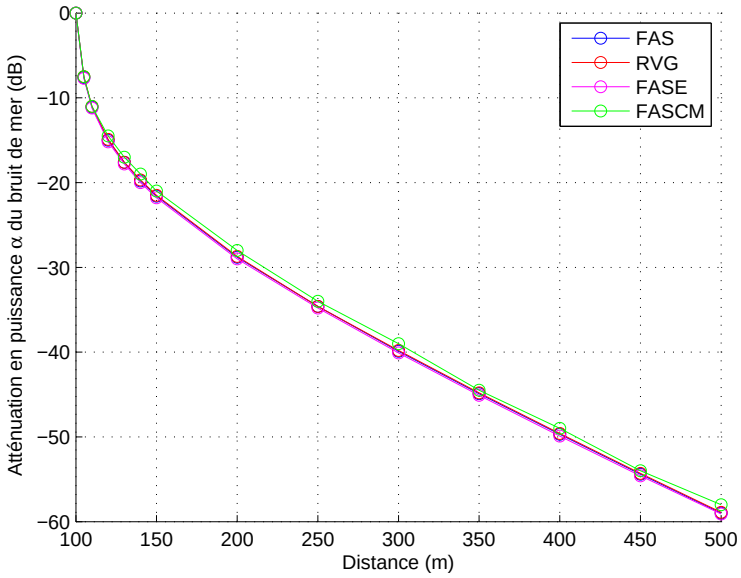

FIG. 3.20 – Evolution du coefficient d'atténuation en puissance en fonction de la distance pour la seconde courbe COR de référence - Cas de plusieurs bruits perturbateurs.

On représente également encore l'évolution du RSB après compensation en puissance en fonction de la distance et au niveau de l'hydrophone.

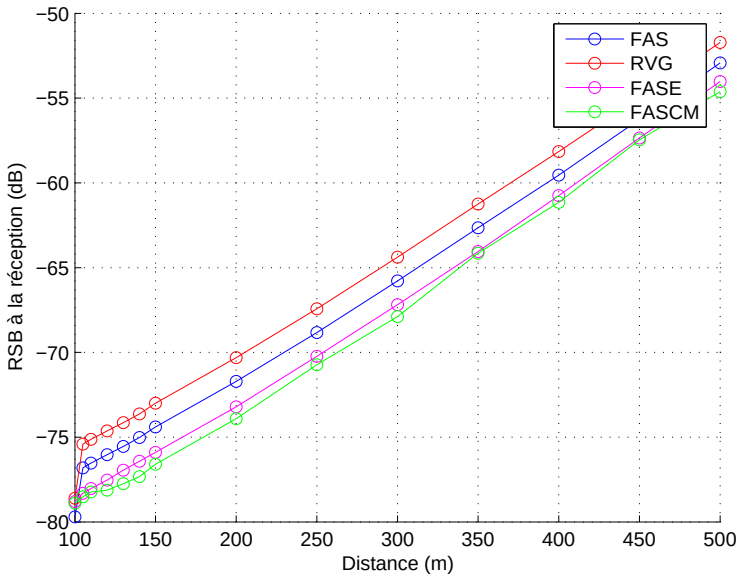

FIG. 3.21 – Evolution du RSB après compensation en puissance en fonction de la distance pour la première courbe COR de référence - Cas de plusieurs bruits perturbateurs.

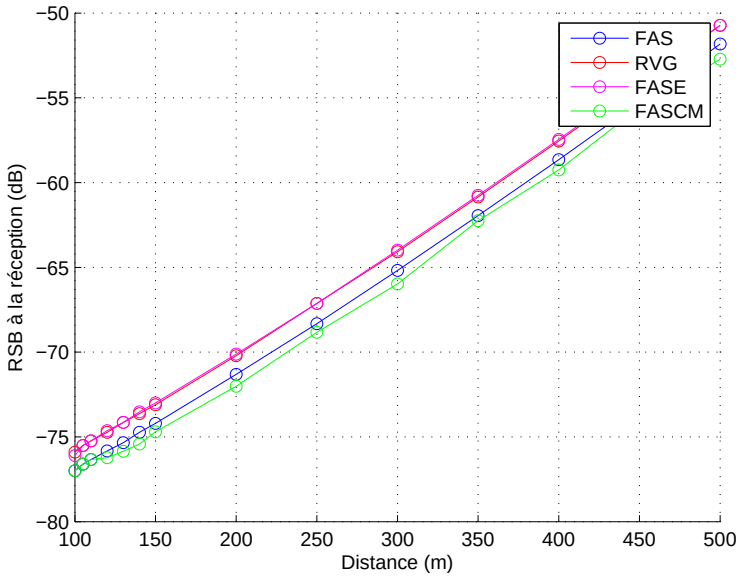

FIG. 3.22 – Evolution du RSB après compensation en puissance en fonction de la distance pour la seconde courbe COR de référence - Cas de plusieurs bruits perturbateurs.

Pour les méthodes de détection précédentes (RVG, FAS et dérivés), les courbes présentent une allure similaire à celles de la figure 3.17, mais on remarque une perte sur le RSB nécessaire pour obtenir les performances fixées *a priori*. Cette perte s'explique par le

changement d'hypothèses sur le bruit. En effet, plusieurs bruits peuvent maintenant perturber le signal, or pour toutes ces méthodes, la connaissance du bruit se fait par le biais d'une matrice de covariance unique. Celle-ci étant estimée à partir du bruit "moyen", elle caractérise moins bien chacun de ces bruits lorsqu'il apparaît seul.

La méthode du FASCM, dont la connaissance de chaque bruit se fait par le biais d'autant de matrices de covariance, présente de légèrement meilleurs résultats au vu de sa courbe qui se place en dessous de toutes les autres.

3.10 Extensions et perspectives

Méthode du FAS et dérivés

Telle que nous l'avons présenté dans ce chapitre, un détecteur basé sur la méthode du FAS (ou d'un de ses dérivé FASE ou FASC) se met sous la forme suivante :

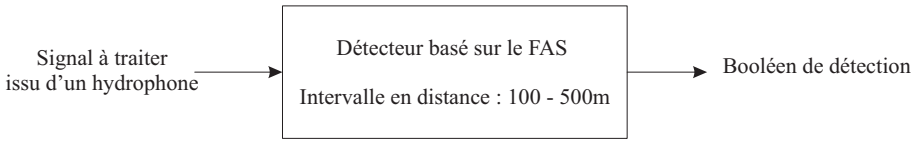

FIG. 3.23 – Utilisation du détecteur basé sur la méthode du FAS (ou l'une de ses dérivées) étudiée dans ce chapitre.

Celui-ci a été étudié pour réagir à des signaux acoustiques générés pour des distances allant de D_{min} et D_{max} (ici 100 et 500m).

Afin d'essayer d'améliorer les performances de détection, on pourrait imaginer un ensemble de détecteurs travaillant en parallèle chacun dans une plage de distance déterminée. Un exemple d'une telle structure est proposé sur la figure 3.24 :

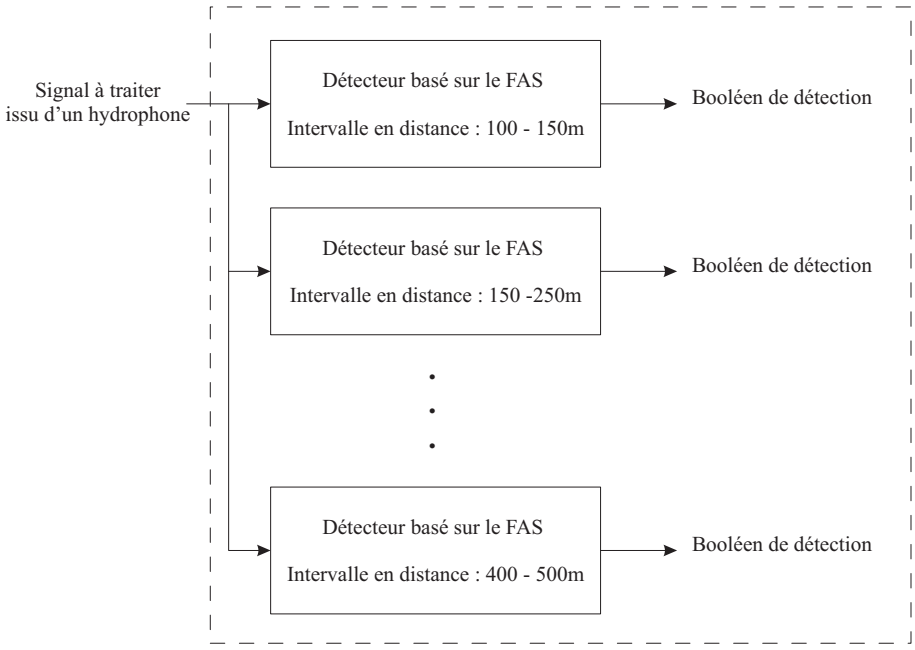

FIG. 3.24 – Proposition d'une structure de détecteurs basés sur la méthode du FAS (ou l'une de ses dérivées) étudiée dans ce chapitre. Les valeurs des intervalles de distance sont données à titre indicatif pour illustrer la figure.

Le choix des intervalles en distances pourrait être basé sur l'évolution du coefficient de corrélation en fonction de la distance. Les matrices de covariance du signal de chaque bloc de détection sont donc estimées dans l'intervalle de distance correspondant.

Si l'on fait l'hypothèse que les filtres sont calculés *a priori*, toutes ces méthodes se prêtent bien à une application temps réel car elles ne consistent qu'en un filtrage linéaire. Si l'on veut améliorer encore les performances de détection, il est possible de calculer plusieurs filtres, correspondant aux différents états de bruits de mer que l'on peut rencontrer (calme, agitée, etc.) ou même correspondant aux différentes périodes de l'année (en effet, l'influence des populations d'animaux marins n'est pas la même tout au long de l'année).

Méthode du FASCM

Les extensions précédentes sont valables dans le cas multi-bruits.

En présence de plusieurs bruits, nous avons considéré que ceux-ci avaient la même puissance (et donc que les RSB correspondants étaient égaux). Pour se rapprocher plus de la réalité, il serait intéressant de pondérer les différents bruits biologiques par des coefficients "dynamiques". Ceux-ci pourraient en effet être modifiés par l'évolution de certains paramètres (heure, période de l'année, conditions climatiques, etc.). Par exemple, la densité des populations de crevettes claqueuses n'est pas la même en été ou en hiver. L'étude de

l'évolution de ces coefficients mène naturellement à une autre extension de ce travail de thèse : l'étude et l'utilisation de "vrais" bruits biologiques issus de mesures réelles. Les recherches effectuées durant cette thèse pour trouver de tels signaux n'ont pas abouti, ceux-ci ne respectant par les contraintes (notamment au niveau de la fréquence d'échantillonnage) nécessaires à nos simulations.

L'étude et l'utilisation d'une vraie méthode de classification dans le cas Antares est une extension intéressante ce de travail de thèse. En effet, la comparaison et le recoupement des résultats obtenus en classant les différents signaux que l'on peut rencontrer (signal acoustique et bruits biologiques) permettraient sans doute d'améliorer la détection.

3.11 Conclusion

Le but de ce chapitre était d'étudier différentes méthodes de détection basées sur une approche statistique des signaux mis en jeu. Dans un premier temps, un seul signal perturbateur a été pris en compte : le bruit de mer. Dans ces conditions, les méthodes de détection étudiées ont été :

- ⋆ L'analyse linéaire discriminante, qui vise à ranger chaque mesure dans la classe "*La mesure est constituée de signal et de bruit*" ou dans la classe "*La mesure est constituée de bruit uniquement*".

- ⋆ La méthode du FAS, qui consiste à projeter la mesure sur un sous-espace de dimension 1 généré par le vecteur propre solution de l'équation caractéristique du FAS associé à la plus grande valeur propre.

- ⋆ La méthode du FASE, extension de la méthode du FAS, qui consiste à projeter la mesure sur un sous-espace généré par les vecteurs propres associés aux plus grandes valeurs propres solutions de l'équation caractéristique du FAS, et à seuiller une fonctionnelle basée sur la puissance de ce projeté.

- ⋆ La méthode du FASC, autre extension du FAS. On recherche ici un sous-espace de dimension p définie *a priori* dans lequel le rapport signal à bruit est maximal, et sur lequel on projette le vecteur de mesure, dont on seuille un critère basé sur la puissance.

Au vu de tous les résultats précédents, il apparaît que la méthode de détection basée sur le FASE pour un sous-espace de dimension 2 donne les meilleurs résultats dans le cas monobruit. Nous décidons donc de retenir cette méthode de détection dans toute la suite

de cette thèse, notamment pour résoudre le problème d'estimation des temps d'arrivée du front d'onde sur l'hydrophone, comme nous le verrons dans le chapitre suivant.

Méthode de détection retenue : FASE pour un SEV de dimension 2

Dans un second temps, nous nous sommes intéressés au cas multi-bruits. En effet, plusieurs autres signaux (d'origine biologiques, comme le cri des crevettes, ou d'origine humaine, comme les bruits d'embarcations par exemple) occupent la même bande passante que le signal acoustique que l'on cherche à détecter et peuvent perturber le détecteur. Pour essayer de s'affranchir de ces nouvelles contraintes, une méthode originale de détection multibruits, extension de la méthode du FASC étudiée précédemment, a été étudiée. Cette méthode se base sur la connaissance des différents bruits mis en jeu par le biais de leurs matrices de covariance et consiste à trouver la base d'un sous-espace de dimension p définie *a priori* dans lequel la moyenne géométrique des différents rapports signal à bruit est maximisée.

Les performances de toutes ces méthodes de détection ont ensuite été comparées, et on constate qu'en présence de plusieurs bruits, la méthode du FASCM pour un sous-espace de dimension 1 donne les meilleurs résultats. Néanmoins, on peut noter que la méthode du FASE retenue précédemment dans le cas mono-bruit donne des performances assez proches. Ce résultat ne peux pas être étendu au cas général mais représente le cas Antares étudié dans cette thèse.

Chapitre 4

Structure d'antenne et estimation de la trajectoire

4.1 Introduction

Après avoir étudié les caractéristiques de l'onde de pression dans le chapitre 1, puis abordé les problèmes de détection de celle-ci dans les chapitres 2 et 3, nous allons maintenant étudier les problèmes liés à l'estimation de la trajectoire du neutrino. Les informations de trajectoire sont importantes en astrophysique pour obtenir des indications sur l'origine de ces particules. En effet, depuis quelques années, les astrophysiciens s'accordent à penser que la plus grande partie de la matière de l'univers est constituée de "matière noire" non lumineuse. Ce phénomène devant mettre en oeuvre une masse bien plus importante que celle constituée par les étoiles et gaz de celles-ci, un bon candidat théorique pour cette matière est une particule stable neutre électriquement : le neutrino. L'estimation de la trajectoire d'un neutrino détecté est donc très précieuse.

4.2 Résumé

Nous allons considérer comme hypothèse pour ce chapitre que l'on dispose d'un système capable de détecter la présence du signal acoustique noyé dans du bruit de mer. Un tel système a été décrit dans les chapitres précédents. Dans un premier temps, nous allons étudier un estimateur (sommaire) du temps de retard entre deux capteurs. Nous proposerons ensuite une structure d'antenne simple qui essayera de répondre aux contraintes physiques du projet (disposition des lignes sous-marines existantes) et aux contraintes mathématiques (observabilité du phénomène). Une fois cette structure définie, nous nous intéresserons à l'estimation de la trajectoire du neutrino détecté. Cet estimateur sera testé dans de nombreuses situations et ses performances seront évaluées.

4.3 Estimation des temps de retard

Dans un premier temps, nous allons nous intéresser à l'estimation des dates de détection du front d'onde au niveau de chaque hydrophone. Dans un second temps, nous étudierons l'estimation des différences de dates d'arrivée du front d'onde entre deux hydrophones distincts. Il est important de noter que l'étude de ces estimateurs est volontairement sommaire et n'a pour but que de montrer la faisabilité de l'estimation. Des méthodes plus poussées pourront être étudiées dans les possibles extensions de ces travaux.

Les notations suivantes seront utilisées dans cette section :

- ⋆ $x(t)$: signal disponible en sortie de l'hydrophone.
- ⋆ $\varepsilon(t)$: bruit de mer ambiant de matrice de variance covariance $\mathbf{B}$ connue (ou estimée précédemment).
- ⋆ T_a : durée du signal acoustique, supposée connue.
- ⋆ T_d : date à laquelle le front d'onde atteint le détecteur.

Si l'on suppose qu'un hydrophone détecte le front d'onde à la date T_d, alors pour des dates inférieures à T_d et supérieures à $T_d + T_a$, le signal $x(t)$ n'est constitué que du bruit $\varepsilon(t)$, alors qu'entre les dates T_d et $T_d + T_a$, il est constitué de la somme du bruit de mer et du signal acoustique. La figure 4.1 illustre ce propos.

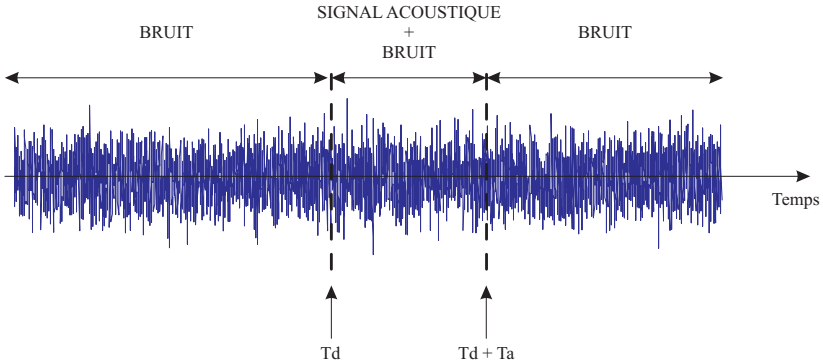

FIG. 4.1 – Exemple de signal sortant d'un hydrophone ayant détecté le front d'onde.

4.3.1 Estimation des dates d'arrivée du front d'onde sur un hydrophone

Hypothèse de base : dans toute la suite de cette étude, nous allons considérer comme négligeable la probabilité que deux cascades d'énergie apparaissent simultanément (ou dans un intervalle de temps très bref, de l'ordre de la micro seconde) dans la zone géographique de détection.

La détection (cf. chapitre précédent) et l'estimation du temps d'arrivée sont basées sur un filtrage de type FASE pour un sous-espace de projection de dimension 2. L'idée est de filtrer le signal reçu sur chaque hydrophone par le filtre issu du FASE $p = 2$ (travaillant sur N points) dans une approche "temps réel" grâce à une fenêtre glissante (figure 4.2). Le filtrage consiste au calcul de la puissance du projeté du signal sur le SEV (ici de dimension 2) généré par les vecteurs du FASE. La fenêtre de temps réel travaille sur les N derniers points du signal, le signal filtré est donc retardé par rapport au temps réel d'une durée égale à la longueur du filtre (64 μs pour nos simulations).

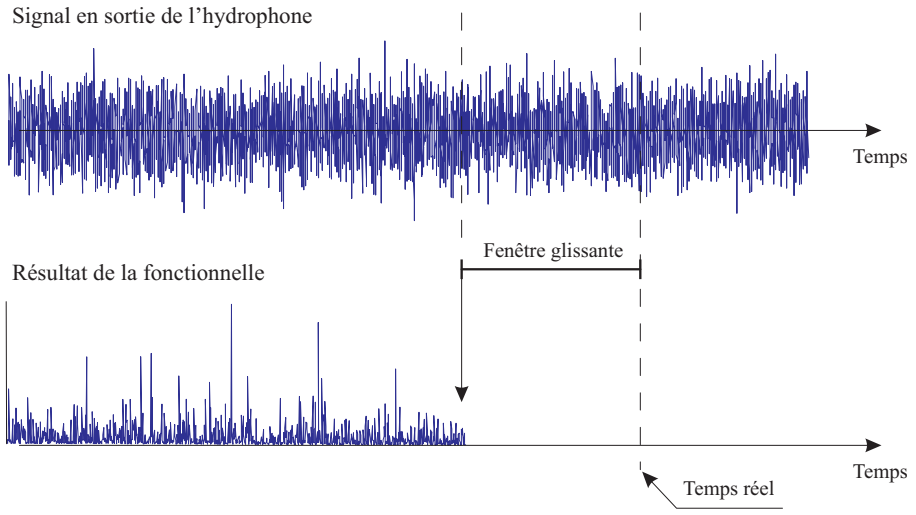

FIG. 4.2 – Illustration de la notion de fenêtre glissante.

Sur la figure 4.3 est représenté un exemple de résultat de la fonctionnelle du FASE appliqué à un signal issu d'un hydrophone. On se fixe ensuite un seuil et on décide que tous les points dépassant celui-ci sont représentatifs d'une détection.

Ces points sont rangés dans deux catégories : les points dit *isolés* (comme les points 1 et 2 de la figure 4.3) et les points *non isolés* (comme le point 3). Un point sera dit isolé si, dans une fenêtre temporelle de taille $2T_A$ centrée autour de celui-ci, moins de deux autres points détectés sont présents. Les points détectés isolés sont systématiquement abandonnés, cette situation correspondant presque sûrement à un artefact de détection.

Les points détectés non isolés permettent d'estimer la date d'arrivée du front d'onde sur l'hydrophone. Cette date est celle du point d'amplitude maximum parmi tous les points détectés non isolés.

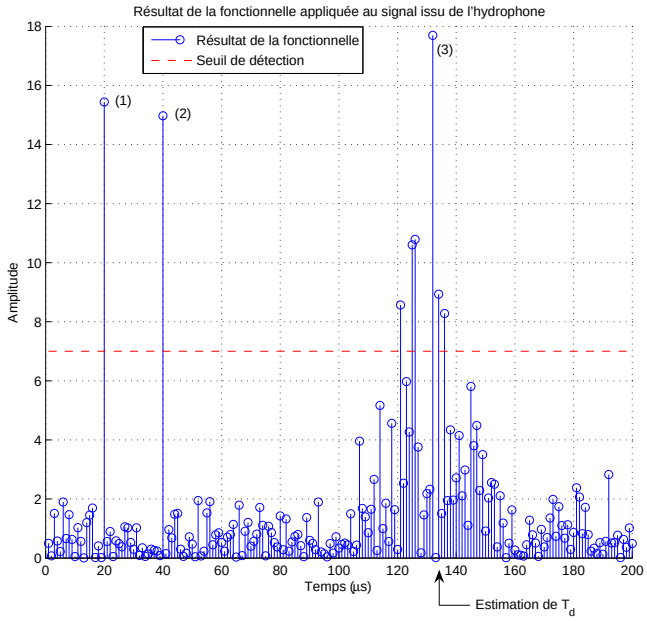

FIG. 4.3 – Exemple du résultat de la fonctionnelle de détection appliquée à un signal issu d'un hydrophone. Les signaux et valeurs numériques utilisées ici ne servent qu'à illustrer le propos et ne reflètent pas un cas réel.

4.3.1.1 Résultats expérimentaux

Pour illustrer les résultats de cet estimateur, on utilise une méthode de Monte-Carlo, dans laquelle les réalisations du bruit de mer changent avec chaque itération. La valeur de $\widehat{T_d}$ est alors estimée par la méthode précédente.

Les paramètres de simulations sont les suivants :

- ⋆ $E = 10^{19}$ eV : énergie de la cascade d'énergie.
- ⋆ $L_{eff} = 10$ m : longueur effective de la cascade.
- ⋆ $D_{max} = 500$ m : distance maximale de détection.
- ⋆ $Nb_{exp} = 1000$: nombre d'expériences de Monte-Carlo.
- ⋆ $T_d = 100$ μs : valeur réelle de T_d.

Pour une valeur de RSB de -65 dB, on représente sur la figure 4.4 l'histogramme des valeurs de $\widehat{T_d}$. L'évolution de l'écart-type σ_{Td} en fonction du RSB est décrite dans le tableau 4.1.

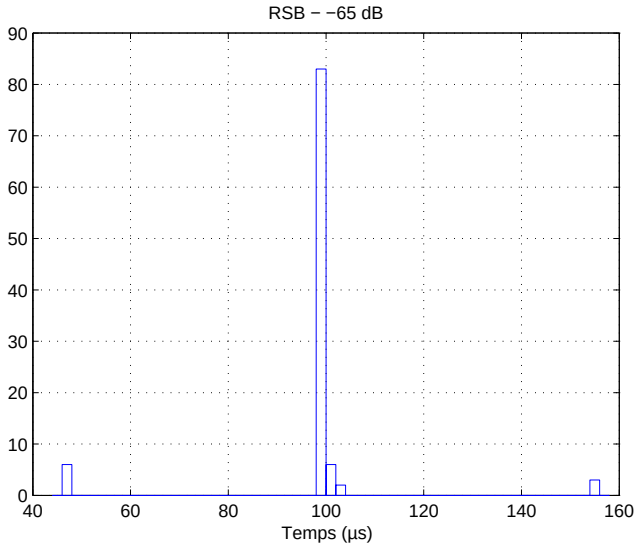

FIG. 4.4 – Histogramme des valeurs de $\widehat{T_d}$ pour un RSB de -65 dB

RSB (dB)	$\widehat{T_d}$ **moyen** (μs)	σ_{Td} (μs)
-54	98.945	1.676
-56	98.938	1.679
-58	98.875	2.375
-60	98.648	4.785
-62	98.575	10.236
-64	98.593	15.948
-66	98.647	20.720
-68	99.937	25.637
-70	101.54	31.239
-72	104.58	43.584

TAB. 4.1 – Evolution de la moyenne et de l'écart type des estimés de T_d pour différents RSB.

Le tracé de l'évolution de cet écart-type en fonction du RSB est mené sur la figure 4.5.

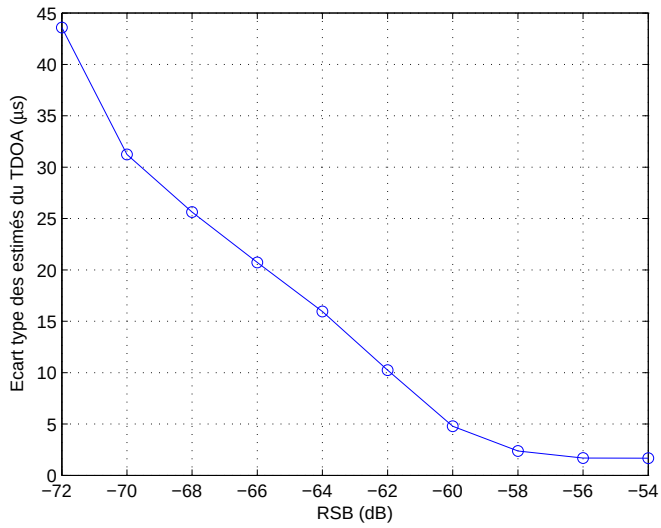

FIG. 4.5 – Evolution de l'écart type de $\widehat{T_d}$ en fonction du RSB

4.3.2 Estimation de la différence du temps d'arrivée sur un couple d'hydrophone

L'estimateur de la trajectoire, étudié plus tard dans ce chapitre, sera basé sur la connaissance des estimés des différences de date d'arrivée du front d'onde entre deux hydrophones. Cette valeur sera noté TDOA (Time Difference Of Arrival). Pratiquement, on se propose d'appairer les hydrophones, formant ainsi des couples de capteurs.

Pour estimer les TDOA, deux méthodes sont proposées. La première est basée sur une simple différence des dates d'arrivée et la seconde sur un calcul d'intercorrélation.

4.3.2.1 Description des deux méthodes d'estimation du TDOA

Méthode 1

L'estimation du TDOA se fait simplement par le calcul des différences des dates d'arrivée du front d'onde sur chacun des hydrophones constituant un couple. Elle est basée sur l'estimateur de la date d'arrivée (d'écart type $\sigma_{\widehat{t}}$) étudié précédemment. Dans la mesure où $\widehat{t_1}$ et $\widehat{t_2}$ sont décorrélés, l'écart type de l'estimateur du TDOA est ainsi donné par :

$$\sigma_{\text{TDOA}}^{(1)} = \sqrt{\sigma_{\widehat{t_1}}^2 + \sigma_{\widehat{t_2}}^2} \tag{4.1}$$

Méthode 2

Cet estimateur est basé sur un calcul classique de corrélation. On dispose ici des signaux issus de chacun des d'hydrophones d'un couple. L'idée est de projeter ces signaux sur le

SEV généré par les vecteurs du FASE $p = 2$. Dans ce SEV, le RSB est maximisé. On calcule ensuite l'intercorrélation (cf. annexe K) entre les deux signaux projetés et L'estimé du TDOA correspond alors la valeur maximale de ce signal de corrélation.

Test de rejet

Si c'est le même front d'onde qui est à l'origine du pic de détection sur les deux hydrophones d'un même couple, alors l'estimé du TDOA doit suivre la relation :

$$\text{TDOA} \leq \frac{d_H}{C_s}$$

où d_H est la distance entre hydrophone dans un couple et C_s est la célérité du son dans l'eau. La quantité $\frac{d_H}{C_s}$ correspond à la durée maximale que peut mettre le front d'onde pour aller d'un hydrophone à l'autre.

Si cette relation n'est pas respecté, l'estimé sera rejeté.

4.3.2.2 Les contraintes liées aux deux méthodes

Nous avons vu que l'estimation du TDOA nécessitait d'appairer les hydrophones en couples. Plusieurs contraintes sont à respecter pour décider d'une distance d_H entre les deux hydrophones d'un couple. Certaines de ces contraintes sont indépendantes de la méthode d'estimation du TDOA considérée :

- ⋆ Plus la distance entre les deux hydrophones d'un couple est grande, plus la durée du trajet du front d'onde augmente. Le nombre de fausse alarme pouvant apparaître sur chacun des capteurs croît avec cette distance. Pour minimiser cette P_{fa}, la distance doit donc être la plus petite possible. Néanmoins, plus on augmente cette distance, plus la dynamique sur les estimés du TDOA augmente. Un compromis devra donc être fait sur cette distance.

D'autres contraintes sont spécifiques aux méthodes étudiées précédemment :

Contraintes liées à la méthode 1

- ⋆ Les bruits de mer présents sur chaque hydrophone doivent être décorellés . En effet, la corrélation du bruit augmente la variance de l'estimateur de la date de détection, et on veut cette variance la plus faible possible. Or, la distance de corrélation du bruit de mer est inversement proportionnelle à sa largeur de bande passante. Nous avions défini dans les chapitres précédents la BP utile à environ 150 kHz, ce qui correspond à une distance de corrélation d'environ 1 cm. Les hydrophones au sein d'un couple

doivent donc être distants d'au moins 1 cm.

Contraintes liées à la méthode 2

- ⋆ Pour que le calcul de corrélation ait un sens, les signaux acoustiques reçus sur chacun des hydrophones doivent avoir une forme similaire. Pour s'assurer de cette hypothèse, on trace sur la figure 4.6 deux exemples d'évolution du coefficient de corrélation du signal acoustique en fonction de la distance. Dans le premier cas, le signal acoustique de référence est pris à une distance de 100m et dans le second cas à une distance de 400m. Dans chacun des cas, le coefficient de corrélation est calculé par rapport à des signaux acoustiques générés à des distances allant de 100 à 500m.

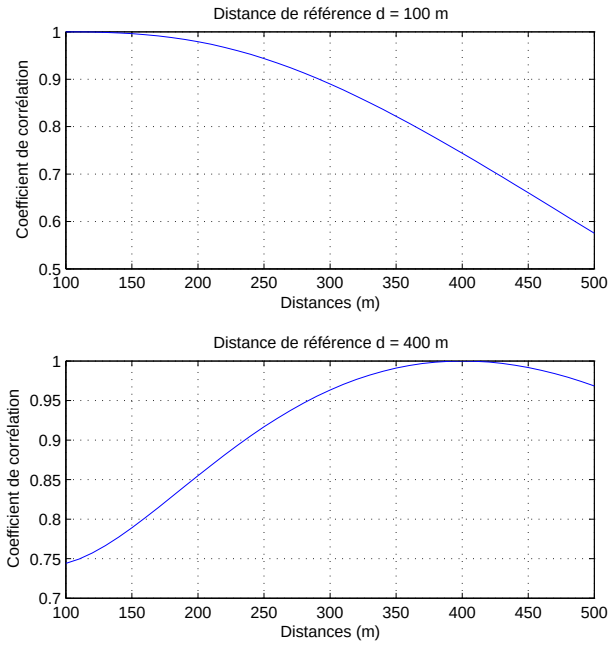

FIG. 4.6 – Coefficient de corrélation en fonction de la distance pour deux exemples signaux acoustiques générés à 100 et 400m.

La lecture de ces courbes nous montre que si l'on éloigne deux hydrophones d'une distance inférieure à 100m, le coefficient de corrélation reste très proche de 1, les signaux correspondant possèdent donc une forme très similaire.

Pour nos simulations, nous décidons d'utiliser la valeur numérique suivante pour la distance inter-hydrophone à l'intérieur du couple :

$$\boxed{d_H = 1.50 \text{ m}}$$

Cette valeur n'est pas optimale au sens d'un critère, mais respecte toutes les contraintes énoncées ci-dessus.

4.3.2.3 Choix de la méthode d'estimation du TDOA

Pour quantifier les performances des deux estimateurs étudiés précédemment, nous avons recours à des expériences de Monte Carlo. Nous sommes ainsi capable d'évaluer les écart-type $\sigma_{\Delta_{\text{TDOA}}}$ sur le TDOA pour différentes valeurs du RSB. Les résultats expérimentaux sont rapportés dans le tableau 4.2. La valeur vraie du TDOA utilisée est de 84.564 μs.

	Méthode 1		Méthode 2	
RSB (dB)	$\widehat{T_d}$ moyen (μs)	σ_{TDOA} (μs)	$\widehat{T_d}$ moyen (μs)	σ_{TDOA} (μs)
- 36	85	0	85.004	0.089443
- 38	85	0	85.004	0.089443
- 40	85	0	85.004	0.089443
- 42	85	0	84.99	0.32575
- 44	84.91	2.0125	84.976	0.45148
- 46	84.91	2.0125	84.934	0.7033
- 48	84.91	2.0125	84.766	1.2738
- 50	84.798	3.3183	84.528	1.7673
- 52	84.842	3.34	84.22	2.213
- 54	84.788	4.2496	83.94	2.5189
- 56	84.214	7.7625	83.692	2.7379
- 58	82.714	12.2245	83.39	2.8953
- 60	80.492	16.4791	83.27	3.8878
- 62	78.214	19.5415	83.054	7.0264
- 64	72.58	24.6726	82.72	10.4545
- 66	66.886	28.8053	82.842	13.1414

TAB. 4.2 – Résumé des résultats

On trace maintenant sur le même graphique l'évolution de l'écart-type en fonction du RSB pour les deux méthodes.

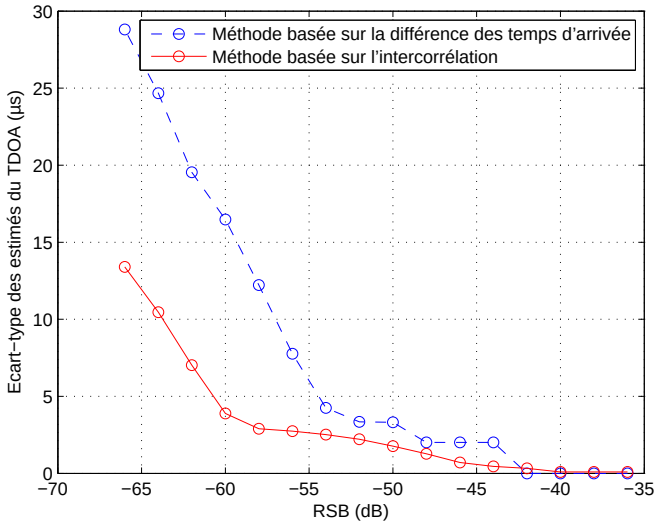

FIG. 4.7 – Evolution de l'écart type des estimés du TDOA en fonction du RSB pour les deux méthodes étudiées.

Pour interpréter la courbe précédente, il est nécessaire de remarquer que pour chacun de ces estimateurs, la précision de l'estimé est liée à la quantification temporelle utilisée, i.e. à la fréquence d'échantillonnage. Dans nos simulations, celle-ci est fixée à $F_e = 1$ MHz, le pas d'échantillonnage est donc de 1 μs. L'estimé du TDOA est donc soumis à un bruit de quantification, que l'on peut identifier à un bruit uniforme de largeur $\Delta t = 1$ μs. L'écart type d'un tel bruit est de $\sigma_u = \frac{\Delta t}{2\sqrt{3}} = 0.2886$ μs.

Ainsi, lorsque la courbe d'évolution de l'écart-type en fonction du RSB passe en dessous de ce seuil (figure 4.8), l'influence de la quantification devient prépondérante sur celle du bruit et l'estimé devient alors "déterministe". Ceci explique pourquoi en dessous de ce seuil, l'estimateur donne toujours la même valeur.

Lorsque la courbe passe au dessus du seuil, le bruit devient alors prédominant par rapport à la quantification et les différents estimés sont donc distribués suivant l'écart type donné par la courbe.

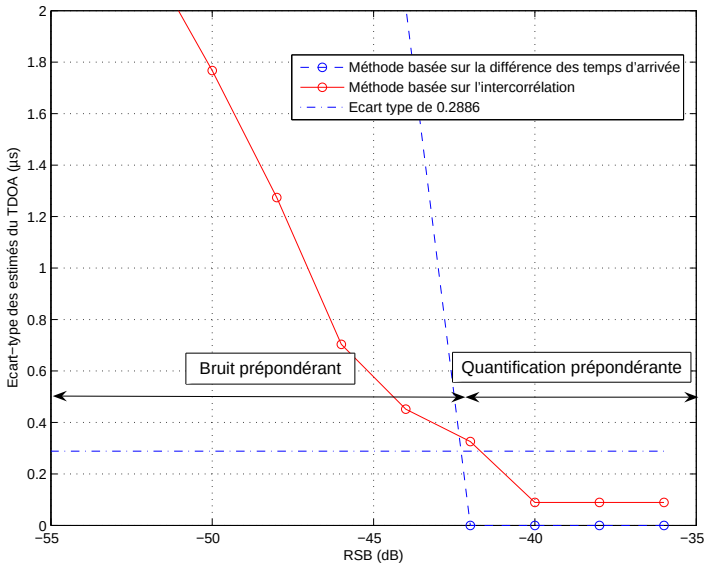

FIG. 4.8 – Evolution de l'écart type des estimés du TDOA en fonction du RSB pour les deux méthodes étudiées et écart type d'une variable aléatoire suivant une loi uniforme de largeur 1 μs.

Les performances des deux méthodes sont équivalentes pour des grandes valeurs de RSB (supérieurs à -40 dB), par contre pour des valeurs plus faibles, les courbes se démarquent nettement et la méthode d'intercorrélation offre des écarts-types nettement plus favorable. Nous choisirons donc comme estimateur du TDOA la seconde méthode, basé sur le calcul de l'intercorrélation entre les signaux issus des hydrophones et filtrés par le FASE $p = 2$.

Il ne faut pas oublier que cette étude de l'estimation des temps de retard a volontairement été succinte et sommaire. Les méthodes étudiés sont simples et ne servent qu'à valider la faisabilité de l'estimation.

4.4 Structure d'antenne : proposition

4.4.1 Contraintes physiques

Le réseau de lignes immergées qui constituent actuellement la structure du téléscope optique Antares forment un volume de base de l'ordre de 0.1 km^3. Sur chaque ligne seront disposés à terme des modules optiques permettant de détecter ce flash lumineux (figure 4.9).

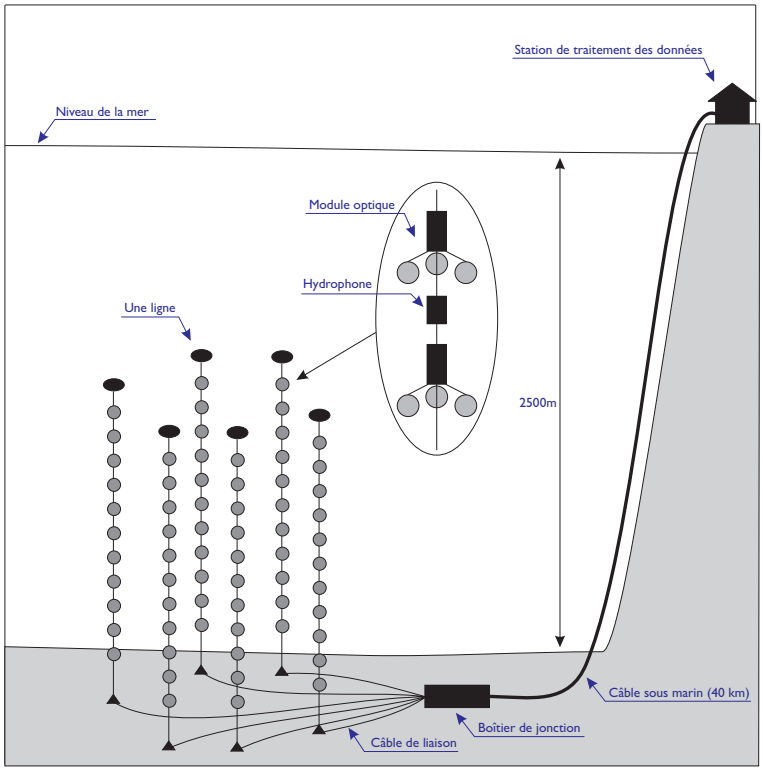

FIG. 4.9 – Structure du téléscope Antares

Nous allons considérer dans ce chapitre que les hydrophones mis en jeu seront disposés sur les mêmes lignes que celles utilisées pour les modules optiques. Cette contrainte "matérielle" limite donc les tailles et formes des réseaux d'hydrophones possibles. La structure sommaire que nous proposons d'étudier est décrite sur la figure 4.10.

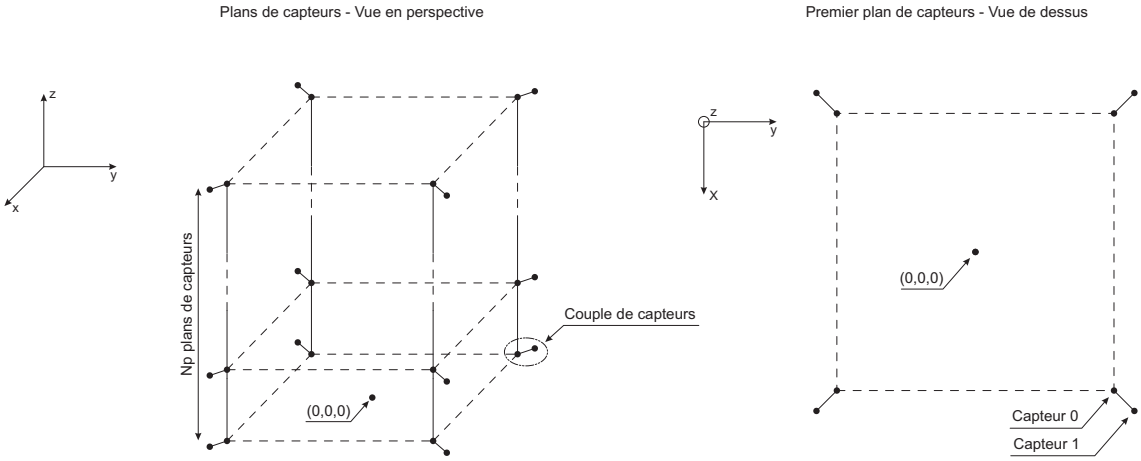

FIG. 4.10 – Structure du réseau d'hydrophones utilisé

Comme convenu plus tôt dans ce chapitre, les hydrophones sont regroupés en couples.

L'origine géographique du repère est choisie arbitrairement au centre de la structure dans le plan d'équation $z = 0$. Ces couples sont disposés en un certain nombre de couches horizontales.

Les ordres de grandeurs de distances suivants seront utilisés pour les simulations : 1.50m pour la distance inter-hydrophone à l'intérieur d'un couple, et 70m pour la distance inter-couple dans un plan horizontal. Les autres valeurs numériques seront détaillées ultérieurement.

Critiques :
Parmi les inconvénients de cette structure, on peut remarquer qu'au vu des grandes distances possibles entre couples de capteurs, les valeurs du RSB entre deux couples peuvent être trés différentes.
La symétrie de la structure peut également poser des problèmes d'observabilité. En effet, on risque d'avoir des problèmes d'estimation pour des positions proches des axes de symétrie de la structure. L'étude de cette observabilité, et par là même, la recherche d'une structure optimale serait donc une très intéressante extension de ces travaux.

4.5 Estimation de la trajectoire du neutrino

4.5.1 Mise en équation du TDOA

La structure d'hydrophones maintenant définie, nous disposons grâce aux sections précédentes d'un ensemble d'estimations de différences de dates d'arrivées. Il y a autant d'estimé du TDOA que de couples de capteurs.

Quelle est l'expression théorique de ce TDOA ?
Nous nous retrouvons dans la situation suivante : soit un repère $\mathbf{R}_1$ définissant l'espace dans lequel se trouve une cascade d'énergie de longueur efficace L_{eff}. Soit A un point quelconque de la cascade. D'aprés la figure 4.11, on peut complètement caractériser la direction et l'orientation de la cascade à partir des informations suivantes :

A	:	un point quelconque de la cascade.
$\alpha \in [0\ 2\pi]$	:	angle d'azimut.
$\varphi \in [-\pi\ \pi]$	:	angle de site.
L_{eff}	:	longueur efficace de la cascade.

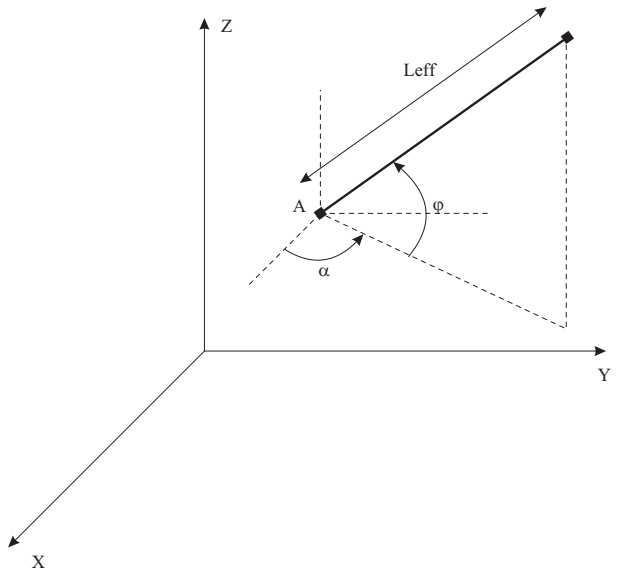

FIG. 4.11 – Description de la cascade d'énergie dans l'espace.

Hypothèse de travail :

Pour la partie théorique suivante, nous allons considérer que la longueur efficace L_{eff} de la cascade est infinie. Cette hypothèse peut se justifier physiquement en considérant des cascade d'énergie issues de neutrinos de très haute énergie (dites cascades électromagnétiques). Dans ce cas, la longueur de la gerbe devient très importante et peut être considérée comme infinie devant les dimensions du télescope [23]. Dans ces conditions, tous les couples de capteurs peuvent donc être atteints par le front d'onde.

Si la cascade n'avait pas une longueur infinie, tous les couples de capteurs ne seraient pas atteints par le front d'onde. Il nous suffira donc dans les simulations de limiter le nombre de couples délivrant une information pour se ramener au cas des cascades hadroniques de longueur efficace finie étudiées dans les chapitres précédents.

Reprenons ici l'hypothèse effectuée dans le chapitre 1, à savoir que la propagation du front d'onde de l'onde acoustique s'effectue de manière orthogonale à la cascade. Ce front d'onde est donc un cylindre dont l'axe est la cascade. Considérons un hydrophone C de coordonnées (x_C, y_C, z_C). Calculons l'expression de la date d'arrivée du front d'onde sur le capteur. Définissons la distance d comme la distance entre l'hydrophone C et la droite (D) générée par la cascade (figure 4.12), c'est donc la distance entre C et son projeté orthogonal C_p sur cette droite. La date d'arrivée est directement proportionnelle à d par la relation $d = c_s.t$, où c_s est la célérité du son dans l'eau (considérée ici comme constante).

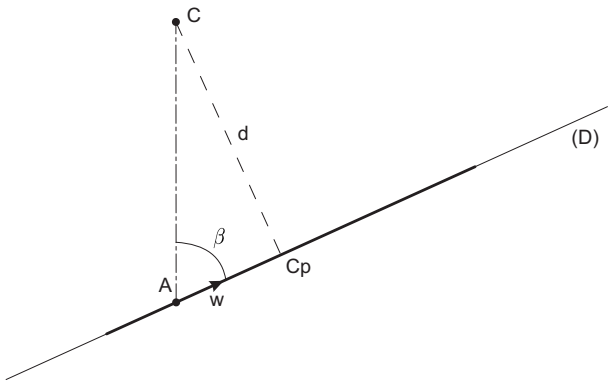

FIG. 4.12 – Distance parcourue par le front d'onde

Les coordonnées du vecteur $\vec{\omega}$ sont :

$$\vec{\omega} = \begin{cases} \cos\alpha\cos\varphi \\ \sin\alpha\cos\varphi \\ \sin\varphi \end{cases} \tag{4.2}$$

Connaissant A, C, α et φ, il est possible de calculer la distance parcourue d. On sait que $\vec{AC} \wedge \vec{\omega} = \vec{C_pC} \wedge \vec{\omega}$ car :

- $\vec{AC}$ et $\vec{C_pC}$ sont dans le même plan.
- $||\vec{AC} \wedge \vec{\omega}|| = ||\vec{AC}||.||\vec{\omega}||.\sin\beta = ||\vec{C_pC}||.||\vec{\omega}||$

où $\beta = (\widehat{\vec{AC}, \vec{\omega}})$ et $||\vec{C_pC}|| = ||\vec{AC}||\sin\beta$.

or $\vec{C_pC} \wedge \vec{\omega} = ||\vec{C_pC}||.||\vec{\omega}||.\sin\left(\frac{\pi}{2}\right) = d||\vec{\omega}||$. On remarque que $||\vec{\omega}|| = 1$ et on peut alors écrire :

$$d = ||\vec{AC} \wedge \vec{\omega}|| \tag{4.3}$$

On remarque également que :

$$||\vec{AC} \wedge \vec{\omega}||^2 + \left(\vec{AC}.\vec{\omega}\right)^2 = ||\vec{AC}||^2.||\vec{\omega}||^2 = ||\vec{AC}||^2 \tag{4.4}$$

On conclut avec le résultat suivant :

$$\boxed{d^2 = ||\vec{AC}||^2 - \left(\vec{AC}.\vec{\omega}\right)^2} \tag{4.5}$$

Notant C_i^0 et C_i^1 les deux capteurs formant le couple i, on peut maintenant donner l'expression de la différence des dates d'arrivées associée. Pour plus de lisibilité, cette valeur de TDOA sera notée τ_i.

$$\tau_i = \frac{d_i^1 - d_i^0}{c_s} \tag{4.6}$$

d_i^0 et d_i^1 sont les distances parcourues par le front d'onde pour atteindre respectivement les capteurs C_i^0 et C_i^1. Utilisant l'équation 4.5, il vient :

$$\boxed{\tau_i = \frac{1}{c_s}\sqrt{||\vec{AC_i^1}||^2 - \left(\vec{AC_i^1}.\vec{\omega}\right)^2} - \frac{1}{c_s}\sqrt{||\vec{AC_i^0}||^2 - \left(\vec{AC_i^0}.\vec{\omega}\right)^2}} \tag{4.7}$$

4.5.2 Mise en équation de l'estimateur

Ayant à disposition un ensemble de N mesures issues de N couple d'hydrophones de position connue, nous allons nous intéresser à l'estimation de la trajectoire du neutrino. Celle-ci est complètement définie par position de la cascade d'énergie correspondante. Les cinq inconnus de ce problème sont les coordonnées du point $A(x_A, y_A, z_A)$ et les deux angles définissant l'orientation de la cascade (α, φ).

Nous ne chercherons pas à estimer la longueur efficace de la cascade mais seulement sa position et son orientation. Celles-ci sont définie par la donnée du point A et des deux angles α et φ. Le point A n'a donc pas a être positionné **sur** la cascade, mais simplement quelque part sur la droite définissant l'orientation de celle-ci.
Nous allons ainsi imposer au point A d'être dans le plan d'équation $z = 0$. Le nombre d'inconnus passe donc de 5 à 4.

Nous allons donc chercher à estimer les quantités suivantes :

$$A = (x_A, y_A)^T \text{ et } \Psi = (\alpha, \varphi)^T$$

que nous regrouperons dans la variable θ :

$$\theta = (A, \Psi)^T = (x_A, y_A, \alpha, \varphi)^T \tag{4.8}$$

A chaque couple i d'hydrophones est associé une mesure τ_i. Il est donc possible de modéliser ces mesures par une fonction dépendant de θ :

$$s_i(\theta) = \frac{1}{c_s}\sqrt{||\vec{AC_i^1}||^2 - \left(\vec{AC_i^1}.\vec{\omega}\right)^2} - \frac{1}{c_s}\sqrt{||\vec{AC_i^0}||^2 - \left(\vec{AC_i^0}.\vec{\omega}\right)^2} \tag{4.9}$$

Définissons également les vecteurs $S(\theta) = \begin{pmatrix} s_1(\theta) \\ s_2(\theta) \\ \vdots \\ s_N(\theta) \end{pmatrix}$ et $\tau = \begin{pmatrix} \tau_1 \\ \tau_2 \\ \vdots \\ \tau_N \end{pmatrix}$

Nous allons mettre en oeuvre une méthode d'estimation basée sur la méthode des moindres carrés [13]. Celle-ci consiste à minimiser le critère :

$$\boxed{M(\theta) = \left\| \frac{\tau - S(\theta)}{\sigma} \right\|^2} \tag{4.10}$$

4.6 algorithme

L'ensemble des calculs menant à l'algorithme d'estimation de la trajection est mené dans l'annexe I.

4.6.1 Structure

La structure de l'algorithme d'estimation des moindres carrés est la suivante :

- ⋆ Initialisation :
  - ∘ choix d'un θ_0
  - ∘ choix d'un ϵ
  - ∘ calcul de $M(\theta_0)$

- ⋆ Itération :
  - ∘ Calcul de $\nabla_\theta M(\theta_n)$.
  - ∘ Calcul de $Hess_M(\theta_n)$.
  - ∘ Calcul du pas de descente optimal p_n.
  - ∘ Calcul de $\theta_{n+1} = \theta_n - p_n \; Hess_M(\theta_n)^{-1} \; \nabla_\theta M(\theta_n)$.
  - ∘ Test d'arrêt : Comparaison de $\frac{\|\theta_{n+1} - \theta_n\|}{\|\theta_{n+1}\|}$ à ϵ.

- ⋆ Si le test est positif, fin de l'algorithme.

- ⋆ Traitement des angles α et φ.

4.6.2 Calcul du pas de descente

Pour que l'algorithme converge vers un des minimums de la fonction, il est nécessaire de bien choisir le pas de descente. Dans l'équation (I.6), le pas de descente est constant et égal à 1 (c'est en fait le coefficient devant le terme $HessM(\theta_n)^{-1}.\nabla_\theta M(\theta_n)$), et peut

entrainer des problèmes numériques dans l'approche du minimum. Pour s'affranchir de ce problème, on remplace cette équation par celle-ci :

$$\theta_{n+1} = \theta_n - p_n \; Hess_M(\theta_n)^{-1} \; \nabla_\theta M(\theta_n) \tag{4.11}$$

où p_n est un nombre compris dans l'intervalle $]0, 1]$.
Le pas optimal p_n^{opt} minimise la quantité :

$$M\left(\theta_n - p_n \; Hess_M(\theta_n)^{-1} \; \nabla_\theta M(\theta_n)\right) \tag{4.12}$$

et est évalué numériquement.

4.6.3 Singularités éventuelles de $Hess_M(\theta)$

Des problémes numériques peuvent également apparaître lors de l'inversion de la matrice du Hessien. En effet, si l'une des valeurs propres de la matrice est trop proche de 0, l'inversion devient numériquement critique. Pour s'en affranchir, on a recours à la technique de Levenberg-Marquardt [15]. Cette méthode consiste à très légèrement déformer la matrice pour la rendre inversible : $Hess_M(\theta)$ est alors remplacée par $Hess_M(\theta) + \gamma \; Id$ où γ est un scalaire pris le plus petit possible.
Dans la pratique, la plus petite des valeurs propres λ_{min} est comparée à une valeur ε fixée et trés petite. La valeur de γ est alors :

$$\gamma = |\lambda_{min} + \epsilon| \tag{4.13}$$

4.6.4 Traitement des angles

Chaque valeur estimée des angles α et φ est ramenée dans l'intervalle $[-\pi, \pi]$. On utilise pour cela les valeurs des composantes en x, y et z du vecteur $\vec{\omega}$ et la fonction de tangente inverse.

4.6.5 Test d'acceptation

Il est nécessaire de construire un test d'acceptation pour les estimés, lorsque les mesures sont perturbées par un bruit. Si le bruit est gaussien (et centré), ce test est défini par la règle de décision suivante :

$$\boxed{M(\theta) = \left\| \frac{\tau - S(\theta)}{\sigma} \right\|^2 = \sum_{i=1}^{N_c} \left(\frac{\tau_i - s_i(\theta)}{\sigma} \right)^2 \gtrless \text{seuil}} \tag{4.14}$$

où σ est l'écart type du bruit, $\dim(\theta) = n_i$ où le nombre d'inconnus (ici $n_i = 4$) et N_c le nombre de couple de capteurs considérés.

Afin de trouver une valeur pertinente du seuil, on s'intéresse à la loi de probabilité de $M(\theta)$ et on montre que c'est une loi du Khi-deux à N_c degrés de liberté, notée $\chi^2_{N_c}$. On montre également que $M(\widehat{\theta})$ suit approximativement une loi du Khi-deux à $N_c - n_i$ degrés de liberté (plus la valeur de N_c sera grande, meilleure sera cette approximation).

Pour une variable aléatoire suivant une loi du χ^2_ν, il est possible de calculer la probabilité que cette variable soit supérieure à une certaine valeur. Cette probabilité sera notée $P = prob(\chi^2_\nu > x)$ et on a :

ν \ P	0.01
1	6.635
2	9.210
3	11.345
4	13.277
5	15.086
6	16.812
7	18.475
8	20.090

TAB. 4.3 – Valeurs de x pour la loi du χ^2_ν

La valeur ν du tableau précédent est reliée au paramètres N_c et n_i par la relation :

$$\nu = N_c - n_i$$

La connaissance du nombre de couples de capteurs nous renseigne donc sur la valeur du seuil à choisir. Par exemple, pour 8 couples d'hydrophones et 4 inconnues, nous avons $\nu = 4$. Le seuil sera choisi à 13.277 et pour un $\widehat{\theta}$ renvoyé par l'algorithme, la probabilité que $T(\widehat{\theta})$ soit supérieur à ce seuil est de 1%.

Le choix du seuil est donc à adapter au nombre de couples de capteurs utilisés.

4.7 Simulations et résultats

L'estimateur précédemment étudié peut maintenant être testé dans plusieurs configurations. Pour chaque hypothèse, une méthode de Monte-Carlo est utilisée pour valider la méthode d'estimation. Suivant le nombre de couples de capteurs atteints, la cascade pourra

être soit électromagnétique, soit hadronique. Si tous les couples délivrent une mesure, la cascade sera forcément électromagnétique, mais si seulement par exemple seuls 60% des couples répondent, la cascade pourra être soit électromagnétique (le RSB était trop défavorable au niveau des couples qui n'ont pas répondus), soit hadronique (les couples qui n'ont pas répondu n'ont pas été atteints physiquement par le front d'onde).

De part la géométrie du réseau considéré (figure 4.11), plusieurs positions distinctes dans l'espace peuvent être définies pour la position de la cascade. Chaque position est caractérisée par l'emplacement du point $A^{(0)}_{vrai}$, intersection de la droite issue de la cascade et du plan de coordonnée $z = 0$. Pour se représenter facilement ces différents cas, on visualisera la structure de capteurs vue de dessus, au niveau du plan d'équation $z = 0$. Plusieurs zones sont ainsi définies dans ce plan sur la figure 4.13

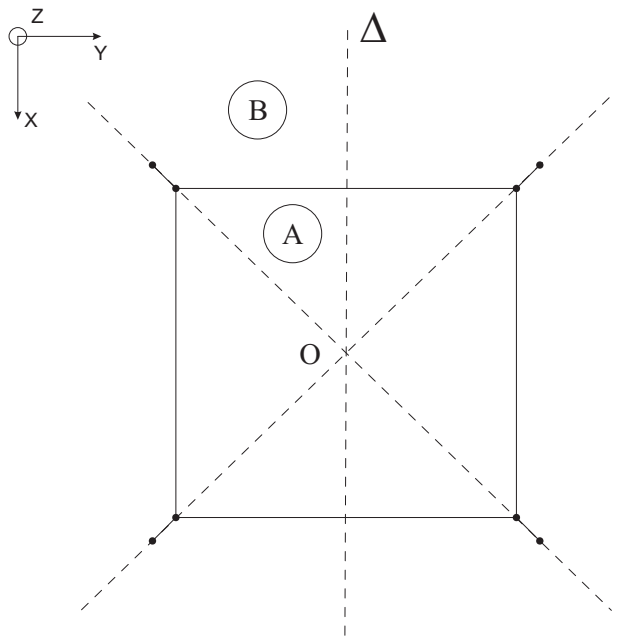

FIG. 4.13 – Découpage du plan d'équation $z = 0$ en différentes zones pour localiser le point $A^{(0)}_{vrai}$ par rapport à la structure de capteurs.

Le point $A^{(0)}_{vrai}$ peut apparaître n'importe où dans le plan d'équation $z = 0$, mais l'étude de seulement 2 zones (représentées par les lettre A et B sur la figure précédente) permet d'obtenir les résultats pour l'ensemble du plan. En effet, par un jeu de rotations et de symétries, il est possible de couvrir toutes les zones du plan.

Remarques : Si l'on se place dans les hypothèses définies dans les chapitres sur la détection, à savoir que la distance minimale de détection est de 100m, le réseau d'hydrophone ne sera pas adapté à une cascade apparaissant dans la zone A. Néanmoins, si l'on extrapole le

modèle de Niess pour des distances inférieures à 100m, on s'attend à avoir des signaux beaucoup plus énergétiques qui serait d'autant plus facilement détectés par les hydrophones. Ce cas d'étude sera donc quand même traité pour valider la méthode d'estimation dans le cas le plus général possible.

4.7.1 Initialisation et valeurs numériques

Le point d'initialisation A_0 est choisi par défaut au barycentre des positions des capteurs dans le plan $z = 0$. Pour les angles α_0 et φ_0, les valeurs d'initialisations seront estimées expérimentalement suivant la région A ou B considérée. On choisira autant que possible des valeurs correspondant à une position initiale de la cascade proche de la verticale.

Considérant que les mesures sont perturbées par un bruit additif, on définit la probabilité de détection P_d comme suit :

$$P_d = \frac{\text{Nombre de couples d'hydrophones ayant détecté la cascade}}{\text{Nombre de couples d'hydrophones total}}$$

En effet, dans l'hypothèse des cascades électromagnétiques, tous les couples d'hydrophones sont atteints par le front d'onde, néanmoins certains peuvent ne pas réagir à son passage (amplitude du signal trop faible, erreur de l'algorithme de détection, etc.).
Dans l'hypothèse des cascades hadroniques, le front d'onde peut ne pas atteindre tous les couples de capteurs, faisant ainsi diminuer la probabilité de détection.
Par exemple, pour une valeur de P_d de 0.1, et si l'on considère une structure à 100 couples d'hydrophones, seul 10 donneront une mesure de temps de retard.

Il faut un minimum de 4 couples de capteurs pour estimer les 4 paramètres qui nous intéressent. La structure utilisée dans ces simulations est constituée de 12 couples (donc 24 hydrophones).

Pour chaque scénario, trois cas différents sont étudiés :

- ⋆ Tous les capteurs détectent ($P_d = 1$).
- ⋆ 9 couples de capteurs sur 12 détectent ($P_d = 0.75$).
- ⋆ 5 couples de capteurs sur 12 détectent ($P_d = 0.416$).

Dans le chapitre précédent consacré à la détection, nous avions choisi d'utiliser un détecteur basé sur le FASE pour un sous-espace de dimension 2. On rappelle que l'évolution du RSB en fonction de la distance de propagation est décrite sur la figure 3.11. Pour l'estimation de l'instant de détection, si nous considérons un RSB de -60 dB, d'après la courbe précédente, la distance maximale de détection est d'environ 500m. On en conclut que pour

cette valeur du RSB, la distance séparant la cascade et l'hydrophone le plus éloigné de la structure ne devra pas dépasser 400m.

Les capteurs sont disposés comme l'indique la figure 4.10 avec les valeurs numériques suivantes :

- ⋆ $N_p = 3$ plans horizontaux contenant 4 couples de capteurs, séparés de $d_z = 10$ m.
- ⋆ Les hydrophones à l'intérieur d'un couple sont distant de 1.50 m.
- ⋆ La distance entre chaque couple d'hydrophones au sein d'un même plan horizontal est d'environ 70m.

Concernant l'écart type du "bruit" affectant les mesures de temps de détection, nous choisissons une valeur compatible avec celle trouvée au début de ce chapitre pour un RSB de -58 dB : $\sigma \approx 2$ μs.

Rappel : la durée du signal Antares est d'environ 60 μs.

Résumé
1000 expériences 24 hydrophones / 12 couples Bruit gaussien d'écart-type = 2 μs
Valeurs Initiales
$A_0 = \begin{pmatrix} 0 \text{ m} \\ 0 \text{ m} \\ 0 \text{ m} \end{pmatrix}$ et $\Psi_0 = \begin{pmatrix} 80° \\ 80° \end{pmatrix}$

Dans chaque cas de figure, on commence par rappeler les valeurs de θ_{vrai} qui sont comparées aux $\widehat{\theta}_{moyen}$, moyennes de tous les résultats obtenus pour les 1000 expériences. Les écarts types et bornes correspondances pour sont ensuite évalués chacune des variables. Le calcul de ces bornes se fait grâce à celui des matrices de Cramer-Rao. Les éléments diagonaux de ces matrices représentent les variances de chacune des composantes du vecteur θ. Les bornes de Cramer-Rao correspondantes aux écarts types des paramètres considérés correspondent donc aux racines de ces variances. Le pourcentage de rejet des estimés et le nombre moyen d'itérations de l'algorithme sont ensuite indiqués. On affichera ensuite les histogrammes du nombre d'itérations par expérience et des valeurs du critère $T(\theta)$.

Pour chaque couple d'estimés (x_A, y_A) et (θ, φ), on trace sur la même figure le point

de départ de l'algorithme, tous les estimés (non rejetés) ainsi que les ellipses de confiance correspondantes. Sur le premier graphique (x_A, y_A), sont également indiqués les positions des couples de capteurs du le plan $z = 0$ pour mieux se représenter physiquement le phénomène. Nous tracerons enfin les histogrammes correspondant à chacun des estimés.

4.7.2 $A_{vrai}^{(0)}$ à l'intérieur de la structure de capteurs Zone A

4.7.2.1 Tous les hydrophones détectent - $P_d = 1$

Tous les couples d'hydrophones détectent le front d'onde et délivrent une mesure.
Les résultats obtenus sont :

$P_d = 1$	θ_{vrai}	$\widehat{\theta}_{moyen}$	Ecart type	Borne
x_A (m)	-10	-9.996	0.13796	0.13297
y_A (m)	20	19.9966	0.10891	0.10656
α (°)	130	129.9397	0.75991	0.75483
φ (°)	60	60.0015	0.097323	0.098852

TAB. 4.4 – Caractéristiques des estimés.

Note : En théorie, la borne calculée théoriquement indique la valeur minimale que peut prendre l'écart-type. Dans les tableaux regroupant les résultats, on se retrouve parfois avec un écart-type inférieur à la borne. Ceci est dû au fait que cet écart-type est lui même estimé et donc soumis à des fluctuations. Cette incertitude numérique fait parfois passer la valeur en dessous de la borne.

Nombre moyen d'itérations :	8
Estimés rejetés :	0.9 %

TAB. 4.5 – Résumé des résultats.

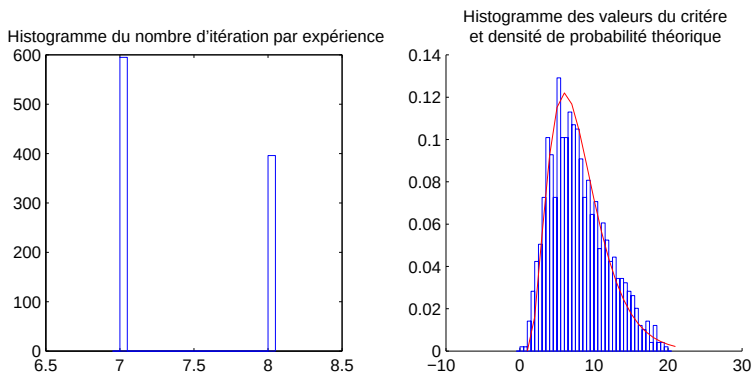

FIG. 4.14 – Informations sur l'algorithme.

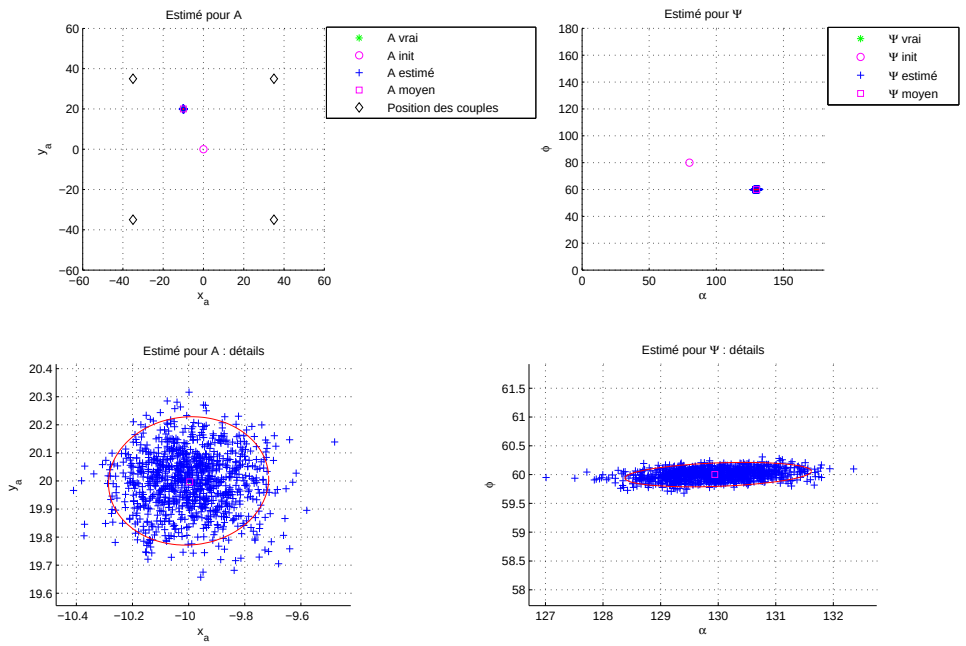

FIG. 4.15 – Représentation des estimés.

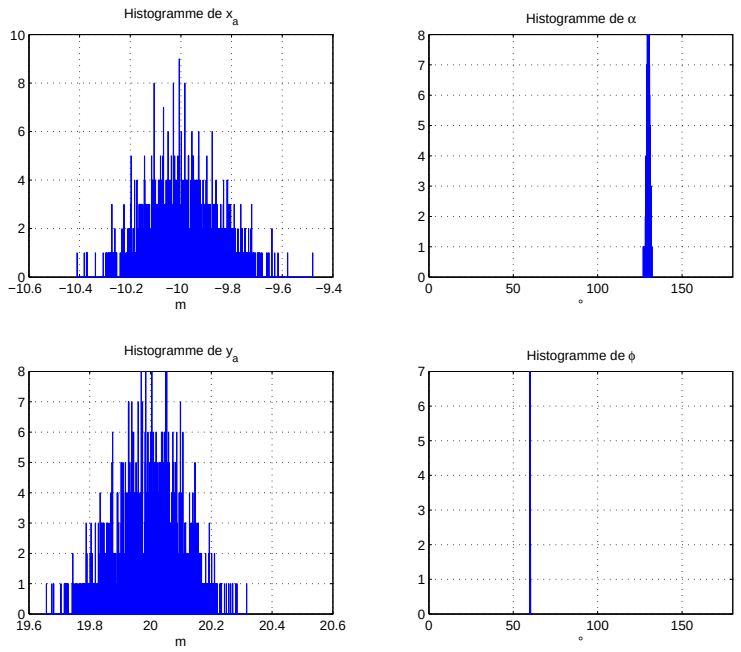

FIG. 4.16 – Histogramme des estimés.

4.7.2.2 9 couples détectent - $P_d = 0.75$

Seuls 9 couples d'hydrophones sur 16 réagissent au passage du front d'onde.
Les résultats obtenus sont les suivants :

$P_d = 0.75$	θ_{vrai}	$\widehat{\theta}_{moyen}$	Ecart type	Borne
x_A (m)	-10	-9.9976	0.1824	0.18997
y_A (m)	20	19.9983	0.11592	0.11482
α (°)	130	129.9973	0.75044	0.75745
φ (°)	60	59.9948	0.15952	0.15656

TAB. 4.6 – Caractéristiques des estimés.

Nombre moyen d'itérations :	7
Estimés rejetés :	1 %

TAB. 4.7 – Résumé des résultats.

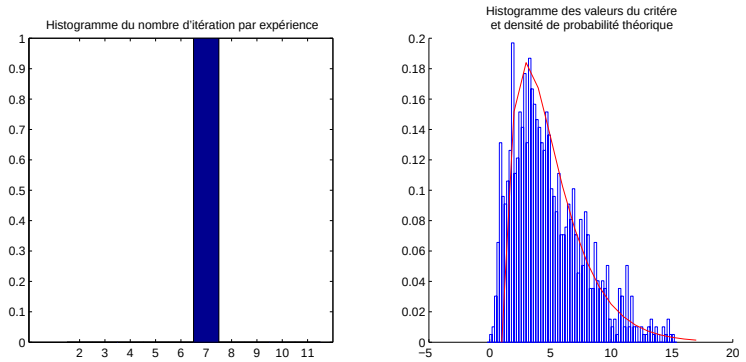

FIG. 4.17 – Informations sur l'algorithme.

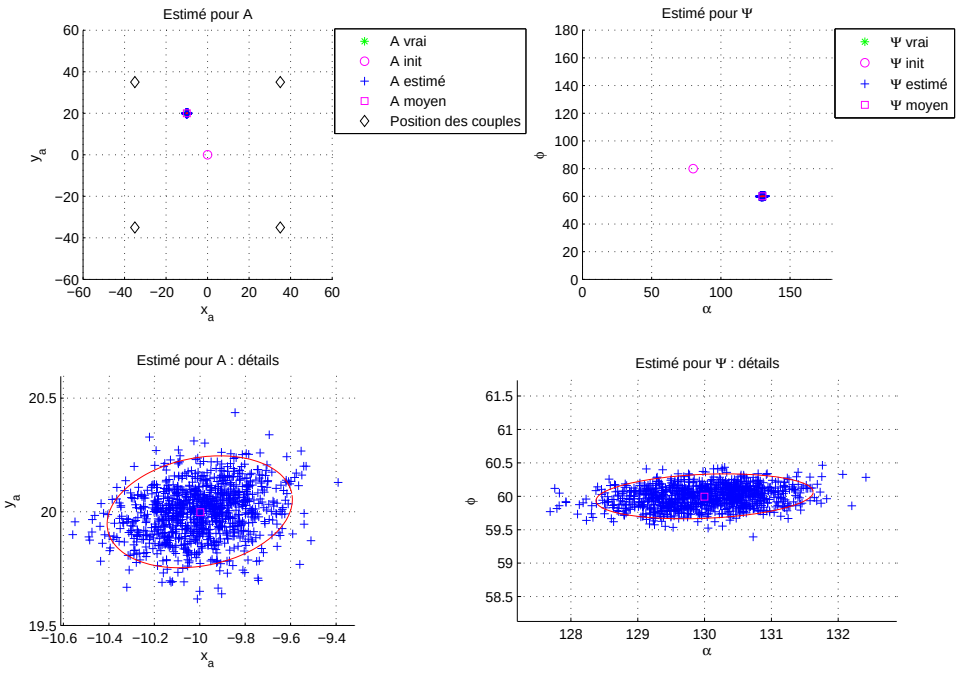

FIG. 4.18 – Représentation des estimés.

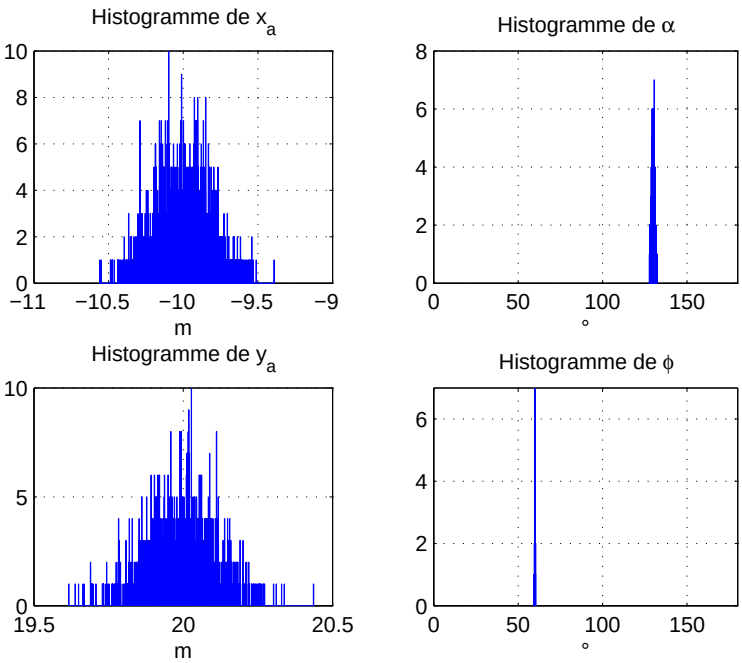

FIG. 4.19 – Histogramme des estimés.

4.7.2.3 5 couples détectent - $P_d = 0.416$

Seuls 5 couples de capteurs sur 12 sont ici atteints par le front d'onde. Nous obtenons les résultats suivants :

$P_d = 0.416$	θ_{vrai}	$\widehat{\theta}_{moyen}$	Ecart type	Borne
x_A (m)	-10	-9.787	2.0372	1.9183
y_A (m)	20	20.0455	0.46888	0.42414
α (°)	130	129.6829	1.6366	1.2992
φ (°)	60	60.1897	1.5525	1.4465

TAB. 4.8 – Caractéristiques des estimés.

Nombre moyen d'itérations :	9
Estimés rejetés :	0.9 %

TAB. 4.9 – Résumé des résultats.

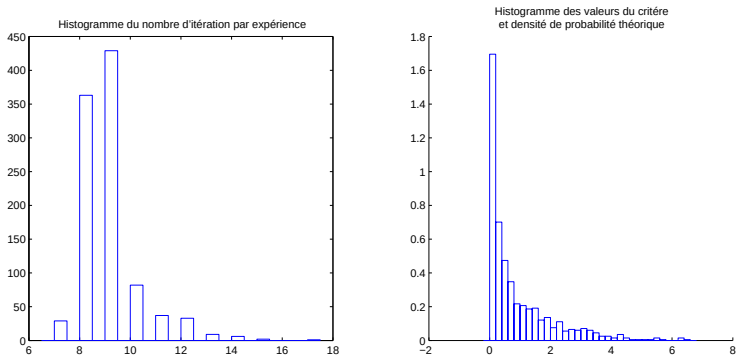

FIG. 4.20 – Informations sur l'algorithme.

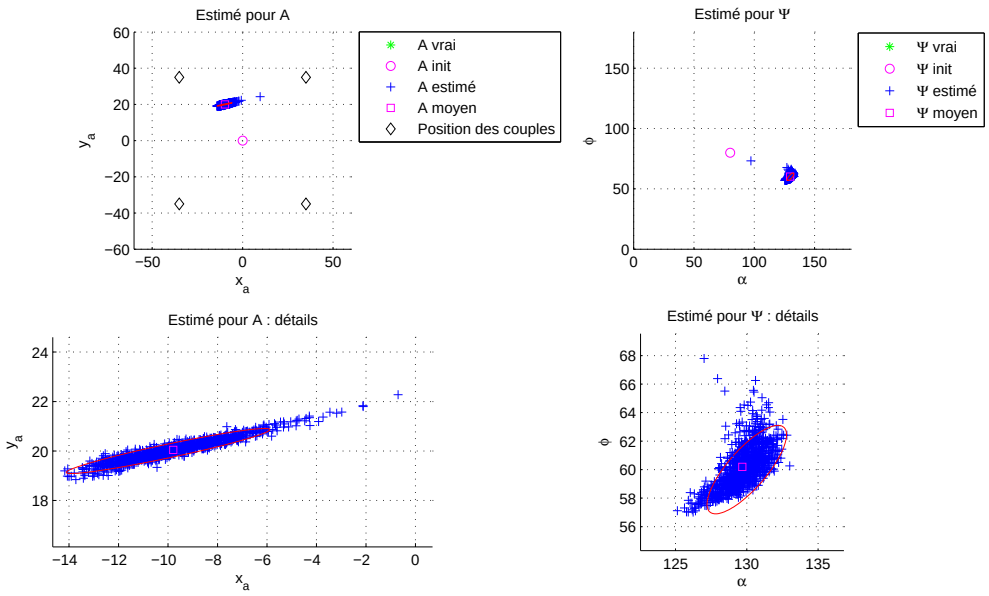

FIG. 4.21 – Représentation des estimés.

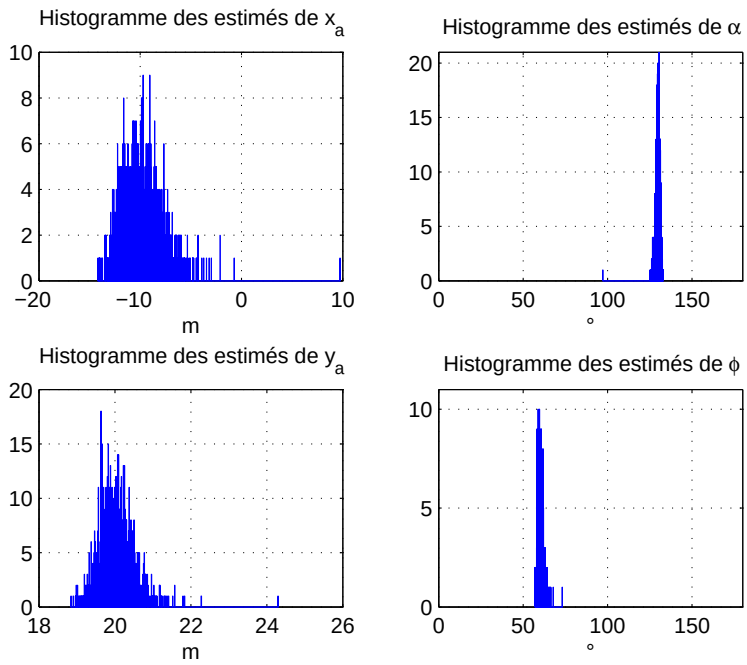

FIG. 4.22 – Histogramme des estimés.

L'algorithme commence ici à montrer ses faiblesses, pour l'estimation du couple (x_A, y_A), l'allure du nuage de points reflète un début de problème d'observabilité. Néanmoins, il ne faut pas perdre de vue que ce cas est critique.

4.7.3 $A_{vrai}^{(0)}$ à l'extérieur de la structure de capteurs Zone B

4.7.3.1 Tous les hydrophones détectent - $P_d = 1$

Tout les couples d'hydrophones délivrent donc une mesure.

Nous obtenons les résultats suivants :

$P_d = 1$	θ_{vrai}	$\widehat{\theta}_{moyen}$	Ecart type	Borne
x_A (m)	-20	-20.001	0.075977	0.077551
y_A (m)	90	89.9996	0.18594	0.18868
α (°)	130	129.9926	0.34512	0.35268
φ (°)	60	60.0015	0.09412	0.095139

TAB. 4.10 – Caractéristiques des estimés

Nombre moyen d'itérations :	8
Estimés rejetés :	0.8 %

TAB. 4.11 – Résumé des résultats.

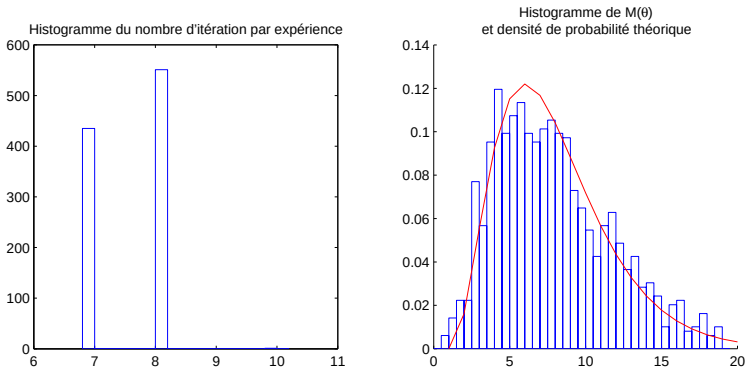

FIG. 4.23 – Informations sur l'algorithme.

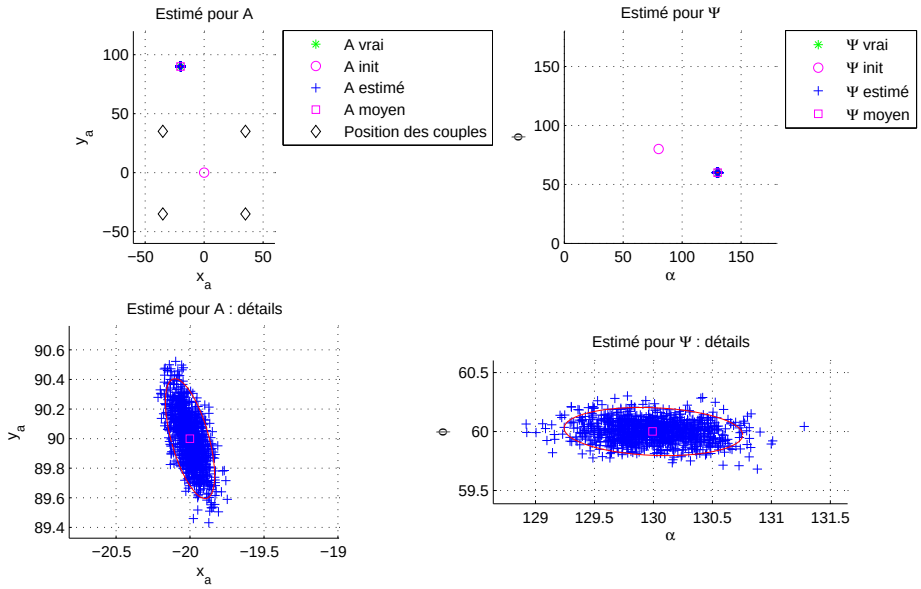

FIG. 4.24 – Représentation des estimés.

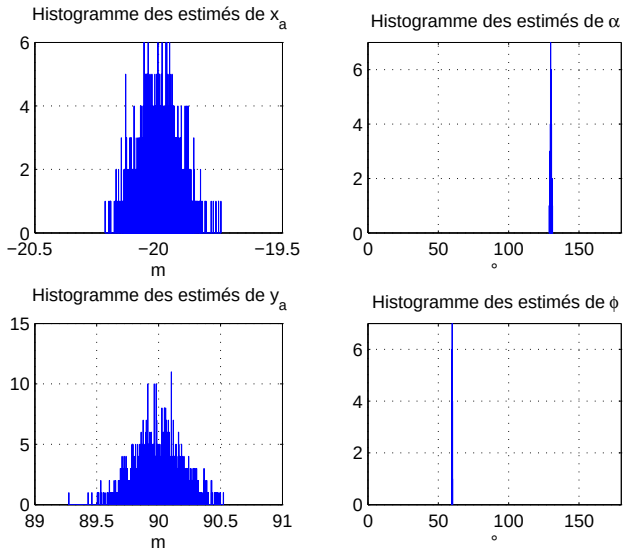

FIG. 4.25 – Histogramme des estimés.

4.7.3.2 9 couples détectent - $P_d = 0.75$

Seuls 9 couples d'hydrophones sur 12 délivrent une mesure.

Nous obtenons les résultats suivants :

$P_d = 0.75$	θ_{vrai}	$\widehat{\theta}_{moyen}$	Ecart type	Borne
x_A (m)	-20	-20.0019	0.10148	0.1034
y_A (m)	90	90.0037	0.19701	0.19914
α (°)	130	130.0025	0.3205	0.32661
φ (°)	60	59.996	0.1607	0.15899

TAB. 4.12 – Caractéristiques des estimés.

Nombre moyen d'itérations :	10
Estimés rejetés :	1.1 %

TAB. 4.13 – Résumé des résultats.

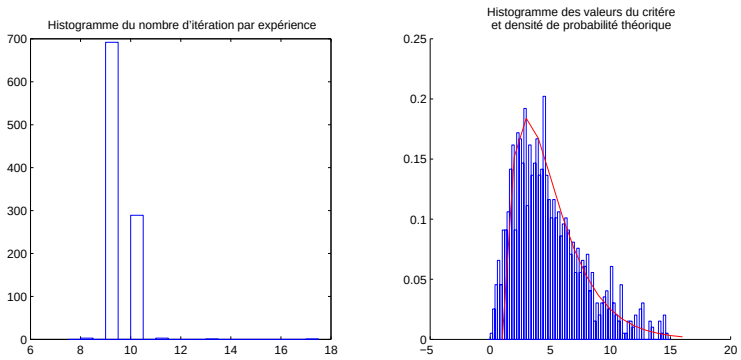

FIG. 4.26 – Informations sur l'algorithme.

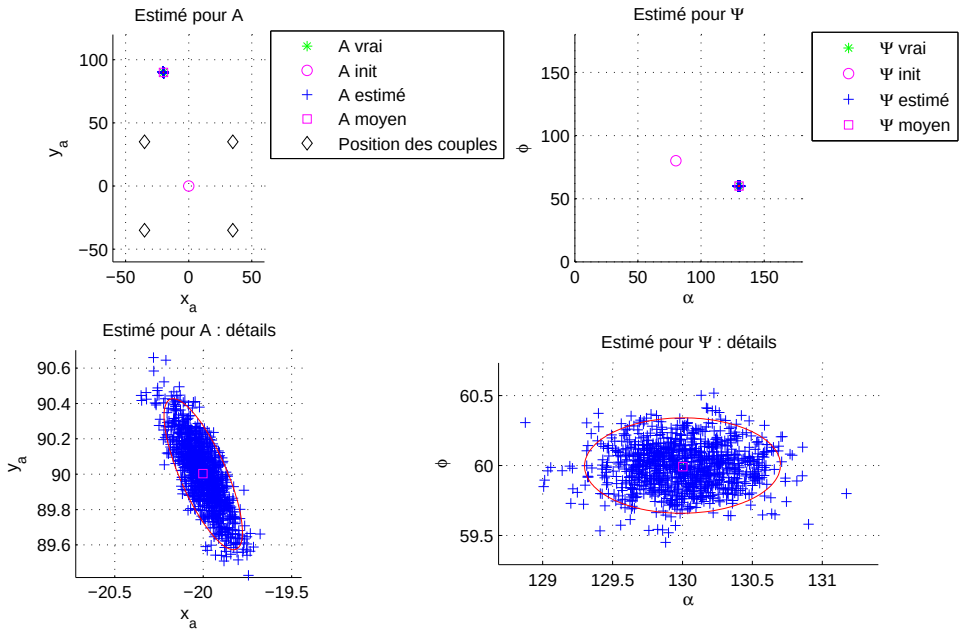

FIG. 4.27 – Représentation des estimés.

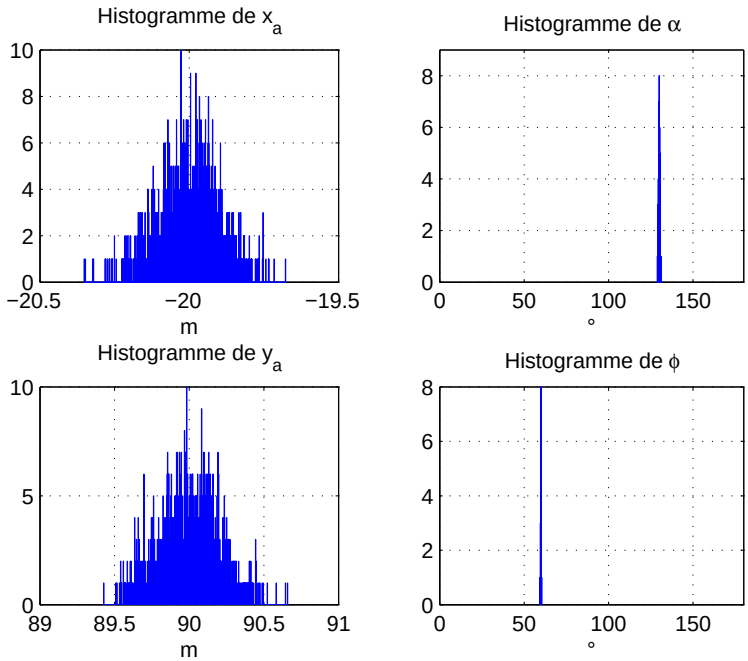

FIG. 4.28 – Histogramme des estimés.

4.7.3.3 5 couples détectent - $P_d = 0.416$

Seuls 5 couples d'hydrophones sur 16 délivrent une mesure.

Avec les conditions initiales précédentes sur les angles, l'algorithme semble ne pas bien converger (95% des estimés sont rejetés). En changeant celles-ci par les valeurs suivantes :

$$\begin{aligned} \alpha_0 &= 0^\circ \\ \varphi_0 &= 0^\circ \end{aligned}$$

le problème de convergence est résolu et on obtient les résultats suivants :

$P_d = 0.416$	θ_{vrai}	$\widehat{\theta}_{moyen}$	Ecart type	Borne
x_A (m)	-20	-20.6599	2.053	1.7218
y_A (m)	90	90.6534	2.7236	2.7819
α (°)	130	130.537	1.2074	1.3014
φ (°)	60	59.049	3.8539	3.7701

TAB. 4.14 – Caractéristiques des estimés.

Nombre moyen d'itérations :	18
Estimés rejetés :	2.2 %

TAB. 4.15 – Résumé des résultats.

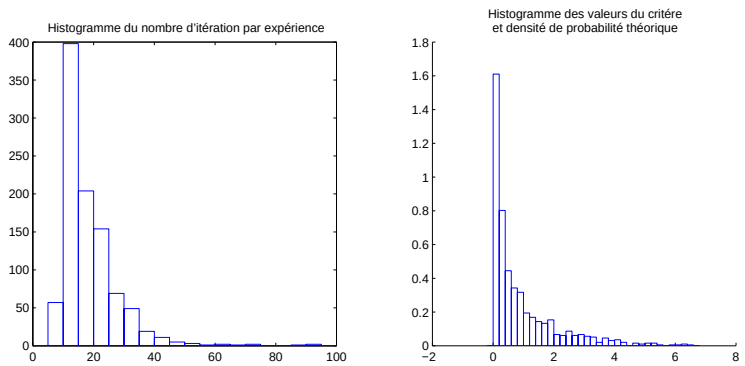

FIG. 4.29 – Informations sur l'algorithme

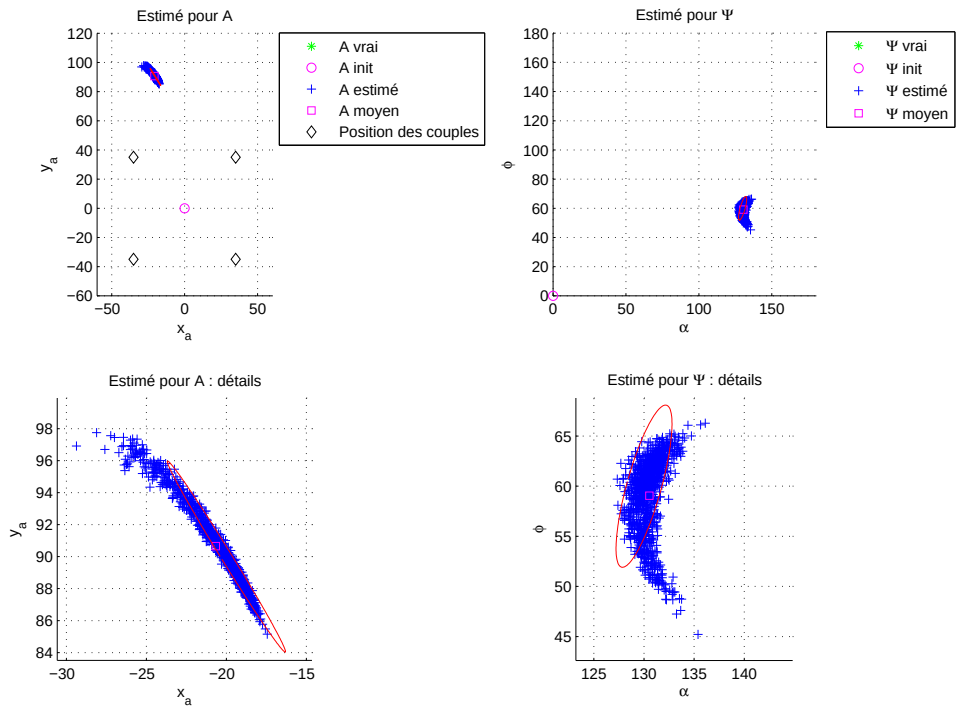

FIG. 4.30 – Représentation des estimés

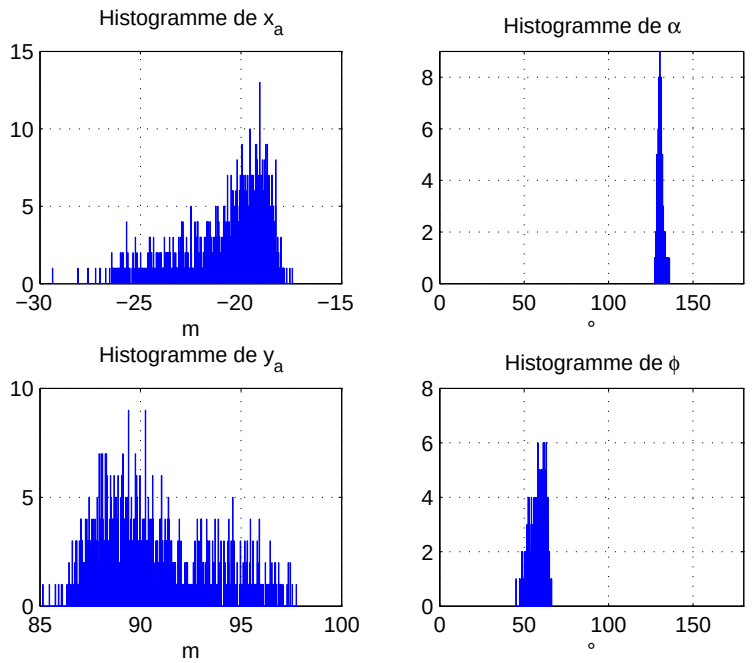

FIG. 4.31 – Histogramme des estimés

Là encore, la forme du nuage de points nous indique que l'on est en limite d'observabi-

lité. Il semble donc que dans les cas un peu extrême comme celui-ci, les conditions initiales jouent un rôle important dans la convergence de l'algorithme.

4.7.4 Résumé et interprétations

Le tableau (4.16) résume les résultats obtenus lors des expériences précédentes.

			$P_d = 1$			$P_d = 0.75$		
		θ_{vrai}	$\widehat{\theta}_m$	$\frac{\text{ecart}}{\text{type}}$	borne	$\widehat{\theta}_m$	$\frac{\text{ecart}}{\text{type}}$	borne
Zone **A**	x_A (m)	-10	-9.996	0.13796	0.13297	-9.9976	0.1824	0.18997
	y_A (m)	20	19.998	0.10891	0.10656	19.9983	0.11592	0.11482
	α (°)	130	129.9397	0.75991	0.75483	129.9973	0.75044	0.75745
	φ (°)	60	60.0015	0.097323	0.098852	59.9948	0.15952	0.15656
Zone **B**	$x_A(m)$	-20	-20.001	0.075977	0.077551	-20.0019	0.10148	0.1034
	y_A (m)	90	89.9996	0.18594	0.18868	90.0037	0.19701	0.19914
	α (°)	130	129.9926	0.34512	0.35268	130.0025	0.3205	0.32661
	φ (°)	60	60.0015	0.09412	0.095139	59.996	0.1607	0.15899

			$P_d = 0.416$		
		θ_{vrai}	$\widehat{\theta}_m$	$\frac{\text{ecart}}{\text{type}}$	borne
Zone **A**	x_A (m)	-10	-9.787	2.0372	1.9183
	y_A (m)	20	20.0455	0.46888	0.42414
	α (°)	130	129.6829	1.6366	1.2992
	φ (°)	60	60.1897	1.5525	1.4465
Zone **B**	$x_A(m)$	-20	-20.6599	2.053	1.7218
	y_A (m)	90	90.6534	2.7236	2.7819
	α (°)	130	130.537	1.2074	1.3014
	φ (°)	60	59.049	3.8539	3.7701

TAB. 4.16 – Résumé des résultats

Pour les positions envisagées de la cascade par rapport à la structure d'hydrophones, le cas $P_d = 1$ est toujours favorable et l'algorithme converge toujours vers la bonne valeur de θ. Le taux de réjection des estimés est de l'ordre de quelques % et le nombre moyen d'itération oscille dans l'intervalle 8 - 20. Le choix des valeurs initiales pour les paramètres à estimer ne semble pas critique. Plusieurs valeurs ont été testées et l'algorithme converge bien dans chaque cas.

Lorsque la probabilité de détection P_d chute à 0.75, (seuls 9 couples d'hydrophones sur 12 fournissent une mesure), les résultats se dégradent un peu mais restent de bonne qualité. L'écart type sur chacune des inconnues augmente légèrement mais reste proche des bornes correspondantes. Le taux de réjection augmente également un peu et le choix des valeurs initiales ne semble toujours pas critique.

Pour une P_d de 0.416, 5 couples seulement fournissent une mesure (c'est le minimum théorique pour que l'algorithme puisse converger). Les résultats se dégradent et on s'aperçoit que l'algorithme ne converge plus vers une valeur de θ unique. Ce problème apparait également lors du calcul de la matrice d'information de Fisher, ce qui est manifeste d'un défaut d'observabilité. On s'aperçoit également que pour certaines valeurs initiales des paramètres à estimer, l'algorithme ne converge pas.

Une extension de cette étude pourrait être une étude approfondie de l'observabilité de la situation. Cette étude se baserait sur une approche géométrique, en tenant compte des caractéristiques du front d'onde issu d'une cascade de longueur efficace finie (comme celle considérée dans les chapitres précédents). On pourrait ainsi affiner le choix des hydrophones lorsque la probabilité de détection P_d n'est pas égale à 1. Nous avions ici considéré que dans ce cas là, les hydrophones atteints étaient les plus proches de la cascade d'énergie (en distance orthogonale). Il est évident que lorsque la cascade a une longueur finie, et que $P_d < 1$ (cas le plus probable), les hydrophones atteint pas le front d'onde ne sont pas les plus proches de la cascade. Néanmoins les résultats précédents ont permis de valider la méthode d'estimation de la trajectoire de la cascade.

4.8 Conclusion

Après la détection, la seconde problématique abordée dans cette thèse a été l'estimation des paramètres de position et d'orientation de la cascade d'énergie. Dans un premier temps, il a fallu estimer la date de l'instant de détection. Un estimateur sommaire, basée sur un filtrage adapté et sur un calcul d'intercorrélation a été étudié et validé. Une extension possible de ces travaux est de poursuivre l'étude de cet estimateur en utilisant des méthodes plus poussées.

Dans un second temps, nous nous sommes intéressés à l'estimation des paramètres de position et d'orientation de la cascade d'énergie. Cette estimation se base sur une structure à base de couples d'hydrophones. Celle-ci a été choisie volontairement simple dans cette première approche et rendue la plus symétrique possible par rapport aux différentes positions d'apparition des phénomènes.

Les nombreuses simulations ont validé les résultats théoriques, et si un nombre minimum de capteur réagissent au passage du front d'onde, les paramètres sont correctement évalués. Il reste néanmoins certaines positions pour lesquelles l'algorithme ne converge pas. Une étude plus poussée de l'observabilité du phénomène constituerait une intéressante extension de

ces travaux. Cette étude prendrait en compte le cas des cascades d'énergie étudiées dans les chapitres dédiées à la détection, à savoir des cascades de longueur finies (d'environ 10m).

Il est également important de remarquer que l'étude de cet estimateur peut assez facilement être adaptée à des réseaux d'hydrophones de forme et taille différentes, mais également à des cascades d'énergie de caractéristiques différentes. Cette étude s'est donc voulue assez générale, et a permis de démontrer la faisabilité de l'estimation dans le cas du télescope Antares.

Chapitre 5

Simulation globale

5.1 Introduction

Ce dernier (et court) chapitre est là pour présenter les résultats d'une simulation dite "globale" mettant en oeuvre les techniques de détection et d'estimation décrites dans les chapitres précédents. Cette simulation se veut effectuée dans des conditions réalistes.

5.2 Contexte physique de la simulation

5.2.1 La structure d'hydrophones

L'idée est ici de mettre en place la chaîne entière de traitement permettant la détection et l'estimation de la trajectoire d'un neutrino ayant généré une cascade d'énergie dont le front d'onde aurait rencontré la structure de capteur décrite ci-dessous.

Pour cette simulation, on se propose de considérer une structure d'hydrophones similaire à celle décrite dans le chapitre précédent, avec les paramètres suivants :

- ⋆ $N_p = 30$: Nombre de plans de capteurs
- ⋆ $D_z = 3$ m : Distance entre chaque plan
- ⋆ $N_{\text{couple}} = 4$: Nombre de couples d'hydrophones par plan
- ⋆ $D_{\text{couple}} = 1.5$ m : Distance entre deux hydrophones dans un couple
- ⋆ $D = 70$ m : Distance entre deux couples d'hydrophones dans un plan

Ces paramètres sont résumés à la figure 5.1.

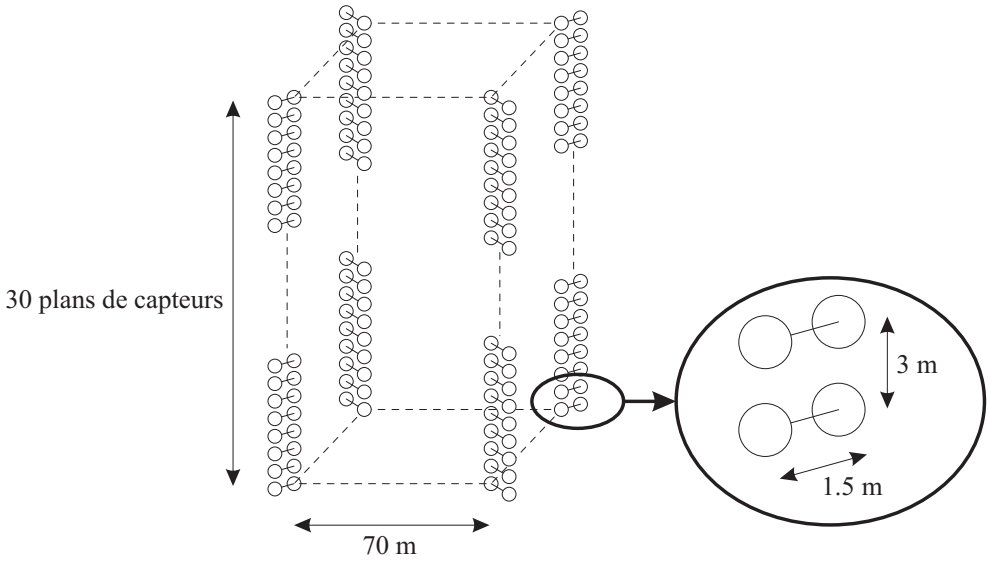

FIG. 5.1 – Description de la structure d'hydrophones utilisée dans cette simulation.

La position de cette structure d'hydrophones par rapport au fond de l'océan n'est pas critique. En effet, celle-ci peut être placée indifféremment sur le sol marin ou sur une position intermédiaire sur les lignes existantes, car seule la position de la cascade d'énergie a une importance dans le traitement effectué.

5.2.2 La cascade d'énergie

La cascade d'énergie et sa position sont caractérisées par plusieurs paramètres :

- ⋆ $E = 10^{19}$ eV : Energie de la cascade.
- ⋆ $L_{\text{eff}} = 10$ m : Longueur efficace de la cascade.
- ⋆ M_1 : Point représentant une extrémité de la cascade.
- ⋆ α : Angle de site de la cascade.
- ⋆ φ : Angle d'azimut de la cascade.

Ces paramètres sont résumés à la figure 5.2.

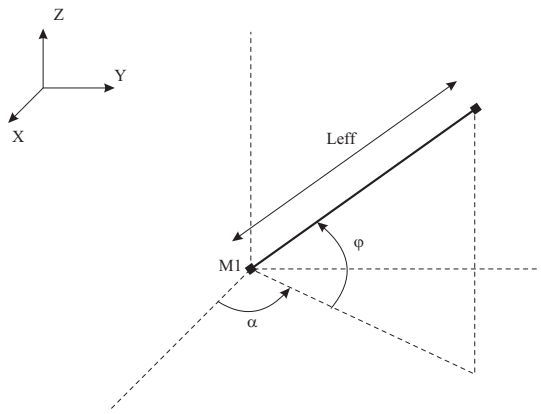

FIG. 5.2 – Description de l'orientation de la cascade d'énergie utilisée dans cette simulation.

Plusieurs cas expérimentaux vont maintenant être étudiés. Dans chacun des cas, la position de la cascade d'énergie est fixée par les données suivantes :

$$M_1 \begin{cases} x = -150 \text{ m} \\ y = 20 \text{ m} \\ z = 0 \text{ m} \end{cases}$$

et

$$\begin{cases} \alpha = 130° \\ \varphi = 75° \end{cases}$$

Comme dans le chapitre 4, l'origine du référentiel est choisie au centre du premier plan de capteur (celui situé à la plus grande profondeur).

5.3 Protocole et résultats expérimentaux

La simulation suit les étapes suivantes :

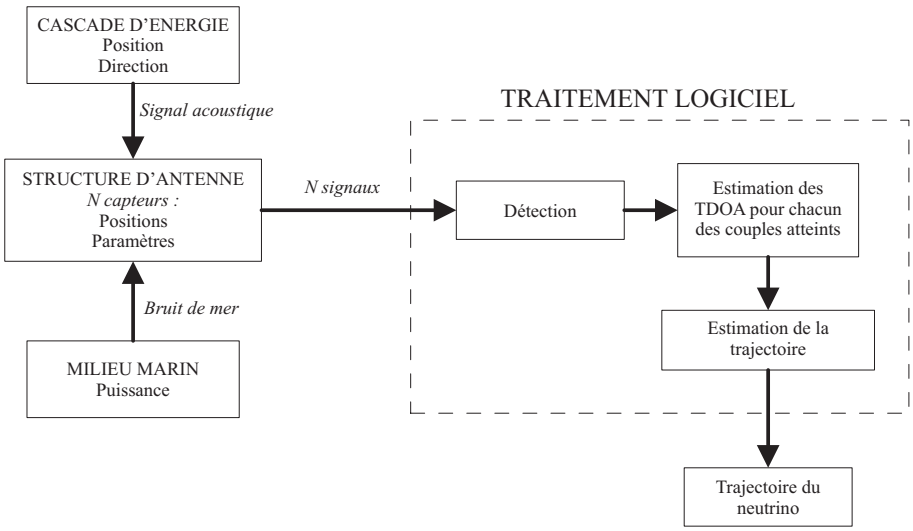

FIG. 5.3 – Les différentes étapes de la simulation.

Notion de rapport signal-à-bruit

La distance [1] entre chaque hydrophone de la structure et une cascade d'énergie de position fixée n'est pas constante. Afin d'utiliser la notion de rapport signal-à-bruit, il faut référencer celui-ci par rapport à un point fixe. Ce point fixe a été arbitrairement choisi comme le barycentre de tous les capteurs de la structure. Le rapport signal-à-bruit est donc le rapport de la puissance du signal acoustique ayant parcouru la distance barycentre / cascade et de la puissance du bruit de mer.

Dans la première expérience, le RSB sera choisi à -45 dB, et dans la seconde à -55dB. Dans ce cas là, le bruit de mer ambiant plus important pourrait correspondre à une mer plus agitée, ou à une profondeur différente pour la structure d'hydrophones.

Bruit de quantification

Nous avons vu dans le chapitre précédent que, pour un couple, l'estimation du TDOA passe par le calcul de l'intercorrélation entre les deux signaux issus de chacun des hydrophones du couple. Le pas du calcul de cette corrélation est de 1 échantillon, ce qui implique un bruit de quantification de l'ordre de $T_e = \frac{1}{F_e}$. La précision maximale de l'estimé du TDOA est donc de l'ordre de $T_e = 1\ \mu s$.

[1] distance entre l'hydrophone et son projeté orthogonal sur la droite engendrée par la cascade d'énergie

5.3.1 Validation des éléments du système

Avant de procéder à une simulation basée sur le système complet, il est nécessaire d'en tester indépendamment chaque module. Le premier module s'occupe de la détection du front d'onde et de la formation des couples correspondants, tandis que le second s'occupe de l'estimation de la trajectoire du neutrino.

5.3.1.1 Modules de détection

Dans un premier temps, le module de détection est soumis à au passage du front d'onde issu de la cascade d'énergie dans une configuration réaliste. La méthode de détection mise en place est celle choisie dans le chapitre 3 et est basée sur un filtrage de type FASE pour un sous-espace de dimension 2.

Une fois identifiés les capteurs ayant réagi au front d'onde, ceux-ci sont regroupés par couples conformément à la figure 5.1. Les couples dont seul l'un des hydrophones a réagi sont alors systématiquement éliminés.

Le RSB étant référencé par rapport au point barycentre des hydrophones de la structure, il est intéressant d'évaluer la valeur du RSB réellement vue par chaque hydrophone. Par exemple, pour un RSB au barycentre de -45 dB, on regroupe ces valeurs sous forme d'histogramme sur la figure 5.4. La dynamique des RSB au niveau de chaque hydrophone va de -37 dB à environ -49 dB.

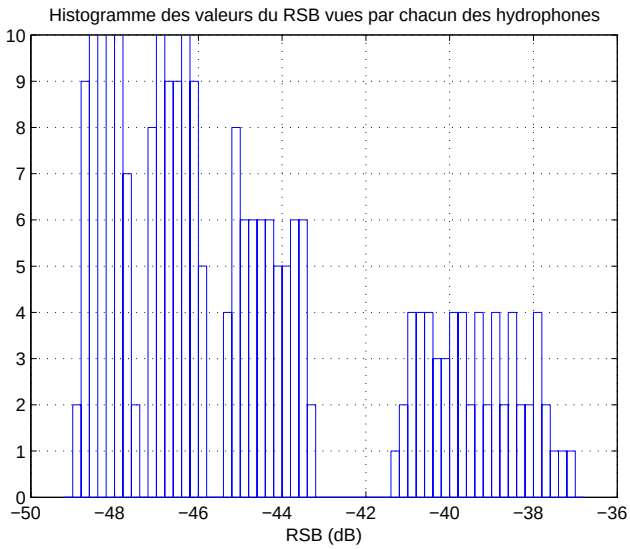

FIG. 5.4 – Histogramme des différents RSB rencontrés sur tous les hydrophones.

Le protocole de test est le suivant : pour un RSB moyen de -45 dB, la structure d'antenne est soumise à la propagation du front d'onde issue de la cascade d'énergie dont la position a été définie précédemment. La valeur du seuil de test de détection est pour le moment choisie manuellement, mais dépend en fait de la valeur du RSB moyen. Une étude du choix de la valeur de ce seuil en fonction de la valeur du RSB moyen serait une intéressante extension de ces travaux. On représente sur la figure 5.5 les résultats de l'algorithme de détection sur la structure d'antenne en trois dimensions. Les hydrophones ayant détectés en théorie et en pratique sont représentés d'une couleur différente.

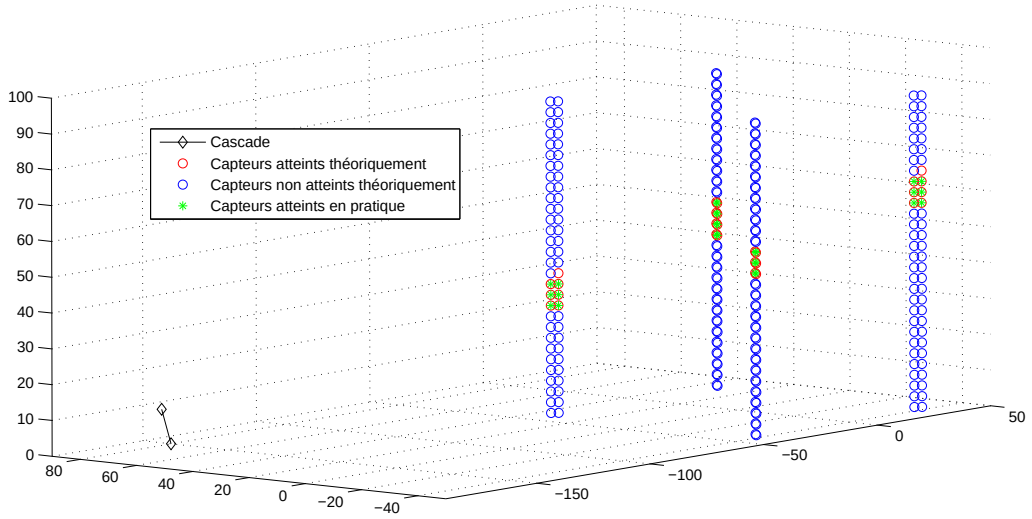

FIG. 5.5 – Représentation de l'état des hydrophones aprés passage du front d'onde.

Au vu de ces résultats, on peut affirmer que le module de détection du front d'onde est opérationnel.

5.3.1.2 Module d'estimation du TDOA

Le second module du système s'occupe d'estimer la différence des temps d'arrivée du front d'onde sur les deux hydrophones constituant chaque couple. La mise en place de cet estimateur est la même que celle utilisée dans le chapitre 4 pour en évaluer les performances (pour plusieurs valeurs de RSB). On rappelle donc simplement celles-ci sur la figure 5.6.

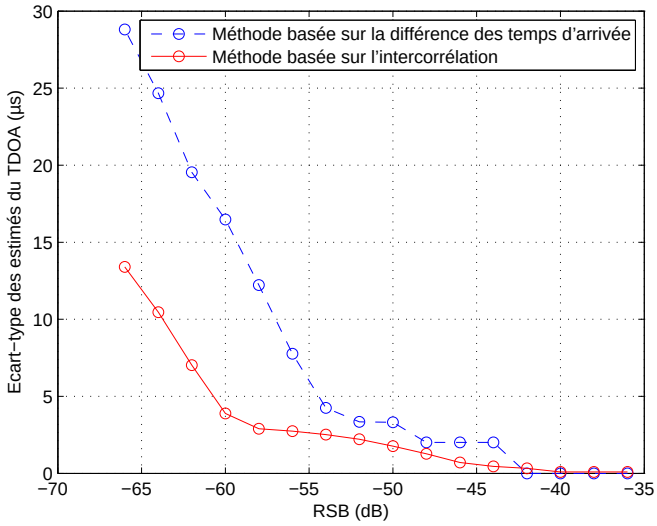

FIG. 5.6 – Evolution de l'écart type des estimés du TDOA en fonction du RSB pour les deux méthodes étudiées.

Les performances d'estimation du TDOA dans la simulation globale dépendront donc de la valeur du RSB au barycentre et de la dispersion de celui-ci pour chacun des hydrophones.

5.3.1.3 Module d'estimation de la trajectoire

Le dernier module du système s'occupe d'estimer la trajectoire initiale du neutrino à partir des valeurs de TDOA estimées précédemment. Pour tester celui-ci, les mesures passées à l'estimateur seront constituées des valeurs vrai du TDOA perturbées par un bruit uniforme de largeur $T_e = 1$ μs (correspondant au bruit de quantification). Ainsi, notant τ_i une mesure passée à l'estimateur, nous avons :

$$\tau = \text{TDOA}_{\text{vrai}} + \varepsilon_i$$

avec $\varepsilon_i \sim \mathtt{U}_{[-\frac{T_e}{2},\frac{T_e}{2}]}$.

Pour 1000 expériences de Monte-Carlo, les résultats sont résumés sur les figure 5.7 et 5.8.

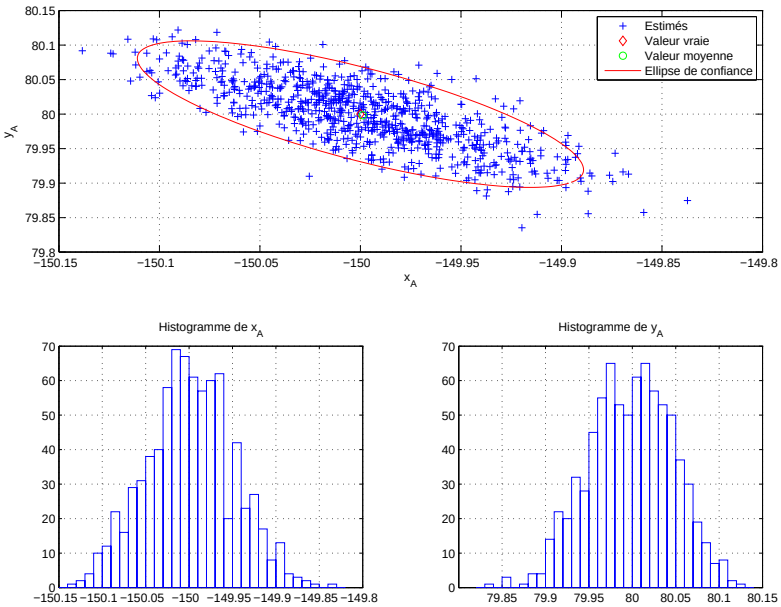

FIG. 5.7 – Résultats du test de l'estimateur de la trajectoire pour un bruit uniforme de largeur le pas de quantification - Positions x_A et y_A.

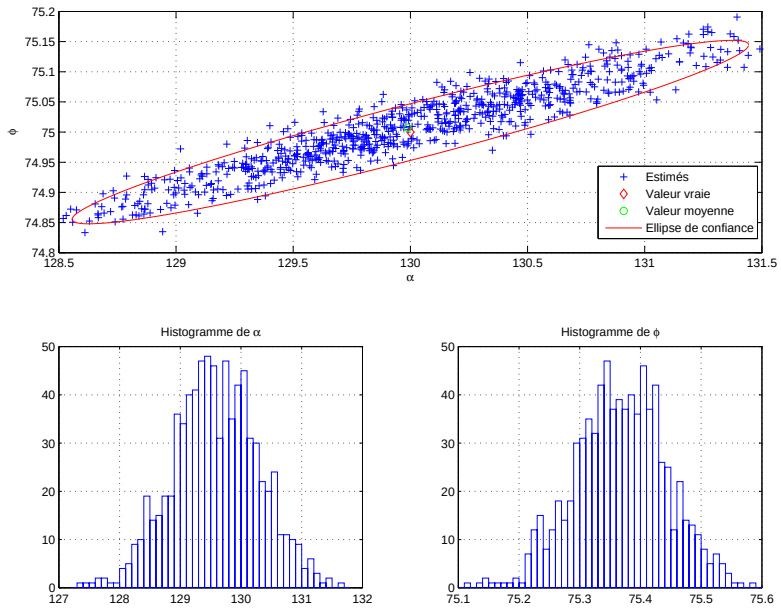

FIG. 5.8 – Résultats du test de l'estimateur de la trajectoire pour un bruit uniforme de largeur le pas de quantification - Angles α et φ.

On rapporte dans le tableau 5.1 les valeurs moyennes, écarts-types et bornes de Cramer-Rao de cet estimateur :

	θ_{vrai}	$\widehat{\theta}_{moyen}$	Ecart type	Borne
x_A (m)	-150	-149.9988	0.051292	0.050738
y_A (m)	80	79.9984	0.049548	0.048512
α (°)	130	129.5879	0.70751	0.673
φ (°)	75	75.3672	0.075436	0.0709

TAB. 5.1 – Caractéristiques des estimés.

Au vu de ces résultats, le module d'estimation de la trajectoire est validé dans cette configuration. Notons que la grande différence avec les simulations présentées au chapitre 4 est la perturbation des mesures par un bruit uniforme et non plus gaussien.

5.3.2 Simulation complète

Avant d'étudier deux cas pratiques différents, il est nécessaire de revenir un instant sur la notion de RSB. Ne disposant pas à ce stade d'études sérieuses sur les niveaux du bruit de mer à des profondeurs de l'ordre de 2500m en Méditerranée, il a été décidé de se fixer les valeurs du RSB et de l'énergie initiale de la cascade (et donc la puissance du signal acoustique). Ces données imposent donc une puissance de bruit de mer, qui pourrait être comparée dans une extension de ces travaux à des valeur réelles de puissance de bruit mesurées sur site.

5.3.2.1 Cas n°1 : RSB = -45 dB

Le premier cas étudié est un cas "favorable" au niveau du RSB. La valeur de celui-ci est choisie à -45 dB. Afin de valider le système complet, on a recours à des simulations de Monte-Carlo sur 700 expériences. Seul le bruit de mer change entre chaque itération, la position de la cascade d'énergie reste la même et les résultats sont les suivants :

	θ_{vrai}	$\widehat{\theta}_m$	écart-type
$x_{M_1^{(0)}}$ (m)	-150	-149.0485	0.065315
$y_{M_1^{(0)}}$ (m)	80	79.4558	0.11561
α (°)	130	132.3129	2.609
φ (°)	75	75.4444	0.12894

On représente sur les figures 5.9 et 5.10 les estimés des coordonnées du point M_1 et des angles α et φ, ainsi que les histogrammes correspondants. Il est également intéressant de

tracer sur la figure 5.11 l'histogramme du nombre de couples d'hydrophones ayant réagis au passage du front d'onde pour toutes les expériences de Monte-Carlo.

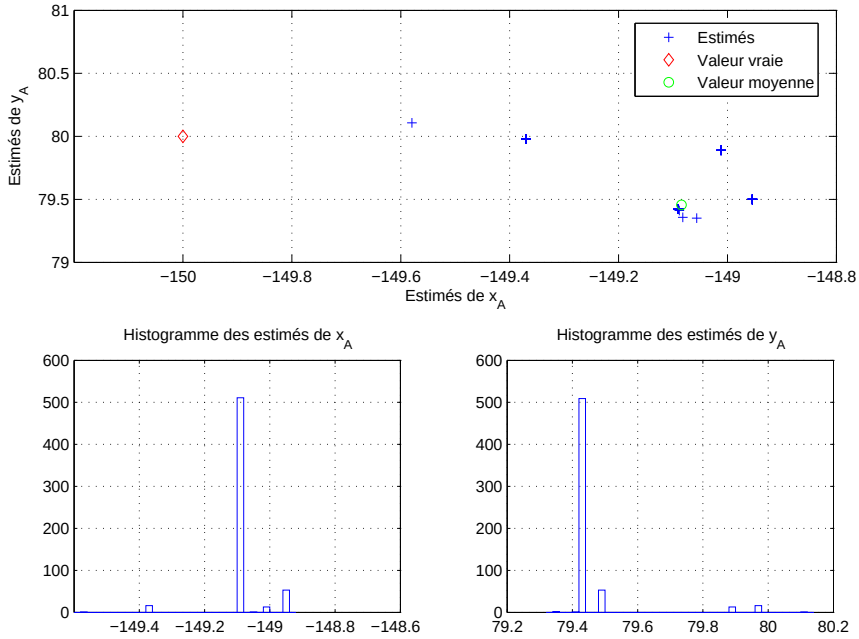

FIG. 5.9 – Représentation et histogrammes des estimés des coordonnées du point M_1.

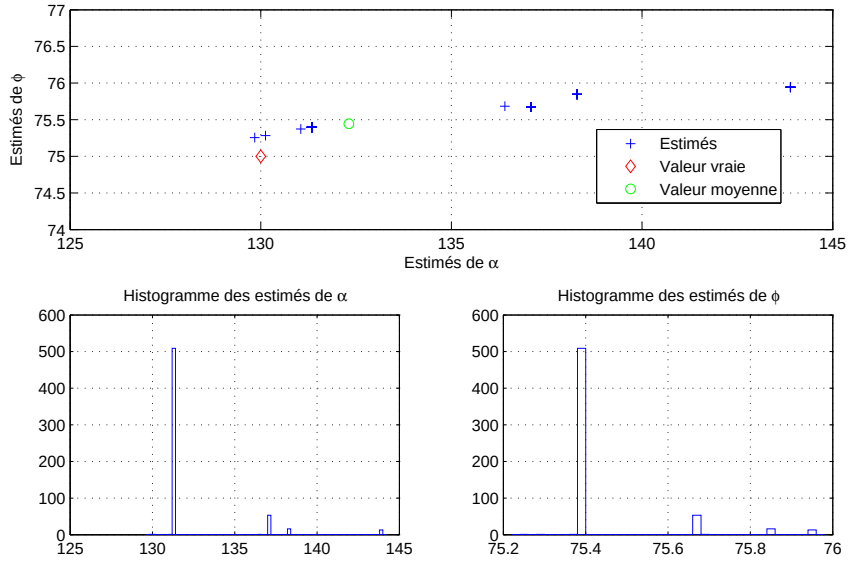

FIG. 5.10 – Représentation et histogrammes des estimés des angles α et φ.

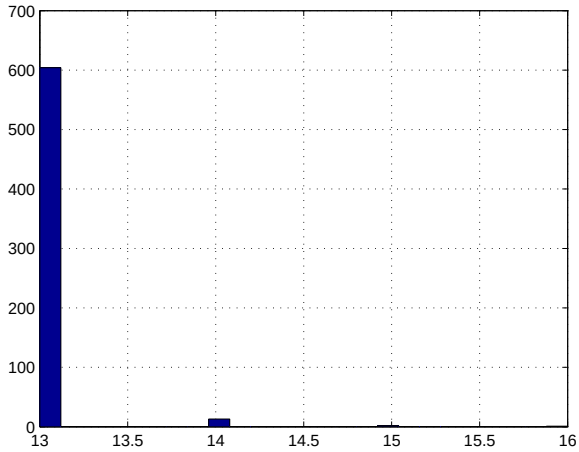

FIG. 5.11 – Histogramme du nombre de couples de capteurs atteints par le front d'onde.

Pour un tel RSB, le taux de rejet des estimés est d'environ 4%. On pourrait s'étonner du faible nombre de points affichés sur les figures précédentes, mais ceci vient du fait que pour une telle valeur de RSB, l'estimateur du TDOA travaille de manière "déterministe". En effet, à cause de la période d'échantillonnage T_e, la mesure τ_{mes} du TDOA s'écrit :

$$\tau_{\text{mes}} = \left\lceil \frac{\tau_{\text{vrai}} + \varepsilon}{T_e} \right\rceil T_e$$

où $\lceil x \rceil$ est la partie entiére de x. Ainsi, suivant nous nous retrouvons dans l'un des cas suivants :

$$\text{Si } \sigma_\varepsilon << \frac{T_e}{\sqrt{12}} \quad \text{alors} \quad \tau_{\text{mes}} = \left\lceil \frac{\tau_{\text{vrai}}}{T_e} \right\rceil T_e$$

$$\text{sinon} \quad \tau_{\text{mes}} \text{ appartient à un ensemble fini de valeurs}$$

Les deux cas précédents sont illustrés sur les figures 5.12 et 5.13. La valeur vraie du TDOA est ici de 80.75μs et la densité de probabilité du bruit supposée gaussienne.

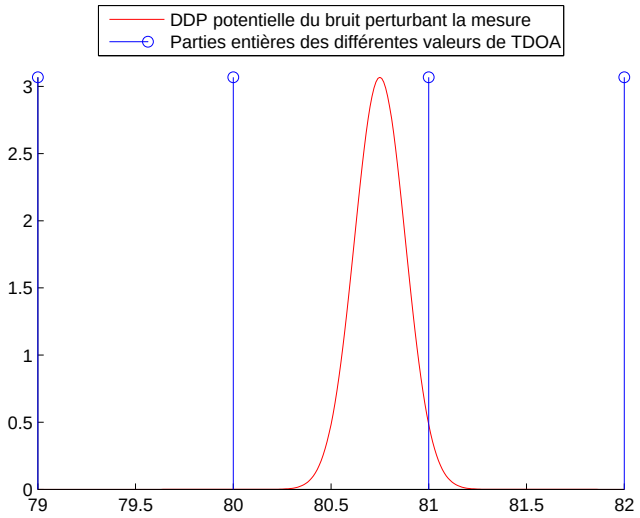

FIG. 5.12 – Premier cas : l'écart type du bruit de mesure est beaucoup plus faible que $\frac{T_e}{\sqrt{12}}$. La valeur possible pour l'estimé est donc $\left\lceil \frac{\tau_{\text{vrai}}}{T_e} \right\rceil T_e$ (ici 81 μs).

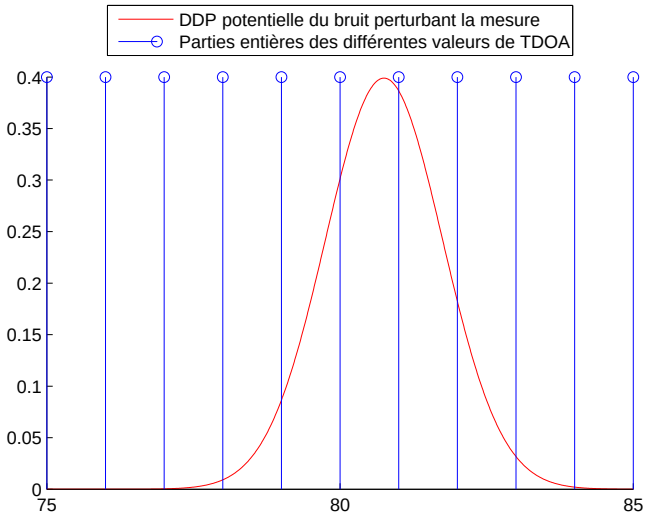

FIG. 5.13 – Second cas : l'écart type du bruit de mesure est plus grand que $\frac{T_e}{\sqrt{12}}$. Les valeurs possibles pour l'estimé appartiennent donc à un ensemble fini de valeurs (représentées en bleu).

5.3.2.2 Cas n°2 : RSB = -55 dB

Le protocole de simulation est le même que celui de la simulation précédente, à la différence que la valeur du rapport signal-à-bruit passe à -55dB. Physiquement ce cas

pourrait se justifier par des conditions de mer plus difficiles, une profondeur plus grande, etc. Les résultats sont les suivants :

	θ_{vrai}	$\widehat{\theta}_m$	écart-type
x_{M_1} (m)	-150	-149.218	1.2592
y_{M_1} (m)	80	79.651	0.6928
α (°)	130	133.9592	11.1947
φ (°)	75	74.8473	11.292

On représente là encore sur les figures 5.14 et 5.15 les estimés des coordonnées du point M_1, des angles α et φ, ainsi que les histogrammes correspondants. Le taux de rejet passe ici à environ 18.7 % et on retrouve le biais dû à la quantification du TDOA. Le phénomène d'estimation identique des TDOA, déjà présent pour un RSB de -45 dB, apparaît encore ici, mais le niveau de bruit étant supérieur, la quantification prédomine moins sur le bruit de mer. Les points représentant les estimés sont donc plus distribués autour de la valeur moyenne, même si de nombreux points sont encore identiques.

On représente enfin sur la figure 5.16 l'histogramme du nombre de couples atteints par le front d'onde pour chaque expérience.

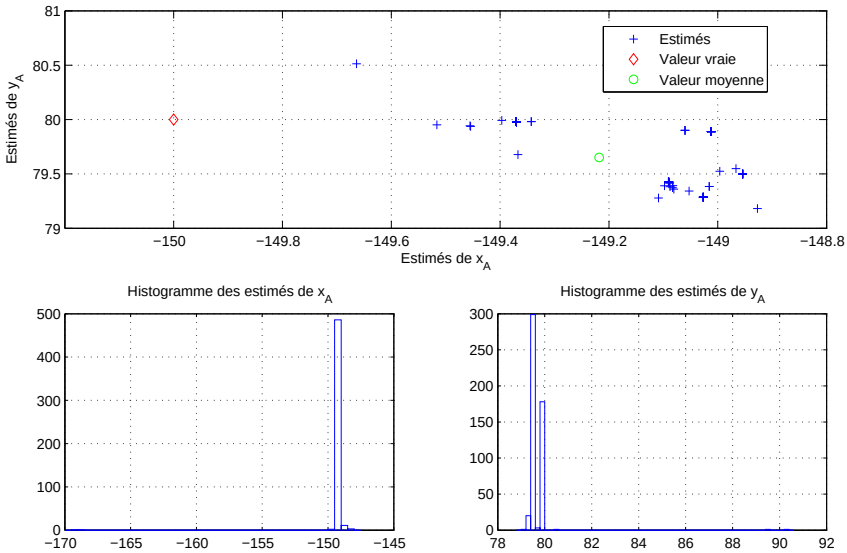

FIG. 5.14 – Représentation et histogrammes des estimés des coordonnées du point M_1.

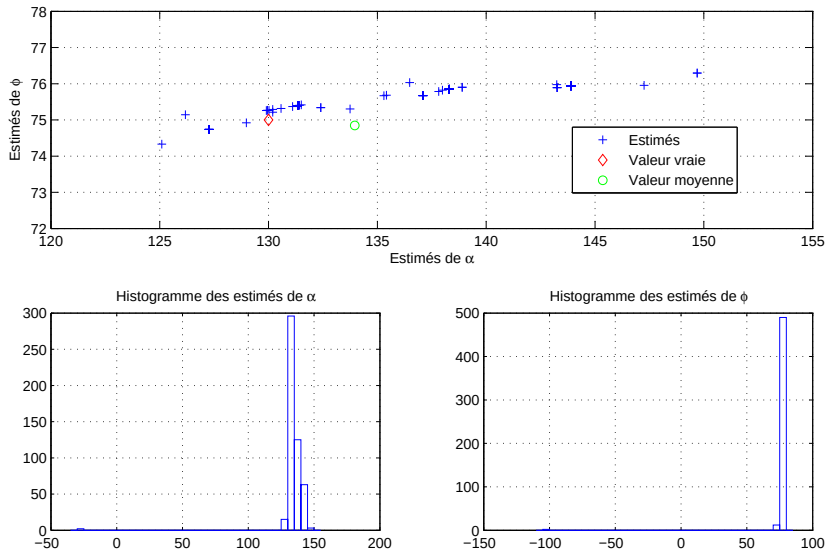

FIG. 5.15 – Représentation et histogrammes des estimés des angles α et φ.

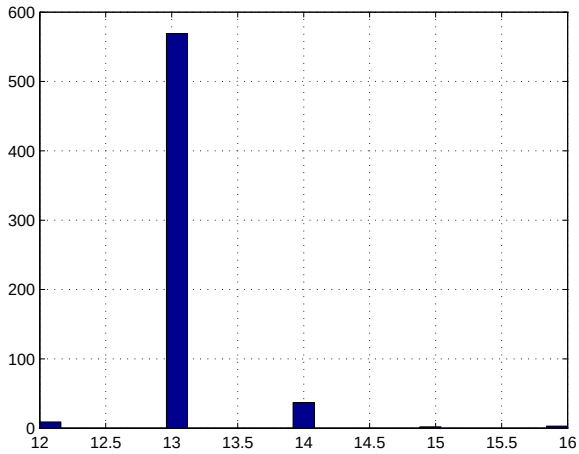

FIG. 5.16 – Histogramme du nombre de couples de capteurs atteints par le front d'onde.

5.4 Conclusion et extensions

Ce chapitre avait pour but de résumer les travaux effectués dans cette thèse en une simulation globale. Celle-ci repose matériellement sur une structure d'hydrophones prédéfinie et logiciellement sur les algorithmes de détection et d'estimation de trajectoire mis au point dans les chapitres précédents. Ceux-ci ont été validés dans cette simulation.

Les extensions possibles de ces simulations sont nombreuses. L'étude de l'influence de la distance de la cascade par rapport à la structure de capteurs permettrait d'affiner le choix de la forme de cette structure. Une étude de l'influence des différentes orientations de la cascade permettrait de mettre en évidence les positions d'inobservabilité et de chercher les moyens de s'en affranchir.

L'estimation du TDOA est certainement la partie de ces simulations la plus susceptible d'être améliorée en utilisant des méthodes plus fines, par exemple les méthodes SCOT, de filtrage d'Eckart, de filtrage à transformation de phase (PHAT), etc. [17].

La littérature actuelle ne nous ayant pas permis de quantifier précisément les niveaux de bruit mis en jeu sur le site du télescope, des mesures calibrées de bruit marin réel à de fortes profondeurs (ici de l'ordre de 2500 m) permettraient d'affiner les résultats de cette simulation et de pouvoir les confronter à la réalité.

Enfin, nous proposons d'étudier un autre estimateur de la trajectoire qui permettrait de diminuer le taux de rejet des estimés. Notant $\boldsymbol{\tau} = (\tau_1, \cdots, \tau_N)$ le vecteur de mesure (ici les estimés des TDOA), il faudrait considérer que chaque mesure τ_i est soit une bonne mesure de détection, soit une mesure de fausse alarme. Un tel estimateur a été étudié [12] et permettrait d'améliorer les performances de l'estimateur de la trajectoire.

Dans les conditions utilisées pour ces simulations, les méthodes de détection et d'estimation étudiées dans cette thèse donnent des résultats satisfaisants. La faisabilité du projet a donc été validée et une première approche a été proposée et pourra servir de base à de futures études.

Conclusion et perspectives

Le but de cette thèse était d'étudier un système capable d'estimer la trajectoire d'un neutrino ayant interagi avec une molécule d'eau dans un environnement marin à forte profondeur. Cet impact aboutit à la création d'une cascade d'énergie qui fait se propager une onde lumineuse couplée à une onde acoustique. C'est sur cette onde acoustique que se sont basés nos travaux.

Modélisation du signal d'intérêt

Dans un premier temps, nous avons étudié différentes modélisations de ce signal. La première était basée sur les travaux d'Askariyan [9]. L'approche employée était analytique et consistait en la résolution de l'équation décrivant la propagation d'une onde acoustique. Askariyan a également utilisé dans son approche une modélisation succincte de l'influence du milieu marin sur cette propagation. Premier modèle disponible au début de cette thèse, celui-ci nous a permis de nous familiariser avec ce genre de signaux courts, mais les limitations du modèle de propagation utilisé et les récentes avancées dans l'étude des cascades d'énergie nous a mené à étudier un seconde modèle basé sur les travaux de Niess [23]. Parmi les deux cascades différentes étudiées par Niess (cascades hadroniques courtes (environ 10m) et cascades électromagnétiques longues (plusieurs centaines de mètres)), nous n'avons étudié que les cascades hadroniques. Le modèle de propagation du front d'onde a été basé sur des études récentes et était plus abouti que celui utilisé par Askariyan.
Dans chacun des deux cas précédent, nous somme capable de tracer l'allure du signal utile reçu pris à une distance d de la cascade. Si cette distance est aléatoire, le signal utile devient donc un signal aléatoire.

Dans la modélisation précédente, nous avons considéré le front d'onde comme plan, ce qui implique que deux signaux acoustiques générés par deux points distincts de la même cascade d'énergie sont les mêmes. Ce n'est pas le cas en réalité, et l'amplitude des signaux générés par les différents points de la cascade suit une distribution proche d'une gaussienne (cf. Chapitre 1). Une première extension de ces travaux pourrait donc être la modification du modèle de signal acoustique reçu en tenant compte de ce phénomène.

Détection du signal acoustique

Le signal d'intérêt maintenant défini, la première étape a été de détecter celui-ci par le biais d'hydrophones. Le signal acoustique détecté par un hydrophone est perturbé par le bruit marin ambiant. Dans un premier temps, nous avons synthétisé ce bruit en le considérant comme gaussien. Dans un second temps, nous avons utilisé du bruit de mer réel issu de campagnes de mesures effectuées en Méditerranée en 2001. Les propriétés de ce bruit ont été étudiées. Différentes méthodes de détection ont ensuite été mises en place, de la plus classique (rapport de vraisemblance) aux plus récentes (maximisation du rapport signal-à-bruit par filtrage adapté stochastique, classification, etc.), et leurs performances ont été quantifiées par le tracé de courbes COR. Parmi toutes ces méthodes, et au vu des résultats expérimentaux obtenus, nous avons décidé de retenir la méthode basée sur le filtrage adapté stochastique étendu (FASE) pour un sous-espace de projection de dimension 2.

Une extension naturelle de cette partie pourrait être l'étude des hydrophones utilisés pour la détection. En effet, ceux-ci induisent un autre bruit dû a l'instrumentation électronique. Ce bruit sensiblement blanc généré par la céramique piézoélectrique du capteur n'est plus négligeable en comparaison du minimum de bruit de mer ambiant, aux fréquences supérieures à 30 kHz [23]. Il faudra donc porter un soin particulier à l'aspect électronique à mettre en place pour les hydrophones.

Comme nous l'avons vu dans le chapitre 3, il semblerait qu'en milieu marin, existent des sources de bruits impulsionnels susceptibles de se confondre avec la signature d'un neutrino. On peut citer par exemple certaines sources biologiques comme les crevettes claqueuses ou les dauphins. Une étude mériterait d'être effectuée pour vérifier que ces bruits perturbent bien l'algorithme de détection (bande passante et amplitudes du même ordre que celles du signal acoustique d'intérêt). Si c'est le cas, on pourrait utiliser le fait que ces "clicks" ne sont pas émis de façon isolée. Il serait alors possible de mettre en place un test d'exclusion basé sur cette "périodicité". De plus, la gestion de ces nouveaux bruits serait à rapprocher des méthodes de classifications de signaux étudiées dans le chapitre 3.

Estimation de la trajectoire du neutrino

Maintenant capable de détecter le signal acoustique au niveau d'un hydrophone, nous avons développé un estimateur de la trajectoire initiale du neutrino. Pour cela, nous avons défini une structure spatiale simple d'hydrophones. Les hydrophones sont regroupés en couples et chaque couple donne une information de temps de retard d'arrivée du front

d'onde.

Dans un premier temps, nous avons étudié un estimateur de la date d'arrivée du front d'onde sur chaque hydrophone, puis un estimateur de la différence des temps d'arrivée (TDOA) de ce même front d'onde sur les deux hydrophones constituant un couple.
Disposant maintenant des informations de TDOA, nous avons développé un algorithme d'estimation de la trajectoire initiale du neutrino. Celle-ci est complètement définie par la donnée d'un point et d'une direction (2 angles). Cet estimateur est basé sur une méthode des moindres carrés. Plusieurs simulations, basées sur différentes positions de la cascade d'énergie par rapport à la structure d'hydrophones, sont venues valider notre estimateur.

La principale extension de la partie estimation est l'amélioration de l'estimateur de TDOA. Par manque de temps, nous avons en effet décidé d'utiliser pour nos simulations un estimateur simpliste basé sur un calcul d'intercorrélation. De nombreuses autres méthodes ont été développées pour améliorer les performances d'estimation de TDOA [17] [10].

Simulation globale

Les résultats d'une simulation dite "globale" occupent le dernier chapitre. Cette simulation a pour but de réunir en une seule expérience tous les résultats obtenus dans chacun des chapitres de cette thèse. Une fois la structure d'hydrophones définie, et une cascade d'énergie correctement positionnée dans l'espace, l'algorithme commence par détecter sur chaque hydrophone le passage du front d'onde. Les signaux issus des couples ayant détecté servent ensuite à estimer le TDOA et ces données sont ensuite traitées par l'estimateur de position. Le système fournit donc en sortie une estimation de la direction du neutrino avant collision. Les performances de ce système ont été évaluées statistiquement grâce à des expériences de Monte-Carlo.

Là encore, plusieurs extensions seraient à envisager. Tout d'abord, une étude plus poussée de l'influence de la forme de la structure de capteurs permettrait d'optimiser les performances d'estimation. Il serait également intéressant d'étudier les zones d'apparition de la cascade menant à une inobservabilité du phénomène et donc à une impossibilité d'estimation. Il semblerait que ces zones soient localisées le long des axes de symétrie de la structure d'hydrophones.

Annexe A

Etude en champ lointain du signal d'Askaryian

On rappelle que dans le domaine fréquentiel, les expressions de la pression acoustique et de la distribution d'énergie sont données par les équations :

$$P_w = iw\frac{\alpha}{4\pi C_p}\chi(w)\int_\Omega \frac{1}{|r_e - r_c|}\ q(r_e)\ e^{i\frac{w}{C_s}|r_e - r_c|}\ \mathrm{d}V$$

$$q(Z_e, R_e) = \frac{1}{2\pi}\sum_{k=1}^{N} A_k(Z_e)e^{-R_e/\lambda_k(Z_e)} \tag{A.1}$$

L'expression de l'énergie totale de la cascade est alors :

$$E_0 = \int_0^\infty \int_0^{2\pi} \int_{-\infty}^{\infty} q(Z_e, R_e)Re\ \mathrm{d}R_e\ \mathrm{d}\varphi_e\ \mathrm{d}Z_e \tag{A.2}$$

Dans l'hypothèse d'une grande distance du récepteur, il est raisonnable d'écrire $R_c \gg \lambda_{eff}$, où λ_{eff} est la dimension effective transverse de la région du dépôt d'énergie. On peut également approximer le terme

$$\int_0^\infty q(Z_e, R_e)\ R_e\ \mathrm{d}Re \tag{A.3}$$

par

$$\int_0^{\lambda_{eff}} q(Z_e, R_e)\ R_e\ \mathrm{d}Re \tag{A.4}$$

car la fonction $f(R_e) = q(Z_e, R_e)$.
R_e est rapidement décroissante et devient négligeable lorsque R_e est supérieure à λ_{eff}. On a donc $Re < \lambda_{eff} \ll R_c$. On peut maintenant réécrire l'équation 1.12 en coordonnées

cylindriques. Pour cela, il est facile de montrer que :

$$|r_e - r_c| = R_c\sqrt{1 + \left(\frac{R_e}{R_c}\right)^2 - 2\frac{R_e}{R_c}\cos(\theta_c - \theta_e) + \frac{(Z_c - Z_e)^2}{R^2}} \tag{A.5}$$

Posons $\gamma = 1 + \frac{(Z_c - Z_e)^2}{R_c^2}$, on obtient alors :

$$|r_e - r_c| = R_c\sqrt{\gamma}\sqrt{1 + \frac{1}{\gamma}\left(\frac{R_e}{R_c}\right)^2 - \frac{2R_e}{R_c\gamma}\cos(\theta_c - \theta_e)} \tag{A.6}$$

Le développement limité à l'ordre 2 de cette expression avec l'hypothèse $R_c \gg R_e$ est :

$$|r_e - r_c| \simeq R_c\sqrt{\gamma} - \frac{R_e}{\sqrt{\gamma}}\cos(\theta_c - \theta_e) + \frac{1}{2}\left(\frac{1}{\sqrt{\gamma}} - \frac{\cos(\theta_c - \theta_e)^2}{2\gamma^{3/2}}\right)\frac{R_e^2}{R_c} \tag{A.7}$$

Or nous savons que $R_e/R_c \ll 1$ et $R_e < \lambda_{eff} < 1$ car la dimension transverse effective de la gerbe est de l'ordre de quelques centimètres. Il vient $R_e^2/Rc \ll 1$, le second terme du développement est donc négligeable. Le développement limité de $|r_e - r_c|$ s'écrit donc :

$$|r_e - r_c| \simeq R_c\sqrt{\gamma}\left(1 - \frac{R_e}{Rc\,\gamma}\cos(\theta_c - \theta_e)\right) \tag{A.8}$$

De la même manière, il faut calculer le développement limité de $\frac{1}{|r_e - r_c|}$. Le même type de raisonnement mène au résultat suivant :

$$\frac{1}{|r_e - r_c|} \simeq \frac{1}{R_c\sqrt{\gamma}} \tag{A.9}$$

Reprenons maintenant l'équation 1.12 et utilisons les deux développements limités précédents :

$$P_w = iw\frac{\alpha}{4\pi C_p}\int_0^\infty\int_0^{2\pi}\int_{-\infty}^\infty \frac{\chi(w)}{R_c\sqrt{\gamma}}\, e^{i\frac{w}{C_s}R\sqrt{\gamma}\left(1 - \frac{R_e}{R_c\gamma}\cos(\theta)\right)}\, q(Z_e, R_e)\, R_e\, \mathrm{d}R_e\, \mathrm{d}\theta\, \mathrm{d}Z_e \tag{A.10}$$

où $\theta = \theta_c - \theta_e$. En remplaçant $q(Z, R_e)$ par son expression A.1, on obtient :

$$P_w = iw\frac{\alpha}{4\pi C_p}\int_0^\infty\int_0^{2\pi}\int_{-\infty}^\infty \frac{\chi(w)}{R_c\sqrt{\gamma}}\, e^{i\frac{w}{C_s}R\sqrt{\gamma}\left(1 - \frac{R_e}{R_c\gamma}\cos(\theta)\right)}\, \frac{1}{2\pi}\sum_{k=1}^N A_k(Z_e)e^{-R_e/\lambda_k(Z_e)}\, R_e\, \mathrm{d}R_e\, \mathrm{d}\theta\, \mathrm{d}Z_e \tag{A.11}$$

Pour simplifier cette expression, il est possible d'intégrer par rapport à R_e et θ. Pour cela, il faut calculer l'intégrale I_k ainsi définie [1] :

$$I_k = \int_0^\infty\int_0^{2\pi} R_e\, e^{R_e\left(-i\frac{w}{C_s}\frac{1}{\sqrt{\gamma}}\cos(\theta) - \frac{R_e}{\lambda_k}\right)}\, \mathrm{d}R_e\, \mathrm{d}\theta \tag{A.12}$$

[1] les $\lambda_k(Z)$ seront notés λ_k dans les calculs qui suivent pour en améliorer la lisibilité

Posant $\beta_k^2 = \left(\frac{w}{C_s}\right)^2 \frac{\lambda_k^2}{\gamma}$, on peut écrire :

$$I_k = \lambda_k^2 \int_0^{2\pi} \frac{1}{\left(1 + i\beta_k \cos(\theta)\right)^2} d\theta \tag{A.13}$$

Le calcul de cette intégrale mène au résultat suivant :

$$I_k = 2\pi \frac{\lambda_k^2}{(1+\beta_k^2)^{3/2}} \tag{A.14}$$

et donc :

$$P_w = iw \frac{\alpha}{4\pi C_p} \int_{-\infty}^{\infty} \frac{\chi(w)}{R_c\sqrt{\gamma}} \sum_{k=0}^{N} \frac{A_k(Z_e)\lambda_k^2(Z_e)}{(1+\beta_k^2)^{3/2}} \; e^{i\frac{w}{C_s}R_c\sqrt{\gamma}} \; \mathrm{d}Z_e \tag{A.15}$$

Il faut maintenant appliquer la transformée de Fourier inverse sur ce résultat pour obtenir l'expression du champ acoustique en fonction du temps :

$$P(r_c, t) = i \frac{\alpha}{4\pi C_p} \int_{-\infty}^{\infty} \frac{1}{2\pi} \int_{-\infty}^{\infty} w \frac{\chi(w)}{R_c\sqrt{\gamma}} \sum_{k=0}^{N} \frac{A_k(Z_e)\lambda_k^2(Z_e)}{(1+\beta_k^2)^{3/2}} \; e^{iw\left(\frac{R_c\sqrt{\gamma}}{C_s} - t\right)} \; \mathrm{d}Z_e \; \mathrm{d}w \tag{A.16}$$

Le problème étant symétrique suivant le signe de w, l'expression du champ de pression, que l'on notera désormais $s_{\text{Askaryian}}$ est donc :

$$s_{\text{Askaryian}}(r_c, t) = i \frac{\alpha}{4\pi^2 C_p} \int_{-\infty}^{\infty} \int_{0}^{\infty} w \frac{\chi(w)}{R_c\sqrt{\gamma}} \sum_{k=0}^{N} \frac{A_k(Z_e)\lambda_k^2(Z_e)}{(1+\beta_k^2)^{3/2}} \; e^{iw(T_0 - t)} \; \mathrm{d}Z_e \; \mathrm{d}w \tag{A.17}$$

avec $T_0 = \frac{R_c\sqrt{\gamma}}{C_s}$

Annexe B

A propos de l'échantillonnage

Pour un nombre impair n de mesures, la matrice $\mathbf{A}_\delta(\Delta t)$ peut se mettre sous la forme :

$$\mathbf{A}_\delta(\Delta t) = \begin{bmatrix} 1 & -\frac{n}{2}\triangle t & (-\frac{n}{2}\triangle t)^2 & \dots & (-\frac{n}{2}\triangle t)^\delta \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \triangle t & \triangle t^2 & \dots & \triangle t^\delta \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \frac{n}{2}\triangle t & (\frac{n}{2}\triangle t)^2 & \dots & (\frac{n}{2}\triangle t)^\delta \end{bmatrix}$$

et donc :

$$\mathbf{A}_\delta(\Delta t) = \begin{bmatrix} 1 & -\frac{n}{2} & (-\frac{n}{2})^2 & \dots & (-\frac{n}{2})^\delta \\ 1 & -\frac{n}{2}+1 & (-\frac{n}{2}+1)^2 & \dots & (-\frac{n}{2}+1)^\delta \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & \dots & 1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \frac{n}{2} & (\frac{n}{2})^2 & \dots & (\frac{n}{2})^\delta \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & \dots & 0 \\ 0 & \triangle t & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \triangle t^\delta \end{bmatrix}$$

$\mathbf{A}_\delta(\Delta t)$ peut donc se mettre sous la forme $\mathbf{A}_\delta(\Delta t) = \mathbf{A}_\delta(1)\mathbf{D}_\delta(\triangle t)$.

Les conséquences sont les suivantes :

$$\begin{aligned} T(X) &= X^T\mathbf{A}_\delta(\Delta t)\left[\mathbf{A}_\delta^T(\Delta t)\mathbf{A}_\delta(\Delta t)\right]^{-1}\mathbf{A}_\delta^T(\Delta t)X \\ &= X^T\mathbf{A}_\delta(1)\mathbf{D}_d(\Delta t)\left[\mathbf{D}_\delta(\Delta t)\mathbf{A}_\delta^T(1)\mathbf{A}_\delta(1)\mathbf{D}_\delta(\Delta t)\right]^{-1}\mathbf{D}_\delta(\Delta t)\mathbf{A}_\delta^T(1)X \\ &= X^T\mathbf{A}_\delta(1)\left[\mathbf{A}_\delta^T(1)\mathbf{A}_\delta(1)\right]^{-1}\mathbf{A}_\delta^T(1)X \end{aligned}$$

Conclusion : la matrice test $\mathbf{A}_\delta(1)\left[\mathbf{A}_\delta^T(1)\mathbf{A}_\delta(1)\right]^{-1}\mathbf{A}_\delta^T(1)$ ne dépend pas de $\triangle t$.

Annexe C

Variances et écarts-types du détecteur classique dans le cas d'un bruit gaussien centré de variance connue et constante

λ	**Variance** σ^2	**Ecart type** σ	S_A^{max}/σ
1	0.1431×10^{-5}	1.1962×10^{-3}	0.2605
2	0.0715×10^{-5}	0.8458×10^{-3}	0.3685
3	0.0477×10^{-5}	0.6906×10^{-3}	0.4513
4	0.0358×10^{-5}	0.5981×10^{-3}	0.5211
5	0.0286×10^{-5}	0.5349×10^{-3}	0.5826
6	0.0238×10^{-5}	0.4883×10^{-3}	0.6382
7	0.0204×10^{-5}	0.4521×10^{-3}	0.6893
8	0.0179×10^{-5}	0.4229×10^{-3}	0.7369
9	0.0159×10^{-5}	0.3987×10^{-3}	0.7816
10	0.0143×10^{-5}	0.3783×10^{-3}	0.8239
11	0.0130×10^{-5}	0.3607×10^{-3}	0.8641
12	0.0119×10^{-5}	0.3453×10^{-3}	0.9025
13	0.0110×10^{-5}	0.3318×10^{-3}	0.9394
14	0.0102×10^{-5}	0.3197×10^{-3}	0.9748
15	0.0095×10^{-5}	0.3089×10^{-3}	1.0091
16	0.0089×10^{-5}	0.2990×10^{-3}	1.0422
17	0.0084×10^{-5}	0.2901×10^{-3}	1.0742
18	0.0079×10^{-5}	0.2819×10^{-3}	1.1054
19	0.0075×10^{-5}	0.2744×10^{-3}	1.1357
20	0.0072×10^{-5}	0.2675×10^{-3}	1.1652

TAB. C.1 – Correspondances entres valeurs de λ, variances et écarts types dans le cas du détecteur classique pour un bruit gaussien centré et de variance connue et constante

Annexe D

Calcul de la loi de $\frac{1}{\sigma}(\mathrm{M}_d^T\mathrm{M}_d)^{-1/2}\mathrm{M}_d^T X$ dans le cas de l'hypothèse H_0

On part de l'hypothèse : $X \sim G\left(0, \sigma^2 \mathbf{I}_n\right)$, on a ainsi :

$$\mathbf{M}_d^T X \sim G\left(\mu, \sigma^2 \mathbf{M}_d^T \mathbf{M}_d\right)$$

avec $\mu = 0$ sous l'hypothèse H_0 et $\mu \neq 0$ dans l'hypothèse H_1.

Posons $\mathbf{S} = \sigma^2 \mathbf{M}_d^T \mathbf{M}_d$.
$\mathbf{S}$ est une matrice symétrique et définie positive. On peut donc la mettre sous la forme $\mathbf{S} = \mathbf{P}^T\mathbf{D}\mathbf{P}$, où $\mathbf{P}$ est une matrice orthogonale ($\mathbf{P}^T\mathbf{P} = \mathbf{P}\mathbf{P}^T = \mathbf{I}_d$) et $\mathbf{D}$ une matrice diagonale dont les termes diagonaux sont les valeurs propres de $\mathbf{S}$. On a donc $\mathbf{D} = diag(\lambda_i)_{i=1\ldots d}$.
On peut donc écrire :

$$\mathbf{S} = \mathbf{P}^T \triangle \triangle \mathbf{P}$$

avec $\triangle = diag(\sqrt{\lambda_i})_{i=1\ldots d}$. On posera dans la suite :

$$\mathbf{S}^{1/2} = \mathbf{S}^T \triangle \mathbf{P} \quad et \quad \mathbf{S}^{-1/2} = \mathbf{P}^T \triangle^{-1} \mathbf{P}$$

L'écriture est cohérente car :

$$\begin{aligned} \mathbf{S}^{1/2}\mathbf{S}^{1/2} &= (\mathbf{P}^T\triangle\mathbf{P})(\mathbf{P}^T\triangle\mathbf{P}) = \mathbf{P}^T\triangle\mathbf{P}\mathbf{P}^T\triangle\mathbf{P} \\ &= \mathbf{P}^T\triangle^2\mathbf{P} = \mathbf{P}^T\mathbf{D}\mathbf{P} = \mathbf{S} \end{aligned}$$

Vérifions maintenant que $\mathbf{S}^{1/2}Y \sim G(\mathbf{S}^{1/2}\mu, \mathbf{I}_d)$.
Nous avons ainsi les propriétés suivantes :

- ⋆ $\mathbf{S}^{1/2}Y$ est gaussien
- ⋆ $\mathsf{E}(\mathbf{S}^{1/2}Y) = \mathbf{S}^{1/2}\mathsf{E}(Y) = \mathbf{S}^{1/2}\mu$

$$
\begin{aligned}
\star \quad Cov(\mathbf{S}^{1/2}Y) &= \mathbf{S}^{1/2}\mathsf{E}(YY^T)\mathbf{S}^{1/2} = \mathbf{S}^{1/2}\mathbf{S}\mathbf{S}^{1/2} \\
&= \mathbf{P}^T\triangle^{-1}\mathbf{P}\mathbf{P}^T\mathbf{D}\mathbf{P}\mathbf{P}^T\triangle^{-1}\mathbf{P} = \mathbf{P}^T\triangle^{-1}\mathbf{D}\triangle^{-1}\mathbf{P} \\
&= \mathbf{P}^T\triangle^{-1}\triangle\triangle\triangle^{-1}\mathbf{P} = \mathbf{P}^T\mathbf{P} = \mathbf{I}_d
\end{aligned}
$$

Annexe E

Dépendance de $X^T \Pi_1 X$ et $X^T X$

Montrons que les variables aléatoires $X^T \Pi_1 X$ et $X^T X$ sont dépendantes.
Pour cela, considérons le vecteur

$$Y = \begin{bmatrix} Y_1 = \Pi_1 X \\ Y_2 = X \end{bmatrix} = \begin{bmatrix} \Pi_1 \\ \mathbf{I}_N \end{bmatrix} X = \Pi \; X$$

X étant un vecteur gaussien par définition on a, pour les deux hypothèses :

$$\begin{aligned} Cov(Y) &= E\left\{\Pi X X^T \Pi^T\right\} \\ &= \Pi \sigma^2 \mathbf{I}_N \Pi^T \\ &= \sigma^2 \Pi \Pi^T \end{aligned}$$

Détaillons $\Pi \; \Pi^T$:

$$\Pi = \begin{pmatrix} \Pi_1 \\ \mathbf{I}_N \end{pmatrix} \quad \Pi^T = (\Pi_1^T, \mathbf{I}_N) = (\Pi_1, \mathbf{I}_N)$$

donc

$$\begin{aligned} \Pi \; \Pi^T &= \begin{bmatrix} \Pi_1^2 & \Pi_1 \\ \Pi_1 & \mathbf{I}_N \end{bmatrix} \\ &= \begin{bmatrix} \Pi_1 & \Pi_1 \\ \Pi_1 & \mathbf{I}_N \end{bmatrix} \end{aligned}$$

On en déduit donc que les vecteurs $\Pi_1 X$ et X sont corrélés. Comme ils sont corrélés, ils sont dépendants. Ce sera donc également le cas pour $\|\Pi_1 X\|^2 = X^T \Pi_1 X$ et $\|X\|^2 = X^T X$.

Annexe F

Indépendance de $X^T(\mathbf{I}_N - \Pi_1)X$ et $X^T\Pi_1 X$

Montrons que les variables aléatoires $X^T(\mathbf{I}_N - \Pi_1)X$ et $X^T\Pi_1 X$ sont indépendantes. Pour cela, considérons le vecteur

$$Y = \begin{bmatrix} Y_1 = \Pi_1 X \\ Y_2 = (\mathbf{I}_N - \Pi_1)X \end{bmatrix} = \begin{bmatrix} \Pi_1 \\ (\mathbf{I}_N - \Pi_1) \end{bmatrix} X = \Pi\ X$$

X étant un vecteur gaussien par définition on a, pour les deux hypothèses :

$$\begin{aligned} Cov(Y) &= E\left\{\Pi X X^T \Pi^T\right\} \\ &= \Pi\sigma^2 \mathbf{I}_N \Pi^T \\ &= \sigma^2 \Pi\Pi^T \end{aligned}$$

Détaillons $\Pi\ \ \Pi^T$:

$$\Pi = \begin{pmatrix} \Pi_1 \\ 1 - \Pi_1 \end{pmatrix} \quad \Pi^T = \left(\Pi_1^T, (\mathbf{I}_N - \Pi_1)^T\right)$$

donc

$$\Pi\ \Pi^T = \begin{bmatrix} \Pi_1\Pi_1^T & \Pi_1(\mathbf{I}_N - \Pi_1^T) \\ (\mathbf{I}_N - \Pi_1)\Pi_1 & (\mathbf{I}_N - \Pi_1)(\mathbf{I}_N - \Pi_1^T) \end{bmatrix} = \begin{bmatrix} \Pi_1 & 0 \\ 0 & \mathbf{I}_N - \Pi_1 \end{bmatrix}$$

On en déduit donc que les sous-vecteurs $\Pi_1 X$ et $(\mathbf{I}_N - \Pi_1)X$ sont décorrélés. Comme ils sont gaussiens, ils sont indépendants. Par conséquent, toute fonction du premier est indépendante de toute fonction du second. C'est donc le cas pour $\|\Pi_1 X\|^2 = X^T\Pi_1 X$ et $\|(\mathbf{I}_N - \Pi_1)X\|^2 = X^T(\mathbf{I}_N - \Pi_1)X$

Annexe G

Evolution des performances de détection en fonction de la distance

Nous allons ici étudier l'évolution des performances d'une méthode de détection en fonction de la distance d parcourue par front d'onde avant d'atteindre l'hydrophone considéré.
Dans toute la suite, on considérera que le signal acoustique correspondant suit le modèle de Niess.

On se fixe *a priori* deux courbes COR dite de référence. Ces courbes représentent les performances de détection que l'on s'impose. La première correspond à des performances moyennes et la seconde à des performances hautes. Elles sont représentées sur les figures G et G.

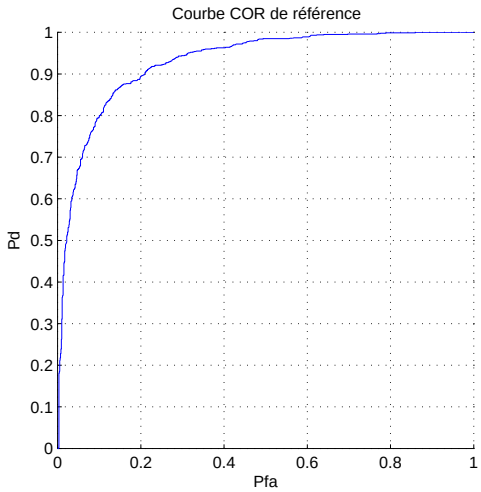

FIG. G.1 – Courbes COR de référence du détecteur correspondant à des performances moyennes.

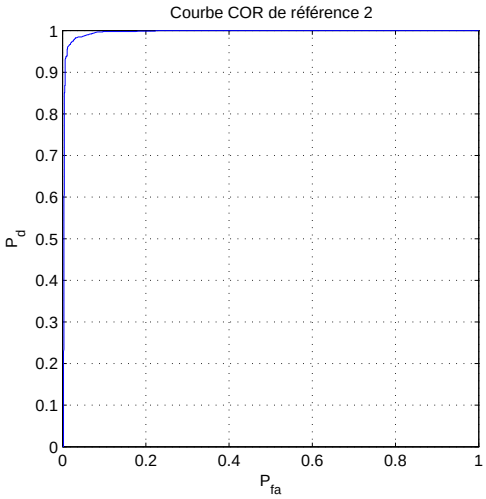

FIG. G.2 – Courbes COR de référence du détecteur correspondant à des performances hautes.

G.1 Atténuation du bruit en fonction de la distance pour des performances de détection constantes

Notons S_{100} le signal acoustique ayant parcouru une distance $d_0 = 100$m, sa puissance sera notée $P_{S_{100}}$ (on rappelle que c'est la distance pour laquelle le modèle de Niess est valable). On décide de référencer les résultats par rapport à S_{100}. Pour cela, on choisit une des deux courbes COR précédentes comme courbe COR de référence et on cherche la puissance du bruit notée P_b pour laquelle la courbe COR de détection soit identique à la courbe de référence. Connaissant les puissances de S_{100} et du bruit, on en déduit le rapport signal-à-bruit pour d_0.

$$RSB_{d_0} = 10\log_{10}\left(\frac{P_{S_{100}}}{P_b}\right)$$

La puissance du bruit est maintenant fixée à P_b et restera constante pour la suite des simulations.

D'après les résultats du chapitre 1, un signal S_d ayant parcours un distance $d > d_0$ aura une puissance plus faible que celle de S_{100} :

$$P_{S_d} < P_{S_{100}}$$

La courbe COR associée à ce signal S_d sera donc inférieure à la courbe COR de référence. L'idée est ici d'évaluer la valeur du coefficient d'atténuation en puissance noté α que l'on devrait appliquer au bruit pour obtenir les performances de références (et donc les mêmes

courbes COR). La nouvelle puissance du bruit est notée P'_b et on a :

$$P'_b = \alpha \; P_b$$

On s'intéresse à l'évolution de α en fonction de d et on trace sur la figure G.3 un exemple de son évolution .

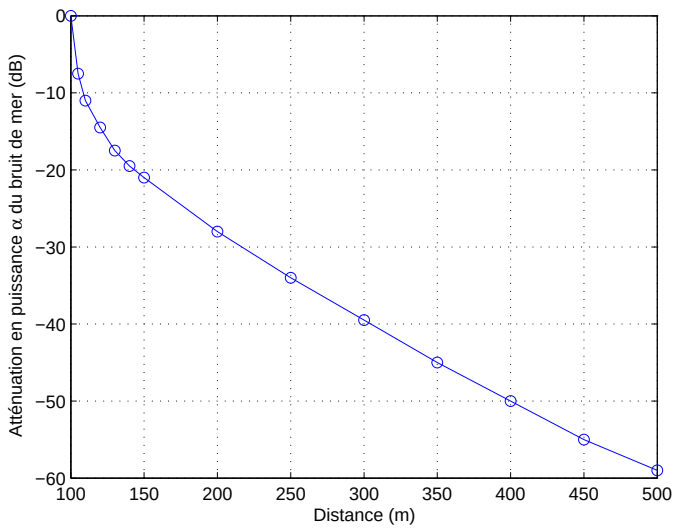

FIG. G.3 – Exemple de l'évolution du coefficient d'atténuation en puissance α en fonction de la distance parcourue par le front d'onde pour des performances de détection fixées.

Les courbes d'évolution des coefficients α en fonction de la distance associées aux différentes méthodes de détection seront commentées dans les chapitres correspondants.

G.2 Evolution du RSB en fonction de la distance pour des performances de détection constantes

On s'intéresse maintenant à l'évolution du rapport signal-à-bruit en fonction de la distance parcourue par le front d'onde. Ce RSB correspond à des performances de détection fixées par la courbe COR de référence choisie et est calculé *après compensation de la puissance du bruit par le coefficient* α. Un exemple de ce type de courbe est représenté sur la figure G.4

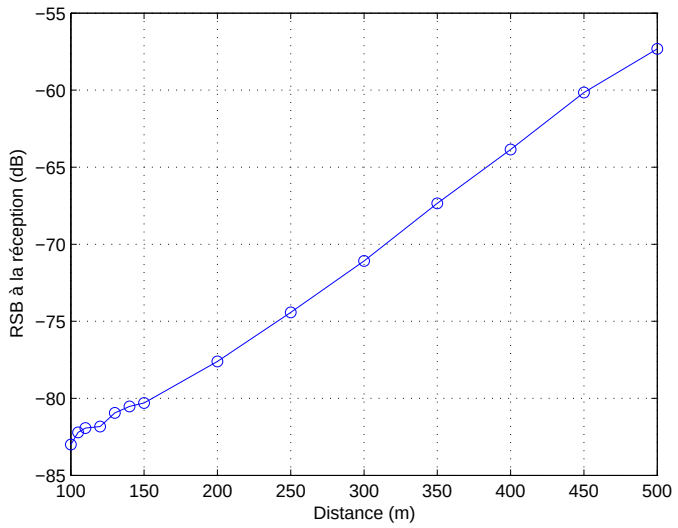

FIG. G.4 – Exemple de l'évolution du RSB au niveau de l'hydrophone en fonction de la distance parcourue par le front d'onde pour des performances de détection fixées.

Les courbes d'évolution du RSB en fonction de la distance associées aux différentes méthodes de détection seront commentées dans les chapitres correspondants.

G.3 Vérification des résultats

Afin de vérifier que les résultats obtenus sont cohérents, on se base sur le calcul suivant :

Le RSB, exprimé en dB, au niveau de l'hydrophone aprés compensation de la puissance du bruit s'écrit :

$$\begin{aligned} RSB_{\text{dB}} &= 10\log_{10}\left(\frac{P_{S_d}}{P'_b}\right) \\ &= 10\log_{10}\left(\frac{P_{S_d}}{\alpha\ P_b}\right) \\ &= -10\log_{10}(\alpha) + 10\log_{10}\left(\frac{P_{S_d}}{P_b}\right) \\ &= -10\log_{10}(\alpha) + 10\log_{10}\left(\frac{P_{S_d}}{P_{S_{100}}}\right) + 10\log_{10}\left(\frac{P_{S_{100}}}{P_b}\right) \end{aligned}$$

Le RSB exprimé en dB se décompose donc en trois termes :

⋆ Le premier terme $10\log_{10}(\alpha)$ correspond au coefficient α exprimé en dB.

⋆ Le second terme $10\log_{10}\left(\frac{P_{S_d}}{P_{S_{100}}}\right)$ correspond au rapport de perte de puissance ex-

primé en dB entre un signal pris à une distance d et le signal pris à 100m.

- ⋆ Le troisième terme $10\log_{10}\left(\frac{P_{S_{100}}}{P_b}\right)$ correspond au terme constant RSB_{d_0} défini précédemment.

La courbe décrivant l'évolution de $10\log_{10}\left(\frac{P_{S_d}}{P_{S_{100}}}\right)$ en fonction de la distance d a été décrite dans le chapitre 1. On en conclut que trace la fonction suivante :

$$RSB_{\mathrm{dB}} + 10\log_{10}(\alpha) - 10\log_{10}\left(\frac{P_{S_{100}}}{P_b}\right)$$

on doit retomber sur le tracé 1.12.

Cette vérification sera effectuée pour chaque méthode de détection utilisée.

Annexe H

Calcul du rapport de vraisemblance dans le cas Antares

Les hypothèses de détection sont les suivantes :

$$\begin{aligned} H_0 &: X = b_m \\ H_1 &: X = S_A + b_m \end{aligned}$$

Si chaque processus mis en jeu est supposé gaussien, il est possible de donner les expressions des densités de probabilité de b_m et S_A :

$$p_{b_m}(e) = \frac{1}{\sqrt{\det(2\pi\mathbf{B})}} \exp^{-\frac{1}{2}e^T \ \mathbf{B}^{-1} \ e} \tag{H.1}$$

$$p_{S_A}(s) = \frac{1}{\sqrt{\det(2\pi\mathbf{A})}} \exp^{-\frac{1}{2}(s-\overline{s_A})^T \ \mathbf{A}^{-1} \ (s-\overline{s_A})} \tag{H.2}$$

On peut en déduire les densités de probabilité associées à Z dans les deux hypothèses précédentes. Pour cela, déterminons les matrices de covariances $\mathbf{R}_0$ et $\mathbf{R}_1$ de Z dans chacune des deux hypothèses :

Sous H_0 :

$$\mathbf{R}_0 = \mathsf{E}(ZZ^T) = \mathsf{E}(\mathbf{H}^T b_m b_m^T \mathbf{H}) = \mathbf{H}^T\mathbf{B}\mathbf{H} \tag{H.3}$$

Sous H_1 :

$$\mathbf{R}_1 = \mathsf{E}\left(ZZ^T\right) = \mathsf{E}\left(\mathbf{H}^T(S_A + b_m)(S_A + b_m)^T\mathbf{H}\right) \tag{H.4}$$

$$= \mathsf{E}\left(\mathbf{H}^T(S_A S_A^T + b_m b_m^T + 2\ S_A\varepsilon^T)\mathbf{H}\right) \tag{H.5}$$

$$= \mathbf{H}^T\left(\mathbf{A} + \mathbf{B}\right)\mathbf{H} \tag{H.6}$$

On en déduit donc les densités de probabilité :

$$p_Z(z/H_0) = p_{b_m}(z) = \frac{1}{\sqrt{\det\left(2\pi\mathbf{R}_0\right)}} \exp^{-\frac{1}{2}z^T\ \mathbf{R}_0^{-1}\ z} \tag{H.7}$$

et

$$\begin{aligned} p_Z(z/H_1) &= \left(p_{b_m} * p_{S_A}\right)(z) && \text{(H.8)} \\ &= \frac{1}{\sqrt{\det\left(2\pi\mathbf{R}_1\right)}} \exp^{-\frac{1}{2}\left(z-\mathbf{H}^T\overline{s_A}\right)^T\ \mathbf{R}_1^{-1}\ \left(z-\mathbf{H}^T\overline{s_A}\right)} && \text{(H.9)} \end{aligned}$$

L'expression du rapport de vraisemblance est donc :

$$\Lambda(z) = \frac{\sqrt{2\pi\det(\mathbf{R}_0)}}{\sqrt{2\pi\det(\mathbf{R}_1)}} \exp^{-\frac{1}{2}\left[(z-\mathbf{H}^T\overline{s_A})^T\mathbf{R}_1^{-1}(z-\mathbf{H}^T\overline{s_A})\ -\ z^T\mathbf{R}_0^{-1}z\right]} \tag{H.10}$$

Cette expression est ensuite comparée à un seuil pour prendre la décision sur la présence ou non du signal. En passant au logarithme népérien, l'expression du test est :

$$-\frac{1}{2}\left[(z-\mathbf{H}^T\overline{s_A})^T\mathbf{R}_1^{-1}(z-\mathbf{H}^T\overline{s_A})\ -\ z^T\mathbf{R}_0^{-1}z\right] \underset{d_0}{\overset{d_1}{\gtrless}} \text{seuil} \tag{H.11}$$

$$(z-\mathbf{H}^T\overline{s_A})^T\mathbf{R}_1^{-1}(z-\mathbf{H}^T\overline{s_A})\ -\ z^T\mathbf{R}_0^{-1}z \underset{d_0}{\overset{d_1}{\lessgtr}} \text{seuil} \tag{H.12}$$

$$z^T\mathbf{R}_1^{-1}z - z^T\mathbf{R}_0^{-1}z - 2\ z^T\mathbf{R}_1^{-1}\mathbf{H}^T\overline{s_A} \underset{d_0}{\overset{d_1}{\lessgtr}} \text{seuil} \tag{H.13}$$

Remplaçant $\mathbf{R}_0$ et $\mathbf{R}_1$ par leurs expressions, il vient :

$$z^T\left[\mathbf{H}^T(\mathbf{A}+\mathbf{B})\mathbf{H}\right]^{-1}z - z^T\left(\mathbf{H}^T\mathbf{B}\mathbf{H}\right)^{-1}z - 2\ z^T\left[\mathbf{H}^T(\mathbf{A}+\mathbf{B})\mathbf{H}\right]^{-1}\mathbf{H}^T\overline{s_A} \underset{d_0}{\overset{d_1}{\lessgtr}} \text{seuil}$$

$$z^T\left[\left(\mathbf{H}^T\mathbf{A}\mathbf{H}+\mathbf{H}^T\mathbf{B}\mathbf{H}\right)^{-1} - \left(\mathbf{H}^T\mathbf{B}\mathbf{H}\right)^{-1}\right]z - 2\ z^T\left[\mathbf{H}^T\mathbf{A}\mathbf{H}+\mathbf{H}^T\mathbf{B}\mathbf{H}\right]^{-1}\mathbf{H}^T\overline{s_A} \underset{d_0}{\overset{d_1}{\lessgtr}} \text{seuil}$$

On utilise les propriétés suivantes déduites de la théorie du FASE [3] :

$$\mathbf{H}^T\mathbf{B}\mathbf{H} = \mathbf{I}_N$$

et

$$\mathbf{H}^T\mathbf{A}\mathbf{H} = \lambda$$

où λ est la matrice diagonale contenant les valeurs propres λ_i. L'expression du test devient alors :

$$z^T\left[(\lambda+\mathbf{I}_N)^{-1} - \mathbf{I}_N\right]z - 2\ z^T\left(\lambda+\mathbf{I}_N\right)^{-1}\mathbf{H}^T\overline{s_A} \underset{d_0}{\overset{d_1}{\lessgtr}} \text{seuil} \tag{H.14}$$

Le test se simplifie et on aboutit à l'expression suivante :

$$\sum_{k=1}^{N} \frac{\lambda_k}{1+\lambda_k} \left(h_k^T x\right)^2 + 2 \sum_{k=1}^{N} \frac{1}{1+\lambda_k} \left(h_k^T x\right) \left(h_k^T \overline{s_A}\right) \underset{d_0}{\overset{d_1}{\gtrless}} \text{seuil} \tag{H.15}$$

Ce test est donc optimal dans le cas gaussien pour une sommation sur tous les N éléments.

Annexe I

Détail des calculs de l'estimateur de la trajectoire du neutrino

Le critère à minimiser est le suivant :

$$\boxed{M(\theta) = \left\| \frac{\tau - S(\theta)}{\sigma} \right\|^2} \tag{I.1}$$

Il faut pour cela résoudre l'équation $\nabla_\theta M(\theta) = 0$. Il vient :

$$\nabla_\theta M(\theta) = -2.\nabla_\theta S(\theta).\,(\tau - S(\theta)) \tag{I.2}$$

Un développement limité à l'ordre 2 de $M(\theta)$ autour du point θ_0 donne :

$$M(\theta) \overset{2}{\cong} M(\theta_0) + (\theta - \theta_0)^T \nabla_\theta M(\theta_0) + \frac{1}{2}(\theta - \theta_0)^T Hess_M(\theta_0)(\theta - \theta_0) \tag{I.3}$$

où $Hess_M(\theta)$ est le hessien de la fonction M, dont les composantes sont

$$[Hess_M(\theta_0)]_{ij} = \left.\frac{\partial^2 M(\theta)}{\partial\theta_i \partial\theta_j}\right|_{\theta=\theta_0}$$

On cherche un point θ_1 solution de $\nabla_\theta M(\theta_1) = 0$. On a

$$\nabla_\theta M(\theta_1) = \nabla_\theta M(\theta_0) + Hess_M(\theta_0).\,(\theta_1 - \theta_0) = 0 \tag{I.4}$$

et donc

$$\theta_1 = \theta_0 - Hess_M(\theta_0)^{-1}.\nabla_\theta M(\theta_0) \tag{I.5}$$

Il est donc possible d'écrire une formule itérative :

$$\theta_{n+1} = \theta_n - Hess_M(\theta_n)^{-1}.\nabla_\theta M(\theta_n) \tag{I.6}$$

Le test d'arrêt de la formule itérative précédente sera obtenu en seuillant la quantité :

$$\frac{\|\theta_{n+1} - \theta_n\|}{\|\theta_{n+1}\|} \tag{I.7}$$

Le choix du point de départ de l'algorithme sera discuté dans les sections suivantes.

Il nous reste à ce stade à calculer les quantités $\nabla_\theta M(\theta)$ et $Hess_M(\theta)$.

I.1 Calcul de $\nabla_\theta M(\theta)$

Pour calculer $\nabla_\theta M(\theta)$ (expression I.2), le seul terme inconnu est $\nabla_\theta S(\theta)$.

I.1.1 Calcul de $\nabla_\theta S(\theta)$

Nous avons :

$$\nabla_\theta S(\theta) = \begin{bmatrix} \nabla_\theta s_1(\theta) & \nabla_\theta s_2(\theta) & \dots & \nabla_\theta s_N(\theta) \end{bmatrix} = \begin{bmatrix} \nabla_A s_1(\theta) & \nabla_A s_2(\theta) & \dots & \nabla_A s_N(\theta) \\ \nabla_\Psi s_1(\theta) & \nabla_\Psi s_2(\theta) & \dots & \nabla_\Psi s_N(\theta) \end{bmatrix} \tag{I.8}$$

Posons maintenant pour le couple i de capteurs indicés C_i^k, $k = 0$ ou 1 :

$$\boxed{t_i^k(\theta) = \frac{1}{c_s}\sqrt{\left\|\vec{AC_i^k}\right\|^2 - \left(\vec{AC_i^k}.\vec{\omega}\right)^2} = \frac{1}{c_s} d_k(\theta)} \tag{I.9}$$

Il vient ainsi $\nabla_\theta s_i(\theta) = \nabla_\theta t_i^1(\theta) - \nabla_\theta t_i^0(\theta)$.

On démontre les relations suivantes :

$$\frac{\partial t_i^k}{\partial x_A}(\theta) = \frac{-(x_i^k - x_A) + \cos\alpha\cos\varphi.\left(\vec{AC_i^k}.\vec{\omega}\right)}{c.d_k(\theta)} \tag{I.10}$$

$$\frac{\partial t_i^k}{\partial y_A}(\theta) = \frac{-(y_i^k - y_A) + \sin\alpha\cos\varphi.\left(\vec{AC_i^k}.\vec{\omega}\right)}{c.d_k(\theta)} \tag{I.11}$$

$$\frac{\partial t_i^k}{\partial \alpha}(\theta) = \frac{\cos\varphi\left[(x_i^k - x_A)\sin\alpha - (y_i^k - y_A)\cos\alpha\right]}{c.d_k(\theta)}\left(\vec{AC_i^k}.\vec{\omega}\right) \tag{I.12}$$

$$\frac{\partial t_i^k}{\partial \varphi}(\theta) = \frac{\sin\varphi\left[(x_i^k - x_A)\cos\alpha + (y_i^k - y_A)\sin\alpha\right] - (z_i^k - z_A)\cos\varphi}{c.d_k(\theta)}\left(\vec{AC_i^k}.\vec{\omega}\right) \tag{I.13}$$

Il est donc possible d'écrire :

$$\nabla_\theta M(\theta) = -2.\nabla_\theta S(\theta).\,(\tau - S(\theta))$$

avec

$$\nabla_\theta S(\theta) = \begin{bmatrix} \nabla_\theta s_1(\theta) & \nabla_\theta s_2(\theta) & \dots & \nabla_\theta s_N(\theta) \end{bmatrix} \tag{I.14}$$

et

$$\nabla_\theta s_i(\theta) = \begin{bmatrix} \frac{\partial t_i^1}{\partial x_A}(\theta) - \frac{\partial t_i^0}{\partial x_A}(\theta) \\ \frac{\partial t_i^1}{\partial y_A}(\theta) - \frac{\partial t_i^0}{\partial y_A}(\theta) \\ \frac{\partial t_i^1}{\partial \alpha}(\theta) - \frac{\partial t_i^0}{\partial \alpha}(\theta) \\ \frac{\partial t_i^1}{\partial \varphi}(\theta) - \frac{\partial t_i^0}{\partial \varphi}(\theta) \end{bmatrix} \tag{I.15}$$

I.2 Calcul de $Hess_M(\theta)$

Le calcul du hessien de la fonction $M(\theta)$ sera simplifié par l'utilisation de l'approximation de Gauss-Newton [19]. Celle-ci consiste à négliger les termes du second ordre dans le calcul de ce hessien. On peut écrire :

$$Hess_M(\theta) = 2.\nabla_\theta M(\theta).\nabla_\theta^T M(\theta) + \text{termes du second ordre} \tag{I.16}$$

Avec l'approximation de Gauss-Newton, le calcul de ce terme se ramène donc à l'équation :

$$\boxed{Hess_M(\theta) = 2\ \nabla_\theta M(\theta).\nabla_\theta^T M(\theta)} \tag{I.17}$$

Cette expression s'évalue directement avec le terme $\nabla_\theta M(\theta)$ calculé dans la section précédente.

I.3 Régions de confiance

Lorsque les mesures sont perturbées par un bruit, il est intéressant de calculer les régions de confiance des estimés. Considérons un vecteur x, réalisation du vecteur aléatoire X dont la loi dépend de $\theta \in \Theta \subset \mathbb{R}$, vecteur contenant les inconnues du problème. Soit également $\mathbf{R}_\theta^\alpha(x)$ un sous-ensemble de Θ. Alors $\mathbf{R}_\theta^\alpha(x)$ est une région de confiance pour θ de niveau α si :

$$P\left\{\theta \in \mathbf{R}_\theta^\alpha(x)\right\} = \alpha \tag{I.18}$$

Dans le cas gaussien, on démontre que :

$$\mathbf{R}_\theta^\alpha(x) = \left\{z \in \mathbb{R}^d, (\theta - z)^T \mathbf{I}_\theta(x)(\theta - z) \leq k_\alpha\right\} \tag{I.19}$$

où k_α est tel que :

$$\int_0^{k_\alpha} p_{\chi_d^2}(t)dt = \alpha \tag{I.20}$$

$\mathbf{I}_\theta(x)$ est la matrice d'information de Fisher et $p_{\chi_d^2}(t)$ est la densité de probabilité d'une variable aléatoire du khi-deux à d degrés de liberté.

Remarque : $\mathbf{R}_\theta^\alpha(x)$ est un ellipsoïde.

I.3.1 Calcul de la matrice d'information de Fisher

Plaçons nous dans le cas où la mesure τ est perturbée par un bruit gaussien, on peut alors écrire :

$$\tau \sim G\left(s(\theta), \mathbf{B}\right) \tag{I.21}$$

où $\mathbf{B}$ est la matrice de variance covariance du bruit.
On démontre que la matrice d'information de Fisher s'écrit :

$$\mathbf{I}_\theta(\tau) = \nabla_\theta M(\theta)\ \mathbf{B}^{-1}\ \nabla_\theta^T M(\theta) \tag{I.22}$$

I.3.2 Région de confiance pour $d = 2$: ellipse de confiance

Lorsque $d = 2$, le tracé de la région de confiance revient au tracé d'une ellipse. Dans le problème qui nous intéresse, il y a quatre paramètres à estimer. Pour se ramener à des tracé de régions facile à mettre en place et à interpréter, on va considérer ces paramètres deux par deux : (x_A, y_A) et (α, φ). Intéressons nous maintenant au cas général à deux inconnues. Dans ce cas, l'équation de la frontière de la région est :

$$\left(\widehat{\theta}_d - z\right)^T \mathbf{I}_{\theta_d}(\tau) \left(\widehat{\theta}_d - z\right) = k_\alpha \tag{I.23}$$

où θ_d est le vecteur contenant les deux inconnues. $\mathbf{I}_{\theta_d}$ est ici une matrice de dimension 2×2. Celle-ci peut être décomposée en fonction de ses vecteurs et valeurs propres :

$$\mathbf{I}_{\theta_d} = \mathbf{H}^T \mathbf{D} \mathbf{H} \tag{I.24}$$

$\mathbf{H}$ est la matrice des vecteurs propres V_1 et V_2 de $\mathbf{I}_{\theta_d}$ et D est une matrice diagonale contenant les valeurs propres λ_1 et λ_2 associées.
Posons maintenant :

$$\zeta(z) \triangleq \sqrt{\mathbf{D}}\ \mathbf{H}\ \left(\widehat{\theta}_d - z\right) \tag{I.25}$$

D'aprés l'équation I.23, on remarque qu'on peut écrire $\|\zeta(z)\|^2 = k_\alpha = \text{cte}$. L'ensemble des $\zeta(z)$ est donc un cercle de rayon $\sqrt{k_\alpha}$ dont l'équation paramétrique est :

$$\begin{cases} \zeta_1(\gamma) = \sqrt{k_\alpha}\ \cos\gamma \\ \zeta_2(\gamma) = \sqrt{k_\alpha}\ \sin\gamma \end{cases},\ \gamma \in [0, 2\pi[\tag{I.26}$$

Il est possible d'exprimer $\widehat{\theta}_d - z$ en fonction de $\zeta(z)$:

$$\begin{aligned} \widehat{\theta}_d - z &= \mathbf{H}^T(\sqrt{\mathbf{D}})^{-1}\zeta(z) && \text{(I.27)} \\ &= \mathbf{H}^T(\sqrt{\mathbf{D}})^{-1}\begin{pmatrix} \sqrt{k_\alpha}\ \cos\gamma \\ \sqrt{k_\alpha}\ \sin\gamma \end{pmatrix},\ \gamma \in [0, 2\pi[&& \text{(I.28)} \end{aligned}$$

On peut maintenant exprimer z en fonction de γ :

$$\begin{aligned} z(\gamma) &= \widehat{\theta}_d - \sqrt{k_\alpha}\ [V1, V2]\begin{bmatrix} \frac{1}{\sqrt{\lambda_1}} & 0 \\ 0 & \frac{1}{\sqrt{\lambda_2}} \end{bmatrix}\begin{bmatrix} \cos\gamma \\ \sin\gamma \end{bmatrix} && \text{(I.29)} \\ &= \widehat{\theta}_d - [V1, V2]\begin{bmatrix} \sqrt{\frac{k_\alpha}{\lambda_1}}\cos\gamma \\ \sqrt{\frac{k_\alpha}{\lambda_2}}\sin\gamma \end{bmatrix}, \gamma \in [0, 2\pi[&& \text{(I.30)} \end{aligned}$$

Cette équation est bien celle d'une ellipse. On démontre que pour avoir une ellipse de confiance à 90 %, il faut $k_\alpha = 4.601$.
Dans la section simulations et résultats, il sera donc possible de tracer les ellipses de confiance associées aux couples de paramètres suivants : (x_A, y_A) et (α, φ).

I.3.3 Borne de Cramer-Rao

Pour quantifier les performances de l'estimateur mis en place, il faut calculer la borne de Cramer-Rao [21] notée $\mathbf{B}_\theta$. Dans le cas d'estimateurs sans biais, on démontre que l'expression de cette borne est :

$$\mathbf{B}_\theta(\tau) = \mathbf{I}_\theta^{-1}(\tau) \tag{I.31}$$

où $\mathbf{I}_\theta(\tau)$ est la matrice d'information de Fisher calculée précédemment.

Annexe J

Protocole et simulations de la détection appliquée au cas réel - Cas monobruit

Cette annexe décrit les protocoles de simulations mis en place pour traiter le cas "pseudo-réel". Le signal acoustique est synthétisé à partir du modèle décrit par Valentin Niess [23]. Le bruit de mer est issu de campagnes de mesures effectuées au large de l'île du Levant, en mer Méditerranée au cours de l'année 2001 par des membres de l'équipe GESSY du laboratoire SIS.

Nous nous trouvons dans le cas classique de détection suivant : un signal d'intérêt noyé dans un bruit ambiant. L'algorithme utilisé est le suivant :

1 signal - 1 bruit

Calcul des courbes COR

- Définition des paramètres de la cascade d'énergie.
- Générations de N_r réalisations du signal acoustique.
- Génération de $2N_r$ réalisations du bruit de mer.
- Evaluation des matrices de covariance du signal et du bruit.
- Evaluation des filtres servant à la détection (dépend de la méthode considérée).
- Génération des signaux de test.
- Filtrage des signaux de test.
- Calcul des probabilités P_d et P_{fa}.
- Affichage des courbes COR.

Chaque étape est maintenant décrite.

J.1 Paramètres des simulations

Sauf mention contraire, les valeurs numériques suivantes ont été utilisées pour toutes les simulations :

- ⋆ Energie : $E = 10^{18}$ eV.
- ⋆ Longueur efficace : $L_{eff} = 10$m.
- ⋆ Distance maximale de détection : $d_{max} = 500$m.
- ⋆ Fréquence d'échantillonnage : $Fe = 1$MHz.
- ⋆ Nombre de réalisations : $Nr = 1000$.
- ⋆ Longueur d'une réalisation : $N = 64$.

J.2 Génération du signal Antares

Le modèle utilisé pour représenter est celui décrit par l'équation 1.20. L'idée est de considérer ce signal acoustique comme une réalisation d'un signal aléatoire défini par :

$$S_A(t, D) = k.E.g_L(D).\mathrm{Re}\left(H_D * s_0\right)(t) \tag{J.1}$$

Dans l'expression de S_A, seul la variable D est aléatoire et de densité de probabilité théorique connue. Tout le reste est déterministe. Il suffit donc, pour générer différentes réalisations du signal acoustique, de générer une réalisation de D et d'en déduire la réalisation de S_A correspondante.

Nous avons ensuite décidé de filtrer ces signaux dans une bande de fréquence dite "bande utile". On utilise pour cela un passe-bas et un passe-haut à des fréquences respectives de 7 kHz et 150 kHz. Ces valeurs sont déduites de l'occupation spectrale maximale que peut occuper le signal acoustique. Celle-ci est obtenue lorsque le signal a subit le moins d'atténuation, c'est à dire pour la distance minimale $d_{min} = 100$m. On affiche sur la figure J.1 la bande passante du signal acoustique pris à la distance d_{min} de la cascade d'énergie.

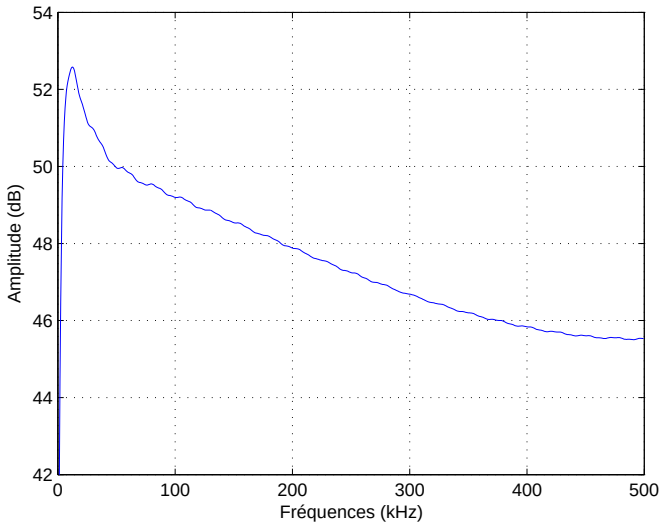

FIG. J.1 – Bande passante du signal acoustique pris à une distance de 100m.

Lors des simulations dans le cas pseudo-réel, les signaux ont été générés de la manière suivante :

Génération des signaux Antares

- Générations de N_r réalisations de la variable aléatoire D représentant la distance cascade/hydrophone.
- Génération des N_r signaux acoustiques correspondants.
- Rééchantillonage des signaux à la fréquence de travail.
- Filtrage de chaque signal dans la bande passante utile (BP = 7kHz - 150 kHz).
- Récupération de la partie utile du signal sur N points.
- Rangement de ces réalisations dans une matrice.

J.3 Génération du bruit de mer

Le bruit de mer "réel" est issu de campagnes de mesures effectuées en mer Méditerranée au large de l'île du Levant en 2001. Ce bruit est filtré dans la bande passante utile de la même manière que le signal acoustique.

Générations du bruit de mer

- Récupération de bruits de mer réels issus de campagnes de mesures.
- Rééchantillonage à la fréquence de travail.
- Filtrage dans la bande passante utile.
- Découpage en réalisations de N points.
- rangement de ces réalisations dans une matrice.

Dans certaines situations, d'autres bruits perturbateurs sont utilisés. Le détail de chacun d'entre eux est donné dans la section correspondante.

J.4 Notion de rapport signal à bruit

La notion de rapport signal à bruit (RSB) est fondamentale dans le problème de détection qui nous intéresse. Si l'on considère un signal s échantillonné sur N points, sa puissance est $P_s = \mathsf{E}(s^T s)$. Le rapport signal à bruit est le rapport de la puissance du signal sur la puissance du bruit : $RSB = \frac{P_s}{P_b}$. Celui-ci est souvent exprimé en dB : $RSB_{dB} = 10 \log_{10}(RSB)$
Pour les tests de détection, il a été décidé de travailler avec des RSB moyens.
On définit le RSB moyen comme le rapport de la puissance moyenne du signal sur la puissance moyenne du bruit.
La puissance moyenne du signal s est :

$$P_s^{moy} = \mathsf{E}\left(s^T s\right) = \frac{1}{N_r} \sum_{i=1}^{N_r} s_i^T s_i$$

où s_i est la réalisation n°i de s.
La puissance moyenne du bruit b est :

$$P_b^{moy} = \mathsf{E}\left(b^T b\right) = \frac{1}{N_r} \sum_{i=1}^{N_r} b_i^T b_i$$

où b_i est la réalisation n°i de b.

J.5 Calcul des matrices de covariance

Pour un signal aléatoire s échantillonné de dimension N, la matrice de variance/covariance $\mathbf{M}_s$ de s est calculée grâce à l'estimateur sans biais suivant :

$$\mathbf{M}_s = \mathsf{E}\left(ss^T\right) = \frac{1}{N_r - 1} \sum_{i=1}^{N_r} s_i s_i^T$$

où s_i est la réalisation n°i de s.

J.6 Les signaux de test

Pour calculer les courbes COR, des signaux de test sont générés. Ces signaux couvrent toutes les hypothèses que l'on peut rencontrer. Dans le cas d'un signal perturbé par un bruit, il y a 2 hypothèses : dans la première, la mesure est constituée par le bruit de mer et dans la seconde par la somme du signal acoustique et du bruit de mer. Le signal de test est décrit sur la figure J.2.

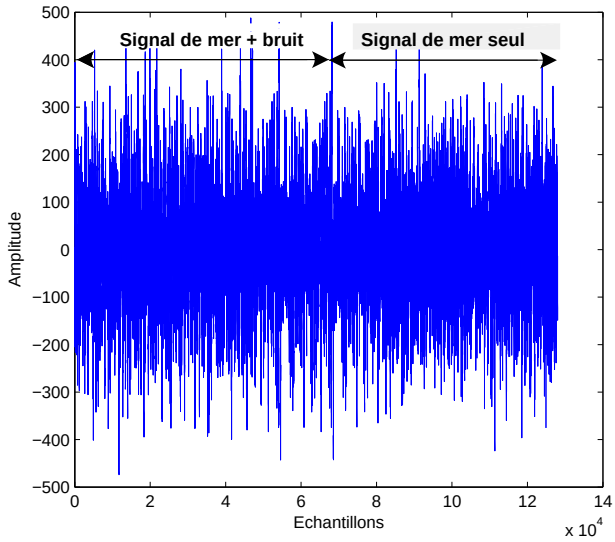

FIG. J.2 – Signal de test pour un signal et un bruit.

J.7 Test de détection

Le test de détection dépend de la méthode choisie. Pour chaque méthode, nous avons défini une fonctionnelle de test. C'est le résultat de cette fonctionnelle qui est comparé à un seuil prédéfini et le résultat de cette comparaison nous indique la décision à prendre quand à la présence du signal d'intérêt.

J.8 Calcul des P_d et P_{fa} et affichage des courbes COR

Connaissant la structure du signal de test et pour un seuil donné, on évalue le couple $\{P_d, P_{fa}\}$ en comptant le nombre de fois où le signal dépasse la valeur du seuil. Il suffit ensuite de faire varier le seuil entre 0 et le maximum du signal filtré pour balayer toutes les valeurs possibles de P_{fa}. Les résultats sont affichés sous forme de courbe COR.

Annexe K

Calcul de l'intercorrélation de deux signaux filtrés par le FASE

Cette annexe décrit la méthode utilisée pour calculer l'intercorrélation entre deux signaux filtrés par le FASE. Nous disposons de deux signaux échantillonnés représentés par des vecteurs X et Y de longueur N_T. Nous décidons de travailler dans un espace vectoriel E de dimension N ($N < N_T$).

K.1 Filtrage par le FASE

On note $\Psi = \{\psi_1, \ldots, \psi_p\}$ la matrice contenant tous les vecteurs issus du FASE. Ces p vecteurs de dimension N engendrent un sous-espace vectoriel E_p^* (dit *optimal*) de dimension $p < N$. Le filtrage temporel se fait via une fenêtre glissante de largeur N se déplaçant d'un échantillon à chaque itération. Si cette fenêtre est notée $X^{(N)}$, alors son filtré par le FASE de dimension p est $X_p^{(N)}$ et on a :

$$X_p^{(N)} \quad = \quad \Psi^T \, . \, X^{(N)}$$

Les étapes de filtrage sont résumées sur la figure K.1.

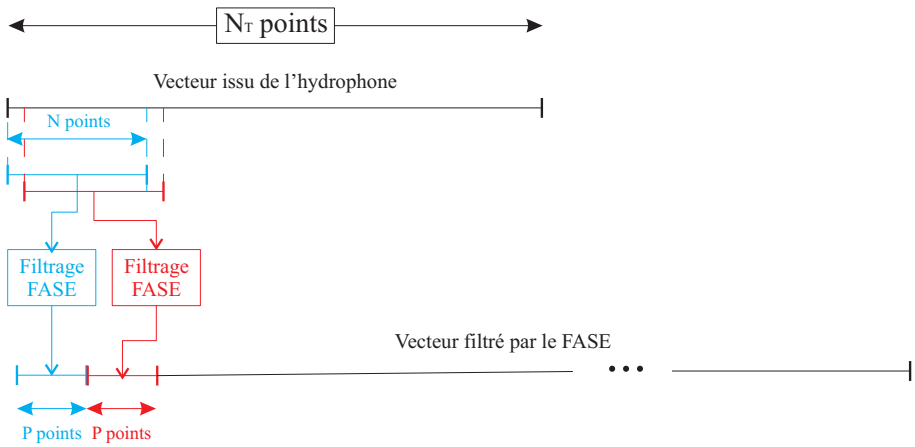

FIG. K.1 – Illustration du filtrage temporel par fenêtre glissante associé au filtrage par le FASE.

K.2 Calcul de l'intercorrélation

Nous disposons maintenant de deux signaux de longueur $(N_T - N) * p$. Le calcul de l'intercorrélation $R(m)$ entre ces deux signaux doit se faire en passant de bloc en blocs (de taille p) se fait suivant la relation :

$$R(m) = \sum_{n=1}^{N_T - N} X_p\Big((n-1)p + 1 : np\Big)\; Y_p\Big((n-1)p + 1 + m : np + m\Big)$$

Glossaire

A	:	Un point de la cascade d'énergie.
$\mathbf{A}$	:	Matrice de variance-covariance du signal acoustique.
$\mathbf{B}$	:	Matrice de variance-covariance du bruit de mer.
C_s	:	[m.s^{-1}] Célérité du son dans l'eau de mer.
d_H	:	[m] Distance entre deux hydrophones dans un couple.
$d_{\min}$	:	[m] Distance minimale de détection.
E	:	[eV] Energie de la cascade.
$\mathsf{E}(.)$	:	Espérance mathématique.
F_e	:	[Hz] Fréquence d'échantillonnage.
$F(a_1, a_2)$	:	Loi de Fisher Snédécor de paramètres a_1 et a_2.
$g_L(d)$	:	Atténuation radiale du signal acoustique dont le front d'onde a parcouru une distance d. Modèle de Niess.
$g_z(z)$	:	Atténuation en altitude le long de la cascade d'énergie. Modèle de Niess.
$h_d(t)$	:	Filtre temporel caractérisant l'atténuation du milieu marin pour un front d'onde ayant parcouru une distance d. Modèle de Niess.
$\mathbf{I}_n$	:	Matrice identité de taille (n, n).
L_{eff}	:	[m] Longueur efficace de la cascade d'énergie.
$\mathbf{M}^T$	:	Transposée de la matrice $\mathbf{M}$.
$\mathbf{M}^{-1}$	:	Inverse de la matrice $\mathbf{M}$.
$\mathbf{M}_d$	:	Matrice fixe permettant l'approximation polynômiale du signal acoustique.

N	:	Taille d'une réalisation de signal acoustique ou de bruit.
N_c	:	Nombre de couples d'hydrophones dans la structure.
N_p	:	Nombre de plans de capteurs horizontaux dans la structure.
N_r	:	Nombre de réalisations utilisées dans les simulations de détection.
$P(r,t)$	:	[Pa] Expression de la pression acoustique prise à une distance r de la cascade d'énergie et à l'instant t. Modèle d'Askariyan.
P_b	:	Puissance du bruit b.
P_d	:	Probabilité de détection.
P_{fa}	:	Probabilité de fausse alarme.
P_s	:	Puissance du signal s.
RSB	:	Rapport signal-à-bruit, RSB $= \frac{P_s}{P_b}$.
S_a	:	Signal acoustique échantillonné.
$s_{\text{Askaryian}}(r,t)$	:	[Pa] Expression du signal acoustique pour le modèle d'Askaryian pour un hydrophone placé à une distance r et à l'instant t.
$s_{\text{Niess}}(r,t)$	:	[Pa] Expression du signal acoustique pour le modèle de Niess pour un hydrophone placé à une distance r et à l'instant t.
$s_0(t)$	:	[Pa] Signal acoustique de référence calculé à la distance minimale de détection d_{min}. Modèle de Niess.
SEV	:	Sous-espace vectoriel.
$tr(\mathbf{M})$	:	Trace de la matrice $\mathbf{M}$.
T_a	:	[s] Durée du signal acoustique.
T_d	:	[s] Date à laquelle le front d'onde atteint l'hydrophone.
$TDOA$	:	Différence des temps d'arrivée sur deux hydrophones (Time Difference Of Arrival).
α	:	Angle d'azimut de l'orientation de la cascade d'énergie.
φ	:	Angle de site de l'orientation de la cascade d'énergie.

Bibliographie

[1] Lefèvre-Balleydier A. *Petit atlas des mers et océans.* Petite encyclopédie Larousse, 2003.

[2] Migdal A.B. Bremsstrahlung and pair production in condensed media at high energies. *Phys. Rev.*, 103(6), september 1956. pp.267-278.

[3] Borloz B. *Estimation, détection, classification, par maximisation du rapport signal-à-bruit : le filtrage adapté stochastique sous contrainte.* PhD thesis, Université du Sud Toulon Var, UFR Sciences et techniques, juin 2005.

[4] Xerri B. Borloz B. Filtre adapté stochastique sous contrainte. *GRETSI Symposium 2001, 12-15 Sept. 2001*, 2001. Toulouse.

[5] Van Der Waerden B.L. *Statistique mathématique.* Dunod, 1967. pp. 238.

[6] Xerri B. Cavassilas J.F. Extension de la notion de filtre adapté. Contribution à la détection de signaux courts en présence de termes perturbateurs. *Revue Traitement du signal*, 10(3), 1993. pp.215-221.

[7] Garrison G.R. François R.E. Sound absorption based on ocean measurements. Part ii : Boric acid contribution and equation for total absorption. *The Journal of the Acoustical Society of America*, 72(6), décembre 1982. pp. 1879-1890.

[8] Askariyan G.A. Hydrodynamical emission of tracks of ionising particles in stable liquids. *Atomnaya Energiya*, 3(8), 1957.

[9] Askariyan G.A., Dolgoshein B.A., Kalinovsky A.N., and Mokhov N.V. Acoustic detection of high energy particle showers in water. *Nuclear instruments and methods*, 164, 1979. pp.267-278.

[10] Bennett G.H. *Acoustic transient TDOA estimation and discrimination.* PhD thesis, Naval Postgraduate School, Monterey, California, september 2000.

[11] Poor H.V. *An introduction to signal detection and estimation.* Springer-Verlag, 1988.

[12] Bar-Shalom Y. Jauffret C. Target motion analysis in the presence of false alarms. *J. Acoust. Soc. Am*, 1(3), 1993. pp. 1372-1375.

[13] Dennis J.E. A user's guide to nonlinear optimization algorithms. *Proceedings of the IEEE*, 72(12), december 1984.

[14] Learned J.G. Acoustic radiation by charged atomic particles in liquid : an analysis. *Phys. Rev D*, 19(11), june 1979.

[15] Moré J.J. The levenberg-marquardt algorithm : implementation and theory. *Numerical Analysis, G.A. Watson Ed., Lecture notes in math. 630, New York*, 1977. pp. 105-116.

[16] Fukunaga K. *Introduction to statistical pattern recognition.* Computer Science and Scientific Computing ISBN 0-12-269851-7, 1990.

[17] Carter G.C. Knapp C.H. The generalized correlation method for estimation of time delay. *IEEE Transaction on acoustic speech and signal processing*, ASSP24(4), 1976. pp. 320-326.

[18] Liebermann L. Sound propagation in chemically active media. *Phys. Rev.*, 76(10), novembre 1949. pp. 1520-1524.

[19] Argonne National Labs. Newton's method. *Mathematical and computer science rep.*, ANL-82-8, 1982.

[20] Dedenko L.G. and al. Prospects for radio-wave and acoustic detection of ultra and superhigh energy cosmic neutrinos. *Phys. Part. Nuclei*, 29(3), may-june 1998.

[21] Kay S.M. *Fundamentals of statistical signal processing.* PTR, 1993.

[22] UNESCO. Algorithms for computation of fundamental properties of seawater. *UNESCO technical papers in marine sciences*, 44, 1983.

[23] Niess V. *Détection acoustique sous-marine de neutrinos de ultra haute énergie dans le cadre de l'expérience Antares.* PhD thesis, Université de la Méditerranée, Aix-Marseille II, UFR de physique, september 2005.

Résumé

Le sujet de cette thèse s'inscrit dans le cadre du projet international Antares dont l'objectif est la construction d'un télescope à neutrinos situé dans un environnement marin au large de Toulon. A de fortes profondeurs, un neutrino a d'autant plus de chance de rentrer en collision avec une molécule d'eau, générant ainsi une réaction en chaîne générant un flash lumineux et une onde sonore. L'objectif de cette thèse est d'étudier cette onde sonore en vue de développer un système capable de détecter le front d'onde correspondant et d'estimer la direction originelle du neutrino

Dans un premier temps, l'étude se porte sur le signal acoustique. Deux descriptions issues de la littérature et de récents travaux effectués au CPPM sont confrontées et aboutissent à une modélisation mathématique du signal et du front d'onde.

Dans un second temps, plusieurs méthodes de détection sont étudiées, de la plus classique (étude du rapport de vraisemblance) à des méthodes plus récentes (filtrage adapté, classification, etc.). La comparaison expérimentale en situation semi réelle de celles-ci aboutit au choix de la méthode de détection suivante : le FASE (Filtrage Adapté Stochastique Etendu).

Enfin, la position et la direction du neutrino sont estimées par un algorithme dérivé de Gauss-Newton. Cet estimateur se base sur la modélisation du déplacement du front d'onde acoustique et sur les informations temporelles de détection fournies par les hydrophones du télescope. De nombreuses configurations sont testées et les performances du système sont évaluées.

Une structure d'hydrophone est proposée et une simulation dite "globale" finalise cette thèse. Dans celle-ci, les étapes de détection et d'estimation sont basées sur les résultats obtenus précédemment. Les bruits de mer sont des bruits réels issus de campagnes de mesure et les résultats obtenus valident les travaux de cette thèse.