



HAL
open science

Conception optimisée d'antenne pour de la localisation passive de sources acoustiques

Milan Courcoux-Caro

► **To cite this version:**

Milan Courcoux-Caro. Conception optimisée d'antenne pour de la localisation passive de sources acoustiques. Acoustique [physics.class-ph]. ENSTA Bretagne - École nationale supérieure de techniques avancées Bretagne, 2022. Français. NNT : 2022ENTA0001 . tel-03848121

HAL Id: tel-03848121

<https://theses.hal.science/tel-03848121>

Submitted on 10 Nov 2022

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.

THÈSE DE DOCTORAT DE

l'École Nationale Supérieure des Techniques Avancées
BRETAGNE - ENSTA BRETAGNE

ÉCOLE DOCTORALE N° 601
*Mathématiques et Sciences et Technologies
de l'Information et de la Communication*
Spécialité : *Signal, Image, Vision*

Par

Milan COURCOUX-CARO

Conception optimisée d'antenne pour de la localisation passive de sources acoustiques

Algorithme de sélection séquentielle bayésienne orientée par les données

Thèse présentée et soutenue à Brest, le 4 Avril 2022
Unité de recherche : Lab-STICC UMR CNRS 6285

Rapporteurs :

Régis MARCHIANO
Pascal LARZABAL

Professeur des universités, Sorbonne Université
Professeur des universités, ENS Paris-Saclay

Composition du Jury :

Président :	Emanuel RADOI	Professeur des universités, Université de Bretagne Occidentale
Examineurs :	Mai Quyen PHAM Dominique FATTACCIOLI Manuel MELON	Enseignante-Chercheuse, IMT Atlantique Expert, DGA Techniques Navales Professeur des universités, Université du Maine
Dir. de thèse :	Alexandre BAUSSARD	Professeur des universités, Université de Technologie de Troyes
Encadrant :	Charles VANWYNSBERGHE	Enseignant-Chercheur, ISEN Nantes
Invité :	Cédric HERZET	Chercheur, INRIA Rennes

RÉSUMÉ

L'objectif de ces travaux de thèse est de proposer une approche permettant de concevoir une antenne pour de la localisation de sources acoustiques aériennes ou sous-marines. Dans un premier temps, nous rappelons le principe de la propagation des ondes émises par les sources dans un milieu. Puis nous présentons les différentes méthodes permettant d'estimer les positions des sources à partir des mesures acquises par les capteurs et du modèle de propagation. Les performances de ces méthodes sont directement liées à la géométrie de l'antenne et donc à la position des capteurs. De nombreux travaux se sont donc intéressés à la conception d'antenne, et plus précisément sur la sélection de positions de capteurs.

Dans ce manuscrit, nous proposons une méthode de sélection séquentielle bayésienne orientée par les données. Cette approche prend en compte l'information contenue dans les mesures acquises précédemment par l'antenne, afin de sélectionner les futurs capteurs. L'application de cette approche est innovante dans un contexte de localisation de sources acoustiques.

Dans un deuxième temps, des premiers résultats obtenus sur données synthétiques sont prometteurs quant à l'application de cette approche dans un cadre de localisation de sources acoustiques sous-marines. Les analyses fournies dans ce manuscrit permettent de juger la pertinence de cette approche dans le contexte testé, ainsi que d'évaluer et de comparer la performance de cette conception d'antenne par rapport à la littérature.

Nous appliquons aussi cette approche sur des données réelles provenant d'une expérience réalisée lors de la thèse. Le milieu de propagation est cette fois-ci aérien et réverbérant avec un grand nombre de capteurs disponibles. Ces données réelles permettent l'évaluation des performances de notre approche et la proposition d'une nouvelle approche de conception d'antenne. Celle-ci est conçue à partir de données synthétiques et est appliquée ensuite sur les données réelles.

Mots clés : localisation de sources acoustiques, aérien, sous-marin, conception d'antenne, données expérimentales, sélection séquentielle, approches bayésiennes.

TABLE DES MATIÈRES

Page de garde	i
Résumé	iii
Table des matières	viii
Liste des acronymes	ix
Notations	xi
Nomenclature	xiii
Table des figures	xx
Liste des tableaux	xxi
Introduction	1
I État de l’art et présentation des algorithmes : Localisation de sources et sélection de positions de capteurs	4
1 LA LOCALISATION DE SOURCES	5
1.1 Introduction	5
1.2 Propagation de l’onde	8
1.2.1 Généralités sur la scène acoustique	9
1.2.2 Propagation en champ proche : espace libre	10
1.2.3 Propagation en champ proche : milieu réverbérant	11
1.2.4 Propagation en champ lointain	12
1.3 Ecriture du problème de localisation de sources	13
1.3.1 Problème direct	14
1.3.2 Problème inverse	15
1.4 Méthodes classiques	16
1.4.1 Formation de voies	16

1.4.2	Problèmes liés au dimensionnement d'antenne	17
1.4.3	Méthodes de super résolution classiques	19
1.5	Méthodes parcimonieuses	19
1.5.1	Approches gloutonnes	20
1.5.2	Approches par relaxation convexe	21
1.6	Approches Bayésiennes	22
1.6.1	Maximum de vraisemblance	23
1.6.2	Maximum a posteriori	23
1.6.3	Inférence bayésienne parcimonieuse (SBI)	24
1.7	Extension des problèmes au cas de mesures multi-snapshots	26
1.8	Critères d'évaluation	28
1.9	Conclusion	32
2	LA SELECTION DE CAPTEURS	35
2.1	Introduction	35
2.2	Ecriture du problème	38
2.3	Les méthodes aléatoires	38
2.3.1	Le tirage aléatoire	38
2.3.2	Algorithme génétique	39
2.4	Les critères d'optimisation par conception optimale	44
2.4.1	L'information de Fisher	45
2.4.2	Matrice de covariance	48
2.5	Sélection séquentielle de capteurs	50
2.5.1	Sélection D-optimale	51
2.5.2	Sélection A-optimale	53
2.6	Sélection orientée par les données	54
2.6.1	Sélection séquentielle bayésienne orientée par les données	54
2.6.2	Discussion	55
2.7	Conclusions	56
II	Sélection séquentielle bayésienne orientée par les données dans un contexte simulé de localisation de sources acoustiques sous-marines	57
3	Sélection séquentielle bayésienne orientée par les données pour la localisation passive de sources acoustiques sous-marines	59
3.1	Introduction	60
3.2	Illustration par des scénarios élémentaires	61

3.2.1	Configuration des scénarios	61
3.2.2	Illustration du critère en fonction du nombre de capteurs sélectionnés	61
3.2.3	Illustration du critère en fonction de la fréquence	63
3.2.4	Illustration du critère en fonction d'autres paramètres	64
3.3	Généralisation des résultats par simulations de Monte Carlo	67
3.3.1	Configuration des simulations de Monte Carlo	68
3.3.2	Relation entre la valeur du critère de sélection et performances de localisation	69
3.3.3	Analyse par rapport au bruit	71
3.4	Comparatif des résultats des simulations de Monte Carlo par rapport à l'état de l'art	72
3.4.1	Présentation de l'étude et paramètres de la simulation	72
3.4.2	Présentation des antennes de l'état de l'art	73
3.4.3	Résultats	74
3.5	Études paramétriques annexes	78
3.5.1	Comparaison d'estimateurs de positions de sources	78
3.5.2	Comparaison des critères d'optimalités	79
3.5.3	Extension multi-snapshots	80
3.6	Conclusion	81

III Application en localisation de sources aériennes en salle réverbérante à partir d'expérimentations 84

4	Mise en place d'une expérience 85
4.1	Introduction 85
4.2	Présentation de la salle MODO 86
4.2.1	Histoire du projet 86
4.2.2	Le microphone MEMS 87
4.2.3	Description de la salle 88
4.2.4	Géométrie de l'antenne 89
4.2.5	Système d'acquisition 89
4.2.6	Calibration géométrique de l'antenne 90
4.3	Expérience 91
4.3.1	Source monopolaire 91
4.3.2	Positionnement de la source 93
4.3.3	Les signaux émis 95
4.3.4	Température et célérité du son 96
4.3.5	Pré-traitements 96

TABLE DES MATIÈRES

4.4	Localisation de sources acoustiques en milieu réverbérant	97
4.4.1	Estimation de l'ordre de réflexion	97
4.4.2	Estimation des coefficients de réflexion	101
4.5	Conclusion	101
5	Application de la sélection de capteurs aux données expérimentales	103
5.1	Introduction	103
5.2	Évaluation des performances en fonction du nombre de sources	104
5.3	Impact de l'ordre de réflexion	106
5.3.1	Données simulées	107
5.3.2	Données expérimentales	109
5.4	Comparaison avec d'autres méthodes	111
5.5	Etudes complémentaires	115
5.5.1	Extension multi-snapshots	115
5.5.2	Choix du critère d'optimalité	116
5.6	Applications en conception d'antenne	118
5.7	Conclusion	121
	Conclusion	123
	Annexes	a
	Liens entre les critères d'optimalité	b
	Divergence de Kullback-Leibler	b
	L'entropie de Shannon	c
	Acquisitions réalisées dans la salle MODO	e
	Source dipolaire	e
	Sources monopolaires cohérentes	f
	Calibration géométrique	f
	Aspirateur	g
	Cocktail party	g
	Liste des publications	I
	Références	III

LISTE DES ACRONYMES

BP	<i>Basis Pursuit</i>
BPDN	<i>Basis Pursuit De-Noising</i>
BF	<i>BeamForming / Formation de Voies</i>
EM	<i>Espérance-Maximisation</i>
ESPRIT	<i>Estimation of Signal Parameters via Rotational Invariance Techniques</i>
FISTA	<i>Fast Iterative Shrinkage Thresholding Algorithm</i>
MAP	<i>Maximum A Posteriori</i>
MEMS	<i>MicroElectroMechanical Systems</i>
ML	<i>Maximum Likelihood / Maximum de Vraisemblance</i>
MODO	<i>les Murs Ont Des Oreilles</i>
MP	<i>Matching Pursuit</i>
MSL	<i>Maximum Sidelobes Level</i>
MSSBI	<i>Multiple Snapshots Sparse Bayesian Inference</i>
MUSIC	<i>MUltiple SIgnal Classification</i>
MVDR	<i>Minimum Variance Distortionless Response</i>
OMP	<i>Orthogonal Matching Pursuit</i>
SBI	<i>Sparse Bayesian Inference / Inférence Bayésienne Parcimonieuse</i>
SNR	<i>ratio signal à bruit</i>
SSBOD	<i>Sélection Séquentielle Bayésienne Orientée par les Données</i>

NOTATIONS

Unités

dB décibel

Vecteurs et Matrices

x scalaire

\mathbf{x} vecteur

\mathbf{X} matrice

$\tilde{\mathbf{x}}$ estimation du vecteur \mathbf{x}

\mathcal{M} ensemble

$\mathbf{x}_{\mathcal{M}}$ vecteur \mathbf{x} appartenant ou calculé à partir de l'ensemble \mathcal{M}

$\mathbf{x}_{\mathcal{M}_i}$ i -ième vecteur \mathbf{x} appartenant ou calculé à partir de l'ensemble \mathcal{M}

$\mathbf{X}_{\mathcal{M},\mathcal{G}}$ matrice \mathbf{X} appartenant ou calculé à partir des ensembles \mathcal{M} et \mathcal{G}

\mathbf{x}^T transposée de \mathbf{x}

\mathbf{x}^H hermitien de \mathbf{x} (transposée conjuguée)

$\Re(\mathbf{x})$ partie réelle du vecteur \mathbf{x}

$\mathbf{X} \odot \mathbf{Y}$ produit de Schur-Hadamard de \mathbf{X} et \mathbf{Y}

\mathbf{I}_N matrice identité de dimensions $N \times N$

\mathbf{X}^{-1} inverse de \mathbf{X}

$|x|$ valeur absolue de x

$|\mathbf{x}|$ module de \mathbf{x} élément par élément

$|\mathbf{X}|$ module de \mathbf{X} élément par élément

$\|\mathbf{x}\|$ norme l_2 de \mathbf{x}

$\|\mathbf{x}\|_p$ norme l_p de \mathbf{x}

$\|\mathbf{x}\|_F$ norme de frobenius de \mathbf{x}

Probabilités et statistiques

$\mathcal{U}([0, 1])$	distribution uniforme entre 0 et 1
$\mathcal{CN}(\mu, \sigma^2)$	distribution complexe circulaire gaussienne de moyenne μ et de variance σ^2
$\Gamma(c, d)$	distribution gamma avec c le paramètre de forme et d le paramètre d'intensité
$\mathcal{P}(x y)$	probabilité de x sachant y
$E[\mathbf{x}]$	espérance de \mathbf{x}
$\text{tr}(\mathbf{X})$	trace de \mathbf{X}
$\det(\mathbf{X})$	déterminant de \mathbf{X}
$\text{VP}_{\min}(\mathbf{X})$	plus petite valeur propre de \mathbf{X}
$\text{VP}_{\max}(\mathbf{X})$	plus grande valeur propre de \mathbf{X}
$\frac{\partial \mathbf{x}}{\partial \theta}$	dérivée partielle de \mathbf{x} par rapport à θ
$\text{diag}(\mathbf{x})$	matrice diagonale dont les termes diagonaux sont égaux à \mathbf{x}
$\sum_{i=1}^N x_i$	somme des N scalaires x_i
$\prod_{i=1}^N x_i$	somme des N scalaires x_i

Opérateurs de relations

$x > y$	x supérieur à y
$x \geq y$	x supérieur ou égal à y
$x \gg y$	x très grand devant y
$\mathbf{X} \succeq 0$	\mathbf{X} est une matrice semi-définie positive
$\mathbf{X} \succeq \mathbf{Y}$	\mathbf{X} est supérieur à \mathbf{Y} dans le sens des matrices semi-définies positives
$\mathcal{M} \subset \mathcal{G}$	sous-ensemble \mathcal{M} de \mathcal{G}
$\mathbf{x} \in \mathcal{M}$	vecteur \mathbf{x} appartenant à l'ensemble \mathcal{M}
$\mathbf{x} \notin \mathcal{G}$	vecteur \mathbf{x} n'appartenant pas à l'ensemble \mathcal{G}
$\mathcal{M} \cup \mathcal{G}$	union des deux ensembles \mathcal{M} et \mathcal{G}
$\mathcal{M} \cap \mathcal{G}$	intersection des deux ensembles \mathcal{M} et \mathcal{G}
$x \propto y$	proportionnalité entre x et y

Fonctions

\mathbb{R}	ensemble des nombres réels
\mathbb{C}	ensemble des nombres complexes
j	nombre complexe
cos	cosinus
sin	sinus
exp	exponentielle
log	logarithme

NOMENCLATURE

Paramètres acoustiques

c	Célérité du son
F	Fréquence de la source
λ	Longueur d'onde de la source
$Temp$	Température du milieu de propagation

Géométrie de l'espace

X	Axe X de l'espace
Y	Axe Y de l'espace
Z	Axe Z de l'espace
θ	Angle theta
ϕ	Angle Phi

Propagation du son

N_M	Nombre de capteurs disponibles
$\mathbf{p}_{\mathcal{M}_i}$	Position d'un capteur disponible i en 3 dimensions
$x_{\mathcal{M}_i}$	Position d'un capteur disponible i selon l'axe X
$y_{\mathcal{M}_i}$	Position d'un capteur disponible i selon l'axe Y
$z_{\mathcal{M}_i}$	Position d'un capteur disponible i selon l'axe Z
\mathcal{M}	Ensemble contenant les positions de capteurs disponibles
N_S	Nombre de sources
$\mathbf{p}_{\mathcal{S}_j}$	Positions de la source j en 3 dimensions
$x_{\mathcal{S}_j}$	Position de la source j selon l'axe X
$y_{\mathcal{S}_j}$	Position de la source j selon l'axe Y
$z_{\mathcal{S}_j}$	Position de la source j selon l'axe Z
\mathcal{S}	Ensemble contenant les positions des sources
d	Distance entre 2 capteurs
$r_{prop_{i,j}}$	Distance que l'onde parcourt entre la position de référence et un capteur i pour une source j
$\tau(\mathbf{p}_{\mathcal{M}_i}, \mathbf{p}_{\mathcal{S}_j})$	Temps de propagation de l'onde entre la source j et le capteur i
D	L'ouverture de l'antenne ou la distance maximale entre 2 capteurs de l'antenne
$a_{i,j}$	Modèle de propagation entre un capteur i et une position j
$\mathbf{a}_{\mathcal{M},i}$	Vecteur ligne du modèle de propagation entre les capteurs et une position j
$\mathbf{A}_{\mathcal{M},\mathcal{S}}$	Matrice du modèle de propagation
$s_{\mathcal{M},j}$	Amplitude complexe de la source j
$\mathbf{s}_{\mathcal{M},\mathcal{S}}$	Vecteur d'amplitudes complexes des sources
$\mathbf{m}_{\mathcal{M},\mathcal{S}}$	Vecteur de mesures acquises par les capteurs
$\mathbf{w}_{\mathcal{M}}$	Vecteur de bruit de mesure
α_0	Variance du bruit gaussien

Propagation en milieu réverbérant

N_{SR}	Nombre de surfaces prises en compte dans la réverbération
N_{SI}	Nombre de sources images prises en compte
$\mathbf{p}_{\mathcal{I}_n}$	Position de la n -ième source image
\mathcal{I}	Ensemble contenant toutes les positions des sources images
Od	Ordre de réflexion
Cr	Coefficient de réflexion de chaque surface

Problème inverse discrétisé

N_G	Nombre d'inconnues
g	Distance entre 2 positions de la grille, ici entre 2 angles d'arrivées
$\mathbf{p}_{\mathcal{G}_j}$	Position j de la grille discrétisée
\mathcal{G}	Ensemble contenant toutes les positions de la grille
$\tilde{s}_{\mathcal{M},j}$	Amplitude de la source j estimée
$\tilde{\mathbf{s}}_{\mathcal{M},S}$	Vecteur d'amplitudes des sources estimées

Multi-Snapshots

N_T	Nombre de snapshots
$\mathbf{S}_{\mathcal{G}}$	Matrice d'amplitude des sources pour chaque snapshot
$\tilde{\mathbf{S}}_{\mathcal{G}}$	Matrice d'amplitude des sources estimées avec les N_T snapshot
$\mathbf{M}_{\mathcal{M},\mathcal{G}}$	Matrice de mesures acquises par les capteurs avec les N_T snapshots
$\mathbf{W}_{\mathcal{M}}$	Matrice de bruit de mesure avec les N_T snapshots

Inférence Bayésienne Parcimonieuse (SBI)

Σ	Matrice de covariance par rapport à $\mathbf{s}_{\mathcal{M},\mathcal{G}}$
$\Sigma_{j,j}$	Élément de la i -ème ligne et j -ème colonne de la matrice Σ
α	Vecteur de variance de la distribution complexe normale par rapport à $\mathbf{s}_{\mathcal{M},\mathcal{G}}$
α_j	j -ème élément du vecteur α
η	Moyenne pour la distribution gamma sur $\mathbf{s}_{\mathcal{M},S}$
ζ	Variance pour la distribution gamma sur $\mathbf{s}_{\mathcal{M},S}$
ρ	Variance pour la distribution gamma sur $\mathbf{w}_{\mathcal{M}}$
τ	Critère de convergence utilisé pour SBI

Evaluation

r_a	Rayon d'admission
J	Indice de Jaccard
N_E	Nombre de sources estimées par l'algorithme de localisation évalué
$\mathbf{p}_{\mathcal{E}_j}$	Position de la source j estimée par l'algorithme de localisation évalué
\mathcal{E}	Ensemble contenant les positions des sources estimées par l'algorithme de localisation évalué
\mathcal{J}	Ensemble estimant le nombre de faux positifs
$\mathbf{p}_{\mathcal{J}_i}$	i -ème position de source estimée par l'algorithme ne correspondant pas à une position de source réelle
$\tilde{\mathcal{S}}$	Ensemble estimant le nombre de vrais positifs

Autres

ϵ	Constante positive non déterminée
\mathbf{v}	Résidu utilisé pour les approches gloutonnes de localisation de sources
\mathcal{H}	Ensemble contenant les positions sélectionnées par les approches gloutonnes

Sélection de capteurs

N_{K_0}	Nombre de capteurs sélectionnés à l'initialisation
N_K	Nombre de capteurs sélectionnés
$\mathbf{p}_{\mathcal{K}_i}$	Position de capteur i sélectionnée
\mathcal{K}	Ensemble contenant les positions de capteurs sélectionnées
$\boldsymbol{\omega}$	Vecteur binaire permettant de sélectionner les positions de capteurs
ω_i	Valeur binaire sélectionnant ou non la position de capteur i

Algorithme génétique

N_{ant}	Nombre d'antennes aléatoires (<i>individus</i>)
N_{gen}	Nombre d'itérations (<i>générations</i>)
N_K	Nombre de positions de capteurs sélectionnées (<i>gènes</i>)
N_M	Nombre de positions de capteurs disponibles (<i>pool de gènes</i>)
ς	Valeur aléatoire suivant une loi uniforme entre 0 et 1
χ	Coefficient de mutation entre 0 et 1
MSL	Critère d'optimisation de l'algorithme génétique (<i>Maximum Sidelobes level</i>)

Information de Fisher

\mathbf{v}	Variable aléatoire
\mathbf{v}	Variable aléatoire multidimensionnelle
ψ	Paramètre influant sur une variable aléatoire
$\boldsymbol{\psi}$	Paramètres influant sur une variable aléatoire
N_γ	Nombre de paramètres
$f(\mathbf{v}; \boldsymbol{\psi})$	Distribution de vraisemblance d'une variable aléatoire multidimensionnelle \mathbf{v} paramétrée par $\boldsymbol{\psi}$
Ω	Domaine d'intégration de la variable aléatoire
F_i	Information de Fisher
\mathbf{F}_i	Matrice d'information de Fisher

TABLE DES FIGURES

1.1	Évolution du front d'onde en fonction de la distance de propagation.	8
1.2	Front d'onde et angles d'arrivées. Cas avec 1 source et 2 capteurs.	9
1.3	Le modèle source-image : cas du problème plan avec une salle rectangulaire. . . .	11
1.4	Localisation de sources en champ lointain pour une antenne linéaire.	13
1.5	Définition des angles d'arrivées θ et ϕ pour une antenne en 3 dimensions centrée en l'origine.	14
1.6	Allure typique d'une formation de voies.	17
1.7	Modèle hiérarchique des lois utilisées dans l'Inférence Bayésienne Parcimonieuse.	25
1.8	Illustration de la définition des de vrais positifs, faux positifs et faux négatifs pour l'indice de Jaccard en fonction du rayon d'admission r_a	31
2.1	Localisation de sources par formation de voies en champ proche par deux antennes de 32 capteurs de géométrie aléatoire différentes.	39
2.2	Les différentes étapes de l'algorithme génétique.	40
2.3	Schéma explicatif de la sélection séquentielle bayésienne orientée par les données.	55
3.1	Calcul du critère d'information en fonction du nombre de capteurs considérés. En rouge sont les positions de capteurs sélectionnés et en vert celui qui sera sélectionné pour l'itération suivante.	62
3.2	Évolution de l'amplitude maximale du critère δh au cours de la sélection des positions de capteurs. Les 6 premières valeurs correspondent aux valeurs des points verts de la figure 3.1.	63
3.3	Allure du critère de sélection de positions de capteurs δh en fonction de la fréquence de la source.	64
3.4	Allure du critère de sélection de positions de capteurs δh en fonction de l'angle d'arrivée θ de la source.	65
3.5	Allure du critère de sélection de positions de capteurs δh en fonction du nombre de sources présentes.	66
3.6	Allure du critère de sélection de positions de capteurs en fonction de la fréquence du rapport signal à bruit.	67
3.7	Application des 4 seuils pour les antennes à basse (bleu foncé), à moyenne-basse (bleu clair), à moyenne-haute (jaune) et à haute (rouge) quantité d'information.	70

3.8	Indice de Jaccard moyen en fonction du nombre de positions de capteurs sélectionnées pour 6 antennes différentes : l'antenne optimale (vert), aléatoire (noir), à basse information (bleu foncé), à moyenne-basse information (bleu clair), à moyenne-haute information (orange) et à haute information (rouge).	71
3.9	Indice de Jaccard moyen en fonction du nombre de positions de capteurs sélectionnées pour l'antenne optimale à différents rapport signal à bruit : 12 dB (rouge), 9 dB (orange), 6 dB (vert), 3 dB (bleu clair) et 0 dB (bleu foncé).	72
3.10	Indice de Jaccard moyen en fonction du nombre de positions de capteurs sélectionnées pour différentes conceptions d'antennes : antenne suivant notre algorithme (bleu), l'antenne aléatoire (orange), l'antenne linéaire uniforme (violet) et l'antenne suivant la règle de Golomb (jaune).	75
3.11	Indice de Jaccard sous forme de boîtes à moustaches pour voir la variance avec le nombre maximal de positions de capteurs sélectionnées pour différentes conceptions d'antennes : optimale, aléatoire, de Golomb, linéaire et génétique.	76
3.12	Indice de Jaccard sous forme de boîtes à moustaches en fonction du nombre de positions de capteurs sélectionnées pour l'antenne optimale.	76
3.13	Positions des capteurs sélectionnées par les différentes conceptions d'antennes testées sous forme d'histogramme au cours des 100 réalisations	77
3.14	Niveau des lobes secondaires après formation de voies sous forme de boîtes à moustaches avec le nombre maximal de positions de capteurs sélectionnées pour différentes conceptions d'antennes : optimale, aléatoire, de Golomb, linéaire et génétique.	78
3.15	Indice de Jaccard moyen en fonction du nombre de positions de capteurs sélectionnées pour différentes conceptions d'antenne selon différents algorithmes de localisation de sources : inférence bayésienne parcimonieuse (rouge foncé), basis poursuit de-noising (rouge clair), orthogonal matching poursuit (bleu clair) et formation de voies (bleu foncé).	79
3.16	Indice de Jaccard moyen en fonction du nombre de positions de capteurs sélectionnées pour différents critères d'optimalité de minimisation de la matrice de covariance : le critère D-optimal (bleu), le critère A-optimal (vert), le critère E-optimal (rouge).	80
3.17	Indice de Jaccard moyen en fonction du nombre de positions de capteurs sélectionnées pour l'antenne optimale à différents niveaux de RSB : 12 dB (rouge), 9 dB (orange), 6 dB (vert), 3 dB (bleu clair) et 0 dB (bleu foncé). Chaque sous-figure correspond à un nombre de snapshots différent.	82
4.1	Le microphone MEMS.	88
4.2	Géométrie de la salle MODO.	88

4.3	Ensemble de faisceaux.	89
4.4	Patron de la salle MODO et géométrie de l'antenne.	90
4.5	Signaux temporels mesurés par chaque microphone.	90
4.6	Système d'acquisition.	91
4.7	Source monopolaire.	92
4.8	Hauteur des sources dans la salle.	93
4.9	Télémètre-laser.	93
4.10	Positionnement du centre de la source en z_S via le télémètre laser.	94
4.11	Positions dans le plan XY (à la hauteur z_S) de la source. L'orientation de la source est donnée par la flèche noire.	94
4.12	Spectrogrammes des signaux.	96
4.13	Capteur de température positionné sur le support cubique de la source.	97
4.14	Champs acoustiques estimés à partir des mesures expérimentales par différents algorithmes de localisation de sources en considérant un ordre de réflexion (Od) égal à 0 ou 3.	99
4.15	Champ acoustique estimé par formation de voies en champ libre sur un espace étendu. L'ordre de réflexion est illustré par les délimitations de différentes couleurs : 0 (noire), 1 (vert), 2 (rouge), 3 (magenta) et 4 (cyan).	100
5.1	Discretisation du domaine de recherche. Chaque intersection de la grille en noir correspond à une position de l'ensemble \mathcal{G} . La zone rouge correspond à l'espace de 0,5 mètres autour des murs qui n'est pas pris en compte.	105
5.2	Évolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées par notre algorithme pour différents nombres de sources présentes : 1 (bleu foncé), 2 (bleu), 3 (bleu clair), 4 (jaune), 5 (orange) et 6 sources (rouge).	107
5.3	Evolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées par notre algorithme (bleu) et par une sélection aléatoire (rouge) pour différents nombres de sources présentes dans le problème.	108
5.4	Évolution du pourcentage de vrais positifs sur données simulées en fonction du nombre de positions de capteurs sélectionnées par notre algorithme pour différents ordres de réflexion considérés : 0 (bleu foncé), 1 (bleu clair), 2 (orange), et 3 (rouge).	109
5.5	Evolution du pourcentage de vrais positifs sur données simulées en fonction du nombre de positions de capteurs sélectionnées par notre algorithme (bleu) et par une sélection aléatoire (rouge) pour différents ordres de réflexions considérés.	110

TABLE DES FIGURES

5.6	Evolution du pourcentage de vrais positifs sur données expérimentales en fonction du nombre de positions de capteurs sélectionnées par notre algorithme pour différents ordres de réflexion considérés : 0 (bleu foncé), 1 (bleu clair), 2 (orange), et 3 (rouge).	111
5.7	Evolution du pourcentage de vrais positifs sur données expérimentales en fonction du nombre de positions de capteurs sélectionnées par notre algorithme (bleu) et par une sélection aléatoire (rouge) pour différents ordres de réflexions considérés.	112
5.8	Evolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées pour différentes méthodes de sélection de positions de capteurs : sélection optimisée (trait bleu), sélection aléatoire (tirets oranges), sélection génétique (plus violets) et sélection relaxée (croix jaunes).	114
5.9	Les 50 positions de capteurs sélectionnées généralement en premier par notre algorithme. Le score est donné par la couleur du capteur.	114
5.10	Evolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées par notre algorithme (bleu) et par une sélection aléatoire (rouge) pour différents le nombre de snapshots utilisés.	117
5.11	Evolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées selon le critère d’optimalité choisi : A-Optimal (bleu), D-Optimal (vert) et E-Optimal (rouge).	118
5.12	Evolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées selon différentes conceptions d’antennes : l’antenne <i>optimisée sur données réelles</i> (bleu foncé), l’antenne <i>aléatoire</i> (bleu clair), l’antenne <i>optimisée à partir d’un jeu de données simulées</i> (orange) et l’antenne <i>optimisée à partir de l’ensemble des jeux de données simulées</i> (rouge).	120
5.13	Source dipolaire.	f
5.14	Sources monopolaires cohérentes.	g
5.15	Positions dans le plan XY des sources mesurées : au laser (vertes) ou à la règle (bleue).	h

LISTE DES TABLEAUX

4.1	Corrélation normalisée entre données expérimentales et données simulées avec différents ordres de réflexion considérés.	101
4.2	Coefficients de réflexion estimés pour chaque surface.	101
5.1	Paramètres de l'étude.	106

INTRODUCTION

Cette thèse a pour but de concevoir une antenne qui cherche à estimer les positions de sources acoustiques, sous-marines ou aériennes, à partir des mesures prises par l'antenne. Ces sources émettent des ondes qui se propagent dans un milieu et sont captées par des microphones ou des hydrophones. La position des capteurs permettent par la suite d'estimer la position de ces sources dans l'espace. Cette estimation est possible par des connaissances au préalable sur le milieu de propagation et en émettant quelques hypothèses sur les caractéristiques du milieu, des sources et des capteurs. La qualité de l'estimation dépend fortement de la disposition de l'antenne et donc des positions des capteurs. Chercher où placer les capteurs afin de localiser les sources avec la plus haute précision possible est le projet du manuscrit.

Contributions de la thèse :

Un premier objectif de la thèse est de proposer une conception d'antenne innovante en localisation de sources acoustiques. Ce travail est présenté dans les chapitres 1 et 2 et est constitué d'une étude bibliographique sur la propagation acoustique, la localisation de sources et la conception d'antenne. La conception d'antenne peut se traduire par un problème d'optimisation avec de multiples possibilités de résolutions. La solution choisie dans ce manuscrit, devenant la principale contribution, est une sélection séquentielle de capteurs utilisant les mesures acquises précédemment comme information dans la conception de l'antenne.

Un deuxième objectif de la thèse est l'application et l'évaluation des performances de la conception d'antenne proposée dans un contexte de localisation de sources acoustiques sous-marines. Puisque la conception d'antenne est innovante dans ce cadre acoustique, la procédure suivie a été de définir des qualités et des limites de l'algorithme proposé, en fonction des paramètres acoustiques de la localisation de sources. L'évaluation des performances de la conception d'antenne proposée a permis la comparaison avec d'autres approches de la littérature. Ce travail est présenté dans le chapitre 3.

Un dernier objectif de la thèse est l'application et l'évaluation des performances de la conception d'antenne sur des données expérimentales. Une expérience a pu être mise en place dans un milieu aérien, fermé et réverbérant. Le changement de milieu de propagation et l'ajout d'erreurs de modèles dans les mesures acquises par les capteurs (l'imprécision sur la positions des sources et des capteurs, des hypothèses émises dans certains algorithmes qui ne sont plus vérifiées), impactent les performances des algorithmes de localisation et de conception d'antennes. Mais la

question est de savoir si notre conception d'antenne, qui utilise l'information contenue dans les mesures, peut réduire cet impact et gagner en performances par rapport à d'autres méthodes de conception d'antenne. Répondre à cette question est le but du chapitre 5, le chapitre 4 se concentrant sur la présentation de la mise en place de l'expérience.

Organisation du manuscrit :

En dehors de ce chapitre d'introduction, la thèse est organisée autour de 6 chapitres. Les deux premiers chapitres présentent deux domaines différents de la littérature, ainsi que leurs méthodes :

- Le chapitre 1 rappelle la propagation acoustique, aérienne et sous-marine, en champ lointain ou proche, en milieu libre ou réverbérant. À partir du modèle de propagation, on décrit le problème de localisation de sources et les approches permettant la résolution de celui-ci : la formation de voies, les méthodes parcimonieuses et les approches bayésiennes. Ce chapitre contient aussi les critères qui permettront l'évaluation des algorithmes de ce manuscrit.
- Le chapitre 2 introduit la notion de sélection de positions de capteurs. De la sélection aléatoire à la sélection définie par une optimisation d'un critère, on expose différentes méthodes de conception d'antenne, dont principalement les méthodes provenant de la conception optimale ("*Optimal Design*" en anglais et dans la littérature). L'algorithme que nous proposons, la contribution principale de ce manuscrit, provient de cette conception optimale.

Le chapitre suivant applique ces méthodes dans un contexte de localisation de sources acoustiques sous-marines :

- Le chapitre 3 évalue l'algorithme de conception d'antenne par ses performances de localisation de sources. Dans ce cas, nous nous intéressons à un contexte de localisation de sources acoustiques sous-marines où les mesures sont simulées numériquement. La simulation nous permet notamment d'étudier l'impact des caractéristiques des sources sur la conception, comme la fréquence émise ou le niveau sonore du bruit présent dans les mesures. Des premières comparaisons avec des conceptions d'antennes provenant de la littérature seront proposées pour conclure ce chapitre.

Les deux chapitres d'après présentent l'application des méthodes sur des mesures expérimentales de localisation de sources acoustiques aériennes :

- Le chapitre 4 présente la mise en place de l'expérience menée dans la salle "les Murs Ont Des Oreilles", appartenant à Sorbonne Universités et située à St-Cyr l'École. Cette salle contient un nombre important de capteurs placés sur les murs, où l'algorithme peut

permettre de réduire le nombre de capteurs nécessaire à la localisation de sources. Des pré-traitements seront obligatoires afin d'utiliser les mesures acquises lors de l'expérience. Ce chapitre permettra notamment de fixer les paramètres du pré-traitement. La salle étant réverbérante, l'impact des réflexions sur les performances des algorithmes de localisation de sources est étudié.

- Le chapitre 5 développe l'application de l'algorithme sur les mesures expérimentales. Des études sur le gain de capteurs donné par notre algorithme par rapport à d'autres méthodes sont donnés en fonction du nombre de sources ou du nombre de réflexions prises en compte. En plus, ce chapitre permet de proposer une nouvelle conception d'antenne. L'antenne est conçue par des données simulées et évaluées ensuite sur des données expérimentales : elle est donc optimisée par des données sensiblement différentes que celles où l'évaluation est réalisée. Ce chapitre permettra de montrer que cette nouvelle antenne a des performances inférieures à notre précédent algorithme mais satisfaisantes et supérieures à d'autres approches de conceptions d'antennes de la littérature.

Enfin, la conclusion synthétise les contributions et indique les recherches qui semblent pertinentes ou possibles à étudier dans des travaux futurs.

PREMIÈRE PARTIE

État de l'art et présentation des algorithmes : Localisation de sources et sélection de positions de capteurs

LA LOCALISATION DE SOURCES

Sommaire

1.1	Introduction	5
1.2	Propagation de l'onde	8
1.2.1	Généralités sur la scène acoustique	9
1.2.2	Propagation en champ proche : espace libre	10
1.2.3	Propagation en champ proche : milieu réverbérant	11
1.2.4	Propagation en champ lointain	12
1.3	Ecriture du problème de localisation de sources	13
1.3.1	Problème direct	14
1.3.2	Problème inverse	15
1.4	Méthodes classiques	16
1.4.1	Formation de voies	16
1.4.2	Problèmes liés au dimensionnement d'antenne	17
1.4.3	Méthodes de super résolution classiques	19
1.5	Méthodes parcimonieuses	19
1.5.1	Approches gloutonnes	20
1.5.2	Approches par relaxation convexe	21
1.6	Approches Bayésiennes	22
1.6.1	Maximum de vraisemblance	23
1.6.2	Maximum a posteriori	23
1.6.3	Inférence bayésienne parcimonieuse (SBI)	24
1.7	Extension des problèmes au cas de mesures multi-snapshots	26
1.8	Critères d'évaluation	28
1.9	Conclusion	32

1.1 Introduction

Les antennes permettent à partir des signaux mesurés par les capteurs (acoustiques, électromagnétiques, optiques ...), de fournir une cartographie du champ. Les signaux reçus par les capteurs peuvent provenir d'une source indépendante de l'antenne de réception : on parle alors

d'acoustique passive. L'alternative, appelée acoustique active, est d'émettre un signal avec l'antenne de capteurs, puis de recevoir les échos de ce même signal. Ces échos sont créés quand l'onde est réfléchiée par un objet ou une surface. À partir de ces échos pour l'acoustique active, ou du signal émis par une source pour l'acoustique passive, il nous est possible de fournir la cartographie acoustique du champ. Ces images peuvent être utilisées à divers fins : détecter, localiser, caractériser ou classifier une source mais aussi de déterminer les propriétés du milieu de propagation. Les applications peuvent donc être nombreuses, tout comme les domaines d'applications. On peut citer quelques exemples en acoustique [1] :

- L'imagerie médicale pour localiser des problèmes (par exemple des tumeurs) en échographie ;
- La détection d'un défaut en contrôle non destructif ;
- La détection sismique et la localisation de l'épicentre ;
- La caractérisation du rayonnement acoustique d'un moteur ou d'une voiture ;
- La caractérisation d'animaux marins en vu d'un répertoriage ;
- La localisation de sources en acoustique sous-marine et aérienne.

On se place dans le dernier domaine où il faut trouver la position d'une source dans l'espace (après l'avoir détectée). On ne cherche donc pas à caractériser la source, ni à la classifier. Localiser une source acoustique aérienne est une activité que l'humain pratique tous les jours. Nos oreilles sont deux capteurs formant une antenne que l'on déplace dans l'espace. A chaque déplacement, nos oreilles acquièrent une nouvelle mesure et fournissent des informations utiles à la localisation. De même la localisation est une pratique courante des animaux marins, appelée écholocation [2, 3, 4], qui se sert des échos et donc de l'acoustique active afin de chercher de la nourriture. La localisation par sonar ou radar est aussi utilisée pour la navigation des sous-marins, des bateaux ou des robots [5]. Par déplacement de leurs capteurs respectifs, l'antenne augmente son nombre de capteurs et de mesures permettant de localiser plus précisément des sources dans une scène acoustique.

Si la localisation est un phénomène naturel pour les animaux et l'humain, pour les antennes il est nécessaire de concevoir un algorithme pour localiser les sources dans l'espace. Il existe un nombre important d'algorithmes de localisation de sources. Le plus célèbre n'est autre que la formation de voies ou Beamforming [6, 7, 8, 9]. Cette méthode est la plus conventionnelle et continue encore à être largement utilisée. La formation de voies a permis de comprendre un nombre considérable de phénomènes acoustiques et de traitement d'antennes analogique et numérique. Elle a permis de mettre en évidence des problèmes physiques liés au dimensionnement d'antenne, telles que des limites basses et hautes fréquences [10, 11]. À partir de la formation de voies, d'autres méthodes ont émergées, et pallient aux contraintes du dimensionnement d'antenne (l'ouverture d'antenne et le nombre de capteurs), mais ajoutent d'autres contraintes sur

le signal en contrepartie. On peut citer, parmi ces méthodes devenues classiques, la méthode de Capon [12], Music [13] et Esprit [14]. D'autres méthodes ont vu le jour en dehors du domaine appliqué à la localisation acoustique, comme l'acquisition comprimée utilisée dans la reconstruction d'images. Celle-ci a été récemment mise en valeur par Donoho [15] et Candes [16] en 2006. Depuis de nombreux articles ont cherché à exploiter cette découverte dans différents domaines. En localisation acoustique, l'utilisation de cette approche est de plus en plus récurrente [17, 18, 19, 20]. Ces méthodes ont permis d'éviter des problèmes liés à la formation de voies, comme la limite haute fréquence imposée par le théorème de Shannon-Nyquist. Cependant si ces problèmes ne sont plus présents, d'autres contraintes sur le nombre de sources ou sur la cohérence du modèle de propagation de l'onde émise (les colonnes de la matrice de propagation sont supposées être suffisamment incohérentes entre elles) sont apparues [21, 22]. Finalement une dernière catégorie de méthodes permet de résoudre le problème de localisation de sources : les approches bayésiennes [23, 24]. Après avoir traduit les connaissances a priori de l'utilisateur sur le milieu de propagation ou sur les caractéristiques de la source sous forme probabiliste, il est possible par le théorème de Bayès d'estimer la probabilité a posteriori du champ acoustique. Remarquons que ces approches bayésiennes peuvent utiliser des distributions a priori enforçant la parcimonie des sources [25, 26], et s'apparentent donc aux méthodes de l'acquisition comprimée. Il est intéressant de noter que les travaux actuels cherchent à augmenter la précision et donc les performances des algorithmes de localisation de sources. Pour cela, la localisation ne se fait plus sur une grille de positions définie arbitrairement par l'utilisateur, et qui souvent est limitée par la complexité calculatoire, mais dans un intervalle où toutes les positions sont possibles [17, 26, 27].

Dans ce chapitre, avant de décrire la localisation de sources, il est nécessaire d'expliquer comment l'onde acoustique se propage dans le milieu. Nous exposons dans la section 1.2 différents modèles de propagation d'ondes acoustiques qui seront employés dans ce manuscrit : le champ lointain, le champ proche et le milieu réverbérant. Le modèle de propagation de l'onde nous permettra ensuite dans la partie 1.3 de proposer une formulation mathématique au problème direct qui est de savoir comment les mesures obtenues par les capteurs sont écrites. En localisation de sources, on cherche à résoudre ce qu'on appelle le problème inverse. Différents types d'algorithmes existent. Sans faire une liste exhaustive, nous pouvons mentionner les méthodes dites classiques dont la formation de voies, les méthodes parcimonieuses venant de l'acquisition comprimée et les approches bayésiennes. Ces types d'algorithmes sont décrits respectivement dans les sections 1.4, 1.5 et 1.6. Le cas particulier où l'antenne acquiert plusieurs mesures au cours du temps (appelés snapshots) sera présenté dans la partie 1.7. Tous les algorithmes de localisation de sources permettent de donner une estimation du champ acoustique. Il est nécessaire de définir un ou plusieurs critères d'évaluation afin de quantifier les performances de ces algorithmes de localisation. Le choix et la description des critères d'évaluation est dans la partie

1.8. Au final, un premier bilan sur la localisation de sources est proposée en section 1.9.

1.2 Propagation de l'onde

Soit une source acoustique émettant un champ propagatif de pression dans un milieu quelconque où se situent des capteurs sensibles à la pression. Ces capteurs, par leurs positions dans l'espace et le nombre, forment une antenne et vont pouvoir mesurer la pression émise par la source. Le but de la localisation est de retrouver la position de la source dans l'espace à partir des mesures de pressions obtenues par les capteurs. Les mesures des capteurs contiennent une information cruciale à la localisation : le temps d'arrivée de l'onde au capteur. Ce temps d'arrivée est différent pour chaque capteur selon sa position dans l'espace. Il apparaît alors un décalage temporel entre les mesures qui se traduit par un décalage de phase correspondant à la distance que doit parcourir le front d'onde entre les deux capteurs. Cependant ce décalage de phase ne va pas permettre de localiser de la même manière en fonction de si on se trouve en champ proche ou en champ lointain [28]. En champ proche, si la source est ponctuelle ou omnidirectionnelle, le front d'onde est sphérique. En champ lointain, le front d'onde devient plan localement au niveau de l'antenne : on parle alors d'onde plane. L'évolution du front d'onde et la transformation de celui-ci de sphérique en plan est un phénomène connu en acoustique illustré dans le schéma 1.1.

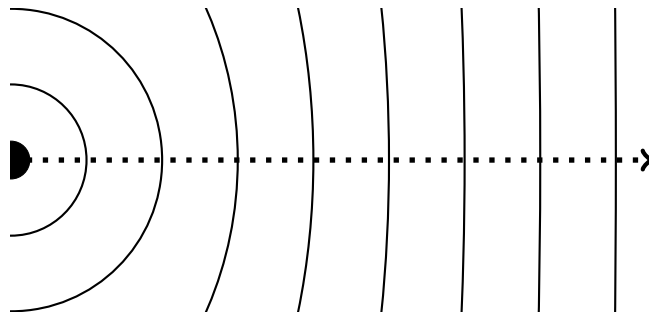


FIGURE 1.1 – Évolution du front d'onde en fonction de la distance de propagation.

Cette différence dans la forme du front d'onde impacte l'angle d'incidence du front d'onde sur le capteur. Dans le cas du champ proche, l'angle d'incidence θ est différent pour chaque capteur (cf. figure 1.2a). À l'inverse en champ lointain, l'angle d'incidence est identique pour chaque capteur (cf. figure 1.2b). Cet angle d'incidence est aussi appelé angle d'arrivée : la localisation en champ lointain est donc une estimation d'angles d'arrivées.

Dans les deux cas présentés, il est possible de complexifier le milieu en prenant en compte par exemple des réflexions contre des parois. Ces parois sont souvent la délimitation entre deux milieux de propagation différents. Par exemple en acoustique sous-marine, le milieu de propagation est l'eau, et il est délimité par l'air (en surface) et le fond marin. Dans le cas de l'acoustique aérienne, les parois sont par exemple des murs formant une rue ou une salle. Ces réflexions pro-

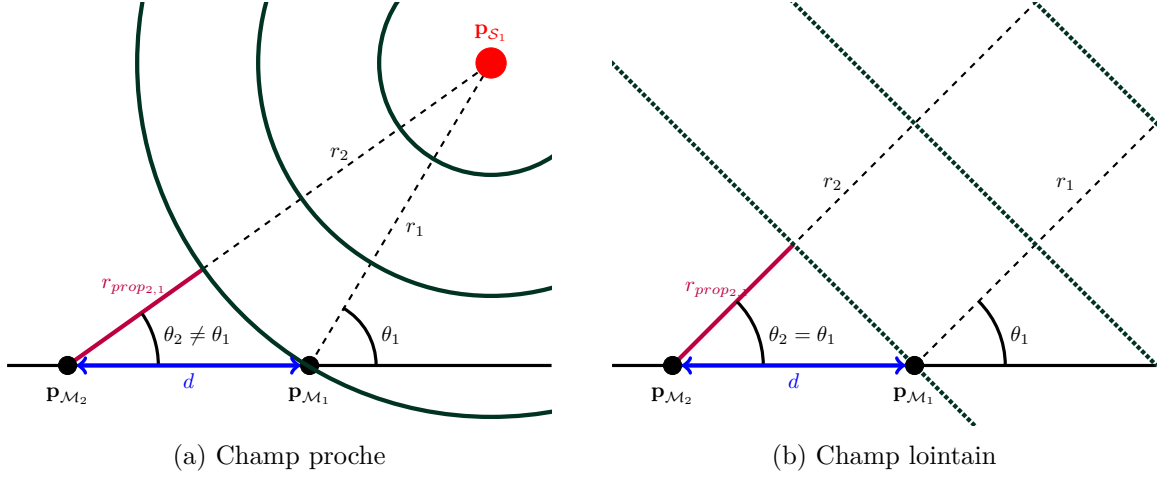


FIGURE 1.2 – Front d'onde et angles d'arrivées. Cas avec 1 source et 2 capteurs.

voquent une propagation multi-trajets. Cette propagation peut même suivre le modèle de guide d'onde si certaines conditions sont respectées (par exemple deux parois et une rupture d'impédance acoustique suffisamment grande) [29]. Dans ce cas, la localisation est mise en second plan pour se concentrer sur la reconstruction des modes qui se propagent afin de mieux connaître le milieu de propagation [30, 31].

Dans cette thèse, le guide d'onde ne sera pas étudié. Cependant dans le cas du champ proche, l'exemple est de prendre une salle fermée où se situent sources et capteurs, et que l'onde se réfléchit sur les murs de la salle [32]. On considère à ce moment que le milieu est réverbérant. On va donc expliquer plus en détail la propagation en champ proche, puis le cas du milieu réverbérant pour la propagation en champ proche et finir sur la propagation en champ lointain. Mais avant de décrire ces 3 situations, nous allons introduire les généralités sur la scène acoustique, les hypothèses physiques et les notations mathématiques.

1.2.1 Généralités sur la scène acoustique

Considérons un milieu de propagation homogène et isotrope (les propriétés du milieu sont les mêmes en tout point de l'espace et dans toutes les directions) en 3 dimensions. La célérité de l'onde c est identique en tout point de l'espace considéré. Dans cet espace, se situent N_M capteurs omnidirectionnels et ponctuels. On définit l'ensemble contenant toutes les positions des capteurs par $\mathcal{M} = \{\mathbf{p}_{\mathcal{M}_i}\}, i \in \llbracket 1, N_M \rrbracket$ avec $\mathbf{p}_{\mathcal{M}_i} \in \mathbb{R}^{1 \times 3}$ la position du capteur i dans l'espace tridimensionnel. Ces capteurs mesurent le signal émis par N_S sources. Les positions des sources sont données par l'ensemble $\mathcal{S} = \{\mathbf{p}_{\mathcal{S}_j}\}, j \in \llbracket 1, N_S \rrbracket$ avec $\mathbf{p}_{\mathcal{S}_j} \in \mathbb{R}^{1 \times 3}$ la position de la source j .

Plusieurs hypothèses vont être émises sur les caractéristiques des sources. Premièrement, le signal reçu par les capteurs est une combinaison linéaire des signaux émis par chacune des sources (hypothèse de linéarité). Deuxièmement, les sources sont ponctuelles, le front d'onde est alors

parfaitement sphérique en champ proche et parfaitement plan en champ lointain. Troisièmement, l'onde émise est monochromatique. Le signal émis par les sources et reçu par les capteurs ne contient alors qu'une seule fréquence F .

Dans la suite, nous n'utiliserons pas le signal temporel reçu par les capteurs, mais le signal fréquentiel correspondant à la fréquence de la source, en passant par une transformée de Fourier discrète. Ce choix permet que chaque capteur ne mesure qu'une seule valeur complexe à la place d'un nombre important par capteur. Il est alors nécessaire de connaître la fréquence de la source. Cette information est soit connue à l'avance, soit estimée en prenant par exemple l'amplitude fréquentielle la plus forte après une transformée de Fourier. Mais travailler dans le domaine fréquentiel au lieu du domaine temporel impose que le milieu, le signal émis et les positions des sources et capteurs, c'est-à-dire la scène acoustique, ne changent pas au cours du temps. En effet, pour travailler dans le domaine fréquentiel, une transformée de Fourier est nécessaire, or celle-ci s'utilise sur un signal temporel où les hypothèses d'invariance temporelle sont faites. Si dans ce manuscrit le temps ne sera pas une variable du problème de localisation (ni de la sélection de capteurs du chapitre 2), une partie de la littérature se concentre sur de la localisation de sources dans un milieu fluctuant [33, 34], ou avec des sources mobiles [35, 36].

1.2.2 Propagation en champ proche : espace libre

Soit $\tau(\mathbf{p}_{\mathcal{M}_i}, \mathbf{p}_{\mathcal{S}_j})$ le temps de propagation de l'onde entre la source j et le capteur i , le modèle de propagation de l'onde dans le domaine fréquentiel $a_{i,j}$ peut s'écrire de la manière suivante :

$$a_{i,j} = \exp^{-j2\pi F\tau(\mathbf{p}_{\mathcal{M}_i}, \mathbf{p}_{\mathcal{S}_j})}. \quad (1.1)$$

La version matricielle du modèle de propagation pour l'ensemble de capteurs \mathcal{M} et l'ensemble de sources \mathcal{S} est noté $\mathbf{A}_{\mathcal{M},\mathcal{S}} \in \mathbb{C}^{N_M \times N_S}$, et $a_{i,j}$ l'élément de la i -ème ligne, j -ème colonne. Si la célérité de l'onde c est identique en tout point, alors le temps de propagation peut s'écrire de la manière suivante :

$$\tau(\mathbf{p}_{\mathcal{M}_i}, \mathbf{p}_{\mathcal{S}_j}) = \frac{1}{c} \|\mathbf{p}_{\mathcal{M}_i} - \mathbf{p}_{\mathcal{S}_j}\|_2, \quad (1.2)$$

avec $\|\mathbf{p}_{\mathcal{M}_i} - \mathbf{p}_{\mathcal{S}_j}\|_2$ la distance euclidienne entre la position du capteur i et la position de la source j .

À cela, il faut ajouter l'atténuation géométrique de l'onde. L'onde étant sphérique, le terme d'atténuation est en $\frac{1}{\|\mathbf{p}_{\mathcal{M}_i} - \mathbf{p}_{\mathcal{S}_j}\|_2}$. Le modèle de propagation dans le cas du champ proche est donc le suivant :

$$a_{i,j} = \frac{1}{\|\mathbf{p}_{\mathcal{M}_i} - \mathbf{p}_{\mathcal{S}_j}\|_2} \exp^{-j\frac{2\pi F}{c} \|\mathbf{p}_{\mathcal{M}_i} - \mathbf{p}_{\mathcal{S}_j}\|_2}. \quad (1.3)$$

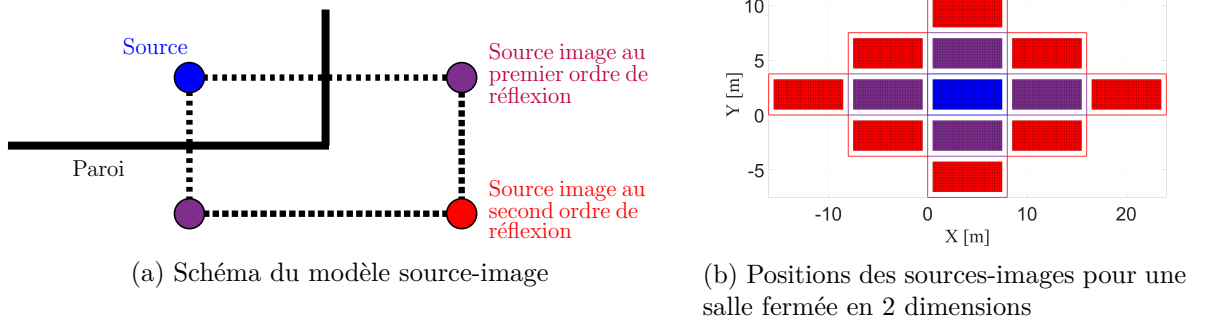


FIGURE 1.3 – Le modèle source-image : cas du problème plan avec une salle rectangulaire.

1.2.3 Propagation en champ proche : milieu réverbérant

En champ proche, si des surfaces réfléchissantes sont présentes dans le milieu de propagation, alors le signal mesuré par les capteurs sera la somme des ondes incidentes (qui suivent le trajet direct) et des ondes réfléchies (une ou plusieurs fois par les surfaces). Le modèle de propagation des ondes incidentes considère le milieu comme étant libre et donc suit l'équation (1.3). Pour les ondes réfléchies, il faut prendre en compte le nombre de réflexions et la position des surfaces réfléchissantes. Afin de simplifier le problème, on considère une salle fermée avec les sources à l'intérieur de cette salle. De plus on émet comme hypothèse que les surfaces soient perpendiculaires entre elles, et que la salle fait un parallépipède rectangle parfait. Pour prendre en compte les réflexions, on utilise le modèle de sources-images [37, 38, 39]. Pour chaque source et chaque surface réfléchissante, une source-image est créée dont sa position est la symétrie de la position de la source par rapport à la surface. Afin de connaître le nombre de sources-images, il est nécessaire de limiter le nombre de réflexions sur les surfaces, c'est-à-dire, de définir l'ordre de réflexion. Par exemple si on ne considère que 2 réflexions sur les surfaces, les positions des sources-images au premier ordre seront les positions symétriques par rapport à chacune des surfaces et au second ordre la combinaison des deux symétries, comme illustré dans la figure 1.3a. Cependant il faudra faire attention à ce que l'onde ne se réfléchisse pas deux fois sur la même surface.

Connaitre le nombre de sources-images N_{SI} , la position de la source et la disposition de la salle permet ensuite de déterminer les positions des sources-images. On définit donc un ensemble contenant les positions des sources-images $\mathcal{I} = \{\mathbf{p}_{\mathcal{I}_n}\}, n \in [1, N_{SI}]$ avec la position de la n -ième source-image $\mathbf{p}_{\mathcal{I}_n} \in \mathbb{R}^{1 \times 3}$. Un exemple avec une salle rectangulaire en 2 dimensions est donné en figure 1.3b où on illustre les positions des sources-images pour différents ordres de réflexions, allant de 0 (couleur bleue) à 2 (couleur rouge). Notons que les zones blanches sont des positions non discrétisées.

Le modèle de sources-images décompose l'onde émise par la source en une multitude d'ondes

sommées dans le signal dont le modèle de propagation est similaire au modèle défini dans l'équation (1.3) et donc à la propagation de l'onde incidente. Cependant une atténuation de l'intensité est présente puisque la surface réfléchissante absorbe une partie de l'énergie de l'onde. Dans le modèle de propagation il est donc nécessaire de connaître le coefficient de réflexion $Cr \in [0, 1]$ de chaque surface. Si plusieurs réflexions sont prises en compte, alors le coefficient de réflexion est la multiplication des deux coefficients correspondants. Le modèle de propagation en champ proche et en milieu réverbérant peut s'écrire de la façon suivante :

$$a_{i,j} = \frac{1}{\|\mathbf{p}_{\mathcal{M}_i} - \mathbf{p}_{\mathcal{S}_j}\|_2} \exp^{-j\frac{2\pi F}{c}\|\mathbf{p}_{\mathcal{M}_i} - \mathbf{p}_{\mathcal{S}_j}\|_2} + \sum_{n=1}^{N_{SI}} Cr_n \frac{\exp^{-j\frac{2\pi F}{c}\|\mathbf{p}_{\mathcal{M}_i} - \mathbf{p}_{\mathcal{I}_n}\|_2}}{\|\mathbf{p}_{\mathcal{M}_i} - \mathbf{p}_{\mathcal{I}_n}\|_2}. \quad (1.4)$$

Cependant ce modèle de propagation a quelques limitations. Les effets de diffusion et la dépendance de l'absorption avec l'angle d'incidence ne sont pas pris en compte. À noter que des études ont permis de définir un modèle de propagation plus réaliste en les prenant en compte [40, 41]. Celui-ci pourrait être pris en compte dans de futurs travaux.

1.2.4 Propagation en champ lointain

La limite entre le champ lointain et le champ proche peut-être définie si la géométrie de l'antenne est connue. Par exemple, pour une antenne linéaire uniforme d'ouverture $D = \max_{i_1, i_2} \|\mathbf{p}_{\mathcal{M}_{i_1}} - \mathbf{p}_{\mathcal{M}_{i_2}}\|_2$, la condition pour considérer le problème comme du champ lointain est [42] :

$$r \geq \frac{2D^2}{\lambda}, \quad (1.5)$$

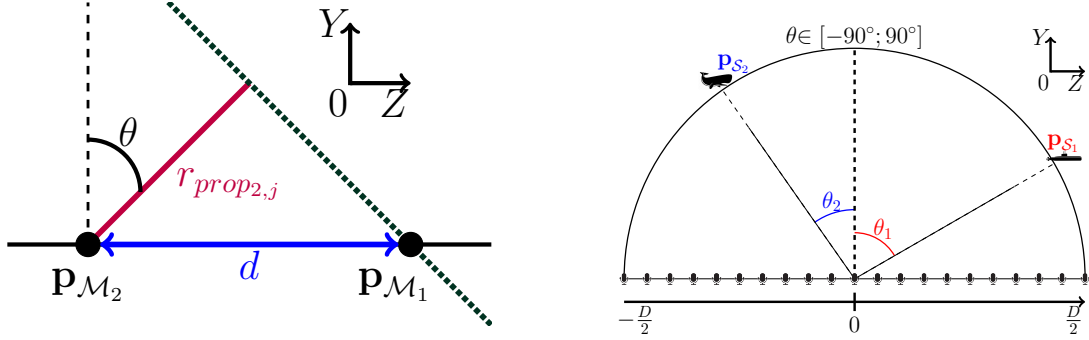
avec r la distance entre la source et l'antenne et $\lambda = \frac{c}{F}$ la longueur d'onde émise par la source. En champ lointain les distances sources-capteurs étant trop grandes, il ne sera donc pas possible de déterminer les positions des sources. Une approximation est de considérer l'onde comme une onde plane monochromatique. Ce sont alors les angles d'arrivées qui seront les inconnues du problème.

On définit dans le cas d'une antenne linéaire, placée sur l'axe Z , avec pour origine le centre de l'antenne et les positions des N_M capteurs $z_{\mathcal{M}_i}$, la distance de propagation $r_{prop_{i,j}}$ que l'onde parcourt entre l'origine et le capteur i pour une source j localisée à un angle d'arrivée θ_j :

$$r_{prop_{i,j}} = z_{\mathcal{M}_i} \sin(\theta_j), \quad (1.6)$$

avec l'angle d'arrivée θ compris entre $-\frac{\pi}{2}$ et $\frac{\pi}{2}$. Ce cas est illustré par la figure 1.4.

Maintenant considérons une antenne à 3 dimensions centrée en l'origine, avec les positions des capteurs $x_{\mathcal{M}_i}$, $y_{\mathcal{M}_i}$ et $z_{\mathcal{M}_i}$ pour respectivement les positions selon l'axe X , Y et Z . Pour cette antenne, les deux angles d'arrivée θ et ϕ sont définis comme illustré dans la figure 1.5. L'équation



(a) Distance de propagation parcouru par l'onde entre 2 capteurs

(b) Définition de l'angle d'arrivée θ

FIGURE 1.4 – Localisation de sources en champ lointain pour une antenne linéaire.

(1.6) se généralise en :

$$r_{prop_{i,j}} = x_{\mathcal{M}_i} \cos(\theta_j) \cos(\phi_j) + y_{\mathcal{M}_i} \cos(\theta_j) \sin(\phi_j) + z_{\mathcal{M}_i} \sin(\theta_j). \quad (1.7)$$

avec i l'indice pour le capteur et j l'indice pour la source.

Le modèle de propagation en champ lointain s'écrit de manière générale par une matrice $\mathbf{A}_{\mathcal{M},S}$ dont les éléments sont :

$$a_{i,j} = \exp^{-j \frac{2\pi F}{c} r_{prop_{i,j}}}. \quad (1.8)$$

La principale différence avec le modèle de propagation du champ proche est l'absence du terme d'atténuation géométrique. Cette absence s'explique par la différence du front d'onde considéré. En propagation en ondes planes, le terme d'atténuation est négligé. Si on considère l'antenne de la figure 1.5, et donc la distance de propagation selon l'équation (1.7), le modèle de propagation en champ lointain dans l'équation (1.8) se précise de la manière suivante :

$$a_{i,j} = \exp^{-j \frac{2\pi}{\lambda} (x_{\mathcal{M}_i} \cos(\theta_j) \cos(\phi_j) + y_{\mathcal{M}_i} \cos(\theta_j) \sin(\phi_j) + z_{\mathcal{M}_i} \sin(\theta_j))}. \quad (1.9)$$

1.3 Ecriture du problème de localisation de sources

Avec le modèle de propagation défini, il faut maintenant définir le problème direct et inverse de la localisation de sources. Le problème direct définit le modèle mathématique des mesures acquises par les capteurs. Quelques hypothèses s'ajouteront dans cette partie. Le problème inverse pose la question suivante : comment estimer les amplitudes et positions (ou angles d'arrivées) des sources, à partir des mesures collectées par l'antenne ? Cette question, et donc le problème inverse, correspondent au problème de localisation de sources.

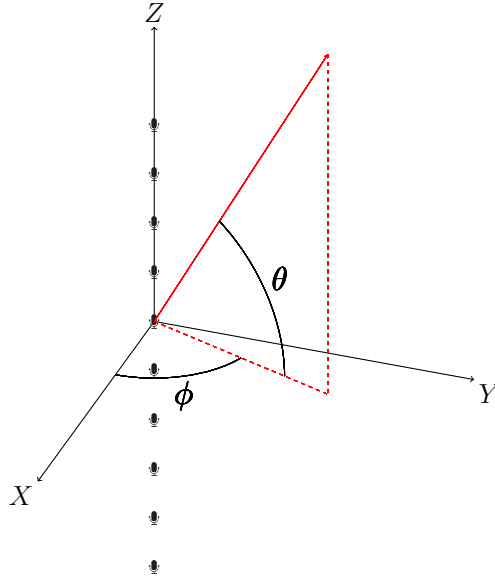


FIGURE 1.5 – Définition des angles d'arrivées θ et ϕ pour une antenne en 3 dimensions centrée en l'origine.

1.3.1 Problème direct

Définissons dans un premier temps le vecteur des amplitudes des sources : $\mathbf{s}_{\mathcal{M},S} \in \mathbb{C}^{N_S \times 1}$. Il y a donc autant de valeurs complexes que de sources. Cela est possible seulement si le signal est monochromatique et donc que le traitement est fait dans le domaine fréquentiel. Avec cette hypothèse, si on a N_M capteurs, alors on acquiert N_M mesures. Quelque soit le modèle de propagation choisi $\mathbf{A}_{\mathcal{M},S}$, le vecteur de mesures $\mathbf{m}_{\mathcal{M},S} \in \mathbb{C}^{N_M \times 1}$ peut s'écrire de la manière suivante :

$$\mathbf{m}_{\mathcal{M},S} = \mathbf{A}_{\mathcal{M},S} \mathbf{s}_{\mathcal{M},S} + \mathbf{w}_{\mathcal{M}}, \quad (1.10)$$

avec $\mathbf{w}_{\mathcal{M}} \in \mathbb{C}^{N_M \times 1}$ un bruit blanc complexe circulaire gaussien de variance α_0^{-1} . Le bruit est également indépendant entre les capteurs. Dépendant de cette variance, le rapport signal sur bruit (SNR pour *Signal-to-Noise Ratio*) exprimé en décibels (*dB*) peut varier. Plus il est élevé, plus le bruit gaussien est faible par rapport au signal émis par les sources. Il se définit ainsi :

$$\text{SNR} = 20 \log \left(\frac{\|\mathbf{A}_{\mathcal{M},S} \mathbf{s}_{\mathcal{M},S}\|_2}{\|\mathbf{w}_{\mathcal{M}}\|_2} \right), \quad (1.11)$$

avec $\|\cdot\|_2$ la norme l_2 . Le niveau de SNR impacte la localisation de sources. En effet, plus le niveau de bruit est important, plus il est difficile de retrouver le signal émis par les sources dans les mesures acquises par les capteurs. L'hypothèse d'un bruit gaussien permet aux méthodes de localisation de fournir une estimation correcte des positions de sources si le nombre de mesures est suffisant. Si le niveau de SNR est très faible, de nombreuses méthodes n'arrivent pas à

localiser les sources. Pour cela, il est possible d'utiliser certaines méthodes permettant de pallier à ce problème précis [43] ou d'utiliser plusieurs mesures par capteurs (cas multi-snapshots) [44]. L'impact du SNR sur la localisation de sources et la sélection de capteurs sera étudié dans le chapitre 3.

1.3.2 Problème inverse

Pour écrire le problème inverse, il faut définir les inconnues du problème. Ces inconnues vont être différentes selon le contexte acoustique. Les inconnues sont, dans le cas du champ lointain, les angles d'arrivées possibles, et dans la cas du champ proche, les différentes positions possibles des sources dans l'espace. Les inconnues sont contenues dans un domaine de recherche. Une approche classique est de discrétiser ce domaine en N_G points. Chaque point correspond donc soit à une position de l'espace pour le champ proche, soit à un angle d'arrivée pour le champ lointain. Cette approche discrète peut poser problème si les sources ne se situent pas sur un point de la grille. Il y aura donc une différence entre la position de la source et celle estimée par les algorithmes de localisation de sources. Si avec des données simulées ce problème peut être évité en plaçant les sources sur des points de cette grille, il est très difficile de faire de même en expérimental. Pour pallier au problème, il est possible d'augmenter le nombre de points de la grille N_G , au prix d'augmenter la complexité calculatoire des algorithmes. Récemment, de nombreuses études cherchent à ne plus résoudre le problème de manière discrète, mais de le résoudre de manière dite continue [45, 17, 46]. Toutefois, dans ce manuscrit, l'approche continue ne sera pas étudiée. Nous éviterons donc l'erreur de modèle en plaçant les sources sur des points de la grille en ce qui concerne les données simulées dans le chapitre 3. Pour les données expérimentales utilisées aux chapitres 4 et 5, cette erreur de modèle sera donc présente.

Dans l'approche discrète, on considère N_G positions dans un espace en 3 dimensions. Similairement aux positions de sources, on définit un ensemble pour les positions discrètes, tel que $\mathcal{G} = \{\mathbf{p}_{\mathcal{G}_j}\}, j \in \llbracket 1, N_G \rrbracket$ avec $\mathbf{p}_{\mathcal{G}_j} \in \mathbb{R}^{1 \times 3}$ la position discrète j . Ensuite il faut transformer le vecteur contenant les amplitudes des sources $\mathbf{s}_{\mathcal{M},\mathcal{S}} \in \mathbb{C}^{N_S \times 1}$ en $\mathbf{s}_{\mathcal{M},\mathcal{G}} \in \mathbb{C}^{N_G \times 1}$ et la matrice de propagation $\mathbf{A}_{\mathcal{M},\mathcal{S}} \in \mathbb{C}^{N_M \times N_S}$ en $\mathbf{A}_{\mathcal{M},\mathcal{G}} \in \mathbb{C}^{N_M \times N_G}$. L'équation (1.10) devient :

$$\mathbf{m}_{\mathcal{M},\mathcal{G}} = \mathbf{A}_{\mathcal{M},\mathcal{G}} \mathbf{s}_{\mathcal{M},\mathcal{G}} + \mathbf{w}_{\mathcal{M}}. \quad (1.12)$$

Si les positions des sources correspondent bien à des points de la grille alors $\mathbf{m}_{\mathcal{M},\mathcal{G}} = \mathbf{m}_{\mathcal{M},\mathcal{S}}$. Sinon des erreurs de modèle seront présentes et ne pourront pas être quantifiées, ou corrigées.

Notons aussi que le vecteur source discret $\mathbf{s}_{\mathcal{M},\mathcal{G}}$ peut être parcimonieux si le nombre de sources est faible par rapport aux nombre d'inconnues, $N_S \ll N_G$ et donc que la présence de valeurs nulles dans le vecteur est importante. Si cette condition est respectée, alors il sera possible d'utiliser les méthodes parcimonieuses qui seront présentées dans la section 1.7.

Le vecteur source discret $\mathbf{s}_{\mathcal{M},\mathcal{G}}$ est inconnu. À partir des mesures et du modèle de propagation choisi, il est possible de l'estimer [47] et ainsi de localiser les sources. Une approche naïve serait de résoudre le problème au sens des moindres carrés tel que :

$$\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}} = \underset{\mathbf{s}_{\mathcal{M},\mathcal{G}}}{\operatorname{argmin}} \|\mathbf{m}_{\mathcal{M},\mathcal{G}} - \mathbf{A}_{\mathcal{M},\mathcal{G}} \mathbf{s}_{\mathcal{M},\mathcal{G}}\|_2^2. \quad (1.13)$$

Pour trouver la solution unique à ce problème, la résolution par moindres carrés nécessite un cadre surdéterminé, c'est-à-dire, que le nombre de mesures N_M est supérieur au nombre d'inconnues N_G : une contrainte non souhaitée dans ce manuscrit. Parmi les 3 familles de méthodes que nous présentons et utilisons dans ce manuscrit, c'est-à-dire, les méthodes dites classiques (comme la formation de voies), les méthodes parcimonieuses et les méthodes bayésiennes, aucune ne résout le problème de l'équation (1.13).

1.4 Méthodes classiques

Dans cette partie, la formation de voies est présentée. Cette approche est très utilisée en acoustique, cependant des problèmes sont présents. Parmi ceux-ci se trouve notamment le dimensionnement de l'antenne ou sa géométrie qui peuvent influencer fortement et négativement l'estimation de localisation de sources. Pour une bonne utilisation de la formation de voies, des contraintes sur la disposition des capteurs en fonction de certaines caractéristiques des sources sont à prendre en compte. Pour pallier à ces contraintes, d'autres méthodes ont émergé, telles que Capon [12], Music [13] et Esprit [14]. Ces méthodes ne seront pas utilisées dans la suite du manuscrit, puisque une grande partie de la littérature a déjà largement contribué au développement de ces méthodes en acoustique. Néanmoins elles méritent d'être mentionnées et introduites aux lecteurs.

1.4.1 Formation de voies

La formation de voies résout le problème de localisation de sources en sommant les signaux temporels mesurés par les capteurs après les avoir décalés dans le temps. Le décalage dans le temps correspond à un déphasage des signaux dans le domaine spectral. L'estimation du vecteur source est obtenue par l'opération suivante, équivalente à un filtrage spatial :

$$\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}} = (\operatorname{diag}(\mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{A}_{\mathcal{M},\mathcal{G}}))^{-1} \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{m}_{\mathcal{M},\mathcal{G}}. \quad (1.14)$$

avec \cdot^H l'hermitien ou complexe conjugué transposé. Si la source est unique, la formation de voies permet de visualiser le diagramme de directivité de l'antenne. Celui-ci est centré autour de la position de la source [48] et est défini par 3 caractéristiques (voir la figure 1.6) :

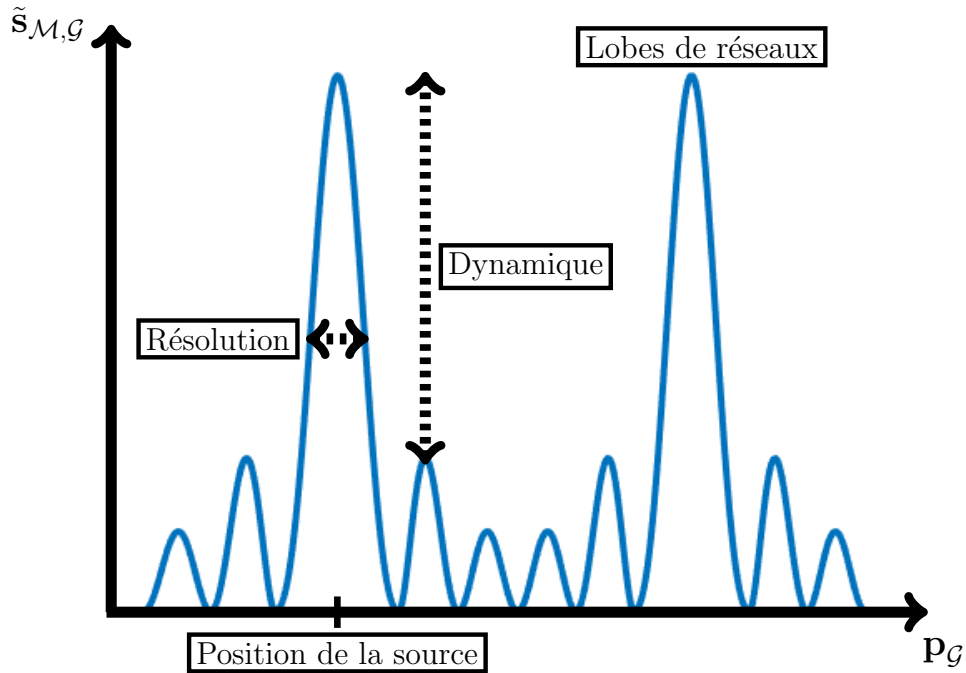


FIGURE 1.6 – Allure typique d’une formation de voies.

- la résolution,
- la dynamique,
- les lobes de réseau.

Chaque caractéristique du diagramme de directivité est influencé par les dimensions et la géométrie de l’antenne. On décrit ces caractéristiques, les problèmes qu’elles provoquent et les solutions utilisées classiquement dans la section suivante.

1.4.2 Problèmes liés au dimensionnement d’antenne

Nous rappelons que les problèmes de dimensionnement d’antenne relevés dans cette partie sont spécifiques à la méthode de la formation de voies. Les autres méthodes de localisation de sources n’auront pas forcément les mêmes problèmes. Nous nous concentrons donc sur les 3 caractéristiques du diagramme de directivité de l’antenne et comment elles participent à la géométrie de l’antenne.

La résolution est définie par la largeur du lobe principal (voir la figure 1.6). Elle dépend de la fréquence émise par les sources et de l’ouverture d’antenne. Elle diminue si l’ouverture d’antenne augmente ou si la fréquence émise augmente.

La résolution impose une distance minimale entre deux sources afin de pouvoir les distinguer. Cela impose une limite basse-fréquence qui dépend de l’ouverture de l’antenne.

Pour pouvoir localiser la source de manière efficace, il est important de réduire la largeur du lobe principal. Pour cela il est nécessaire de respecter la condition suivante [10] :

$$D > \epsilon\lambda, \tag{1.15}$$

avec $\epsilon > 1$ une constante positive dépendant du pouvoir de résolution souhaité, D l'ouverture d'antenne et λ la longueur d'onde du signal émis par la source. Cette constante peut dépendre de l'angle d'arrivée si on considère le champ lointain [49], mais surtout du nombre de sources présentes dans le problème. Si 2 sources sont présentes, afin de les localiser il est nécessaire que leurs positions soient suffisamment séparées. Dans le cas de la formation de voies, les deux lobes principaux ne doivent pas "trop" se superposer. La constante ϵ permet d'ajuster ce niveau de superposition entre les lobes principaux de chaque source.

Les lobes de réseau sont les répliques du lobe principal (voir la figure 1.6). Les lobes de réseau sont une conséquence du phénomène de recouvrement spatial, analogue au repliement spectral en analyse fréquentielle de signaux temporels.

Les lobes de réseau, ayant la même amplitude que le lobe principal, créent des ambiguïtés dans les résultats des algorithmes de localisation de sources. Des positions seront estimées aux positions des lobes de réseau où des sources ne seront pas présentes.

Le recouvrement spatial peut être évité en respectant le théorème d'échantillonnage de Shannon-Nyquist [11]. Pour cela, la distance entre des capteurs régulièrement espacés doit respecter l'inégalité suivante :

$$d \leq \frac{\lambda}{2}. \tag{1.16}$$

Pour une fréquence émise par les sources donnée, il est possible de déterminer une ouverture d'antenne par l'équation (1.15), puis la distance entre les capteurs par l'équation (1.16). Si ensuite, on considère une antenne uniforme linéaire alors le nombre de capteurs nécessaire à la localisation des sources respecte l'inégalité suivante :

$$N_M \geq \frac{D}{d} + 1. \tag{1.17}$$

Au-delà de cette constatation donnée par ces limites théoriques, la réalisation en pratique d'antennes avec une forte densité de capteurs sur une grande ouverture peut être impossible à réaliser. En général, l'ouverture d'antenne est limitée, dû à des moyens techniques et financiers. La fréquence est donc souvent élevée, amenant une distance entre les capteurs très faible et un nombre de capteurs important. Dans un tel cas, il peut être intéressant de changer de méthode d'estimation d'angles d'arrivées.

Enfin, la dynamique est la différence de niveau entre les lobes secondaires et le lobe principal (voir la figure 1.6). La dynamique est influencée par la position des capteurs dans l'antenne.

La présence de lobes secondaires impacte et peut fausser la localisation de sources. Il faut donc maximiser la dynamique, et donc optimiser le placement des capteurs dans l'antenne. Ce sujet est traité dans le chapitre 2 avec les algorithmes génétiques.

1.4.3 Méthodes de super résolution classiques

Des méthodes dites de super résolutions permettent de pallier aux contraintes de la formation de voies. L'impact de la largeur du lobe principal et des lobes secondaires peut être diminué par la méthode de Capon [12] (aussi appelée MVDR pour *Minimum Variance Distortionless Response*). Cette méthode montre une amélioration de la résolution mais demande une estimation empirique de la matrice de covariance. La qualité de l'estimation de la matrice de covariance dépend du nombre de mesures acquises au cours du temps (nombre de snapshots) par les capteurs et l'estimation de la matrice de covariance s'appuie sur une hypothèse de stationnarité spatiale au cours des mesures. Une baisse des performances peut être observée si cette hypothèse n'est pas respectée [50].

D'autres approches exploitent les structures propres des signaux et leurs propriétés supposées. La matrice de covariance des mesures acquises par les capteurs, étant une matrice semi-définie positive, peut être décomposée en une base orthonormée de vecteurs propres. Il est possible alors de séparer les sous-espaces propres relatifs au signal et au bruit. Le sous-espace propre associé au signal appartient au sous-espace défini par la matrice de propagation, tandis que le sous-espace associé au bruit est orthogonal à celui-ci. L'algorithme MUSIC (pour *MUltiple SIgnal Classification*) [13] exploite cette propriété d'orthogonalité. Cependant comme pour la méthode de Capon, cela nécessite l'estimation de la matrice de covariance et donc un nombre de mesures indépendantes important. L'algorithme nécessitant une décomposition en sous-espaces, une complexité calculatoire issue de la décomposition en valeurs et vecteurs propres est à prévoir si le nombre d'inconnues N_G est importante.

Un nombre important d'autres techniques existe, telles que la prédiction linéaire, la minimisation de la norme, la maximisation de l'entropie, la décomposition en harmoniques de Pisarenko, ou encore ESPRIT (pour *Estimation of Signal Parameters via Rotational Invariance Techniques*). Cependant, pour ne pas saturer les lecteurs avec des méthodes non utilisées par la suite, nous ne expliciterons pas plus ces méthodes ainsi que leurs avantages et défauts. Un nombre important d'articles dans la littérature en fait déjà la comparaison [51, 7, 52].

1.5 Méthodes parcimonieuses

Les méthodes parcimonieuses sont des approches qui sont devenues populaires dans la littérature, notamment pour la reconstruction d'image [161]. Mais des conditions sont à respecter pour l'utilisation de ces méthodes. Dans ces conditions, on note que les méthodes parcimonieuses ont

l'avantage de ne plus avoir à respecter le théorème de Shannon-Nyquist (cf. équation (1.15)) ni la limite basse fréquence (cf. équation (1.17)) [161]. Cependant, si les performances de ces méthodes ne sont plus impactées par la fréquence du signal, elles sont dépendantes de deux autres facteurs : la parcimonie du vecteur source (nombre de sources N_S par rapport au nombre d'inconnues N_G) et la cohérence de la matrice de propagation [53]. Le nombre de capteurs nécessaires à la reconstruction du vecteur source n'est donc plus lié à la fréquence, mais au nombre de coefficients non nuls du vecteur source, autrement dit, au nombre de sources présentes et à l'incohérence entre les colonnes de la matrice (autrement appelé dictionnaire dans la littérature) de propagation. Dans un cadre acoustique, les colonnes de la matrice de propagation sont liées à des positions dans l'espace ou des angles d'arrivées. Les colonnes sont donc cohérentes entre elles et cette deuxième condition ne sera donc pas respectée. Mais cette condition étant une condition suffisante et non nécessaire, l'utilisation des méthodes parcimonieuses est possible [54, 18].

Les méthodes parcimonieuses ont pour but de résoudre le problème d'optimisation suivant :

$$\underset{\mathbf{s}_{\mathcal{M},\mathcal{G}}}{\text{minimize}} \|\mathbf{s}_{\mathcal{M},\mathcal{G}}\|_0 \text{ subject to } \mathbf{A}_{\mathcal{M},\mathcal{G}} \mathbf{s}_{\mathcal{M},\mathcal{G}} = \mathbf{m}_{\mathcal{M},\mathcal{G}}. \quad (1.18)$$

Ce problème d'optimisation est similaire à celui de l'équation (1.13), mais impose que le vecteur source $\mathbf{s}_{\mathcal{M},\mathcal{G}}$ soit N_S -parcimonieux. Le problème d'optimisation est combinatoire, aussi appelé NP-difficile. Il n'existe pas à notre connaissance d'algorithme de complexité polynomiale permettant de résoudre toutes les instances de l'équation (1.18). Les méthodes parcimonieuses vont chercher à résoudre un problème d'optimisation approché. Pour cela un grand nombre d'approches et de méthodes existent. Dans ce manuscrit, nous nous sommes restreint aux approches suivantes : les approches gloutonnes ou itératives et les approches par relaxation convexe.

1.5.1 Approches gloutonnes

Les méthodes les plus connues sont le *Matching Pursuit* (MP) [55, 56] et sa variante orthogonale, *Orthogonal Matching Pursuit* (OMP) [57, 58]. Ces méthodes estiment le vecteur source $\mathbf{s}_{\mathcal{M},\mathcal{G}}$ de façon incrémentale.

MP propose une approche simple : elle consiste à sélectionner la colonne de la matrice $\mathbf{A}_{\mathcal{M},\mathcal{G}}$, notée $\mathbf{a}_{\mathcal{M},i^*}$, la plus corrélée avec le signal reçu $\mathbf{m}_{\mathcal{M},\mathcal{G}}$:

$$i^* = \underset{i=1,\dots,N_G}{\operatorname{argmax}} |\mathbf{m}_{\mathcal{M},\mathcal{G}}^H \mathbf{a}_{\mathcal{M},i}|, \quad (1.19)$$

puis d'estimer la composante correspondante du vecteur source :

$$s_{\mathcal{M},i^*} = \mathbf{a}_{\mathcal{M},i^*}^H \mathbf{m}_{\mathcal{M},\mathcal{G}}. \quad (1.20)$$

La contribution de la colonne sélectionnée est soustraite au signal $\mathbf{m}_{\mathcal{M},\mathcal{G}}$ afin de donner le résidu

$\mathbf{v} \in \mathbb{C}^{N_M \times 1}$:

$$\mathbf{v} = \mathbf{m}_{\mathcal{M},\mathcal{G}} - \mathbf{a}_{\mathcal{M},i^*} (\mathbf{a}_{\mathcal{M},i^*}^H \mathbf{m}_{\mathcal{M},\mathcal{G}}). \quad (1.21)$$

Par la suite, le résidu \mathbf{v} est confronté à la matrice $\mathbf{A}_{\mathcal{M},\mathcal{G}}$ pour sélectionner une seconde colonne :

$$i^* = \operatorname{argmax}_{i=1,\dots,N_G} |\mathbf{v}^H \mathbf{a}_{\mathcal{M},i}|, \quad (1.22)$$

L'opération est répétée pour sélectionner d'autres colonnes et estimer d'autres composantes de $\mathbf{s}_{\mathcal{M},\mathcal{G}}$. Le nombre d'itérations est fixé par l'utilisateur. Dans le cadre de notre étude, le nombre d'itérations sera égal au nombre de sources, puisque ce nombre sera connu.

MP donne des résultats convaincants quand les colonnes de la matrice $\mathbf{A}_{\mathcal{M},\mathcal{G}}$ sont orthogonales. Cela n'est pas obligatoirement respecté en localisation de sources acoustiques. Lors du calcul du résidu dans l'équation (1.21), l'algorithme introduit de nouvelles composantes dans les directions des atomes déjà sélectionnés. Ce phénomène est évité en soustrayant au résidu \mathbf{v} la projection du signal $\mathbf{m}_{\mathcal{M},\mathcal{G}}$ sur l'espace engendré par tous les atomes précédemment sélectionnés. Pour cela on introduit dans l'estimation du vecteur source, la projection sur l'espace vectoriel engendré par les atomes sélectionnés, et l'algorithme devient OMP. L'équation (1.20) devient :

$$\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{H}^{(n)}} = (\mathbf{A}_{\mathcal{M},\mathcal{H}^{(n)}}^H \mathbf{A}_{\mathcal{M},\mathcal{H}^{(n)}})^{-1} \mathbf{A}_{\mathcal{M},\mathcal{H}^{(n)}}^H \mathbf{m}_{\mathcal{M},\mathcal{G}}, \quad (1.23)$$

avec $\mathcal{H}^{(n)} \subset \mathcal{G}$ l'ensemble qui contient toutes les positions sélectionnées par l'algorithme à l'itération n . Le calcul du résidu de l'équation (1.21) est aussi modifié et devient :

$$\mathbf{v}^{(n)} = \mathbf{m}_{\mathcal{M},\mathcal{G}} - \mathbf{A}_{\mathcal{M},\mathcal{H}^{(n)}} \tilde{\mathbf{s}}_{\mathcal{M},\mathcal{H}^{(n)}}. \quad (1.24)$$

Dans la suite du manuscrit, nous utilisons seulement OMP. Le pseudo code d'OMP est donné dans l'Algorithme 1.

1.5.2 Approches par relaxation convexe

Les approches relaxées résolvent le problème d'optimisation en modifiant la norme l_0 en norme l_1 . Le problème peut se poser de la manière suivante [59] :

$$\underset{\mathbf{s}_{\mathcal{M},\mathcal{G}}}{\text{minimize}} \|\mathbf{s}_{\mathcal{M},\mathcal{G}}\|_1 \text{ subject to } \mathbf{A}_{\mathcal{M},\mathcal{G}} \mathbf{s}_{\mathcal{M},\mathcal{G}} = \mathbf{m}_{\mathcal{M},\mathcal{G}}. \quad (1.25)$$

Le problème posé s'appelle le *Basis Pursuit* (BP). Celui-ci ne prend pas en compte le bruit. Une autre version de ce problème prenant en compte un bruit gaussien existe et s'appelle le *Basis Pursuit De-Noising* (BPDN) [60] :

$$\underset{\mathbf{s}_{\mathcal{M},\mathcal{G}}}{\text{minimize}} \|\mathbf{s}_{\mathcal{M},\mathcal{G}}\|_1 \text{ subject to } \|\mathbf{A}_{\mathcal{M},\mathcal{G}} \mathbf{s}_{\mathcal{M},\mathcal{G}} - \mathbf{m}_{\mathcal{M},\mathcal{G}}\|_2^2 \leq \alpha_0^{-1}. \quad (1.26)$$

Algorithme 1 Reconstruction du vecteur source $\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}$ par *Orthogonal Matching Pursuit* (OMP)

Entrée : N_S si connu sinon un nombre fixe d'itérations, $\mathbf{A}_{\mathcal{M},\mathcal{G}} \in \mathbb{C}^{N_M \times N_G}$, $\mathbf{m}_{\mathcal{M},\mathcal{G}} \in \mathbb{C}^{N_M \times 1}$, $\mathcal{G} = \{\mathbf{p}_{\mathcal{G}_j}\}, j \in [1, N_G]$ avec $\mathbf{p}_{\mathcal{G}_j} \in \mathbb{R}^{1 \times 3}$ la position discrète j .

1. **Initialisation :**

- résidu $\mathbf{v}^{(0)} = \mathbf{m}_{\mathcal{M},\mathcal{G}}$
- ensemble des positions sélectionnées $\mathcal{H}^{(0)} = \{\}$

2. **pour** $n = 1, 2, \dots, N_S$ **faire :**

- Trouver l'indice i^* de l'ensemble \mathcal{G} qui maximise le problème suivant :

$$i^* = \underset{i=1, \dots, N_G}{\operatorname{argmax}} |(\mathbf{v}^{(n-1)})^H \mathbf{a}_{\mathcal{M},i}|$$

- Mise à jour de l'ensemble des positions sélectionnées :

$$\mathcal{H}^{(n)} = \mathcal{H}^{(n-1)} \cup \{\mathbf{p}_{\mathcal{G}_{i^*}}\}$$

- Estimation du vecteur source :

$$\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{H}^{(n)}} = (\mathbf{A}_{\mathcal{M},\mathcal{H}^{(n)}}^H \mathbf{A}_{\mathcal{M},\mathcal{H}^{(n)}})^{-1} \mathbf{A}_{\mathcal{M},\mathcal{H}^{(n)}}^H \mathbf{m}_{\mathcal{M},\mathcal{G}}$$

- Mise à jour du résidu :

$$\mathbf{v}^{(n)} = \mathbf{m}_{\mathcal{M},\mathcal{G}} - \mathbf{A}_{\mathcal{M},\mathcal{H}^{(n)}} \tilde{\mathbf{s}}_{\mathcal{M},\mathcal{H}^{(n)}}$$

Sortie : $\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}} = \tilde{\mathbf{s}}_{\mathcal{M},\mathcal{H}^{(n)}}$

Pour résoudre ce problème, un nombre important d'algorithmes existe dont le plus célèbre est le *Fast Iterative Shrinkage Thresholding Algorithm* (FISTA) [61]. Nous allons utiliser un outil appelé CVX [62] qui permet de résoudre aisément les problèmes d'optimisation convexe. Cependant il peut nécessiter un temps de calcul considérable, et il est limité par le nombre d'inconnues et de capteurs que l'on peut considéré. Cela limitera certains de nos résultats numériques.

1.6 Approches Bayésiennes

Avant de parler des approches bayésiennes et de comment elles permettent d'estimer le vecteur source ayant généré les observations, il est nécessaire de parler du maximum de vraisemblance (ou *maximum likelihood* ML) et du maximum a posteriori (MAP).

1.6.1 Maximum de vraisemblance

Le maximum de vraisemblance a pour but de reconstruire le vecteur source $\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}$ en maximisant une fonction de vraisemblance, qui selon les paramètres d'un modèle statistique calcule la probabilité des mesures observées. La fonction de vraisemblance $\mathcal{P}(\mathbf{m}_{\mathcal{M},\mathcal{G}}|\mathbf{s}_{\mathcal{M},\mathcal{G}})$ est la probabilité d'observer $\mathbf{m}_{\mathcal{M},\mathcal{G}}$ étant donné un vecteur $\mathbf{s}_{\mathcal{M},\mathcal{G}}$. Le but est d'estimer le vecteur source $\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}$ qui maximise cette probabilité [63] :

$$\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}} = \operatorname{argmax}_{\mathbf{s}_{\mathcal{M},\mathcal{G}}} \mathcal{P}(\mathbf{m}_{\mathcal{M},\mathcal{G}}|\mathbf{s}_{\mathcal{M},\mathcal{G}}). \quad (1.27)$$

Si on considère que le bruit présent dans les mesures est un bruit complexe gaussien de moyenne nulle et de variance α_0^{-1} , alors l'estimation donnée par le maximum de vraisemblance a pour moyenne [64] :

$$\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}} = (\alpha_0 \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{A}_{\mathcal{M},\mathcal{G}})^{-1} \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{m}_{\mathcal{M},\mathcal{G}}, \quad (1.28)$$

et pour covariance [64] :

$$\boldsymbol{\Sigma} = (\alpha_0 \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{A}_{\mathcal{M},\mathcal{G}})^{-1}. \quad (1.29)$$

1.6.2 Maximum a posteriori

Le maximum a posteriori cherche à estimer $\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}$ comme le maximum d'une distribution a posteriori des mesures $\mathbf{m}_{\mathcal{M},\mathcal{G}}$. Le problème du maximum a posteriori s'écrit [26] :

$$\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}} = \operatorname{argmax}_{\mathbf{s}_{\mathcal{M},\mathcal{G}}} \mathcal{P}(\mathbf{s}_{\mathcal{M},\mathcal{G}}|\mathbf{m}_{\mathcal{M},\mathcal{G}}). \quad (1.30)$$

Selon le théorème de Bayes, il est possible de réécrire le problème de la manière suivante :

$$\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}} = \operatorname{argmax}_{\mathbf{s}_{\mathcal{M},\mathcal{G}}} \frac{\mathcal{P}(\mathbf{m}_{\mathcal{M},\mathcal{G}}|\mathbf{s}_{\mathcal{M},\mathcal{G}})\mathcal{P}(\mathbf{s}_{\mathcal{M},\mathcal{G}})}{\mathcal{P}(\mathbf{m}_{\mathcal{M},\mathcal{G}})}. \quad (1.31)$$

Le dénominateur $\mathcal{P}(\mathbf{m}_{\mathcal{M},\mathcal{G}})$ ne dépend pas de $\mathbf{s}_{\mathcal{M},\mathcal{G}}$ et donc ne joue pas de rôle dans le problème d'optimisation. Il convient de ne pas le prendre en compte. On remarque aussi que si la probabilité a priori $\mathcal{P}(\mathbf{s}_{\mathcal{M},\mathcal{G}})$ suit une loi uniforme, alors le maximum a posteriori correspond au maximum de vraisemblance. Si on considère que la probabilité a priori $\mathcal{P}(\mathbf{s}_{\mathcal{M},\mathcal{G}})$ suit une loi complexe normale de moyenne nulle et de variance $\operatorname{diag}(\boldsymbol{\alpha})^{-1}$, avec diag la matrice diagonale dont les éléments diagonaux correspondent au vecteur $\boldsymbol{\alpha} \in \mathbb{C}^{N_G \times 1}$, alors l'estimation par maximum a posteriori a pour moyenne [26] :

$$\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}} = \alpha_0 (\alpha_0 \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{A}_{\mathcal{M},\mathcal{G}} + \operatorname{diag}(\boldsymbol{\alpha})^{-1})^{-1} \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{m}_{\mathcal{M},\mathcal{G}}, \quad (1.32)$$

et pour covariance [26] :

$$\mathbf{\Sigma} = (\alpha_0 \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{A}_{\mathcal{M},\mathcal{G}} + \text{diag}(\boldsymbol{\alpha})^{-1})^{-1}. \quad (1.33)$$

1.6.3 Inférence bayésienne parcimonieuse (SBI)

Si le maximum de vraisemblance ne permet pas de prendre en compte la parcimonie du signal, la probabilité a priori présente dans le maximum a posteriori peut être utilisée pour respecter ce critère. Un algorithme connu est l'Inférence Bayésienne Parcimonieuse (ou SBI pour *Sparse Bayesian Inference*) développé par Tipping en 2001 [25], et utilisé dans un cadre acoustique par Yang en 2013 [26]. Pour mettre en avant la parcimonie, le choix de la distribution a priori sur $\mathbf{s}_{\mathcal{M},\mathcal{G}}$ est de suivre une distribution favorisant la parcimonie. Cependant cela n'est pas possible suivant un maximum a posteriori. Un moyen de contourner le problème est de combiner une distribution a priori complexe normale sur $\mathbf{s}_{\mathcal{M},\mathcal{G}}$:

$$\mathcal{P}(\mathbf{s}_{\mathcal{M},\mathcal{G}}|\boldsymbol{\alpha}) = \mathcal{CN}(\mathbf{0}_{N_G}, \text{diag}(\boldsymbol{\alpha})), \quad (1.34)$$

avec $\mathbf{0}_{N_G}$ la matrice nulle de dimension $\mathbb{R}^{N_G \times N_G}$ et une probabilité a priori suivant une distribution gamma pour chaque valeur de la distribution de probabilité $\boldsymbol{\alpha}$:

$$\mathcal{P}(\boldsymbol{\alpha}|1, \rho) = \prod_{j=1}^{N_G} \Gamma(\alpha_j|1, \rho), \quad (1.35)$$

avec ρ qui tend vers 0. Pour rappel une distribution complexe normale s'écrit :

$$\mathcal{CN}(\mathbf{x}|\boldsymbol{\mu}, \mathbf{\Sigma}) = \frac{1}{\pi^N |\mathbf{\Sigma}|} \exp\{-(\mathbf{x} - \boldsymbol{\mu})^H \mathbf{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})\}, \quad (1.36)$$

avec $\boldsymbol{\mu}$ la moyenne et $\mathbf{\Sigma}$ la covariance et une distribution gamma s'écrit :

$$\Gamma(u|c, d) = [\Gamma(c)]^{-1} d^c u^{c-1} \exp\{-d u\}, \quad (1.37)$$

avec c le paramètre de forme et d le paramètre d'intensité.

La même méthode est appliquée pour la probabilité a priori du bruit. Le bruit $\mathbf{w}_{\mathcal{M}}$ suit donc une distribution complexe normale :

$$\mathcal{P}(\mathbf{w}_{\mathcal{M}}|\alpha_0) = \mathcal{CN}(\mathbf{0}_{N_M}, \alpha_0^{-1} \mathbf{I}_{N_M}), \quad (1.38)$$

où la probabilité α_0 suit elle-même une distribution gamma :

$$\mathcal{P}(\alpha_0|\eta, \zeta) = \Gamma(\alpha_0|\eta, \zeta), \quad (1.39)$$

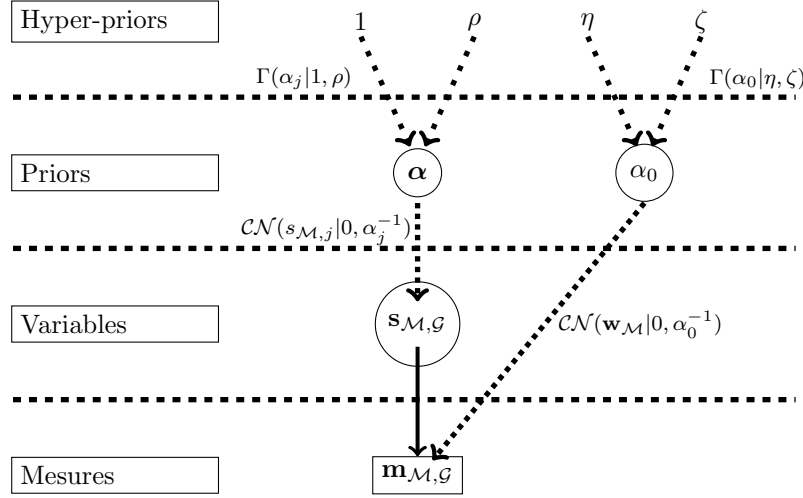


FIGURE 1.7 – Modèle hiérarchique des lois utilisées dans l'Inférence Bayésienne Parcimonieuse.

avec η, ζ qui tendent vers 0. Cette approche a pour avantage de ne pas avoir à connaître la variance de la distribution a priori de $\mathbf{s}_{\mathcal{M}, \mathcal{G}}$ ou du bruit $\mathbf{w}_{\mathcal{M}}$. En revanche, il est nécessaire d'estimer $\boldsymbol{\alpha}$ et α_0 . Pour cela, on suit une procédure d'espérance-maximisation (EM) [25]. Après une initialisation des distributions de probabilité $\boldsymbol{\alpha}$ et α_0 , l'étape d'espérance permet de calculer $\tilde{\mathbf{s}}_{\mathcal{M}, \mathcal{G}}$ et $\boldsymbol{\Sigma}$ suivant respectivement les équations (1.32) et (1.33). Ensuite l'étape de maximisation permet d'estimer les distributions de probabilité $\boldsymbol{\alpha}$ et α_0 . Suivant les lois de probabilités choisies, la distribution de probabilité sur $\mathbf{s}_{\mathcal{M}, \mathcal{G}}$ s'écrit pour chaque élément j du vecteur $\boldsymbol{\alpha}$ [26] :

$$\alpha_j = \frac{\sqrt{1 + 4\rho(\|\tilde{\mathbf{s}}_{\mathcal{M}, \mathcal{G}}\|_2^2 + \Sigma_{j,j})} - 1}{2\rho}, \quad (1.40)$$

et la distribution de probabilité a priori sur $\mathbf{w}_{\mathcal{M}}$ à :

$$\alpha_0 = \frac{N_M + \eta - 1}{\|\mathbf{m}_{\mathcal{M}, \mathcal{G}} - \mathbf{A}_{\mathcal{M}, \mathcal{G}} \tilde{\mathbf{s}}_{\mathcal{M}, \mathcal{G}}\|_F^2 + \alpha_0^{-1} \sum_{j=1}^{N_G} 1 - \alpha_j^{-1} \Sigma_{j,j} + \zeta}. \quad (1.41)$$

avec $\|\cdot\|_F$ la norme de Frobenius. Ces étapes d'espérance et de maximisation sont répétées tant que la procédure n'a pas convergé vers un résultat. La convergence choisie est la suivante :

$$\frac{\|\boldsymbol{\alpha}^i - \boldsymbol{\alpha}^{i-1}\|_2}{\|\boldsymbol{\alpha}^i\|_2} \leq \tau, \quad (1.42)$$

avec i l'indice d'itération et τ le critère de convergence choisi. Pour éviter un nombre d'itérations trop important, nous limitons le nombre d'itérations à 5000 dans la suite de ce manuscrit.

Le résumé de l'algorithme SBI se trouve dans l'Algorithme 2 et le modèle hiérarchique des lois utilisées dans le schéma 1.7.

Algorithme 2 Reconstruction du vecteur source $\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}$ par Inférence Bayésienne Parcimonieuse (SBI)

Entrée : $\mathbf{A}_{\mathcal{M},\mathcal{G}} \in \mathbb{C}^{N_M \times N_G}$, $\mathbf{m}_{\mathcal{M},\mathcal{G}} \in \mathbb{C}^{N_M \times 1}$, α_0^{-1} la variance du bruit gaussien

1. **Initialisation :**

- paramètres des distributions de probabilité $\zeta, \eta, \rho = 10^{-4}$
- distribution de probabilité sur le vecteur source $\boldsymbol{\alpha} = \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{m}_{\mathcal{M},\mathcal{G}}$

2. **pour** $i = 1, \dots, 5000$ **faire :**

- Estimation de la matrice de covariance $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma} = (\alpha_0 \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{A}_{\mathcal{M},\mathcal{G}} + \text{diag}(\boldsymbol{\alpha})^{-1})^{-1}$$

- Estimation du vecteur source $\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}$:

$$\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}} = \alpha_0 (\alpha_0 \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{A}_{\mathcal{M},\mathcal{G}} + \text{diag}(\boldsymbol{\alpha})^{-1})^{-1} \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{m}_{\mathcal{M},\mathcal{G}}$$

- Calcul de la distribution de probabilité α_0 sur le bruit gaussien :

$$\alpha_0 = \frac{N_M + \eta - 1}{\|\mathbf{m}_{\mathcal{M},\mathcal{G}} - \mathbf{A}_{\mathcal{M},\mathcal{G}} \tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}\|_F^2 + \alpha_0^{-1} \sum_{j=1}^{N_G} 1 - \alpha_j^{-1} \Sigma_{j,j} + \zeta}$$

- Calcul de la distribution de probabilité $\boldsymbol{\alpha}$ sur le vecteur source élément par élément :

$$\alpha_j = \frac{\sqrt{1 + 4\rho(\|\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}\|_2^2 + \Sigma_{j,j})} - 1}{2\rho}$$

- Vérification du critère de convergence τ :

$$\frac{\|\boldsymbol{\alpha}^i - \boldsymbol{\alpha}^{i-1}\|_2}{\|\boldsymbol{\alpha}^i\|_2} \leq \tau$$

Sortie : $\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}$, $\boldsymbol{\Sigma}$, α_0 , $\boldsymbol{\alpha}$

1.7 Extension des problèmes au cas de mesures multi-snapshots

Une manière d'améliorer l'estimation de la localisation de source est de considérer plusieurs mesures, appelées snapshots, par capteur. Chaque snapshot est donc une mesure à un temps différent. Cependant la dépendance temporelle peut être évitée en posant une hypothèse de stationnarité où la seule variable modifiée entre les snapshots d'un capteur est le bruit gaussien. Cette considération permet donc d'atténuer l'impact du bruit présent dans les mesures et ainsi de pouvoir localiser plus précisément les sources. Dans la littérature, de nombreuses méthodes ont vu le jour [68, 69, 70, 44]. Cependant, dans ce manuscrit, les snapshots vont être utilisés seulement si ceux-ci peuvent améliorer la sélection de capteurs. L'algorithme SBI peut considérer

plusieurs snapshots [26, 71] et ainsi permettent une meilleure estimation des positions des sources $\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}$. Si l'estimation des positions des sources est plus aboutie, alors la sélection de positions de capteurs peut-être impactée positivement. Cet impact sera étudié dans le chapitre 3.

À des fins de clarté, nous allons seulement présenter l'algorithme SBI incluant plusieurs snapshots, que nous appellerons par la suite MSSBI (pour Multiple Snapshots Sparse Bayesian Inference). On définit dans un premier temps, le nombre de snapshots N_T ainsi que la matrice de mesures $\mathbf{M}_{\mathcal{M},\mathcal{G}} \in \mathbb{C}^{N_M \times N_T}$ qui contient pour chaque capteur N_T mesures. Pour chaque snapshot de chaque capteur, le bruit complexe gaussien est différent et est contenu dans la matrice $\mathbf{W}_{\mathcal{M}} \in \mathbb{C}^{N_M \times N_T}$. Chaque snapshot permet une estimation de la position des sources, et donc une estimation du vecteur d'amplitude des sources. On définit ainsi une matrice contenant tous les vecteurs d'amplitude des sources estimés $\tilde{\mathbf{S}}_{\mathcal{G}} \in \mathbb{C}^{N_G \times N_T}$. On obtient donc autant d'estimations de localisation de sources que de snapshots. Il est alors possible de moyennner cette matrice ou de moyennner les critères d'évaluations sur les N_T vecteurs. Il est plus commun de moyennner la matrice afin de n'obtenir qu'un vecteur et ainsi de pouvoir comparer facilement les résultats par rapport au nombre de snapshots utilisés. Dans nos comparaisons du chapitre 3, nous moyennnerons la matrice d'amplitude de sources estimée afin que le critère d'évaluation soit identique quelque soit le nombre de snapshots utilisés.

L'ajout des snapshots ne modifie pas les dimensions de $\boldsymbol{\alpha}$, α_0 et $\boldsymbol{\Sigma}$ mais modifie les lois de probabilité concernant les probabilités a priori, c'est-à-dire les équations (1.34) et (1.38). Les lois de probabilité pour MSSBI sont les suivantes :

$$\mathcal{P}(\mathbf{S}_{\mathcal{G}}|\boldsymbol{\alpha}) = \prod_{t=1}^{N_T} \mathcal{CN}(\mathbf{s}_{\mathcal{M},\mathcal{G}}(t)|\mathbf{0}_{N_G}, \text{diag}(\boldsymbol{\alpha})), \quad (1.43)$$

$$\mathcal{P}(\boldsymbol{\alpha}|\eta, \zeta) = \prod_{j=1}^{N_G} \Gamma(\alpha_j|\eta, \zeta), \quad (1.44)$$

$$\mathcal{P}(\mathbf{W}_{\mathcal{M}}|\alpha_0) = \prod_{t=1}^{N_T} \mathcal{CN}(\mathbf{w}_{\mathcal{M}}(t)|\mathbf{0}_{N_M}, \alpha_0^{-1}\mathbf{I}_{N_M}), \quad (1.45)$$

$$\mathcal{P}(\alpha_0|1, \rho) = \Gamma(\alpha_0|1, \rho). \quad (1.46)$$

Ces nouvelles considérations modifient aussi le calcul de $\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}$, $\boldsymbol{\Sigma}$, α_0 et $\boldsymbol{\alpha}$, autrement dit les équations (1.32), (1.33), (1.41) et (1.40). Ceux-ci se définissent de la manière suivante :

$$\tilde{\mathbf{S}}_{\mathcal{G}} = \alpha_0(\alpha_0\mathbf{A}_{\mathcal{M},\mathcal{G}}^H\mathbf{A}_{\mathcal{M},\mathcal{G}} + \text{diag}(\boldsymbol{\alpha})^{-1})^{-1}\mathbf{A}_{\mathcal{M},\mathcal{G}}^H\mathbf{M}_{\mathcal{M},\mathcal{G}}, \quad (1.47)$$

$$\boldsymbol{\Sigma} = (\alpha_0\mathbf{A}_{\mathcal{M},\mathcal{G}}^H\mathbf{A}_{\mathcal{M},\mathcal{G}} + \text{diag}(\boldsymbol{\alpha})^{-1})^{-1}, \quad (1.48)$$

$$\alpha_0 = \frac{N_M + \frac{\eta-1}{N_T}}{\left\| \frac{\mathbf{M}_{\mathcal{M},\mathcal{G}}}{\sqrt{N_T}} - \mathbf{A}_{\mathcal{M},\mathcal{G}} \frac{\tilde{\mathbf{S}}_{\mathcal{G}}}{\sqrt{N_T}} \right\|_F^2 + \alpha_0^{-1} \sum_{j=1}^{N_G} 1 - \alpha_j^{-1} \Sigma_{j,j} + \frac{\zeta}{N_T}}, \quad (1.49)$$

$$\alpha_j = \frac{\sqrt{1 + 4 \frac{\rho}{N_T} (\left\| \frac{\tilde{\mathbf{S}}_{\mathcal{G}}}{\sqrt{N_T}} \right\|_2^2 + \Sigma_{j,j})} - 1}{2 \frac{\rho}{N_T}}. \quad (1.50)$$

En dehors de ces modifications, le principe de l'algorithme est identique à SBI. Après une initialisation de α_0 et α , les étapes d'espérance et de maximisation correspondantes aux équations (1.47), (1.48), (1.49) et (1.50) sont répétées jusqu'à convergence. On considère que l'algorithme a convergé suivant la même condition que pour SBI, c'est-à-dire l'équation (1.42). Notons que si le nombre de snapshots considéré est de $N_T = 1$, l'algorithme MSSBI est identique à SBI. L'algorithme MSSBI est résumé dans l'Algorithme 3.

1.8 Critères d'évaluation

Pour évaluer l'estimation de localisation de sources, nous présentons dans un premier temps les critères les plus utilisés. On présente l'erreur quadratique, définie de la manière suivante :

$$MSE = \|\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}} - \mathbf{s}_{\mathcal{M},\mathcal{G}}\|_2. \quad (1.51)$$

L'erreur quadratique étant en décibels, la valeur la plus haute est 0 et un bon estimateur cherchera à avoir l'erreur quadratique la plus basse possible. Un autre critère classique pour l'évaluation des performances en localisation de sources est la corrélation :

$$Corr = \frac{|\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}^H \mathbf{s}_{\mathcal{M},\mathcal{G}}|}{\|\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}\|_2 \|\mathbf{s}_{\mathcal{M},\mathcal{G}}\|_2}. \quad (1.52)$$

L'erreur quadratique et la corrélation sont dépendantes de la localisation mais aussi de l'amplitude estimée. Dans un cadre parcimonieux, les performances de l'estimateur jugé par ces deux critères sont souvent faibles. En effet, chaque petite erreur est sanctionnée par le critère de manière importante. Par exemple si l'estimateur reconstruit la position de la source au mauvais endroit, l'erreur est doublée par la non-détection de la source et par la reconstruction de la source donnée par l'estimateur.

Pour évaluer les performances de localisation dans un cadre parcimonieux, on peut utiliser deux critères : l'indice de Jaccard et la *Flat Metric*. On s'inspire du travail fait en thèse par Quentin Denoyelle [27] qui a travaillé sur la localisation parcimonieuse et continue en utilisant ces deux critères afin d'évaluer son estimateur. L'indice de Jaccard n'évalue que la localisation et non l'estimation de l'amplitude. De plus un nouveau paramètre entre en jeu, un rayon d'admission

Algorithme 3 Reconstruction du vecteur source $\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}$ par Inférence Bayésienne Parcimonieuse à Plusieurs Snapshots (MSSBI)

Entrée : $\mathbf{A}_{\mathcal{M},\mathcal{G}} \in \mathbb{C}^{N_M \times N_G}$, $\mathbf{M}_{\mathcal{M},\mathcal{G}} \in \mathbb{C}^{N_M \times N_T}$, α_0^{-1} la variance du bruit gaussien et N_T le nombre de snapshots

1. **Initialisation :**

- paramètres des distributions de probabilité $\zeta, \eta, \rho = 10^{-4}$
- distribution de probabilité sur le vecteur source $\boldsymbol{\alpha} = \frac{1}{N_T} \sum_{t=1}^{N_T} \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{m}_{\mathcal{M},\mathcal{G}}(t)$

2. **pour** $i = 1, \dots, 5000$ **faire :**

- Estimation de la matrice de covariance $\boldsymbol{\Sigma}$:

$$\boldsymbol{\Sigma} = (\alpha_0 \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{A}_{\mathcal{M},\mathcal{G}} + \text{diag}(\boldsymbol{\alpha})^{-1})^{-1}$$

- Estimation du vecteur source $\tilde{\mathbf{S}}_{\mathcal{G}}$:

$$\tilde{\mathbf{S}}_{\mathcal{G}} = \alpha_0 (\alpha_0 \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{A}_{\mathcal{M},\mathcal{G}} + \text{diag}(\boldsymbol{\alpha})^{-1})^{-1} \mathbf{A}_{\mathcal{M},\mathcal{G}}^H \mathbf{M}_{\mathcal{M},\mathcal{G}}$$

- Calcul de la distribution de probabilité sur le bruit gaussien α_0 :

$$\alpha_0 = \frac{N_M + \frac{\eta-1}{N_T}}{\left\| \frac{\mathbf{M}_{\mathcal{M},\mathcal{G}}}{\sqrt{N_T}} - \mathbf{A}_{\mathcal{M},\mathcal{G}} \frac{\tilde{\mathbf{S}}_{\mathcal{G}}}{\sqrt{N_T}} \right\|_F^2 + \alpha_0^{-1} \sum_{j=1}^{N_G} 1 - \alpha_j^{-1} \Sigma_{j,j} + \frac{\zeta}{N_T}}$$

- Calcul de la distribution de probabilité sur le vecteur source $\boldsymbol{\alpha}$ élément par élément :

$$\alpha_j = \frac{\sqrt{1 + 4 \frac{\rho}{N_T} (\left\| \frac{\tilde{\mathbf{S}}_{\mathcal{G}}}{\sqrt{N_T}} \right\|_2^2 + \Sigma_{j,j})} - 1}{2 \frac{\rho}{N_T}}$$

- Vérification du critère de convergence τ :

$$\frac{\|\boldsymbol{\alpha}^i - \boldsymbol{\alpha}^{i-1}\|_2}{\|\boldsymbol{\alpha}^i\|_2} \leq \tau$$

Sortie : $\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}} = \frac{1}{N_T} \sum_{t=1}^{N_T} \tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}(t)$, $\boldsymbol{\Sigma}$, α_0 , $\boldsymbol{\alpha}$

r_a qui permet d'accepter une erreur sur la localisation.

L'indice de Jaccard évalue la similarité entre deux ensembles en calculant le rapport entre l'intersection et l'union des ensembles considérés [72]. Dans notre cas, les ensembles sont l'ensemble des positions de sources vraies \mathcal{S} et l'ensemble des positions de sources estimées $\mathcal{E} = \{\mathbf{p}_{\mathcal{E}_j}\}$, $j \in [1, N_E]$ avec $\mathbf{p}_{\mathcal{S}_j} \in \mathbb{R}^{1 \times 3}$ la position j estimée de la source et N_E le nombre de sources

estimées par l'algorithme. L'indice de Jaccard est par définition l'expression suivante :

$$J = \frac{|\mathcal{S} \cap \mathcal{E}|}{|\mathcal{S} \cup \mathcal{E}|}, \quad (1.53)$$

avec \cap l'opérateur d'intersection et \cup l'opérateur d'union. Remarquons que actuellement cette définition ne prend pas en compte le rayon d'admission r_a .

L'ensemble \mathcal{E} s'obtient à partir du vecteur source estimé $\tilde{\mathbf{s}}_{\mathcal{M},\mathcal{G}}$. Une manière assez simple est de sélectionner les pics les plus forts par la fonction "*findpeaks*" de Matlab, et ensuite de ne prendre que ceux supérieurs à une valeur fixe et arbitraire ou de n'en prendre qu'un certain nombre (par exemple le nombre de sources si celui-ci est connu). Les deux cas seront considérés : dans le chapitre 3 une valeur seuil sera utilisée, et dans les chapitres 4 et 5 le nombre de sources sera utilisé.

On transforme l'opérateur d'intersection \cap entre une position de source estimée et une position de source vraie, de l'équation (1.53) par l'utilisation du rayon d'admission r_a . Cela permet de déterminer si la source est localisée et considérée comme un vrai positif (ou *TP* pour *True Positive*). Pour être évaluée comme un vrai positif, la distance entre la position de la source estimée i et la position de la source vraie j doit respecter l'inégalité suivante :

$$\|\mathbf{p}_{\mathcal{E}_i} - \mathbf{p}_{\mathcal{S}_j}\|_2 \leq r_a. \quad (1.54)$$

À partir de cette équation il est possible de définir un premier critère qui est le pourcentage de vrais positifs, défini par l'équation suivante :

$$TP_{\%} = \frac{TP}{N_S} \cdot 100. \quad (1.55)$$

Ce critère sera utilisé principalement dans les chapitres 4 et 5. L'évaluation de ce critère est décrite dans l'Algorithme 4.

À partir de l'équation (1.53) et du calcul du nombre de vrais positifs *TP*, il est possible de définir un autre critère d'évaluation qui est l'indice de Jaccard défini de la manière suivante [66] :

$$J = \frac{TP}{TP + FP + FN}, \quad (1.56)$$

avec *FP* le nombre de faux positifs et *FN* le nombre de faux négatifs. La définition des vrais positifs, des faux positifs et des faux négatifs en fonction du rayon d'admission est illustrée dans la figure 1.8. Le nombre de faux positifs et de faux négatifs se déterminent après avoir calculé le nombre de vrais positifs. Les faux négatifs sont simplement les positions de sources qui n'ont pas été estimées par l'algorithme et les faux positifs sont les positions de sources estimées par l'algorithme qui ne sont pas des positions de sources vraies. Un cas particulier est quand

Algorithme 4 Calcul des Vrais Positifs (TP)

Entrée : $\mathcal{E} = \{\mathbf{p}_{\mathcal{E}_j}\}, j \in [1, N_E]$ avec $\mathbf{p}_{\mathcal{E}_j} \in \mathbb{R}^{1 \times 3}$ la position j estimée de la source, $\mathcal{S} = \{\mathbf{p}_{\mathcal{S}_j}\}, j \in [1, N_S]$ avec $\mathbf{p}_{\mathcal{S}_j} \in \mathbb{R}^{1 \times 3}$ la position de la source j , r_a

1. **Initialisation :**

- $\tilde{\mathcal{S}}^{(0)} = \{\}$
- $\mathcal{J}^{(0)} = \mathcal{E}$

2. **pour** $j = 1, \dots, N_S$ **faire :**

- Trouver l'indice i^* de l'ensemble \mathcal{E} qui minimise le problème suivant :

$$i^* = \underset{i=1, \dots, |\mathcal{J}^{(j-1)}|}{\operatorname{argmin}} \|\mathbf{p}_{\mathcal{J}_i^{(j-1)}} - \mathbf{p}_{\mathcal{S}_j}\|_2$$

- **Si** $\|\mathbf{p}_{\mathcal{E}_{i^*}} - \mathbf{p}_{\mathcal{S}_j}\|_2 \leq r_a$ **alors :**

$$\begin{aligned} \rightarrow \tilde{\mathcal{S}}^{(j)} &= \tilde{\mathcal{S}}^{(j-1)} \cup \mathcal{J}_{i^*}^{(j-1)} \\ \rightarrow \mathcal{J}^{(j)} &= \mathcal{J}^{(j-1)} \setminus \mathcal{J}_{i^*}^{(j-1)} \end{aligned}$$

Sortie : $TP = |\tilde{\mathcal{S}}^{(j)}|$

l'algorithme estime deux positions de sources pour une seule source de la vérité terrain. Dans ce cas, la position estimée la plus proche de la vérité terrain est donc considérée comme un vrai positif et l'autre position comme un faux positif. L'indice de Jaccard est décrit dans l'Algorithme 5.

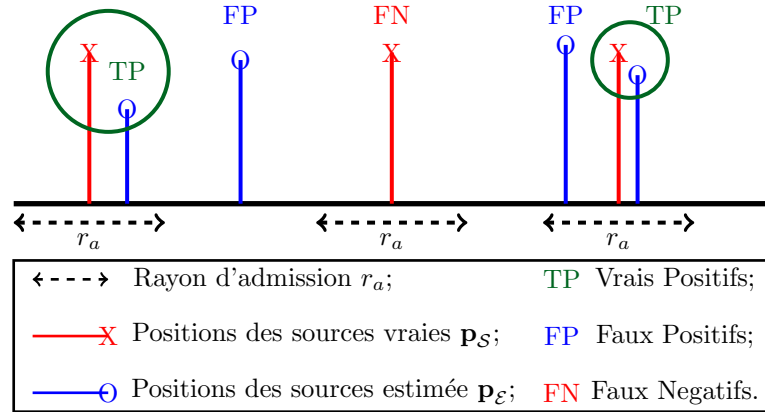


FIGURE 1.8 – Illustration de la définition des de vrais positifs, faux positifs et faux négatifs pour l'indice de Jaccard en fonction du rayon d'admission r_a .

Si le nombre de sources est connu et est utilisé pour définir l'ensemble de positions de sources estimées \mathcal{E} , alors le calcul du nombre de faux positifs et de faux négatifs n'a pas lieu d'être. Il n'est donc pas intéressant d'utiliser l'indice de Jaccard et dans ce cas précis, on se reportera sur

Algorithme 5 Calcul de l'indice de Jaccard (J)

Entrée : $\mathcal{E} = \{\mathbf{p}_{\mathcal{E}_j}\}, j \in [1, N_E]$ avec $\mathbf{p}_{\mathcal{S}_j} \in \mathbb{R}^{1 \times 3}$ la position j estimée de la source, $\mathcal{S} = \{\mathbf{p}_{\mathcal{S}_j}\}, j \in [1, N_S]$ avec $\mathbf{p}_{\mathcal{S}_j} \in \mathbb{R}^{1 \times 3}$ la position de la source j , r_a

1. **Initialisation :**

- $\tilde{\mathcal{S}}^{(0)} = \{\}$
- $\mathcal{J}^{(0)} = \mathcal{E}$

2. **pour** $j = 1, \dots, N_S$ **faire :**

- Trouver l'indice i^* de l'ensemble \mathcal{E} qui minimise le problème suivant :

$$i^* = \underset{i=1, \dots, |\mathcal{J}^{(j-1)}|}{\operatorname{argmin}} \|\mathbf{p}_{\mathcal{J}_i^{(j-1)}} - \mathbf{p}_{\mathcal{S}_j}\|_2$$

- **Si** $\|\mathbf{p}_{\mathcal{E}_{i^*}} - \mathbf{p}_{\mathcal{S}_j}\|_2 \leq r_a$ **alors :**

$$\begin{aligned} \rightarrow \tilde{\mathcal{S}}^{(j)} &= \tilde{\mathcal{S}}^{(j-1)} \cup \mathcal{J}_{i^*}^{(j-1)} \\ \rightarrow \mathcal{J}^{(j)} &= \mathcal{J}^{(j-1)} \setminus \mathcal{J}_{i^*}^{(j-1)} \end{aligned}$$

3. $TP = |\tilde{\mathcal{S}}^{(j)}|$

4. $FP = |\mathcal{J}^{(j)}|$

5. $FN = N_S - TP$

Sortie : $J = \frac{TP}{TP+FP+FN}$

le pourcentage de vrais positifs. Dans le cadre expérimental apporté dans les chapitres 4 et 5, le nombre de sources sera connu, et donc le pourcentage de vrais positifs sera utilisé. Au contraire, dans le cadre des simulations dans le chapitre 3, le nombre de sources sera considéré comme inconnu dans l'évaluation, il sera donc possible d'utiliser l'indice de Jaccard.

1.9 Conclusion

Dans ce chapitre, nous avons amené le contexte de la localisation de sources acoustiques passives en présentant différentes situations : le champ proche et le champ lointain, le milieu libre et le milieu réverbérant. Le problème direct et inverse de localisation de sources ont alors été posés. Le problème inverse est le problème d'optimisation de la localisation de sources acoustiques qu'il nous faut résoudre. Classiquement, la solution est d'utiliser la formation de voies. Cependant cette solution comporte un grand nombre d'inconvénients, notamment sur le dimensionnement d'antenne avec des limites basses et hautes fréquences. Pour éviter ceux-ci, des méthodes dites parcimonieuses, popularisées par l'acquisition comprimée, peuvent être utilisées. La parcimonie dans la distribution des sources est une information que peut posséder l'utilisateur en préparant son expérience. Cette structure parcimonieuse peut s'exprimer sous forme probabiliste, et les

approches bayésiennes peuvent permettre de résoudre le problème de localisation de sources. Les performances des algorithmes utilisés dans la suite du manuscrit seront évaluées par les critères d'évaluation proposées dans ce chapitre.

LA SELECTION DE CAPTEURS

Sommaire

2.1	Introduction	35
2.2	Ecriture du problème	38
2.3	Les méthodes aléatoires	38
2.3.1	Le tirage aléatoire	38
2.3.2	Algorithme génétique	39
2.4	Les critères d'optimisation par conception optimale	44
2.4.1	L'information de Fisher	45
2.4.2	Matrice de covariance	48
2.5	Sélection séquentielle de capteurs	50
2.5.1	Sélection D-optimale	51
2.5.2	Sélection A-optimale	53
2.6	Sélection orientée par les données	54
2.6.1	Sélection séquentielle bayésienne orientée par les données	54
2.6.2	Discussion	55
2.7	Conclusions	56

2.1 Introduction

Les choix effectués dans la manière d'acquérir des mesures lors d'une expérience sont importants, car ils auront un impact sur la qualité des résultats que l'on peut en tirer. Il est nécessaire de savoir quels paramètres du problème sont à estimer, quels paramètres peuvent être contrôlés pour configurer la mesure, et quels paramètres sont aléatoires ou difficilement contrôlables. Cela amène à se poser la question de comment les mesures doivent être effectuées, autrement dit, cela amène à faire un plan d'expérience ("Experimental Design") [73]. Le plan d'expériences est multi-disciplinaire. Il est notamment très présent en psychologie [74, 75], en biologie [76], en sociologie [77], en agriculture [78, 79] ou encore en médecine [80]. Le plan d'expérience permet, dans le domaine médical, de mener des tests cliniques dans les meilleures conditions afin que les variables estimées et les interprétations qui en découlent soient les plus précises possibles.

En acoustique, le facteur décisif est la mesure acquise par les capteurs. Quand on prépare une expérience acoustique, il est nécessaire de placer ses capteurs. On peut chercher à :

- réduire le nombre de données,
- diminuer la complexité calculatoire et ainsi utiliser des algorithmes plus performants,
- pallier à des contraintes matérielles et économiques en diminuant le nombre de capteurs.

La réduction du nombre de capteurs par l'optimisation de leur position est donc un avantage non négligeable lors de la mise en place de l'expérience. La sélection de mesures peut se retrouver dans de nombreuses applications, telles que la propagation modale dans une structure [81], la cartographie de température d'un bâtiment par des mesures thermiques [82], des mesures de bruit d'un moteur pour connaître la directivité et les caractéristiques de celui-ci [83], ou la localisation de sources électromagnétiques ou acoustiques [84].

Dans le cadre de la localisation de sources acoustiques, l'utilisation d'antennes uniformes (ou les capteurs sont espacés de la même distance) reste courant, même s'il existe un nombre important de géométries d'antennes proposées dans la littérature ayant des meilleures performances dans des domaines précis [85, 86]. Cependant il a été remarqué que certaines antennes aléatoires permettent d'avoir des performances plus précises en localisation que des antennes uniformes [87], notamment en réduisant le niveau des lobes secondaires [42]. Il s'avère aussi que l'utilisation d'antennes aléatoires permet de s'approcher des conditions requises en acquisition comprimée [88]. Elles permettent d'améliorer la décohérence entre les mesures. Ceci permet de réduire le nombre de capteurs utilisés en les plaçant aléatoirement [89]. En revanche le choix aléatoire des positions de capteurs ne garantit pas la localisation des sources. Pour éviter des faibles performances en localisation, il est nécessaire de contrôler le positionnement des capteurs et donc de déterminer un critère de sélection de positions de capteurs.

Le critère pour le positionnement des capteurs doit être en lien avec l'algorithme utilisé en localisation de sources afin d'obtenir une meilleure estimation des positions des sources. En localisation de sources acoustiques, le principal algorithme utilisé est la formation de voies. Dans la partie 1.4.2, on a vu 3 critères liés à la formation de voies : le recouvrement spatial, la résolution et la dynamique. Le premier ne peut pas être utilisé, notamment pour les antennes dont les distances inter-capteurs ne sont pas constantes. La résolution dépend de la fréquence de la source et de l'ouverture d'antenne. Or par la suite du manuscrit, nous avons choisi de fixer l'ouverture d'antenne. Un choix cohérent avec un cadre expérimental où l'ouverture d'antenne est souvent contrainte par des limites techniques ou financières. La résolution ne peut donc pas être utilisée comme critère de positionnement des capteurs. La dynamique est un critère fortement impacté par la position des capteurs et donc a été choisi dans de nombreux articles [90, 91, 92]. Réduire le niveau des lobes secondaires permet une meilleure précision dans la localisation, et permet aussi d'identifier plusieurs sources d'amplitudes fortement variées. Pour optimiser ce critère, un algorithme utilisé originalement en biologie est l'algorithme génétique

[93, 94]. Avec ce dernier il est possible de tester un nombre important d'antennes aléatoires mais aussi de ne conserver que les meilleures antennes. Cela permet de converger vers une antenne aléatoire unique qui optimise au mieux le critère choisi : la dynamique [95]. Cependant l'algorithme génétique peut s'avérer long puisque le nombre d'antennes aléatoires à tester est souvent grand. L'autre inconvénient est qu'il est lié à la dynamique et donc à la formation de voies. Si l'algorithme de localisation de sources n'est pas la formation de voies, l'intérêt de cet algorithme devient moindre.

Dans le domaine du plan d'expériences, une catégorie d'approches a retenu notre attention : la conception optimale ("Optimal Design") [96, 97]. Créée en 1918 par Kirstine Smith [98], la conception optimale cherche à optimiser un critère lié à la matrice de covariance de l'estimateur choisi, supposé non biaisé (c'est-à-dire que la valeur moyenne estimée est identique à la vraie valeur). En général, la variance d'un estimateur non biaisé diminue avec le nombre de mesures, mais il est possible de diminuer cette variance plus rapidement en sélectionnant les mesures intelligemment. Sachant que les mesures sont liées aux positions des capteurs en localisation de sources acoustiques, la conception optimale consiste à choisir les positions.

La conception optimale est un problème d'optimisation combinatoire et nécessite pour sa résolution des approximations. Le problème d'optimisation de sélection de capteurs peut être résolu soit par une relaxation du problème [99, 100], soit par une approche séquentielle [64]. Les deux versions seront présentées mais l'approche choisie par la suite du manuscrit sera l'approche séquentielle dans laquelle les positions de capteurs seront sélectionnées une par une. Il est alors possible d'apporter de l'information dans le calcul du critère de sélection de capteurs sous forme probabiliste : on appelle cela le plan d'expérience bayésien ("Bayesian Experimental Design") [101]. L'information apportée concerne les positions probables des sources et permet donc de sélectionner les positions de capteurs qui fournissent le plus d'informations pour les localiser.

L'information apportée dans le plan d'expérience bayésien peut aussi provenir des mesures prises par les capteurs. On parle alors de sélection de capteurs orientée par les données. Ces approches utilisent le modèle de propagation des ondes acoustiques comme les méthodes citées précédemment, mais utilisent en plus les mesures déjà obtenues dans le passé [102, 103]. L'ajout de l'information contenue dans les mesures est discutable au niveau de l'application des approches de sélection de capteurs puisque celles-ci nécessitent donc de sélectionner au préalable des capteurs et d'acquérir des mesures avant d'en placer de nouveaux. Or, il s'avère que ces méthodes permettent de réduire le nombre de capteurs fortement et correspondent à certaines applications acoustiques [104, 83]. Une discussion sur ces applications est présente dans ce chapitre.

Dans la suite du chapitre, nous posons tout d'abord le problème d'optimisation qu'est la sélection de capteur (section 2.2). Les différentes approches de sélection sont ensuite décrites : les méthodes aléatoires avec l'algorithme génétique dans la section 2.3, la conception optimale

avec les différents critères dans la section 2.4, la sélection séquentielle de positions de capteurs dans la section 2.5 et finalement la sélection séquentielle orientée par les données dans la section 2.6. C’est dans cette dernière partie que la conception d’antenne proposée dans ce manuscrit est exposée. Puis nous concluons dans la section 2.7.

2.2 Ecriture du problème

Dans le contexte du chapitre 1, on a implicitement considéré l’utilisation de la totalité des capteurs disponibles $\mathcal{M} = \{\mathbf{p}_{\mathcal{M}_i}\}, i \in \llbracket 1, N_M \rrbracket$ avec $\mathbf{p}_{\mathcal{M}_i} \in \mathbb{R}^{1 \times 3}$. Dans le domaine de la sélection de capteurs, on cherche à sélectionner N_K positions de capteurs dans l’ensemble \mathcal{M} . On définit alors un ensemble de positions de capteurs sélectionnées tel que $\mathcal{K} = \{\mathbf{p}_{\mathcal{K}_i}\}, i \in \llbracket 1, N_K \rrbracket$ avec $\mathbf{p}_{\mathcal{K}_i} \in \mathbb{R}^{1 \times 3}$ et $\mathcal{K} \subset \mathcal{M}$. En toute rigueur, l’ensemble de positions de capteurs sélectionnées \mathcal{K} peut remplacer l’ensemble de positions de capteurs total \mathcal{M} dans toutes les équations du chapitre 1, car les équations et algorithmes fonctionnent de la même manière. Par exemple, le problème de localisation de sources au sens des moindres carrés, vu dans l’équation (1.13), deviendrait :

$$\mathbf{s}_{\tilde{\mathcal{K}}, \mathcal{G}} = \underset{\mathbf{s}_{\mathcal{K}, \mathcal{G}}}{\operatorname{argmin}} \|\mathbf{m}_{\mathcal{K}, \mathcal{G}} - \mathbf{A}_{\mathcal{K}, \mathcal{G}} \mathbf{s}_{\mathcal{K}, \mathcal{G}}\|_2^2. \quad (2.1)$$

La sélection de positions de capteurs a pour but de réduire le nombre de capteurs utilisés (ainsi $N_K < N_M$) tout en dégradant le moins possible les performances en localisation de sources. Il faut donc trouver les N_K positions de capteurs et ainsi construire l’ensemble \mathcal{K} apportant la meilleure estimation du vecteur source. Nous verrons dans le chapitre 3 qu’il est même possible, dans certaines situations, de réduire le nombre de capteurs utilisés tout en conservant la même qualité d’estimation du vecteur source estimé.

2.3 Les méthodes aléatoires

2.3.1 Le tirage aléatoire

Dans la localisation de sources, on a vu que pour une bonne utilisation de la formation de voies, le nombre de capteurs nécessaires dans une antenne régulière doit permettre de respecter le théorème de Shannon-Nyquist (cf. équation (1.17)), ce qui peut amener à un grand nombre de capteurs si la fréquence est élevée. Afin de diminuer ce nombre de capteurs, il a été remarqué qu’utiliser une antenne avec un placement aléatoire des capteurs peut apporter de meilleures performances. En effet, puisque la distance inter-capteur n’est plus constante, le théorème de Shannon-Nyquist présenté dans le chapitre 1 (voir équation (1.16)) ne peut plus être pris en compte. Si l’antenne n’est plus uniforme, la dynamique et la forme des lobes secondaires ne sont plus les mêmes, permettant ainsi une meilleure distinction des sources [105, 106]. De plus

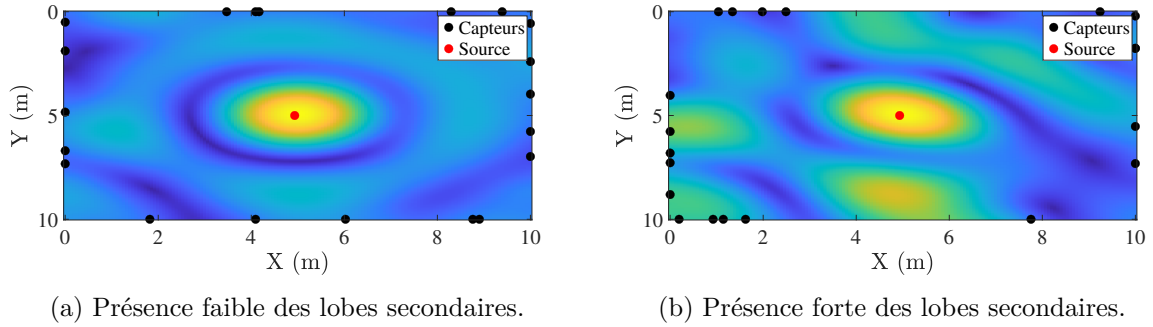


FIGURE 2.1 – Localisation de sources par formation de voies en champ proche par deux antennes de 32 capteurs de géométrie aléatoire différentes.

l'antenne aléatoire est très utilisée dans le domaine de l'acquisition comprimée [88].

Cependant, la différence de performance entre les antennes aléatoires peut s'avérer importante, comme le montre la figure 2.1 avec une localisation de source par formation de voies via deux antennes aléatoires différentes. On pourrait imaginer, en théorie, générer une multitude d'antennes aléatoires, et conserver la plus performante d'entre elles par rapport à un certain critère. Mais cette approche n'a en pratique que peu d'intérêt, car elle est peu efficace et coûteuse en temps de calcul.

2.3.2 Algorithme génétique

Un compromis entre la sélection aléatoire et une sélection optimisée est l'exploitation d'un algorithme génétique [93, 94]. Les algorithmes génétiques sont utilisés dans un nombre important d'applications différentes [107, 108]. On y retrouve par exemple l'identification modale ou de la surveillance de la santé [109, 110]. Mais l'utilisation de cet algorithme en sélection de capteurs pour de la localisation de sources acoustiques est rare. On trouve cependant quelques exemples dans la littérature, notamment en acoustique aérienne [95, 111].

L'algorithme génétique s'initialise par une population d'individus, appelés des parents. Chaque individu possède un ensemble de gènes, tiré aléatoirement parmi un "pool" de gènes disponibles. Chaque individu est ensuite évalué selon un critère. Puis vient l'étape de sélection : dans celle-ci on croise 2 parents, tirés aléatoirement, qui donneront 1 enfant avec son propre ensemble de gènes. Les gènes de l'enfant sont sélectionnés selon les gènes des deux parents (par hérédité) selon l'évaluation de chacun des parents. L'enfant aura tendance à prendre les gènes du parent avec la meilleure évaluation du critère. Puis on considère une mutation dans les gènes de l'enfant. Les gènes de l'enfant peuvent être modifiés, avec une probabilité faible définie comme paramètre d'entrée de la méthode. Les étapes d'héritage et de mutation sont répétées jusqu'à obtenir autant d'enfants que de parents. Cela permet de conserver le même nombre d'individus dans la population. Les parents sont retirés de la population et les enfants deviennent ensuite

les parents. La génération est alors finie. On répète le processus pour un nombre arbitraire de générations. La valeur moyenne du critère d'évaluation des individus a tendance à converger vers celle de l'individu le plus adapté. On parle alors d'uniformisation croissante de la population. L'algorithme finit par une dernière étape de sélection si celui-ci n'a pas convergé vers une population d'individus identiques. Cette dernière étape se résume généralement par une sélection de l'individu avec la meilleure évaluation. Le fonctionnement de l'algorithme est résumé à la figure 2.2.

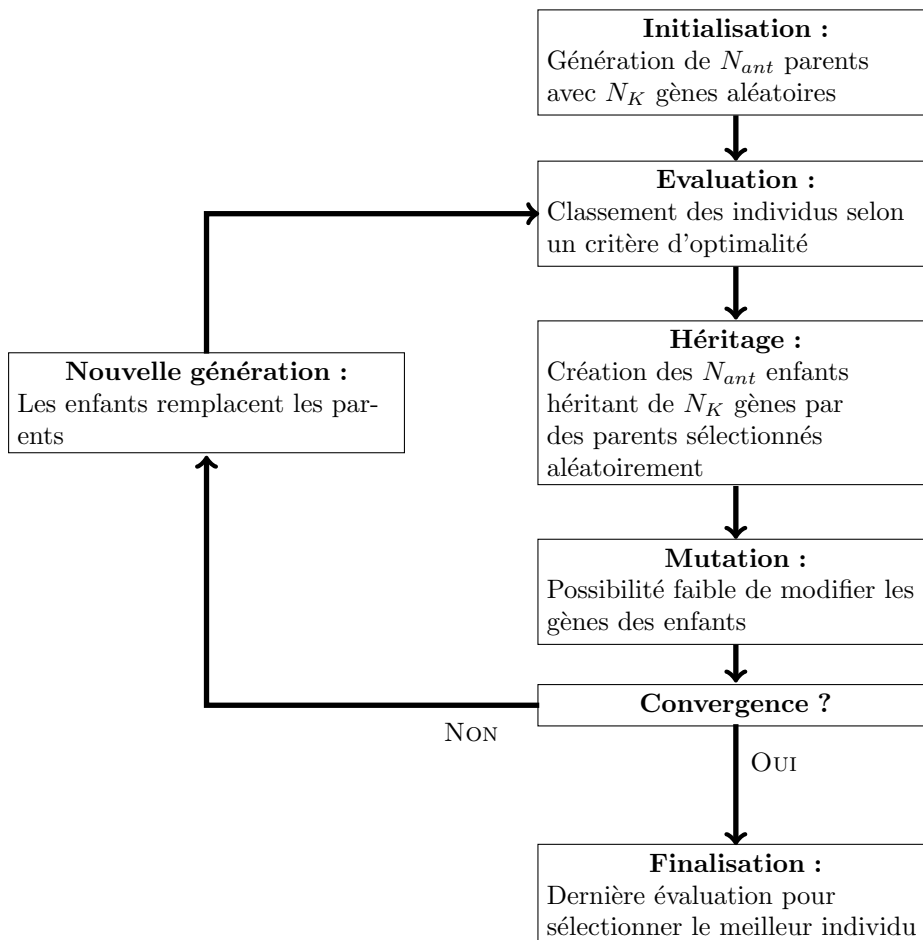


FIGURE 2.2 – Les différentes étapes de l'algorithme génétique.

L'algorithme génétique fera partie des méthodes de l'état de l'art qui sera utilisée par la suite pour comparer les performances des méthodes de sélection de capteurs. Pour cela nous allons définir les étapes de la méthode génétique dans le cadre de la localisation de sources en explicitant chacune des étapes et leurs équations :

Initialisation

On génère un tirage aléatoire de N_{ant} antennes (individus) contenant chacune N_K positions de capteurs (gènes) parmi un groupe de N_M positions de capteurs disponibles (pool de gènes).

Évaluation

Chaque antenne parent est évaluée selon un critère. Un nombre important de critères existe pour évaluer les antennes [112, 113]. Ce critère doit être adapté par rapport à l'application choisie. Pour cela, dans la localisation de sources, on peut penser à la largeur du lobe principal ou le niveau des lobes secondaires en ce qui concerne la formation de voies [95]. Dans la suite du manuscrit, le choix du critère se porte sur la dynamique ou la différence entre le lobe principal et le lobe secondaire lors du calcul de la directivité de l'antenne (autrement appelé *MSL* pour *Maximum Sidelobes Level*). Pour chaque antenne, on calcule sa directivité selon chaque position de source tel que :

$$Dir = |(\mathbf{A}_{\mathcal{K},S})^H \mathbf{A}_{\mathcal{K},S}|. \quad (2.2)$$

Ensuite on en déduit le niveau des lobes secondaires en prenant le niveau du second pic le plus élevé, pour chacune des positions de source (le premier pic étant le lobe principal). On obtient donc un niveau des lobes secondaires pour chaque source, que l'on moyenne par la suite. Puis on normalise l'évaluation *MSL* des antennes entre 0 et 1, avec 1 pour l'antenne qui optimise au mieux le critère (le *MSL*) et 0 pour celle qui l'optimise le moins.

Héritage

À partir de l'évaluation des antennes parents, il est possible de concevoir les antennes enfants qui hériteront leurs positions des capteurs des antennes parents. Afin de concevoir une antenne enfant, on sélectionne aléatoirement deux antennes parents. Puis on sélectionne aléatoirement une position de capteur de chaque antenne parent. L'antenne enfant prend la position de capteur du parent 1 si l'inégalité suivante est respectée :

$$\varsigma \leq MSL_1 \cdot (1 - MSL_2), \quad (2.3)$$

avec ς une valeur aléatoire suivant une loi uniforme entre 0 et 1. MSL_1 et MSL_2 représentent le critère d'évaluation normalisé du parent 1 et 2. Si l'inégalité n'est pas respectée l'antenne enfant prendra la position du capteur du parent 2. Il est donc possible que la nouvelle antenne hérite d'une position de capteur moins performante pour optimiser le critère. On réitère cette opération pour chaque position de capteur de l'antenne enfant avec les deux mêmes antennes parents. Lors de cette démarche, il faut faire attention à ne pas prendre 2 fois la même position de capteur. L'héritage est ensuite répétée pour chaque antenne enfant afin que chaque génération conserve

le même nombre d’antennes.

Mutation

La mutation permet d’explorer d’autres positions de capteur qui ne seraient pas présentes dans l’initialisation. Un coefficient de mutation χ est choisi arbitrairement entre 0 et 1. Plus il sera élevé, plus la position de capteur aura de chance de se changer en une autre. Pour chaque position de capteur de chaque antenne enfant, on compare une valeur aléatoire suivant une loi uniforme entre 0 et 1 et le coefficient de mutation. Si la valeur aléatoire est inférieure au coefficient, alors la position de capteur de l’antenne est modifiée. La modification est simplement une sélection aléatoire d’une position de capteur parmi les N_M disponibles.

Nouvelle génération

Les antennes enfants remplacent les antennes parents. On répète ensuite les étapes d’évaluations, d’héritage et de mutation jusqu’à convergence. On considère que l’algorithme a convergé si la valeur moyenne du critère d’optimisation (dans notre cas le niveau moyen des lobes secondaires) ne varie plus au cours des générations. Cependant, en général, on fixe un nombre maximal de générations N_{gen} .

Finalisation

Généralement les algorithmes génétiques ont tendance à ce que les N_{ant} antennes convergent vers l’antenne la plus adaptée au problème et donc soient identiques. Mais pour atteindre ce résultat, un nombre important de générations peut être nécessaire et donc cela implique un temps de calcul important. Cette dernière étape permet d’éviter ce problème.

Après un nombre fixé N_{gen} de générations, on possède N_{ant} antennes, or il nous en faut qu’une seule. Pour cela on réévalue les N_{ant} antennes de la même manière que dans la partie **Évaluation** de l’algorithme et on sélectionne l’antenne avec la meilleure évaluation. Cette antenne donne donc N_K positions de capteur et l’ensemble \mathcal{K} . Il est maintenant possible de résoudre le problème de localisation de source.

L’algorithme génétique de sélection de capteurs est résumé dans l’Algorithme 6.

Algorithme 6 Sélection de positions de capteurs par algorithme génétique

Entrée : N_{gen} le nombre de générations ou d’itérations de l’algorithme, N_{ant} le nombre d’antennes aléatoires (individus), \mathcal{M} l’ensemble contenant toutes les positions de capteurs disponibles, $\mathbf{A}_{\mathcal{M},\mathcal{S}} \in \mathbb{C}^{N_M \times N_S}$ la matrice de propagation, N_K le nombre de positions de capteurs sélectionnées, χ un coefficient de mutation entre 0 et 1.

1. Initialisation :

- Création de N_{ant} antennes aléatoires avec N_K positions de capteurs sélectionnées aléatoirement dans l'ensemble \mathcal{M} .
2. **pour** $i = 1, 2, \dots, N_{gen}$ **faire** :
- **Évaluation** :
 Pour chaque antenne (parent) :
 - Calcul de la directivité par l'équation (2.2).
 - À partir de la directivité, mesurer le niveau (en dB) du lobe principal et du lobe secondaire.
 - Calcul du MSL par soustraction des deux niveaux.
 - **Normalisation** :
 → Normalisation des MSL en mettant l'antenne avec le MSL le plus élevé (respectivement le plus petit) à 1 (respectivement à 0).
 - **Héritage** :
 Pour chaque nouvelle antenne (enfant) :
 - Sélection aléatoire de 2 antennes (parents)
 - Pour chaque position de capteur (gène) de la nouvelle antenne :
 - (a) Génération d'une valeur aléatoire ς suivant une loi uniforme entre 0 et 1.
 - (b) Comparaison de la valeur aléatoire ς avec les MSL des deux parents. Si l'inégalité de l'équation (2.3) :

$$\begin{cases} \text{est respectée, alors prendre la position de capteur du parent 1} \\ \text{n'est pas respectée, alors prendre la position de capteur du parent 2} \end{cases}$$
 - Vérification que la nouvelle antenne ne possède pas 2 fois la même position de capteur :
 Si cette condition n'est pas respectée, on remplace les doublons par des positions de capteurs sélectionnées aléatoirement dans l'ensemble \mathcal{M} .
 - **Mutation** :
 Pour chaque nouvelle antenne (enfant) :
 - Pour chaque position de capteur (gène) de la nouvelle antenne :
 - (a) Tirage d'une valeur aléatoire ς entre 0 et 1
 - (b) Comparaison de la valeur aléatoire ς avec le coefficient de mutation χ . Si :
 - $\varsigma < \chi$, alors la position de capteur est modifiée en sélectionnant aléatoirement dans l'ensemble \mathcal{M}

– $\varsigma \geq \chi$, alors la position de capteur est conservée.

• **Nouvelle Génération :**

→ Les nouvelles antennes (enfants) remplacent les anciennes (parents).

3. **Finalisation :**

Si toutes les N_{ant} antennes ne sont pas identiques :

- Réitération de l'étape **Évaluation**
- Sélection de l'antenne avec le MSL normalisé à 1.
- Création de l'ensemble \mathcal{K} contenant les positions de capteurs de l'antenne sélectionnée.

Sortie : \mathcal{K}

2.4 Les critères d'optimisation par conception optimale

Si l'algorithme de localisation de sources n'est pas la formation de voies, alors la sélection génétique décrite dans la section précédente n'est pas la plus adéquate. En effet son critère d'optimisation (niveau MSL) est pertinent pour un traitement des mesures par formation de voies, mais pas pour une approche basée sur les méthodes parcimonieuses. De plus l'aspect aléatoire très présent dans l'algorithme génétique peut être contraignant s'il est nécessaire de tester un grand nombre d'antennes. Il semble donc plus intéressant de trouver une méthode donnant les positions de capteurs informatives pour la localisation de sources. Pour cela on présente la conception optimale.

En conception optimale, on considère que l'estimateur de localisation de sources est non biaisé. Un estimateur peut-être caractérisé par sa moyenne (*i.e.* son biais) et sa variance. La moyenne ne peut pas être le facteur déterminant pour la sélection de positions de capteur si l'estimateur est non biaisé. On utilise alors la variance de l'estimateur. La sélection va chercher à réduire cette variance, et idéalement la rapprocher de la borne de Cramer-Rao (borne inférieure sur la variance d'un estimateur non biaisé) [114, 115]. Le but de la sélection de positions de capteurs est donc de rendre l'estimateur le plus efficace possible tout en respectant une contrainte sur le nombre de capteurs à utiliser. Il s'avère ici que deux possibilités sont à présenter : minimiser la variance de l'estimateur, ou maximiser l'information de Fisher, c'est-à-dire, l'inverse de la borne de Cramer-Rao.

Dans cette partie du chapitre, on commence par présenter la matrice de Fisher, son lien avec la borne de Cramer-Rao et comment il est possible de "réduire" cette borne quand il s'agit d'une matrice et non d'un scalaire. Puis le lien avec la minimisation de la matrice de covariance est introduit ainsi que l'avantage de l'utiliser par rapport à la matrice de Fisher.

Notons que des liens avec la maximisation de la divergence de Kullback-Leibler [116] et la minimisation de l'entropie de Shannon [117] existent et sont donnés dans l'Annexe A. D'autres

critères d'optimalité sont présentés dans un article de Koyama et al. [118] qui les compare dans un contexte précis : le contrôle du champ sonore. Mais ceux-ci sont trop éloignés des critères utilisés dans ce manuscrit afin de faire un quelconque lien ou de les présenter.

2.4.1 L'information de Fisher

L'information de Fisher est une notion de statistique introduite par Ronald Aylmer Fisher ?? qui quantifie les données relatives à un paramètre contenues dans une distribution. Elle est définie comme l'espérance d'un score. Dans le cas de plusieurs paramètres, on parle de matrice d'information de Fisher. Commençons par la définition de l'information de Fisher dans le cas avec un seul paramètre.

Soit $f(\mathbf{v}; \psi)$ la distribution de vraisemblance d'une variable multidimensionnelle \mathbf{v} paramétrée par ψ , qui est doublement différentiable par rapport à ψ pour tout \mathbf{v} . Le score est défini comme la dérivée partielle du logarithme de la distribution de vraisemblance par rapport au paramètre ψ :

$$\frac{\partial}{\partial \psi} \log f(\mathbf{v}; \psi) = \frac{1}{f(\mathbf{v}; \psi)} \frac{\partial f(\mathbf{v}; \psi)}{\partial \psi}. \quad (2.4)$$

Il est possible de montrer que l'espérance de la fonction de score est nulle :

$$\begin{aligned} \mathbb{E} \left[\frac{\partial}{\partial \psi} \log f(\mathbf{v}; \psi) | \psi \right] &= \int_{\Omega(\mathbf{v})} \frac{\frac{\partial}{\partial \psi} f(\mathbf{v}; \psi)}{f(\mathbf{v}; \psi)} f(\mathbf{v}; \psi) d\mathbf{v} \\ &= \frac{\partial}{\partial \psi} \int_{\Omega(\mathbf{v})} f(\mathbf{v}; \psi) d\mathbf{v} \\ &= \frac{\partial}{\partial \psi} \int_{\Omega(\mathbf{v})} 1 d\mathbf{v} \\ &= 0. \end{aligned} \quad (2.5)$$

avec $\Omega(\mathbf{v})$ le domaine d'intégration de la variable aléatoire.

L'information de Fisher est définie comme la variance de la fonction de score :

$$\begin{aligned} Fi(\psi) &= \mathbb{E} \left[\left(\frac{\partial}{\partial \psi} \log f(\mathbf{v}; \psi) \right)^2 | \psi \right] \\ &= -\mathbb{E} \left[\frac{\partial^2}{\partial \psi^2} \log f(\mathbf{v}; \psi) | \psi \right]. \end{aligned} \quad (2.6)$$

Si on considère maintenant N_γ paramètres $\boldsymbol{\psi} \in \mathbb{C}^{N_\gamma \times 1}$, l'information de Fisher devient une matrice $\mathbf{Fi} \in \mathbb{C}^{N_\gamma \times N_\gamma}$. L'équation (2.6) devient donc :

$$Fi(\boldsymbol{\psi})_{i,j} = \mathbb{E} \left[\left(\frac{\partial}{\partial \psi_i} \log f(\mathbf{v}; \boldsymbol{\psi}) \right)^H \left(\frac{\partial}{\partial \psi_j} \log f(\mathbf{v}; \boldsymbol{\psi}) \right) | \boldsymbol{\psi} \right]. \quad (2.7)$$

La question à se poser ici est la suivante : à quoi correspond $\boldsymbol{\psi}$ en localisation de sources ? $\boldsymbol{\psi}$ correspond à toutes les inconnues du problème de localisation de sources. Dans le cas d'une source unique, cela correspond généralement à :

- la position de la source et donc à 2 inconnues en champ lointain (θ et ϕ) et à 3 inconnues en champ proche dans le volume (X, Y, Z),
- l'amplitude complexe de la source, pouvant être paramétrée via son module et sa phase.

La matrice de Fisher de l'équation (2.7) est calculée pour une position de capteur. Pour considérer toute l'antenne, on considère que les mesures acquises par chaque capteur sont indépendantes. Par additivité de l'information de Fisher, on somme les matrices de Fisher pour chaque position de capteur de l'antenne :

$$\mathbf{F}_{\mathcal{M}}(\boldsymbol{\psi}) = \sum_{i=1}^{N_M} \mathbf{F}_{\mathbf{i}_i}(\boldsymbol{\psi}). \quad (2.8)$$

Avec la matrice de Fisher, le problème d'optimisation de sélection de position de capteurs se définit de la manière suivante :

$$\mathcal{K} = \underset{\mathcal{K}}{\operatorname{argmax}} f_s \left(\sum_{k=1}^{N_K} \mathbf{F}_{\mathbf{i}_k}(\boldsymbol{\psi}) \right), \quad (2.9)$$

avec $f_s(\cdot)$ un scalaire calculé à partir de la matrice de Fisher $\mathbf{F}_{\mathbf{i}_k}$ correspondant à la position de capteur k de l'ensemble \mathcal{K} . À partir de cette équation, il est important de déterminer à quoi correspond la matrice de Fisher, le scalaire déduit de cette matrice (appelés critère d'optimalité) et comment on résout le problème d'optimisation.

Notons qu'actuellement, la sélection de capteurs proposée dans l'équation (2.9) n'est optimisée que pour un seul vecteur de paramètres $\boldsymbol{\psi}$, c'est-à-dire, que pour une seule position de source et une seule amplitude complexe.

Le calcul de la matrice de Fisher

On considère en localisation de sources que le vecteur de la variable aléatoire correspond au vecteur d'amplitude des sources, $\mathbf{v} = \mathbf{s}_{\mathcal{M}, \mathcal{G}}$. La fonction de vraisemblance, qui elle suit une loi gaussienne multivariée complexe [119, 120], est :

$$f(\mathbf{v}; \boldsymbol{\psi}) = \pi^{-N_G} |\boldsymbol{\Sigma}|^{-1} \exp^{-(\mathbf{v} - \mathbb{E}[\mathbf{v}])^H \boldsymbol{\Sigma}^{-1} (\mathbf{v} - \mathbb{E}[\mathbf{v}])}. \quad (2.10)$$

Si l'estimation du vecteur d'amplitude des sources est réalisé par maximum de vraisemblance (voir équation (1.28)), et que le bruit est complexe gaussien de variance α_0^{-1} , et indépendant

entre les capteurs, alors la fonction de vraisemblance correspond à [119] :

$$f(\tilde{\mathbf{s}}_{\mathcal{K},\mathcal{G}}; \boldsymbol{\psi}) \propto \exp^{-\alpha_0(\tilde{\mathbf{s}}_{\mathcal{K},\mathcal{G}} - \mathbb{E}[\tilde{\mathbf{s}}_{\mathcal{K},\mathcal{G}}])^H(\tilde{\mathbf{s}}_{\mathcal{K},\mathcal{G}} - \mathbb{E}[\tilde{\mathbf{s}}_{\mathcal{K},\mathcal{G}}])}. \quad (2.11)$$

À partir de cette équation, pour calculer la dérivée partielle de l'équation (2.7) il faut choisir si le modèle de propagation suit le modèle de champ proche ou de champ lointain.

Les choix de l'optimalité

En conception optimale, trois critères d'optimalité sont classiquement utilisés dans la littérature : le critère A-optimal, le D-optimal et E-optimal [96, 97]. La différence entre ces critères d'optimalité se trouve dans le choix du scalaire qui définit l'information de la matrice de Fisher. En A-optimal, c'est la trace de la matrice :

$$\mathcal{K} = \operatorname{argmax}_{\mathcal{K}} \operatorname{tr}\left(\sum_{k=1}^{N_K} \mathbf{F}\mathbf{i}_k(\boldsymbol{\psi})\right), \quad (2.12)$$

le déterminant pour le D-optimal :

$$\mathcal{K} = \operatorname{argmax}_{\mathcal{K}} \det\left(\sum_{k=1}^{N_K} \mathbf{F}\mathbf{i}_k(\boldsymbol{\psi})\right), \quad (2.13)$$

et la plus petite valeur propre de la matrice pour le E-optimal :

$$\mathcal{K} = \operatorname{argmax}_{\mathcal{K}} \operatorname{VP}_{\min}\left(\sum_{k=1}^{N_K} \mathbf{F}\mathbf{i}_k(\boldsymbol{\psi})\right). \quad (2.14)$$

La méthode de résolution

Maintenant que l'équation (2.8) est définie en connaissant l'expression de la matrice de Fisher et en choisissant le critère d'optimalité, il faut choisir comment résoudre le problème d'optimisation. Le problème dans l'équation (2.8) est combinatoire puisqu'il faudrait tester toutes les possibilités de sélection de positions de capteurs. Il serait possible de résoudre ce problème par algorithme génétique comme proposé dans la partie 2.3.2. Cependant la complexité calculatoire de la matrice de Fisher est importante et l'algorithme qui nécessite de calculer le critère d'optimisation pour chaque antenne aléatoire à chaque génération risque d'être trop long à calculer. Il est plus intéressant de réécrire le problème sous forme relaxée. Pour cela on introduit une nouvelle variable $\boldsymbol{\omega} \in \{0, 1\}^{N_M}$, un vecteur de pointage binaire permettant de sélectionner dans l'ensemble \mathcal{M} les positions de capteurs. Notons que la condition $\|\boldsymbol{\omega}\|_0 = N_K$ permet de ne sélectionner que le nombre de positions de capteurs souhaité. On réécrit le problème d'optimisation

de sélection de positions de capteurs avec la matrice de Fisher sous la forme suivante :

$$\tilde{\omega} = \underset{\omega=\{0,1\}^{N_M}, \|\omega\|_0=N_K}{\operatorname{argmax}} f_s\left(\sum_{m=1}^{N_M} \omega_m \mathbf{F}i_m(\boldsymbol{\psi})\right). \quad (2.15)$$

Le problème d'optimisation sous cette forme est combinatoire. Par conséquent, on la transforme sous sa forme relaxée :

$$\tilde{\omega} = \underset{\omega=[0,1]^{N_M}, \|\omega\|_1=N_K}{\operatorname{argmax}} f_s\left(\sum_{m=1}^{N_M} \omega_m \mathbf{F}i_m(\boldsymbol{\psi})\right). \quad (2.16)$$

Le vecteur de pointage $\tilde{\omega}$ est ensuite seuillé, permettant ainsi de ne sélectionner que N_K positions de capteur. Résoudre l'équation (2.16) est possible avec la toolbox CVX, par exemple comme le fait Chepuri dans [121]. Cependant deux problèmes se posent : le problème doit être surdéterminé, donc le nombre de capteurs doit être supérieur aux nombre d'inconnues, et CVX nécessite des dimensions assez faibles pour pouvoir effectivement résoudre le problème.

Rappelons que les méthodes présentées jusqu'à ce point sont des approches de l'état de l'art. Une de ces méthodes, l'approche relaxée qui maximise la plus petite valeur propre de la matrice de Fisher [121], sera utilisée par la suite pour comparer les performances avec la méthode proposée dans ce manuscrit.

2.4.2 Matrice de covariance

Cette section traite du problème de sélection de capteurs dans le cas où on cherche à réduire la variance de l'estimateur plutôt que de maximiser l'information capturée par les capteurs. La borne de Cramer-Rao est la variance minimale que peut atteindre un estimateur et équivaut à l'inverse de la matrice de Fisher. Dans un problème multivarié, la matrice de covariance de tout estimateur est supérieure, dans le sens des matrices semi-définies positives, à cette borne :

$$\boldsymbol{\Sigma} \succeq \mathbf{F}i^{-1}(\boldsymbol{\psi}). \quad (2.17)$$

Il est tout à fait possible de s'appuyer sur $\boldsymbol{\Sigma}$ plutôt que sur $\mathbf{F}i(\boldsymbol{\psi})$, dans le cadre de la sélection de capteurs. Le problème d'optimisation de la sélection de positions de capteurs devient donc le suivant :

$$\mathcal{K} = \underset{\mathcal{K}}{\operatorname{argmin}} f_s(\boldsymbol{\Sigma}_{\mathcal{K}}), \quad (2.18)$$

avec $\boldsymbol{\Sigma}_{\mathcal{K}}$ la matrice de covariance estimée avec les positions de capteurs sélectionnées. Similairement à la matrice de Fisher, on peut décomposer cette équation en 3 parties : le calcul de la matrice de covariance, les critères d'optimalité et la méthode de résolution.

Le calcul de la matrice de covariance

Pour calculer la matrice de covariance, dans un contexte ne mettant en jeu que des lois normales multivariées, deux solutions sont proposées : le maximum de vraisemblance et le maximum a posteriori. Ces deux méthodes sont données dans les équations (1.29) et (1.33). La différence entre les 2 approches de calcul de la matrice de covariance est la présence de la distribution de probabilité a priori dans le maximum a posteriori. Dans le cas sous-déterminé, l'inversion présente dans le maximum a posteriori est possible contrairement au maximum de vraisemblance et à la matrice de Fisher. Un avantage de la matrice de covariance par rapport à la matrice de Fisher est que la dimension de la matrice de covariance ne dépend pas du nombre de sources ou de capteurs. Il n'y a pas besoin de calculer la matrice de covariance pour chaque position de capteur et sa dimension ne dépend que du nombre d'inconnues N_G .

Les choix d'optimalité

Les critères d'optimalité sont identiques à ceux de la matrice de Fisher. Le critère A-optimal cherche à minimiser la trace de la matrice de covariance :

$$\mathcal{K} = \underset{\mathcal{K}}{\operatorname{argmin}} \operatorname{tr}(\mathbf{\Sigma}_{\mathcal{K}}), \quad (2.19)$$

le critère D-optimal minimise son déterminant :

$$\mathcal{K} = \underset{\mathcal{K}}{\operatorname{argmin}} \det(\mathbf{\Sigma}_{\mathcal{K}}), \quad (2.20)$$

et le critère E-optimal minimise sa plus grande valeur propre :

$$\mathcal{K} = \underset{\mathcal{K}}{\operatorname{argmin}} \operatorname{VP}_{\max}(\mathbf{\Sigma}_{\mathcal{K}}). \quad (2.21)$$

La méthode de résolution

Il est possible de résoudre le problème dans l'équation (2.18) de plusieurs manières possibles : soit par une approche relaxée ou soit par une approche itérative.

L'approche relaxée est similaire à celle pour la matrice de Fisher, c'est-à-dire, l'introduction d'un vecteur $\boldsymbol{\omega}$:

$$\tilde{\boldsymbol{\omega}} = \underset{\boldsymbol{\omega}=[0,1]^{N_M}, \|\boldsymbol{\omega}\|_1=N_K}{\operatorname{argmin}} f_s\left(\left(\alpha_0 \sum_{i=1}^{N_M} \omega_i \mathbf{a}_{i,\mathcal{G}}^H \mathbf{a}_{i,\mathcal{G}} + \operatorname{diag}(\boldsymbol{\alpha})^{-1}\right)^{-1}\right). \quad (2.22)$$

Remarquons que pour l'équation (2.22), la matrice de covariance peut provenir d'un maximum a posteriori, mais aussi de l'inférence bayésienne parcimonieuse SBI, puisque l'ajout de probabilités

a priori sur les hyperparamètres ne modifie pas la formule de la matrice de covariance.

L’approche itérative cherche à ne sélectionner qu’une position de capteur à la fois. Le critère d’optimisation est mis à jour en fonction des positions de capteurs déjà sélectionnées. En d’autres termes, l’ensemble de positions de capteurs sélectionnées \mathcal{K} est mis à jour à chaque nouvelle sélection et sa dimension augmente jusqu’à atteindre le nombre de capteurs souhaité N_K . Remarquons que l’ensemble \mathcal{K} peut contenir un nombre faible (ou nul) de positions de capteurs avant la sélection. La formulation du problème d’optimisation itératif peut se traduire par :

$$\mathbf{p}_{\mathcal{K}_{i+1}} = \underset{\mathbf{p}_{\mathcal{K}_{i+1}} \in \mathcal{M} \setminus \mathcal{K}}{\operatorname{argmin}} f_s(\boldsymbol{\Sigma}_{\mathcal{K} \cup \mathbf{p}_{\mathcal{K}_{i+1}}}). \quad (2.23)$$

Observons que le calcul de la matrice de covariance est possible par maximum de vraisemblance et maximum a posteriori (et donc par inférence bayésienne parcimonieuse). Cependant si le choix de l’utilisateur se pose sur le maximum de vraisemblance, alors il sera nécessaire d’être dans un cas surdéterminé et donc de commencer la sélection de positions de capteurs avec un nombre de positions capteurs déjà sélectionnées supérieur au nombre d’inconnues du problème de localisation de sources. Mais si l’utilisateur choisit le maximum a posteriori, celui-ci ne doit fournir qu’une information a priori sur la distribution des sources, pouvant être sommaire avec par exemple une loi uniforme.

Remarquons que l’approche bayésienne est simplement la prise en compte de probabilités dans le calcul du critère d’optimisation, c’est-à-dire, le calcul de la matrice de covariance. Donc en calculant la matrice de covariance par maximum a posteriori ou par inférence bayésienne parcimonieuse SBI, la sélection de positions de capteurs devient bayésienne. Notons que que l’approche bayésienne est donc à combiner à l’approche relaxée ou itérative afin de résoudre le problème d’optimisation de sélection de positions de capteurs.

2.5 Sélection séquentielle de capteurs

Dans cette partie, l’objectif est de montrer l’intérêt d’utiliser la sélection séquentielle par rapport aux sélections relaxées. On se concentre uniquement sur le critère qui sera utilisé dans la suite du manuscrit : la matrice de covariance. Dans la section 2.4.2, on a vu que pour utiliser ce critère d’optimisation, 3 questions se posent : comment calculer la matrice de covariance, quel critère d’optimalité choisir et quelle méthode de résolution utiliser. La méthode de résolution est l’approche séquentielle et le calcul de la matrice de covariance a plus d’intérêt à se faire par maximum a posteriori. En effet, par maximum de vraisemblance, il est nécessaire de commencer la sélection séquentielle avec plus de positions de capteurs que d’inconnues dans le problème de localisation de sources.

En partant du problème d’optimisation de la matrice de covariance en séquentiel de l’équation

(2.23), et du calcul de la matrice de covariance par maximum a posteriori de l'équation (1.33), il est possible de diminuer la complexité calculatoire du problème d'optimisation pour deux des trois critères d'optimalité : les critères D-optimal et A-optimal.

2.5.1 Sélection D-optimale

Commençons par rappeler le problème d'optimisation D-optimal séquentiel :

$$\mathbf{p}_{\mathcal{K}_{i+1}} = \operatorname{argmin}_{\mathbf{p}_{\mathcal{K}_{i+1}} \in \mathcal{M} \setminus \mathcal{K}} \det(\boldsymbol{\Sigma}_{\mathcal{K} \cup \mathbf{p}_{\mathcal{K}_{i+1}}}), \quad (2.24)$$

avec la matrice de covariance $\boldsymbol{\Sigma}_{\mathcal{K} \cup \mathbf{p}_{\mathcal{K}_{i+1}}}$ contenant la nouvelle position de capteur $\mathbf{p}_{\mathcal{K}_{i+1}}$ et calculée en considérant une distribution de probabilité complexe normale (voir équation (1.34)) :

$$\begin{aligned} \boldsymbol{\Sigma}_{\mathcal{K} \cup \mathbf{p}_{\mathcal{K}_{i+1}}} &= (\alpha_0 \mathbf{A}_{\mathcal{K}, \mathcal{G}}^H \mathbf{A}_{\mathcal{K}, \mathcal{G}} + \operatorname{diag}(\boldsymbol{\alpha})^{-1} + \alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}})^{-1} \\ &= (\boldsymbol{\Sigma}_{\mathcal{K}}^{-1} + \alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}})^{-1}. \end{aligned} \quad (2.25)$$

Dans un souci de simplification, on cherche à minimiser le logarithme du déterminant de la matrice de covariance tel que :

$$\mathbf{p}_{\mathcal{K}_{i+1}} = \operatorname{argmin}_{\mathbf{p}_{\mathcal{K}_{i+1}} \in \mathcal{M} \setminus \mathcal{K}} \log \left(\det(\boldsymbol{\Sigma}_{\mathcal{K} \cup \mathbf{p}_{\mathcal{K}_{i+1}}}) \right), \quad (2.26)$$

où le logarithme du déterminant de la matrice de covariance équivaut à :

$$\begin{aligned} \log \left(\det(\boldsymbol{\Sigma}_{\mathcal{K} \cup \mathbf{p}_{\mathcal{K}_{i+1}}}) \right) &= \log \left(\det \left((\boldsymbol{\Sigma}_{\mathcal{K}}^{-1} + \alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}})^{-1} \right) \right) \\ &= -\log \left(\det \left(\boldsymbol{\Sigma}_{\mathcal{K}}^{-1} + \alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \right) \right). \end{aligned} \quad (2.27)$$

Supposant une matrice carrée inversible \mathbf{A} et deux vecteurs colonnes \mathbf{u} et \mathbf{v} , le lemme du déterminant de la matrice se définit ainsi [122] :

$$\det(\mathbf{A} + \mathbf{u}\mathbf{v}^H) = (1 + \mathbf{v}^H \mathbf{A}^{-1} \mathbf{u}) \det(\mathbf{A}). \quad (2.28)$$

Le lemme peut être exploité dans l'équation (2.27) [102] :

$$\begin{aligned} \log \left(\det(\boldsymbol{\Sigma}_{\mathcal{K} \cup \mathbf{p}_{\mathcal{K}_{i+1}}}) \right) &= -\log \left((1 + \alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H \boldsymbol{\Sigma}_{\mathcal{K}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}) \det(\boldsymbol{\Sigma}_{\mathcal{K}}^{-1}) \right) \\ &= \log \left(\det(\boldsymbol{\Sigma}_{\mathcal{K}}) \right) - \log \left(1 + \alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H \boldsymbol{\Sigma}_{\mathcal{K}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \right). \end{aligned} \quad (2.29)$$

On remarque dans cette dernière formulation deux termes : le premier est le logarithme du déterminant de la matrice de covariance estimé à partir des positions de capteurs déjà sélectionnées

et le deuxième terme prend en compte la nouvelle position de capteur. On peut donner un nom à chacun de ces termes : la variance de l'estimation de localisation de sources donnée par les positions de capteurs sélectionnées et l'information que peut apporter cette nouvelle position de capteur pour la minimiser. Puisque seul le deuxième terme dépend de la variable d'intérêt (la nouvelle position de capteur) il est plus intéressant de chercher à maximiser ce terme au lieu de minimiser toute l'équation (2.29). Le problème d'optimisation D-optimal séquentiel devient :

$$\mathbf{p}_{\mathcal{K}_{i+1}} = \underset{\mathbf{p}_{\mathcal{K}_{i+1}} \in \mathcal{M} \setminus \mathcal{K}}{\operatorname{argmax}} \log \left(1 + \alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \boldsymbol{\Sigma}_{\mathcal{K}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H \right). \quad (2.30)$$

Pour résoudre ce nouveau problème d'optimisation, la méthode simple est de tester toutes les positions de capteurs appartenant à l'ensemble de positions de capteurs \mathcal{M} , mais qui n'ont pas déjà été sélectionnées et donc qui n'appartiennent pas à \mathcal{K} . La complexité calculatoire de cette équation étant faible, il est possible de le faire de cette façon. Une fois la nouvelle position de capteur trouvée, il suffit de l'inclure dans l'ensemble de positions de capteurs sélectionnées \mathcal{K} , c'est-à-dire, de mettre à jour cet ensemble, puis de réitérer la résolution du problème d'optimisation itératif de l'équation (2.30) jusqu'à atteindre le nombre de capteurs souhaités par l'utilisateur. Remarquons qu'un critère d'arrêt autre que le nombre de capteurs pourrait être utilisé. Lors de cette thèse, j'ai tenté d'évaluer le gain d'information (lié à la grandeur en équation (2.30)) apporté par chaque capteur ajouté. Cependant, l'utilisation d'un nombre de capteurs arbitraire comme critère d'arrêt est sûrement le plus logique expérimentalement. L'utilisateur, en préparant son expérience, aura à sa disposition un nombre de capteurs limités N_K qu'il souhaitera tous placer parmi un grand nombre N_M de positions disponibles.

En ce qui concerne la formule du lemme du déterminant de la matrice dans l'équation (2.28), une version générale existe et permet au final dans l'équation (2.30), non plus d'ajouter une seule position de capteur, mais un nombre quelconque de positions de capteurs. Cela permet par exemple d'accélérer cette approche en ajoutant 3 capteurs à la fois. Une application peut en découler de cette remarque. En effet, on peut imaginer ajouter des sous-antennes à chaque itération, plutôt que d'ajouter un capteur à la fois. Dans le cadre de la détection de sources par un sous-marin en mouvement, il serait possible de choisir des instants d'échantillonnage optimaux par l'antenne tractée. Cependant, il est important de faire attention. L'hypothèse de stationnarité qui considère que la scène acoustique est invariante dans le temps peut se retrouver fautive. Malgré cela, il existe des applications déplaçant des antennes entières, comme l'estimation de la directivité d'une source [83].

2.5.2 Sélection A-optimale

On rappelle le problème d'optimisation A-optimal séquentiel :

$$\mathbf{p}_{\mathcal{K}_{i+1}} = \operatorname{argmin}_{\mathbf{p}_{\mathcal{K}_{i+1}} \in \mathcal{M} \setminus \mathcal{K}} \operatorname{tr}(\boldsymbol{\Sigma}_{\mathcal{K} \cup \mathbf{p}_{\mathcal{K}_{i+1}}}), \quad (2.31)$$

avec :

$$\operatorname{tr}(\boldsymbol{\Sigma}_{\mathcal{K} \cup \mathbf{p}_{\mathcal{K}_{i+1}}}) = \operatorname{tr} \left((\boldsymbol{\Sigma}_{\mathcal{K}}^{-1} + \alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H)^{-1} \right). \quad (2.32)$$

Supposant 4 matrices $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\mathbf{C} \in \mathbb{C}^{k \times k}$, $\mathbf{U} \in \mathbb{C}^{n \times k}$ et $\mathbf{V} \in \mathbb{C}^{k \times n}$, avec \mathbf{A} et \mathbf{C} deux matrices inversibles, l'identité de la matrice de Woodbury [123] s'écrit :

$$(\mathbf{A} + \mathbf{UCV})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{U} (\mathbf{C}^{-1} + \mathbf{VA}^{-1} \mathbf{U})^{-1} \mathbf{VA}^{-1}. \quad (2.33)$$

Dans le cas particulier où la dimension $k = 1$ et donc $\mathbf{C} = 1$, on réécrit l'identité de la matrice de Woodbury avec $\mathbf{A} \in \mathbb{C}^{n \times n}$, $\mathbf{u} \in \mathbb{C}^{n \times 1}$ et $\mathbf{v} \in \mathbb{C}^{1 \times n}$:

$$(\mathbf{A} + \mathbf{uv})^{-1} = \mathbf{A}^{-1} - \mathbf{A}^{-1} \mathbf{u} (1 + \mathbf{v} \mathbf{A}^{-1} \mathbf{u})^{-1} \mathbf{v} \mathbf{A}^{-1}. \quad (2.34)$$

Par cette identité, on transforme l'équation (2.32) en :

$$\operatorname{tr}(\boldsymbol{\Sigma}_{\mathcal{K} \cup \mathbf{p}_{\mathcal{K}_{i+1}}}) = \operatorname{tr} \left(\boldsymbol{\Sigma}_{\mathcal{K}} - \alpha_0 \boldsymbol{\Sigma}_{\mathcal{K}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H (1 + \alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \boldsymbol{\Sigma}_{\mathcal{K}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H)^{-1} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \boldsymbol{\Sigma}_{\mathcal{K}} \right). \quad (2.35)$$

Ensuite en utilisant les propriétés de la trace, on démontre que :

$$\begin{aligned} \operatorname{tr}(\boldsymbol{\Sigma}_{\mathcal{K} \cup \mathbf{p}_{\mathcal{K}_{i+1}}}) &= \operatorname{tr}(\boldsymbol{\Sigma}_{\mathcal{K}}) - \operatorname{tr} \left(\alpha_0 \boldsymbol{\Sigma}_{\mathcal{K}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H (1 + \alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \boldsymbol{\Sigma}_{\mathcal{K}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H)^{-1} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \boldsymbol{\Sigma}_{\mathcal{K}} \right) \\ &= \operatorname{tr}(\boldsymbol{\Sigma}_{\mathcal{K}}) - \frac{1}{1 + \alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \boldsymbol{\Sigma}_{\mathcal{K}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H} \operatorname{tr} \left(\alpha_0 \boldsymbol{\Sigma}_{\mathcal{K}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H \boldsymbol{\Sigma}_{\mathcal{K}} \right) \\ &= \operatorname{tr}(\boldsymbol{\Sigma}_{\mathcal{K}}) - \frac{\alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \boldsymbol{\Sigma}_{\mathcal{K}}^2 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H}{1 + \alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \boldsymbol{\Sigma}_{\mathcal{K}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H}. \end{aligned} \quad (2.36)$$

On retrouve de nouveau un problème se décomposant en deux termes, dont seulement le deuxième est dépendant de la variable d'intérêt. On cherche donc à maximiser :

$$\mathbf{p}_{\mathcal{K}_{i+1}} = \operatorname{argmax}_{\mathbf{p}_{\mathcal{K}_{i+1}} \in \mathcal{M} \setminus \mathcal{K}} \frac{\alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \boldsymbol{\Sigma}_{\mathcal{K}}^2 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H}{1 + \alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}} \boldsymbol{\Sigma}_{\mathcal{K}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}^H}. \quad (2.37)$$

On remarque que le dénominateur correspond au D-optimal dans l'équation (2.30). La complexité de ce critère est donc légèrement plus importante que celui du D-optimal.

2.6 Sélection orientée par les données

Dans les sélections séquentielles proposées dans la section précédente, on remarque la présence de la matrice de covariance Σ estimée par maximum a posteriori. La difficulté du maximum a posteriori est la nécessité de donner une information sur les positions des sources sous forme probabiliste avec α , le vecteur de variance de la distribution complexe normale par rapport à $\mathbf{s}_{\mathcal{M},\mathcal{G}}$. Si l'utilisateur ne peut pas donner d'informations sur les positions des sources, deux possibilités sont réalisables :

- considérer une loi uniforme pour la distribution des positions des sources
- prendre en compte une structure hiérarchique bayésienne avec des hyperparamètres.

L'utilisation d'hyperparamètres peut se faire par l'inférence bayésienne parcimonieuse vue dans la partie 1.6.3. Si l'inférence bayésienne parcimonieuse est adoptée pour l'estimation de la matrice de covariance, des mesures acquises par les capteurs sont nécessaires pour l'exécution algorithmique. Avant de discuter de la pertinence de l'utilisation des mesures dans la sélection de positions de capteurs, le processus de sélection de positions de capteurs que nous proposons est explicité.

2.6.1 Sélection séquentielle bayésienne orientée par les données

Pour commencer la sélection, il est nécessaire d'avoir au préalable la matrice de covariance estimée, mais pour l'inférence bayésienne parcimonieuse des mesures acquises par les capteurs sont obligatoires. Comme initialisation, il faut un nombre faible de capteurs déjà sélectionnés. Dans la localisation de sources, l'estimation de la position des sources se fait en comparant la phase entre les capteurs. Un minimum de 2 capteurs est nécessaire pour discriminer spatialement les sources dans un scénario plan. Arbitrairement ces 2 positions de capteur sont choisies de manière à maximiser l'ouverture d'antenne dès la première étape de sélection. Après l'initialisation, le processus alterne entre 3 étapes :

- l'estimation de la matrice de covariance par inférence bayésienne parcimonieuse, et par la même occasion l'estimation des positions des sources,
- la sélection d'une nouvelle position de capteur par D-optimal,
- la mise à jour de l'ensemble des positions de capteurs sélectionnées, ainsi que l'acquisition de la nouvelle mesure correspondant à la nouvelle position de capteur.

Ce procédé est résumé dans l'Algorithme 7 et dans le schéma 2.3.

On retrouve notamment cette méthode dans [102] avec des mesures et un modèle de propagation aléatoires. Notons aussi que le critère D-optimal peut-être remplacé par le critère A-optimal ou E-optimal, et l'inférence bayésienne parcimonieuse peut-être remplacée par l'inférence bayésienne parcimonieuse à plusieurs snapshots (algorithme 3). Différentes études de performances

Algorithme 7 Sélection séquentielle bayésienne orientée par les données

Entrée : \mathcal{M} , $\mathbf{A}_{\mathcal{M},\mathcal{G}}$, $\mathbf{m}_{\mathcal{M},\mathcal{G}}$, N_K , N_{K_0} le nombre de positions de capteurs sélectionnées à l'initialisation

- Initialisation :

Sélection des N_{K_0} positions de capteurs dans \mathcal{M}

Création de l'ensemble de positions de capteurs sélectionnées \mathcal{K}

Sélection du vecteur de mesures $\mathbf{m}_{\mathcal{K},\mathcal{G}}$ correspondant

Sélection de la matrice du modèle de propagation $\mathbf{A}_{\mathcal{K},\mathcal{G}}$ correspondante

pour $i = N_{K_0} + 1, \dots, N_K$ **faire**

- Estimation des positions des sources $\mathbf{s}_{\tilde{\mathcal{K}},\mathcal{G}}$ et de la matrice de covariance $\Sigma_{\mathcal{K}}$ par inférence bayésienne parcimonieuse (algorithme 2)

- Sélection de la nouvelle position de capteur $\mathbf{p}_{\mathcal{K}_i}$ par D-optimal (équation (2.30))

- Mise à jour :

de l'ensemble de positions de capteurs sélectionnées $\mathcal{K}^i = \mathcal{K}^{i-1} \cup \mathbf{p}_{\mathcal{K}_i}$

du vecteur de mesures $\mathbf{m}_{\mathcal{K},\mathcal{G}}$

de la matrice du modèle de propagation $\mathbf{A}_{\mathcal{K},\mathcal{G}}$

fin pour

Sortie : \mathcal{K}

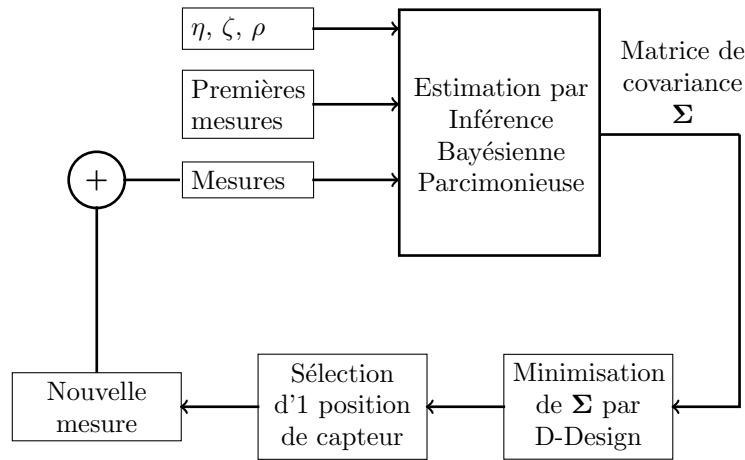


FIGURE 2.3 – Schéma explicatif de la sélection séquentielle bayésienne orientée par les données.

seront effectuées afin d'évaluer la différence du choix de critère d'optimalité, ainsi que du choix de stratégie de mesure.

2.6.2 Discussion

Est-il pertinent d'utiliser les données dans la sélection afin de préparer son expérience ? Cette question a eu des réponses positives [124, 104] et négatives [125]. Pour nous, dans un premier

temps le côté applicatif semble restreint et le gain insuffisant, notamment vu la complexité calculatoire supplémentaire apportée par l'inférence bayésienne parcimonieuse. Cependant pour le côté applicatif, ce n'est pas parce qu'il est restreint, qu'il est inexistant. On voit notamment son utilisation dans l'estimation de la directivité de la source [83] ou la reconstruction d'images [126].

Si la stationnarité de la scène acoustique est maintenue au cours de l'expérience, alors cette méthode est aussi performante que les autres, voire plus puisque celle-ci utilise l'information contenue dans les données. Si la méthode de sélection de positions de capteurs n'utilise que le modèle et que celui-ci se retrouve erroné, ce qui est souvent le cas en expérimental, alors la sélection se retrouvera dégradée. Or en utilisant les données, il est possible de réduire cet impact. Ce point sera notamment abordé dans les chapitres expérimentaux (chapitres 4 et 5).

Par contre, si la scène acoustique n'est spatialement pas stationnaire, alors cette méthode en état est inadéquate pour faire de la sélection "online". Les solutions possibles seraient de :

- prendre toutes les mesures possibles et de sélectionner parmi afin de ne prendre que celles nécessaires,
- concevoir l'antenne avant la prise de mesures.

2.7 Conclusions

La finalité de ce chapitre était d'arriver à cet algorithme qu'est la sélection séquentielle bayésienne orientée par les données. Cet algorithme est notre contribution. Les performances de cet algorithme pour de la localisation de sources acoustiques seront démontrées dans la partie II dans un contexte de sources sous-marines. L'algorithme sera ensuite utilisé sur des données expérimentales provenant de sources acoustiques aériennes en milieu réverbérant dans la partie III.

DEUXIÈME PARTIE

**Sélection séquentielle bayésienne
orientée par les données dans un
contexte simulé de localisation de
sources acoustiques sous-marines**

SÉLECTION SÉQUENTIELLE BAYÉSIENNE ORIENTÉE PAR LES DONNÉES POUR LA LOCALISATION PASSIVE DE SOURCES ACOUSTIQUES SOUS-MARINES

Sommaire

3.1	Introduction	60
3.2	Illustration par des scénarios élémentaires	61
3.2.1	Configuration des scénarios	61
3.2.2	Illustration du critère en fonction du nombre de capteurs sélectionnés	61
3.2.3	Illustration du critère en fonction de la fréquence	63
3.2.4	Illustration du critère en fonction d'autres paramètres	64
3.3	Généralisation des résultats par simulations de Monte Carlo	67
3.3.1	Configuration des simulations de Monte Carlo	68
3.3.2	Relation entre la valeur du critère de sélection et performances de localisation	69
3.3.3	Analyse par rapport au bruit	71
3.4	Comparatif des résultats des simulations de Monte Carlo par rapport à l'état de l'art	72
3.4.1	Présentation de l'étude et paramètres de la simulation	72
3.4.2	Présentation des antennes de l'état de l'art	73
3.4.3	Résultats	74
3.5	Études paramétriques annexes	78
3.5.1	Comparaison d'estimateurs de positions de sources	78
3.5.2	Comparaison des critères d'optimalités	79
3.5.3	Extension multi-snapshots	80
3.6	Conclusion	81

3.1 Introduction

Dans ce contexte d’acoustique sous-marine que l’on se pose, on considère des sources en champ lointain, et la localisation de sources est donc une estimation d’angles d’arrivées. On considère également un problème dans lequel le nombre de source est faible. Dans le cas d’un problème de localisation de source discret sur une grille avec un grand nombre d’angles, il est possible de s’appuyer sur une hypothèse de parcimonie.

Dans le chapitre 1, on a vu comment bénéficier de la parcimonie du problème dans les algorithmes d’estimation de localisation de sources, par des méthodes parcimonieuses ou des approches bayésiennes. Ces dernières, en plus d’estimer les angles d’arrivée, permettent aussi d’obtenir une estimation de la matrice de covariance, un élément clé pour le problème de sélection de positions de capteurs. Dans le chapitre 2, après la présentation de différents critères d’optimisation permettant de résoudre le problème de sélection de positions de capteurs, l’utilisation de la matrice de covariance a posteriori de l’estimateur a été retenu. On a vu qu’elle possède des caractéristiques intéressantes pour la sélection :

- la possibilité d’une sélection séquentielle,
- son utilisation dans un cadre sous-déterminé si la matrice de covariance est estimée par maximum a posteriori ou inférence bayésienne parcimonieuse (et donc la possibilité d’une sélection bayésienne),
- la possibilité d’une sélection orientée par les données.

L’algorithme de sélection séquentielle bayésienne orientée par les données proposé dans ce manuscrit (voir Algorithme 7), est une combinaison de ces trois caractéristiques. Cet algorithme a déjà été proposé dans [102] pour une matrice de propagation dont chaque élément est gaussien et indépendant. Dans notre cas, la physique impose un modèle déterministe de propagation d’ondes planes. Ce chapitre se focalise sur le contexte simulé d’acoustique sous-marine, afin d’en tirer une étude de performances de la méthode de sélection de capteur proposée.

Dans un premier temps dans la partie 3.2, on montre comment varie le critère de sélection de positions de capteurs au cours de la sélection et en fonction des paramètres physiques de la localisation de sources. La section 3.3 répond à la question de pertinence et de performance du critère de sélection, par l’utilisation de différentes sélections séquentielles qui proviennent de l’Algorithme 7 et par l’utilisation de simulations de Monte Carlo. Ensuite dans la même section, on détermine les limites de l’algorithme proposé en observant l’impact du bruit gaussien sur la sélection de capteurs et l’estimation de localisation de sources. Puis, dans la section 3.4, nous comparons les performances de l’antenne conçue par notre algorithme avec des méthodes de l’état de l’art. Avant de conclure dans la partie 3.6, nous proposons une série d’études annexes dans la partie 3.5. Celle-ci comprend :

- une comparaison des estimateurs de positions de sources décrits dans le chapitre 1,

- une comparaison des différents critères d’optimalité de la matrice de covariance donnés dans le chapitre 2,
- et une généralisation multi-snapshots qui est une option intéressante dans les scénarios où les mesures acquises ont un faible rapport signal à bruit.

3.2 Illustration par des scénarios élémentaires

Dans cette section, nous allons illustrer le comportement du critère de sélection de capteurs en fonction des paramètres physiques de la localisation de sources acoustiques. Nous allons utiliser pour cela des scénarios élémentaires qui permettront de visualiser l’allure du critère.

On rappelle que le critère de sélection de capteurs de notre algorithme (voir Algorithme 7) est le suivant :

$$\delta h(\mathbf{p}_{\mathcal{K}_i}) = \log \left(1 + \alpha_0 \mathbf{a}_{\mathbf{p}_{\mathcal{K}_i}, \mathcal{G}} \boldsymbol{\Sigma}_{\mathcal{K}} \mathbf{a}_{\mathbf{p}_{\mathcal{K}_i}, \mathcal{G}}^H \right). \quad (3.1)$$

On présente dans un premier temps la configuration des scénarios élémentaires. Puis on commente l’évolution du critère en fonction de la sélection des positions de capteurs, de la fréquence de la source mais aussi de la position de la source ou du nombre de sources présentes.

3.2.1 Configuration des scénarios

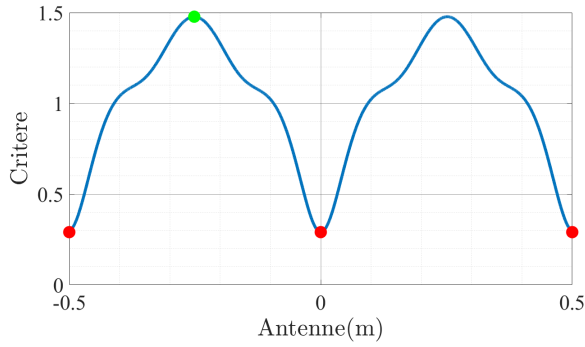
On considère une antenne linéaire uniforme avec une ouverture d’antenne de 1 mètre et qui contient 11 capteurs placés sur l’axe Z . On positionne une source avec un angle d’arrivée à $\theta = 30^\circ$. À partir des mesures obtenues, on utilise l’inférence bayésienne parcimonieuse en discrétisant les angles de $-\frac{\pi}{2}$ à $\frac{\pi}{2}$ avec $N_G = 181$ points. On considère l’amplitude de la source égale à $s_{\mathcal{M}, \mathcal{S}} = 1$, à la fréquence $F = 7500$ Hz. Le milieu étant de l’eau, on suppose que l’onde se propage à une vitesse $c = 1500$ m/s. Le bruit gaussien est de niveau faible et donne un rapport signal à bruit de $\text{SNR} = 30$ dB. Le critère est calculé sur $N_M = 10001$ positions espacées uniformément sur l’antenne. De cette manière, on contraint naturellement l’ouverture de l’antenne à $D = 1$ m.

3.2.2 Illustration du critère en fonction du nombre de capteurs sélectionnés

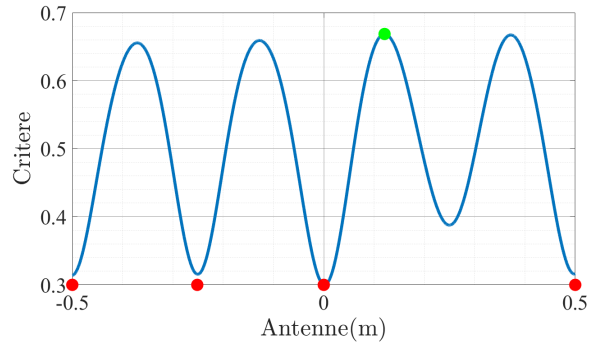
L’algorithme proposé est une sélection séquentielle de positions de capteurs. À chaque itération de l’algorithme et donc à chaque position de capteur sélectionnée, la matrice de covariance est estimée par inférence bayésienne parcimonieuse (équation (1.33)) et le critère d’information (équation (3.1)) est calculé. La matrice de covariance et donc le critère de sélection évoluent au cours de la sélection.

Partons des paramètres utilisés dans la partie 3.2.1, mais en diminuant le nombre de capteurs de départ. La sélection commence avec $N_{K_0} = 3$ capteurs déjà sélectionnés, 2 étant aux extré-

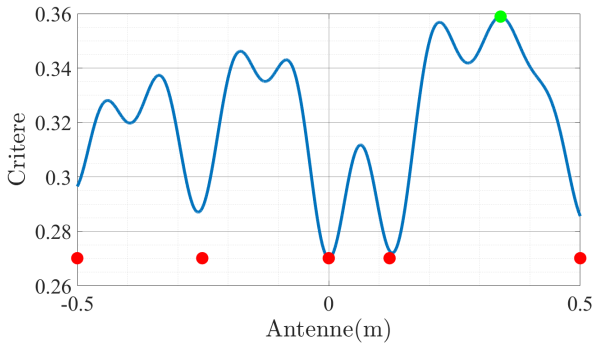
mités de l'antenne et 1 au centre de celle-ci. On calcule alors le critère sur plusieurs itérations et on obtient la figure 3.1.



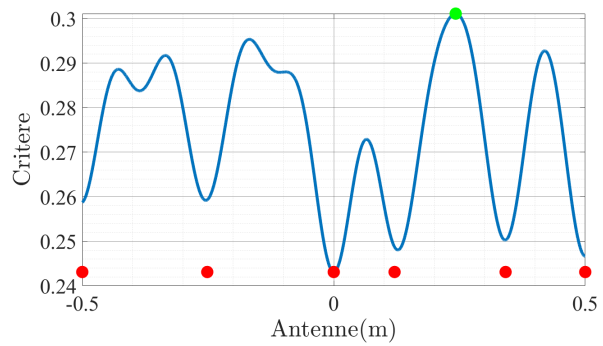
(a) La sélection de la 4^{ème} position de capteur



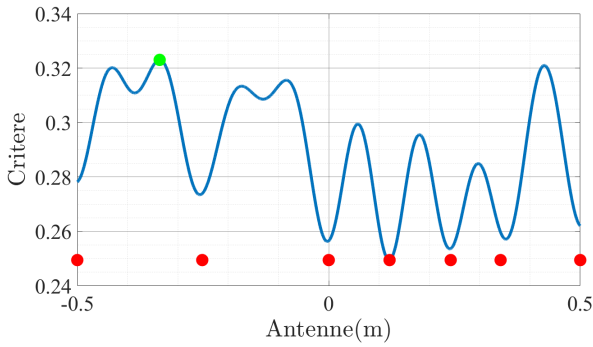
(b) La sélection de la 5^{ème} position de capteur



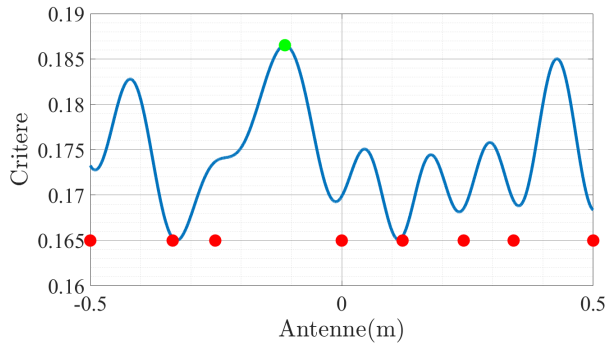
(c) La sélection de la 6^{ème} position de capteur



(d) La sélection de la 7^{ème} position de capteur



(e) La sélection de la 8^{ème} position de capteur



(f) La sélection de la 9^{ème} position de capteur

FIGURE 3.1 – Calcul du critère d'information en fonction du nombre de capteurs considérés. En rouge sont les positions de capteurs sélectionnés et en vert celui qui sera sélectionné pour l'itération suivante.

Observons que l'amplitude du critère est plus faible aux positions de l'antenne où des capteurs sont déjà présents. Cela prouve qu'en conception d'antenne, il est peu intéressant de placer des capteurs trop proches et plus utile de remplir l'espace des positions candidates de façon

homogène. On remarque aussi que l'amplitude globale du critère diminue au cours de la sélection. Cette remarque est confirmée par la figure 3.2 en prenant la valeur maximale du critère à chaque nouvelle position de capteur sélectionnée pour 100 sélections. On observe que l'amplitude du critère diminue avec le nombre de capteurs sélectionnés, et semble tendre vers 0. Le critère est nul si et seulement si la matrice de covariance est nulle et donc que le nombre de capteurs est infini. On peut remarquer que plus il y a des capteurs déjà présents, moins la quantité d'information acquise par un capteur supplémentaire est grand.

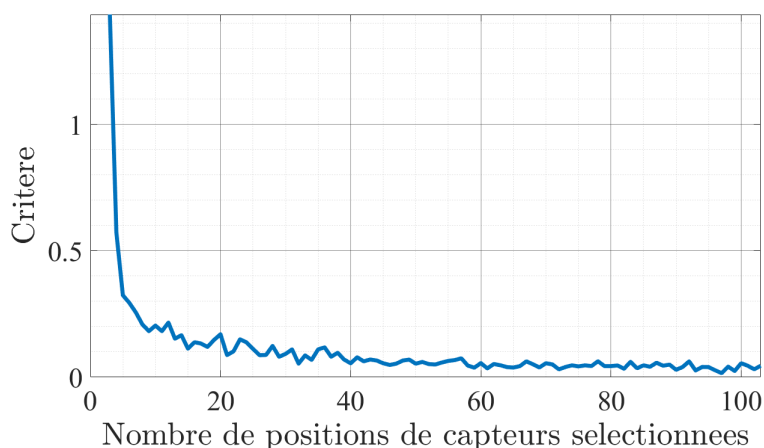


FIGURE 3.2 – Évolution de l'amplitude maximale du critère δh au cours de la sélection des positions de capteurs. Les 6 premières valeurs correspondent aux valeurs des points verts de la figure 3.1.

3.2.3 Illustration du critère en fonction de la fréquence

La fréquence de la source est un facteur important dans la localisation de sources et dans le dimensionnement d'antenne. Pour cela on part des paramètres de la partie 3.2.1 mais on modifie la fréquence de la source en allant de 500 à 15000 Hz. On obtient alors la figure 3.3.

Cette étude nous permet de repérer deux comportements connus dans le dimensionnement d'antenne. Le premier concerne les sources à très hautes fréquences, où la conception d'antenne cherche naturellement à respecter le théorème de Shannon [127]. Si on observe la figure 3.3a, le critère sélectionne une position entre deux positions déjà sélectionnées. Cela permettra alors réduire l'espace inter-capteur et de respecter la limite haute fréquence de l'équation (1.15). Le deuxième comportement connu concerne les sources à très basse fréquence, où la conception d'antenne cherche à augmenter l'ouverture d'antenne le plus possible pour respecter la limite basse fréquence de l'équation (1.17). Le critère dans la figure 3.3d tend à augmenter l'ouverture d'antenne mais comme l'ouverture d'antenne est fixe, la limite ne peut pas être respectée. Le meilleur compromis donné par le critère est alors de sélectionner les positions, déjà sélectionnées,

aux extrémités de l'antenne.

Un point intéressant est que le critère utilisé n'a pas de lien explicite avec la formation de voies, puisque celui-ci utilise l'inférence bayésienne parcimonieuse. Pour autant le comportement induit par la sélection de capteur itérative a une tendance naturelle à améliorer les conditions requises pour améliorer la formation de voies. L'autre point intéressant de cette étude est de connaître le comportement de la conception d'antenne quand la source se situe entre le régime basse fréquence et haute fréquence.

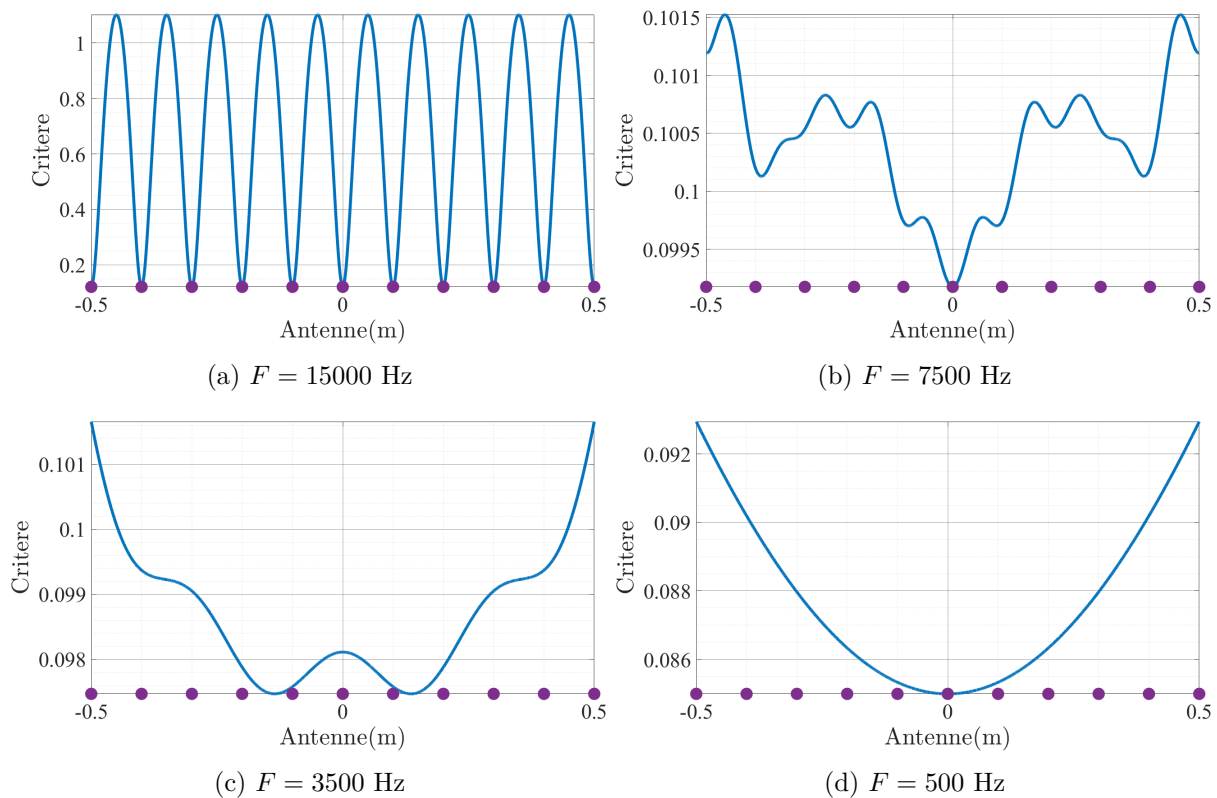


FIGURE 3.3 – Allure du critère de sélection de positions de capteurs δh en fonction de la fréquence de la source.

3.2.4 Illustration du critère en fonction d'autres paramètres

Finalement, 3 paramètres peuvent aussi influencer le critère : l'angle d'arrivée de la source θ , le nombre de sources présentes N_S et le rapport signal à bruit SNR. Ces études permettent d'observer que la conception d'antenne informée est optimale uniquement dans la scène acoustique considérée.

Impact de la position angulaire de la source

Le critère sélectionne des positions de capteurs différentes, si la source n'est pas positionnée au même endroit comme le montre la figure 3.4, où la source est placée à différents angles d'arrivées. On remarque que le critère peut quasiment être identique si on compare les figures 3.4a et 3.4b, ou à l'inverse être complètement différent entre les figures 3.4a et 3.4d. Cette différence importante entre les allures de courbes n'indiquent a priori pas de comportement commun à l'ensemble des scénarios où la source a une origine différente.

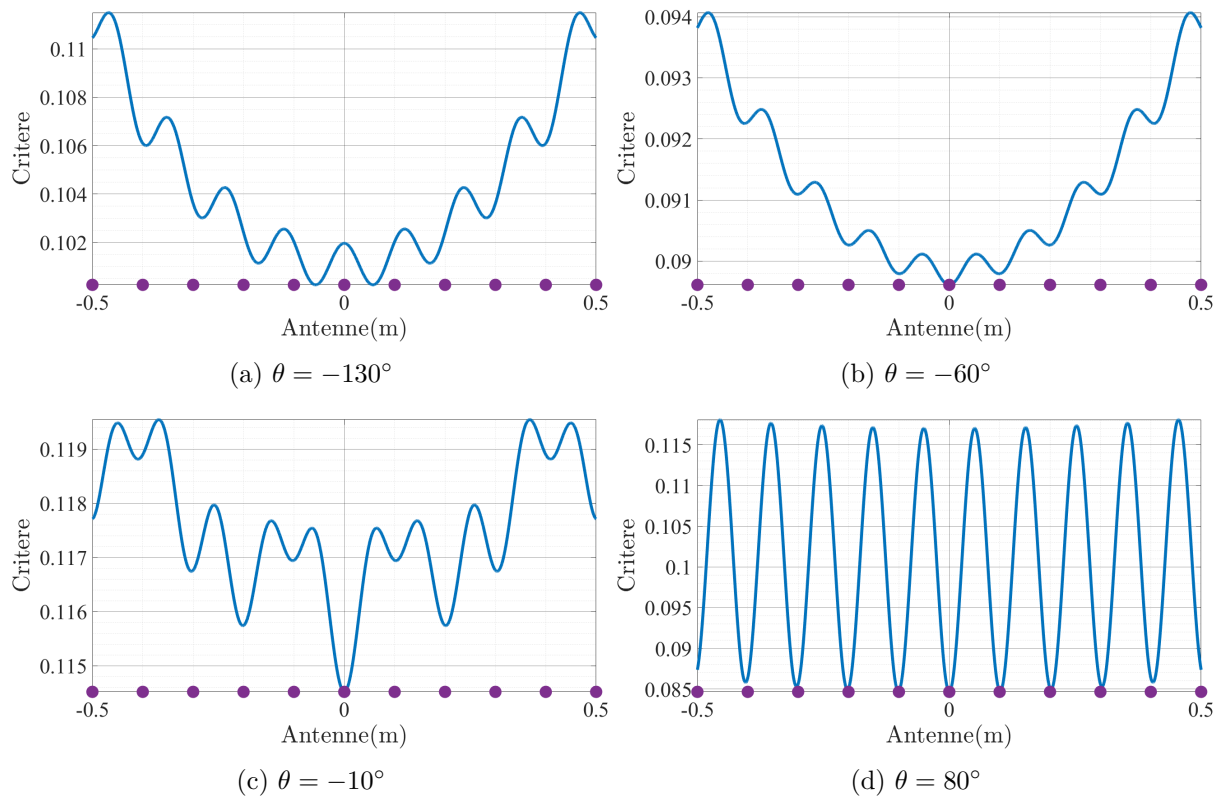


FIGURE 3.4 – Allure du critère de sélection de positions de capteurs δh en fonction de l'angle d'arrivée θ de la source.

Impact du nombre de sources

La figure 3.5 illustre l'impact du nombre de sources sur les profils de critère d'information, à la fréquence $F = 3500$ Hz. Si le nombre de sources augmente, malgré l'indépendance entre les signaux émis par chaque source, les positions de capteurs sélectionnées ne seront pas les mêmes puisque l'information qu'elles apportent doivent permettre de réduire la variance de l'estimateur des positions de toutes les sources.

Si on passe de la figure 3.5a à la figure 3.5b en ajoutant une source, le critère ne sélectionne plus les extrémités. En considérant 4 ou 8 sources, dans les figures 3.5c et 3.5d, le critère considère que les positions proches des extrémités de l'antenne ne sont plus les positions optimales également. On peut donc conclure qu'en fonction du nombre de sources et de leurs positions, la tendance à augmenter l'ouverture de l'antenne (vue dans la section 3.2.3) n'est pas nécessairement vérifiée, car cela dépend du contexte donné. On remarque aussi que la valeur maximale du critère augmente quand le nombre de sources augmente. On en déduit que la mesure est plus informative si le nombre de sources N_S présentes est plus grand.

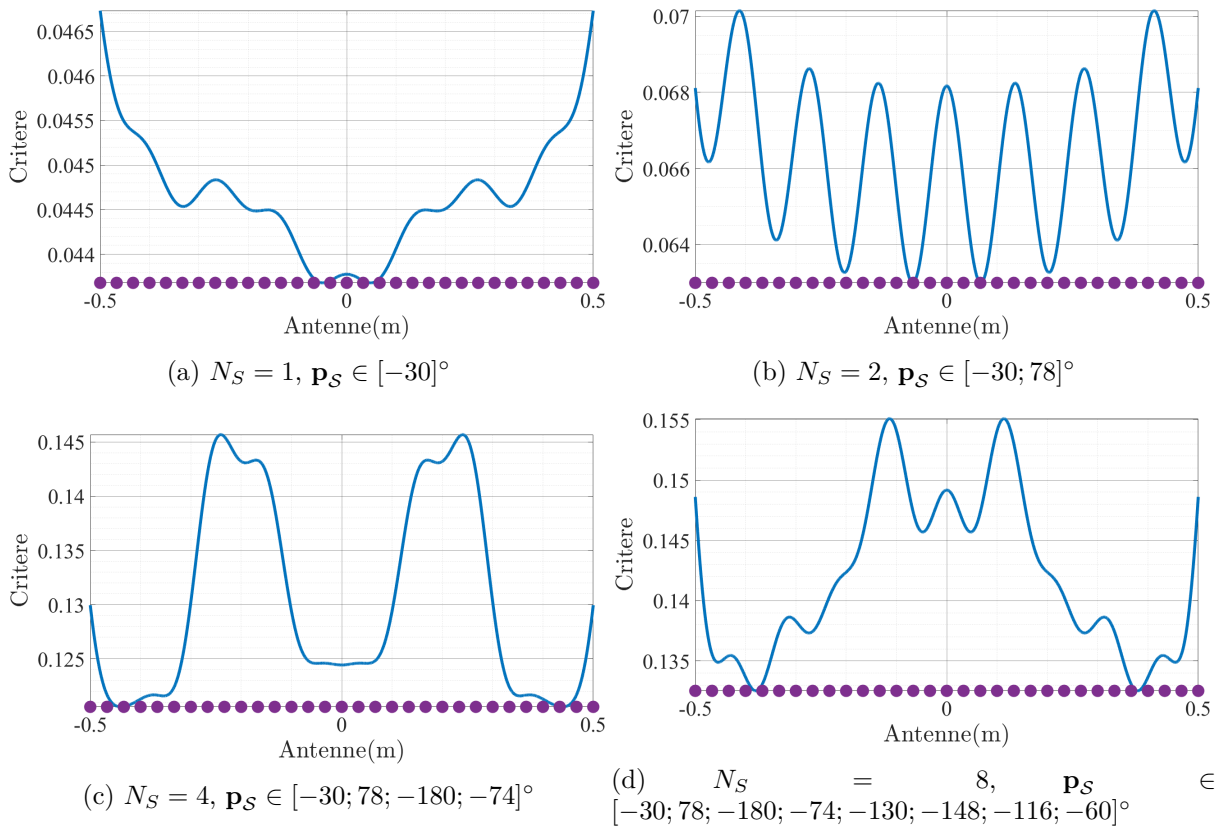


FIGURE 3.5 – Allure du critère de sélection de positions de capteurs δh en fonction du nombre de sources présentes.

Impact du bruit

Dans cette sous-section, on revient aux paramètres initiaux ($N_S = 1$, angle d'arrivée $\theta = 30^\circ$ et $F = 3500$ Hz), et on change l'amplitude du bruit. En comparant les figures 3.6a et 3.6b, on remarque que le critère n'est pas trop impacté si le rapport signal à bruit reste assez élevé. Par contre si le bruit devient trop important alors le critère est modifié, comme dans les figures 3.6c

et 3.6d. Dans ce cas, le critère ne sélectionne plus les mêmes positions de capteurs.

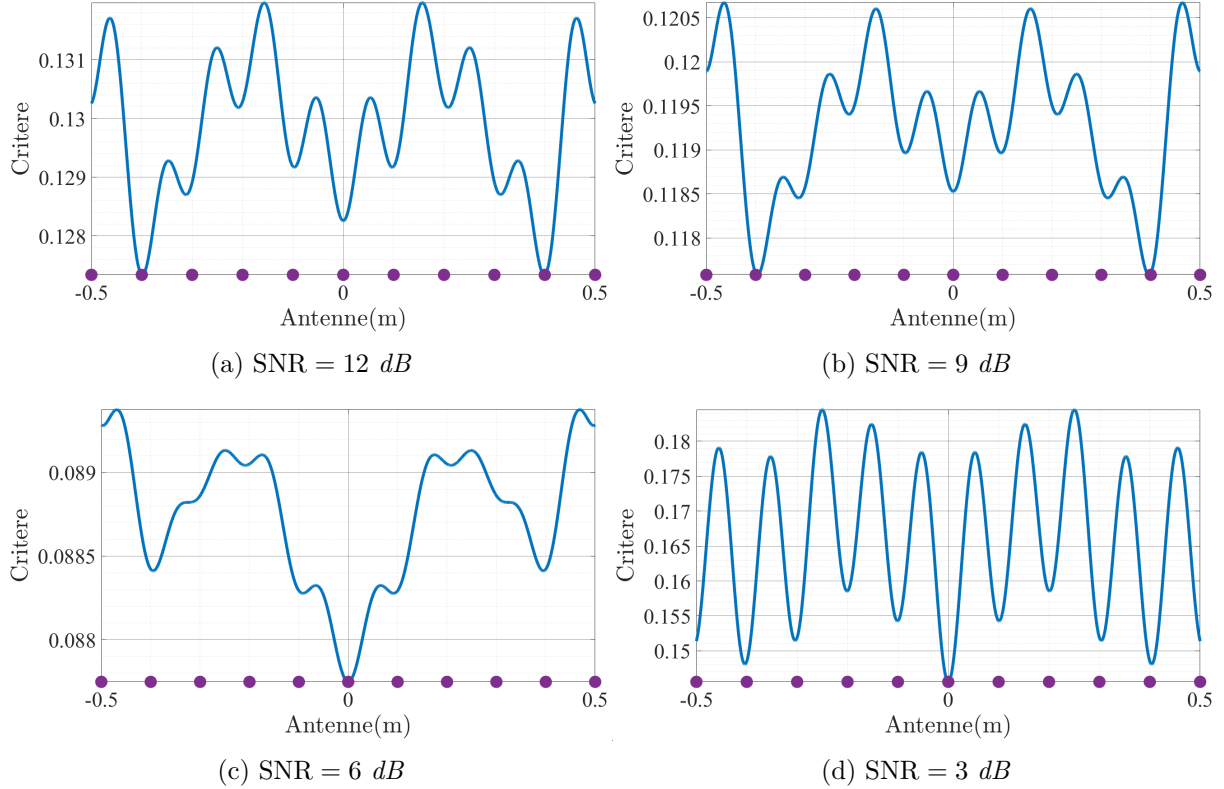


FIGURE 3.6 – Allure du critère de sélection de positions de capteurs en fonction de la fréquence du rapport signal à bruit.

Nous concluons cette section par une remarque sur les différentes analyses issues de scénarios simples. Ces scénarios ont permis de montrer des tendances dans des cas particuliers mais ne permettent pas de tirer des interprétations générales, ou d'observer les performances de l'algorithme proposé dans ce manuscrit. C'est pour cette raison que nous utilisons dans la suite de ce chapitre des simulations de Monte Carlo, où les performances de l'algorithme pourront être évaluées dans un nombre important de scènes acoustiques.

3.3 Généralisation des résultats par simulations de Monte Carlo

À partir de cette section et pour le reste de ce chapitre, nous évaluons les performances de localisations des conceptions d'antennes par des simulations de Monte Carlo. Nous allons donc concevoir un grand nombre de scènes acoustiques définies par des paramètres physiques identiques mais qui diffèrent par une partie aléatoire correspondant au bruit et sur les positions des sources.

Nous définissons dans un premier temps la configuration, c'est-à-dire, les paramètres physiques des scènes acoustiques, des simulations de Monte Carlo. Puis nous évaluons la performance de notre algorithme. Pour cela, nous cherchons à quantifier la relation entre le critère de sélection de capteurs et la performance de localisation. Pour finir, nous analysons la performance de notre algorithme en fonction du niveau de bruit présent dans les mesures.

3.3.1 Configuration des simulations de Monte Carlo

Dans chaque simulation de Monte Carlo, nous générons 100 réalisations qui suivent les paramètres physiques décrits par la suite. Chaque réalisation contient le bruit gaussien \mathbf{w}_K différent. Dans le cas de l'antenne aléatoire un nouveau tirage de N_K capteurs est effectué. Le vecteur d'amplitude des sources $\mathbf{s}_{K,G}$ est identique, c'est-à-dire, que les positions des sources, leurs phases et amplitudes sont identiques à chaque réalisation.

En acoustique sous-marine, on se place en champ lointain avec le modèle de propagation défini dans l'équation (1.9). Cependant, pour simplifier le problème, on ne va pas considérer l'angle ϕ et l'antenne ne sera que sur une seule dimension de l'espace, l'axe Z . L'angle d'arrivée θ est contenu dans l'intervalle $[-\frac{\pi}{2}; \frac{\pi}{2}]$. Cet intervalle est discrétisé en $N_G = 181$ points. Parmi ces N_G angles, sont situées $N_S = 10$ sources indépendantes. Le signal émis par chacune des sources est de fréquence $F = 1500$ Hz et le vecteur d'amplitudes est égal à $\mathbf{s}_{M,S} = \mathbf{1}^{N_S}$.

Le milieu de propagation est l'eau. La célérité du son est donc de $c = 1500$ m/s. Le rapport signal à bruit SNR, défini dans l'équation (1.11), est de 12 dB.

On considère n'avoir que 45 capteurs à notre disposition. Ce nombre est supposé assez grand pour que toutes les antennes arrivent à localiser les sources. Les capteurs sont considérés ponctuels et peuvent donc être très proches les uns des autres. De plus, le calcul du critère de sélection de positions de capteurs est rapide à calculer. Cela permet de définir un grand nombre de positions candidates de capteurs $N_M = 10001$. On cherche donc à sélectionner $N_K = 45$ positions de capteurs parmi $N_M = 10001$ positions de capteurs disponibles.

Pour finir, l'algorithme proposé nécessite des positions de capteurs déjà sélectionnées. On prend alors le minimum de positions de capteurs nécessaire à la localisation, c'est-à-dire, $N_{K_0} = 2$, placés aux extrémités de l'antenne. Ainsi l'ouverture de toutes les antennes est invariante, de manière à marginaliser son impact du test de Monte Carlo. L'ouverture d'antenne est de $D = 10$ mètres.

Pour finir, les performances sont évaluées en localisation de sources pour chaque antenne à chaque nouvelle position de capteur sélectionnée. L'évaluation des performances se fait par l'indice de Jaccard, défini en section 1.8. Rappelons que pour le calcul de l'indice de Jaccard, un paramètre doit être fixé arbitrairement par l'utilisateur : le rayon d'admission r_a . Dans notre cas, le pas de la grille étant de 1° , le rayon d'admission est fixé à $r_a = \pm 1^\circ$. Notons aussi que pour le calcul de l'indice de Jaccard, il est nécessaire de définir l'ensemble de positions des

sources estimées $\mathcal{E} = \{\mathbf{p}_{\mathcal{E}_j}\}, j \in [1, N_E]$ avec $\mathbf{p}_{\mathcal{E}_j} \in \mathbb{R}^{1 \times 3}$ la position j estimée de la source et N_E le nombre de positions de sources estimées. Pour cela, on normalise le vecteur d'amplitude estimé $\tilde{\mathbf{s}}_{\mathcal{M}, \mathcal{G}}$ et tous les pics d'amplitudes (déterminés par la fonction *findpeaks* de *Matlab*) au dessus de 0,3 correspondent à des sources. L'indice de Jaccard est ensuite moyenné sur les 100 réalisations.

3.3.2 Relation entre la valeur du critère de sélection et performances de localisation

Le but de cette partie est de savoir si le critère d'information choisi δh est pertinent pour concevoir l'antenne optimale pour l'estimation de positions de sources acoustiques sous-marines. Dans cette partie, on va donc concevoir différentes antennes optimisant plus ou moins le critère d'information et comparer leurs performances en localisation de sources.

On présentera dans un premier temps l'étude et comment les différentes antennes sélectionnent les positions de capteurs. Puis, les performances de localisation de sources seront montrés en utilisant l'indice de Jaccard. Finalement quelques interprétations seront données afin de comprendre au mieux ces résultats.

Présentation de l'étude

Nous analysons dans quelle mesure le critère de sélection δh (voir équation (3.1)) est pertinent pour améliorer séquentiellement les performances de localisation de sources évaluées par l'indice de Jaccard. Nous avons vu précédemment que l'intervalle de valeurs que peut prendre δh dépend du scénario. Dans cette étude nous allons créer un proxy à partir de δh , de manière à hiérarchiser facilement les positions candidates en fonction leur quantité potentielle d'information.

Tout d'abord, $\delta h(\mathbf{p}_{\mathcal{K}_i})$ est mis à l'échelle de façon linéaire entre 0 et 1, pour obtenir le proxy $\overline{\delta h}(\mathbf{p}_{\mathcal{K}_i})$. Par conséquent, $\overline{\delta h}(\mathbf{p}_{\mathcal{K}_{i^*}}) = 1$ (resp. 0) est la position candidate la plus (resp. la moins) informative, correspondant à la sélection optimale dans l'Algorithme 7. Ensuite, nous exécutons le placement séquentiel et choisissons au hasard les N_K positions du capteur $\mathbf{p}_{\mathcal{K}}$ telles que $\overline{\delta h}(\mathbf{p}_{\mathcal{K}}) \in [0; 0, 25]$. Le même processus est répété 4 fois, des intervalles les moins informatifs aux intervalles les plus informatifs $[0, 25; 0, 5]$, $[0, 5; 0, 75]$, et $[0, 75; 1]$, comme illustré dans la figure 3.7. Notons qu'un placement totalement aléatoire correspond à $\overline{\delta h}(\mathbf{p}_{\mathcal{K}}) \in [0; 1]$, et que le choix optimal est obtenu pour $\overline{\delta h}(\mathbf{p}_{\mathcal{K}}) = 1$.

Résultats

La figure 3.8 illustre les performances obtenues par chaque antenne en fonction du nombre de positions de capteurs sélectionnées. On remarque que pour toutes les conceptions d'antennes (c'est à dire aléatoire, optimale, et en fonction de la quantité d'information), l'ajout d'une mesure

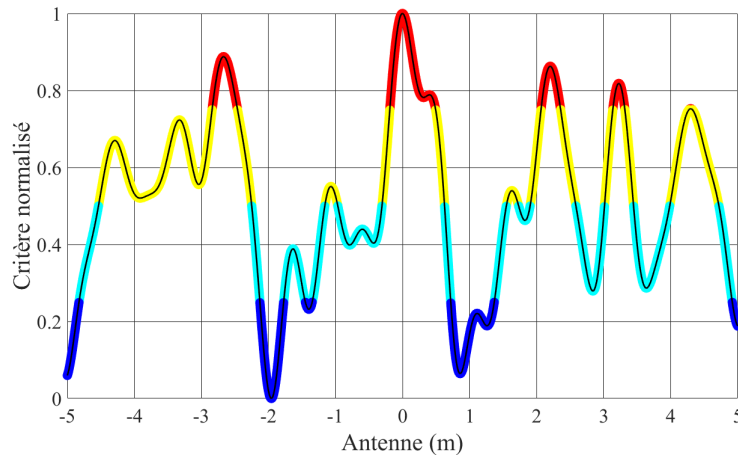


FIGURE 3.7 – Application des 4 seuils pour les antennes à basse (bleu foncé), à moyenne-basse (bleu clair), à moyenne-haute (jaune) et à haute (rouge) quantité d’information.

supplémentaire augmente la valeur de l’indice de Jaccard en moyenne sur les 100 réalisations. Un seuil de performances est atteignable pour toutes les antennes (à l’exception de l’antenne à basse information), c’est-à-dire, qu’il n’est pas possible d’obtenir des meilleures performances même en ajoutant un nombre important de mesures. On observe notamment que ce seuil de performances est presque identique pour toutes les conceptions d’antennes. Le seuil de performances correspond environ à un indice de Jaccard de 0,95.

Pour comparer les conceptions d’antennes, on décide de les classer. Pour cela on regarde le nombre de positions de capteurs sélectionnées nécessaires pour que chaque antenne dépasse un indice de Jaccard de 0,9 (une valeur proche du seuil de performances). On observe que :

- l’antenne optimale nécessite 22 positions de capteurs sélectionnées,
- 24 pour l’antenne à haute information,
- 30 pour l’antenne à moyenne haute information,
- 30 pour l’antenne aléatoire,
- 39 positions de capteurs sélectionnées pour l’antenne à moyenne basse information,
- et l’antenne à basse information ne dépasse pas la valeur demandée.

Cela nous permet donc de classer les antennes, l’antenne optimale étant la meilleure et l’antenne à basse information la moins bonne.

Par ce classement, on remarque que plus le critère est optimisé, plus les performances de l’antenne en localisation de sources sont élevées. On en conclut que le critère est non seulement pertinent, mais en plus performant pour la conception d’antenne.

Notons que le classement effectué n’est pertinent que dans un contexte où le nombre de capteurs est restreint. Si l’utilisateur possède un très grand nombre de capteurs, il n’a pas

forcément d'intérêt à appliquer l'algorithme proposé. La sélection aléatoire lui permettra en moyenne d'atteindre les mêmes performances que toutes les autres antennes. Cependant en acoustique, surtout en milieu marin, les capteurs sont précieux et peu nombreux. Qualifier la conception d'antenne par son nombre de positions de capteurs nécessaire à atteindre une performance est donc logique.

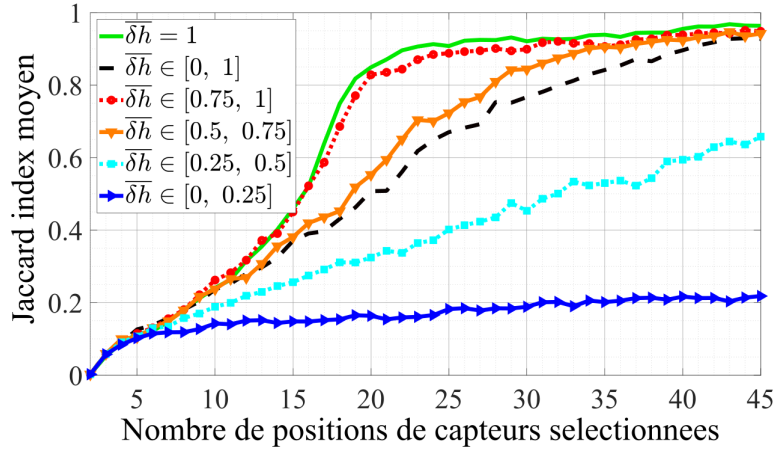


FIGURE 3.8 – Indice de Jaccard moyen en fonction du nombre de positions de capteurs sélectionnées pour 6 antennes différentes : l'antenne optimale (vert), aléatoire (noir), à basse information (bleu foncé), à moyenne-basse information (bleu clair), à moyenne-haute information (orange) et à haute information (rouge).

3.3.3 Analyse par rapport au bruit

Présentation de l'étude

On s'intéresse dans cette étude à l'impact du bruit gaussien dans la conception d'antenne. Dans la partie 3.2.4, il nous a été possible de voir que lorsque le rapport signal à bruit était trop faible, la sélection n'était plus la même. Mais est-ce que les performances de la conception d'antenne sont dégradées ? À partir des simulations de Monte Carlo présentées dans la section 3.3.1, on étudie les performances de l'antenne optimale avec différents rapports signal à bruit, de 0 dB à 12 dB avec un pas de 3 dB.

Résultats

La figure 3.9 montre les performances de la conception d'antenne optimale en fonction du rapport signal à bruit au cours de la sélection de capteurs. La première observation est la dégradation des performances quand le rapport signal à bruit diminue. Ce comportement est normal dans la localisation de sources. On remarque que chaque position de capteur ajouté

permet d’augmenter les performances de localisation. Pour un rapport signal à bruit supérieur ou égale à 6 dB, on observe un gain de performances plus important entre la 15^{ème} et 20^{ème} position de capteurs sélectionnées. Ce gain est peu visible pour un rapport signal à bruit de 3 dB et inexistant pour un rapport signal à bruit de 0 dB. Ce gain nous confirme qu’à partir d’un certain nombre de capteurs, le critère permet de sélectionner les capteurs les plus utiles à la localisation de sources. Cependant ce gain s’efface si le bruit est trop important. On en conclut donc que le bruit perturbe la sélection de capteurs. Un moyen de contrecarrer cet impact négatif serait d’augmenter le nombre de mesures acquises de la scène acoustique, c’est-à-dire de mesurer plusieurs snapshots (cf section 3.5.3).

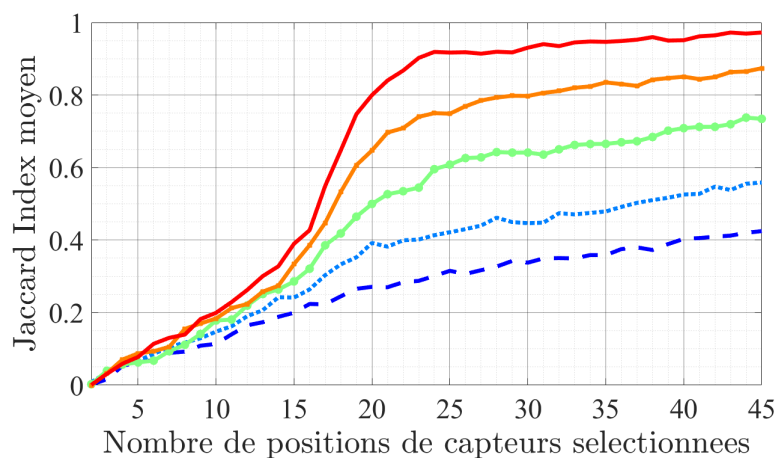


FIGURE 3.9 – Indice de Jaccard moyen en fonction du nombre de positions de capteurs sélectionnées pour l’antenne optimale à différents rapport signal à bruit : 12 dB (rouge), 9 dB (orange), 6 dB (vert), 3 dB (bleu clair) et 0 dB (bleu foncé).

3.4 Comparatif des résultats des simulations de Monte Carlo par rapport à l’état de l’art

3.4.1 Présentation de l’étude et paramètres de la simulation

Dans cette section, nous allons comparer les performances de l’antenne conçue par notre algorithme avec celles provenant de 4 antennes de l’état de l’art :

- l’antenne issue d’une sélection aléatoire,
- l’antenne linéaire uniforme,
- l’antenne suivant la règle de Golomb [130],
- l’antenne conçue par algorithme génétique.

Pour la simulation, on réutilise les paramètres de la section 3.3.1 à une différence près : les positions et les phases des sources sont différentes pour chaque réalisation. Notons que la norme de l'amplitude du signal émis par chacune des sources reste cependant égale à 1.

Après une présentation des antennes provenant de l'état de l'art, nous montrons les performances de chacune des conceptions d'antennes. Les performances seront exposées sous plusieurs angles :

- l'évolution moyenne de l'indice de Jaccard au cours de la sélection pour chaque conception d'antenne (fig. 3.10),
- l'indice de Jaccard sous forme de boîte à moustaches avec le nombre maximal de capteurs sélectionnés pour chaque conception d'antenne (fig. 3.11),
- l'indice de Jaccard sous forme de boîte à moustaches au cours de la sélection pour l'antenne optimale (fig. 3.12),
- un histogramme des positions de capteurs sélectionnés pour l'antenne optimale et l'antenne génétique (fig. 3.13),
- le niveau des lobes secondaires après formation de voies sous forme de boîte à moustaches avec le nombre maximal de capteurs sélectionnés N_K pour chaque conception d'antenne (fig. 3.14).

On remarque l'utilisation de boîtes à moustaches qui nous permet de visualiser la variance des performances et du niveau des lobes secondaires après formation de voies qui a pour but de proposer un autre critère d'évaluation des antennes que l'indice de Jaccard. Notons aussi que l'histogramme des positions de capteurs sélectionnés ne concerne que 2 des 5 conceptions d'antennes. L'histogramme permet de montrer les positions de capteurs les plus sélectionnées par les algorithmes au cours des différentes réalisations. Mais l'antenne linéaire uniforme et l'antenne suivant la règle de Golomb ne varient pas selon les réalisations. Pour l'antenne aléatoire, la sélection de positions de capteurs suit une loi uniforme. Il n'est donc pas utile de visualiser l'histogramme pour ces trois conceptions d'antennes.

3.4.2 Présentation des antennes de l'état de l'art

On cherche à comparer l'antenne optimale qui suit notre algorithme avec des conceptions d'antennes venant de domaines différents : l'antenne aléatoire utilisée communément dans l'acquisition comprimée [128, 129], l'antenne suivant la règle de Golomb utilisée pour concevoir une antenne supposée plus performante que les antennes aléatoires dans l'acquisition comprimée [130], l'antenne linéaire uniforme utilisée en acoustique sous-marine [131, 127] et l'antenne génétique utilisée en localisation acoustique notamment pour la formation de voies [95]. Dans ces antennes, seules les deux premières permettent une sélection séquentielle de positions. Pour les trois autres conceptions d'antennes, la sélection séquentielle n'est pas possible.

L'antenne linéaire uniforme place les capteurs pour que l'espace inter-capteurs soit égal à :

$$d = D/(N_K - 1) \quad (3.2)$$

La règle de Golomb permet une sélection de capteurs à partir d'un maillage de points réguliers, où toutes les distances entre les capteurs sont différentes. Cependant cette conception d'antenne contient un problème majeur : on ne connaît pas de règle de Golomb ayant un nombre de positions supérieur à 27. La conception de cette antenne ne pourra donc pas dépasser 27 positions de capteurs sélectionnées. Pour finir l'antenne issue de l'algorithme génétique ne sera pas utilisée lors de l'évaluation moyenne de l'indice de Jaccard au cours de la sélection. Ce choix a été fait car l'approche est longue et l'utiliser séquentiellement, c'est-à-dire, de sélectionner 2 à 45 positions de capteurs multiplié par le nombre de réalisations nécessiterai un temps de calcul trop important. La sélection génétique utilisée dans cette étude suit les mêmes paramètres d'entrée que l'article [95]. On considère $N_{gen} = 100$ générations avec $N_{ant} = 100$ individus contenant $N_K = 45$ gènes. Le coefficient de mutation est fixé à $\chi = 0,01$.

Remarquons que l'évaluation est l'indice de Jaccard utilisé sur les positions de sources estimées par inférence bayésienne parcimonieuse. Or certaines conceptions d'antennes n'ont pas besoin d'utiliser l'inférence bayésienne parcimonieuse pour la sélection de positions de capteurs. Cependant, afin que l'estimation de localisation de source et l'évaluation soit identique pour toutes les conceptions d'antennes, nous avons décidé d'utiliser l'inférence bayésienne parcimonieuse quelque soit l'antenne considérée.

3.4.3 Résultats

Suivant les paramètres et les conceptions d'antennes présentées, on obtient la figure 3.10. Cette figure nous permet de réaliser le classement utilisé dans la partie 3.3.2, c'est-à-dire, de classer les conceptions d'antennes en fonction du nombre de capteurs nécessaires pour dépasser une valeur arbitraire de l'indice de Jaccard. La meilleure antenne est celle qui nécessite en moyenne le moins de positions de capteurs sélectionnées afin que l'indice de Jaccard dépasse 0,9. L'approche suivant notre algorithme nécessite 23 positions de capteurs, l'antenne linéaire uniforme nécessite autour de 35 positions de capteurs puis l'antenne aléatoire qui nécessite environ 40 positions de capteurs. L'antenne suivant la règle de Golomb quand à elle ne réussit pas à atteindre cette valeur.

L'évolution de l'indice de Jaccard au cours de la sélection de positions de capteurs pour l'antenne suivant la règle de Golomb est chaotique : elle augmente ou diminue de manière abrupte en ajoutant une nouvelle position de capteur. On remarque cependant une tendance globale croissante malgré son allure abrupte qui peut s'expliquer du fait que l'approche est non séquentielle (à chaque ajout de capteur, l'antenne est redéfinie). Cependant cet impact est non

présent dans l'antenne linéaire uniforme qui elle aussi est une approche non séquentielle. On peut alors supposer que l'évolution chaotique de l'indice de Jaccard est causée par la règle de Golomb elle-même.

La performance de l'antenne linéaire uniforme par rapport à celle de l'antenne aléatoire est étonnante. En effet, l'algorithme de localisation de sources est une méthode parcimonieuse. On peut supposer que l'utilisation d'antennes aléatoires soit meilleure que l'antenne linéaire uniforme, comme cela est souvent mentionné dans la littérature [54].

Remarquons que seule l'antenne optimale arrive à dépasser un indice de Jaccard moyen de 0,95 et semble tendre vers un indice de Jaccard moyen de 1.

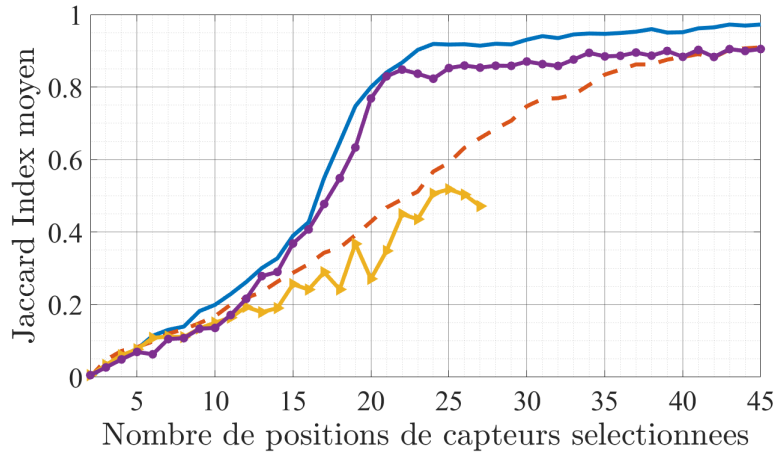


FIGURE 3.10 – Indice de Jaccard moyen en fonction du nombre de positions de capteurs sélectionnées pour différentes conceptions d'antennes : antenne suivant notre algorithme (bleu), l'antenne aléatoire (orange), l'antenne linéaire uniforme (violet) et l'antenne suivant la règle de Golomb (jaune).

On présente l'indice de Jaccard sous forme de boîtes à moustaches avec le nombre maximal de positions de capteurs sélectionnées (45 sauf celle de Golomb qui ne peut que faire 27) dans la figure 3.11. Les traits rouges représentent la médiane, les boîtes bleues représentent le premier et le troisième quartiles, les traits noirs sont les minimums et maximums hors valeurs aberrantes et les plus rouges sont les valeurs aberrantes. Notons que nous rajoutons maintenant la conception d'antenne suivant l'algorithme génétique.

Si on écarte l'antenne suivant la règle de Golomb, on remarque que l'antenne aléatoire, l'antenne linéaire uniforme et l'antenne génétique arrivent à des performances similaires en médiane et en variance de l'indice de Jaccard. Par contre l'antenne optimale a une variance quasi nulle, mis à part quelques valeurs aberrantes. L'antenne optimale à 45 positions de capteurs sélectionnées estime les positions de sources de différentes scènes acoustiques avec la même qualité et précision. Cette remarque sur l'antenne optimale peut-être prolongée en observant l'évolution de

l'indice de Jaccard sous forme de boîtes à moustaches au cours de la sélection, donnée dans la figure 3.12. On remarque qu'avec seulement 38 positions de capteurs sélectionnées, la variance de l'estimation des positions des sources est nulle.

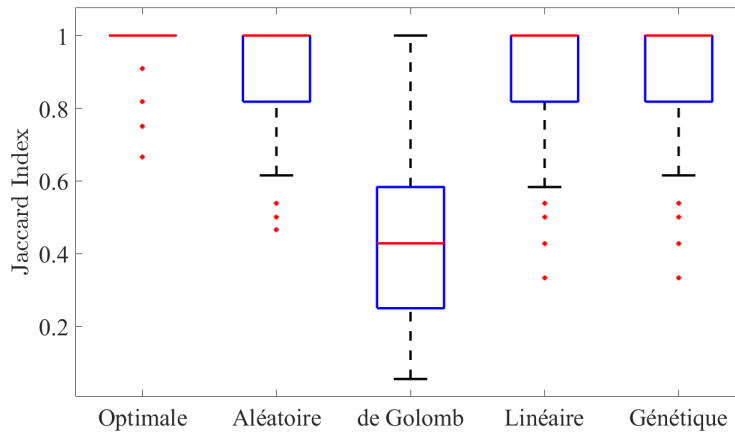


FIGURE 3.11 – Indice de Jaccard sous forme de boîtes à moustaches pour voir la variance avec le nombre maximal de positions de capteurs sélectionnées pour différentes conceptions d’antennes : optimale, aléatoire, de Golomb, linéaire et génétique.

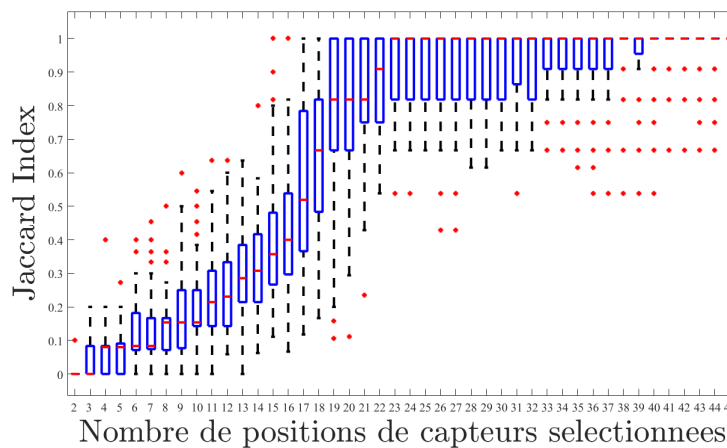


FIGURE 3.12 – Indice de Jaccard sous forme de boîtes à moustaches en fonction du nombre de positions de capteurs sélectionnées pour l’antenne optimale.

Pour expliquer la différence entre les conceptions d’antennes, on peut directement observer les positions spatiales sélectionnées. Pour cela on réalise un histogramme des positions de capteurs sélectionnées parmi les 10001 disponibles sur les 100 réalisations. On montre dans la figure 3.13, deux histogrammes correspondant à l’antenne optimale et l’antenne génétique.

L’antenne optimale, dans la majorité des réalisations, a tendance à sélectionner les extrémités

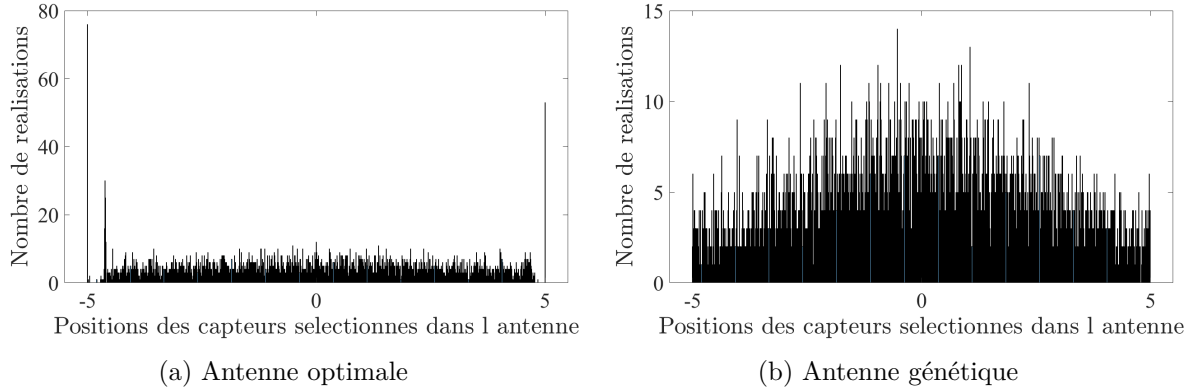


FIGURE 3.13 – Positions des capteurs sélectionnées par les différentes conceptions d’antennes testées sous forme d’histogramme au cours des 100 réalisations

à plusieurs reprises. Hormis ces 2 exceptions, l’histogramme relativement plat ne semble pas révéler de positions sélectionnées redondantes. Pour l’antenne génétique, on semble observer une sélection un peu plus importante au centre de l’antenne.

Enfin, la figure 3.14 présente le niveau maximal du lobe secondaire des diagrammes de directivités centrés aux positions des sources, sous forme de boîtes à moustaches avec le nombre maximal de capteurs sélectionnés pour les différentes conceptions d’antennes. Pour cela, on calcule le diagramme de directivité (voir équation (2.2)) centré sur une des positions des sources. Puis, on récupère le niveau maximal du lobe secondaire, en prenant le second pic le plus élevé du diagramme de directivité (le premier étant le lobe principal). On répète le calcul du diagramme de directivité et le niveau maximal du lobe secondaire pour chaque source (i.e. $N_S = 10$) et chaque réalisation (i.e. 100). On se retrouve avec 1000 niveaux de lobes secondaires par conception d’antenne.

Dans le contexte d’une localisation de source par formation de voies, l’antenne à privilégier minimise le niveau des lobes secondaires. Dans cette simulation, l’antenne optimale est la meilleure et l’antenne aléatoire la pire. L’antenne optimale qui cherche à minimiser la variance d’estimation de l’inférence bayésienne parcimonieuse permet donc de concevoir une antenne performante pour de la formation de voies. Remarquons aussi l’absence de variance dans le niveau des lobes secondaires pour l’antenne linéaire uniforme et l’antenne de Golomb puisque leurs positions de capteurs ne varient pas au cours des réalisations. Les performances de l’antenne aléatoire sont très variables et sont plus faibles que l’antenne linéaire. Enfin, l’antenne génétique a des performances relativement bonnes : cela est logique puisque le niveau de lobe secondaire maximal est le critère optimisé.

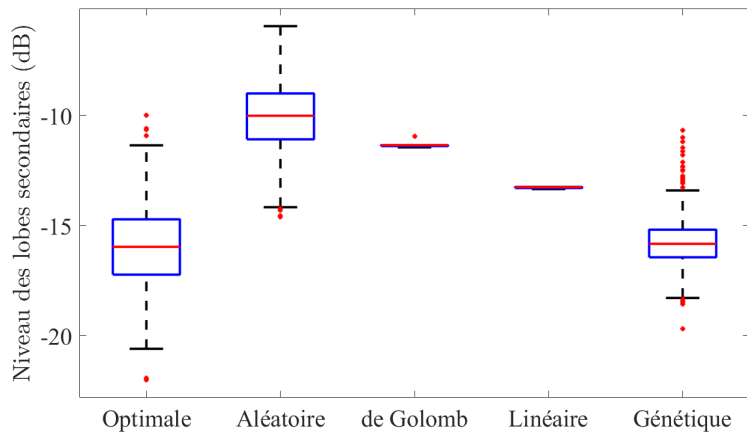


FIGURE 3.14 – Niveau des lobes secondaires après formation de voies sous forme de boîtes à moustaches avec le nombre maximal de positions de capteurs sélectionnées pour différentes conceptions d’antennes : optimale, aléatoire, de Golomb, linéaire et génétique.

3.5 Études paramétriques annexes

Dans cette sections, nous cherchons à analyser des pistes possibles permettant d’améliorer les performances de notre algorithme. Pour cela nous essayons les trois possibilités suivantes :

- choisir une autre méthode d’estimation de positions de sources que SBI,
- modifier le critère d’optimalité pour la sélection de capteurs,
- la prise en compte de plusieurs mesures (snapshots) en utilisant l’inférence bayésienne parcimonieuse à multi-snapshots (Algorithme 3).

Les performances des modifications apportées sur l’algorithme sont évaluées à partir de la même configuration que lors de la comparaison avec les conceptions d’antennes de l’état de l’art (partie 3.4).

3.5.1 Comparaison d’estimateurs de positions de sources

Dans cette étude, notre but est de tester d’autres algorithmes de localisation de sources que SBI, pour l’antenne conçue par notre méthode. Les 4 algorithmes de localisation de sources testés sont l’inférence bayésienne parcimonieuse (partie 1.6.3), la formation de voies (partie 1.4.1), le basis pursuit de-noising (partie 1.5.2) et l’orthogonal matching pursuit (partie 1.5.1). Quelque soit l’algorithme de localisation de sources utilisé, le critère d’évaluation restera le même, c’est-à-dire l’indice de Jaccard. Notons que la formation de voies n’étant pas un algorithme parcimonieux, il est nécessaire d’utiliser un détecteur de pic ou de lobes principaux. Pour cela, on utilise la fonction *findpeaks* de *Matlab*.

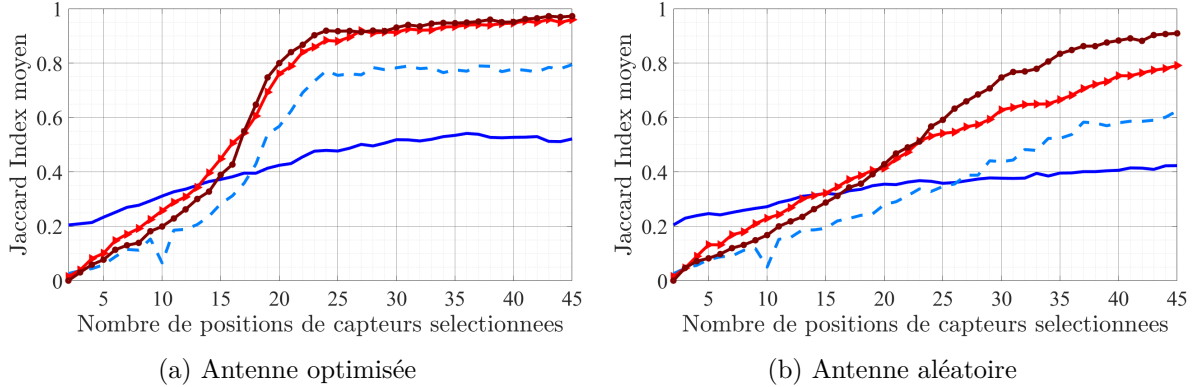


FIGURE 3.15 – Indice de Jaccard moyen en fonction du nombre de positions de capteurs sélectionnées pour différentes conceptions d’antenne selon différents algorithmes de localisation de sources : inférence bayésienne parcimonieuse (rouge foncé), basis pursuit de-noising (rouge clair), orthogonal matching pursuit (bleu clair) et formation de voies (bleu foncé).

La figure 3.15a permet de visualiser l’évolution de l’indice de Jaccard en fonction de l’estimateur de positions de sources pour l’antenne optimisée. L’inférence bayésienne parcimonieuse atteint un indice de Jaccard moyen de 0,95. Le même score est atteint pour le basis pursuit de-noising. L’orthogonal matching pursuit quand à lui n’atteint qu’un indice de Jaccard de 0,8 et la formation de voies oscille autour de 0,5.

Si les performances de la formation de voies s’expliquent notamment par le nombre important de sources et l’incompatibilité du critère d’évaluation avec l’algorithme, les performances de l’orthogonal matching pursuit sont en dessous des autres algorithmes. Notamment on observe une chute de performances de localisation de sources à 10 positions de capteurs sélectionnées non expliquée.

Le seuil de performances maximal atteint par chaque algorithme permet de classer les estimateurs de positions de sources. Dans ce cas, il est plus intéressant d’utiliser l’inférence bayésienne parcimonieuse.

À partir de la figure 3.15a qui représente l’évolution de l’indice de Jaccard en fonction de l’estimateur de positions de sources pour l’antenne aléatoire, on peut noter les mêmes observations et arriver aux mêmes conclusions qu’avec l’antenne optimisée.

3.5.2 Comparaison des critères d’optimalités

On propose une étude où les différents critères d’optimalité, c’est-à-dire, les critères A-optimal, D-optimal et E-optimal sont comparés entre eux.

Notons que les critères A-optimal et D-optimal sont calculés par les équations (2.37) et (2.30). Cela permet d’accélérer le calcul des critères et donc l’algorithme, et permet que l’algorithme soit numériquement stable puisqu’il ne nécessite pas de calculer l’inversion de la matrice de

covariance. Dans cette étude, nous n’exploitons pas de technique d’accélération pour le calcul du critère E-optimal. En effet, lors du calcul du critère E-optimal, il faudra pour chaque position de capteur disponible, inverser la matrice de covariance contenant la nouvelle position de capteur (voir équation (2.25)), puis calculer les valeurs propres afin de sélectionner la position de capteur qui minimisera la plus grande valeur propre.

On obtient la figure 3.16 évaluant la performance des antennes conçues par l’optimisation des 3 critères. Remarquons que quelque soit le critère optimisé, l’ajout de capteurs permet une augmentation des performances en localisation de sources. Si les performances obtenues par les critères D-optimal et A-optimal sont similaires, les performances de l’antenne conçue par le critère E-optimal sont plus faibles.

Les performances des critères D-optimal et A-optimal sont similaires mais le temps de calcul est très différent selon le critère d’optimalité choisi. La sélection par optimisation du critère D-optimal est la plus rapide, suivi du critère A-optimal, pour finir avec celui du critère E-optimal. Il est donc plus intéressant d’utiliser le critère D-optimal dans ce contexte. On rappelle que le critère E-optimal fait intervenir des calculs potentiellement instables numériquement, que ce soit dans l’inversion de la matrice de covariance, ou dans le calcul des valeurs propres.

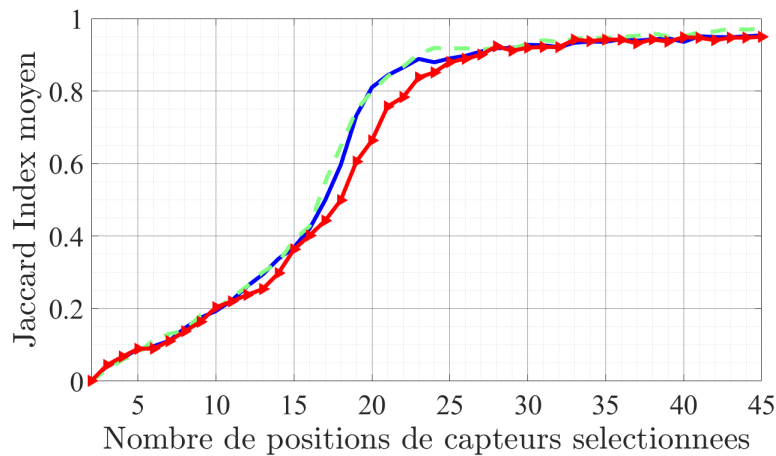


FIGURE 3.16 – Indice de Jaccard moyen en fonction du nombre de positions de capteurs sélectionnées pour différents critères d’optimalité de minimisation de la matrice de covariance : le critère D-optimal (bleu), le critère A-optimal (vert), le critère E-optimal (rouge).

3.5.3 Extension multi-snapshots

Dans cette étude, on a pour but de démontrer que l’utilisation de snapshots permet dans un premier temps une amélioration de l’estimation des positions des sources, mais aussi une augmentation des performances de la sélection de positions de capteurs quand le rapport signal à bruit est bas. Pour cela on considère que chaque nouvelle position de capteur acquiert un

certain nombre de mesures N_T (snapshots), allant de 1 à 20. L'ensemble des snapshots mesure une même scène acoustique (hypothèse de stationnarité). On remplace l'inférence bayésienne parcimonieuse dans l'algorithme par l'inférence bayésienne parcimonieuse à plusieurs snapshots (voir Algorithme 3). Notons que l'ajout des snapshots ne modifie pas les dimensions de la matrice de covariance, le calcul du critère de sélection de positions de capteurs est donc identique. Pour finir, l'étude est similaire à celle menée dans la partie 3.3.3, où l'antenne optimale est testée pour différents RSBs.

L'évolution de l'indice de Jaccard en fonction du nombre de positions de capteurs sélectionnées et du nombre de snapshots utilisés et du rapport signal à bruit est donnée dans la figure 3.17. Remarquons que l'augmentation du nombre de snapshots améliore les performances de localisation de sources. Il est possible d'atteindre un indice de Jaccard moyen de 1 dans certaines conditions : par exemple, si le rapport signal à bruit est supérieur à 6 dB avec 20 snapshots et que l'antenne contient un minimum de 25 capteurs. Notons aussi qu'avec 20 snapshots et un rapport signal à bruit de 0 dB (c'est à dire que la composante signal mesurée a la même énergie que la composante bruit), il est possible d'atteindre un indice de Jaccard moyen de 0,95 avec 35 capteurs. Si on se rappelle qu'il y a 10 sources présentes dans la scène, le nombre de capteurs et de snapshots restent petits.

On remarque que le nombre de capteurs nécessaire pour que l'indice de Jaccard atteigne son seuil de performances maximales diminue avec le nombre de snapshots.

Une autre remarque est la présence d'un nombre de positions de capteurs minimal à avoir pour atteindre le seuil de performances maximal. Si on prend le cas du plus fort rapport signal à bruit, c'est-à-dire, 12 dB , le seuil de performances est atteint à 40, 24, 21 et 20 capteurs pour respectivement 2, 5, 10 et 20 snapshots. L'évolution du nombre de capteurs nécessaires pour atteindre le seuil de performances semble tendre vers 20 capteurs. On peut supposer que en augmentant fortement le nombre de snapshots, ce nombre de capteurs risque de ne plus diminuer.

Pour finir, on observe que pour un nombre de snapshots assez grand ($N_T > 5$), les courbes d'indice de Jaccard sont de moins en moins hétérogènes en fonction des différents niveaux de RSB. Par exemple, dans la figure 3.17e avec 20 snapshots, on observe que l'évolution de l'indice de Jaccard pour un rapport signal à bruit de 12, 9 et 6 dB sont très similaires. La prise en compte de plusieurs snapshots permet donc bien de contrebalancer l'impact du bruit gaussien sur l'estimation des positions de sources.

3.6 Conclusion

Ce chapitre valide l'utilisation de notre algorithme dans un contexte de localisation passive de sources acoustiques sous-marines par des résultats sur données synthétiques. Dans un premier

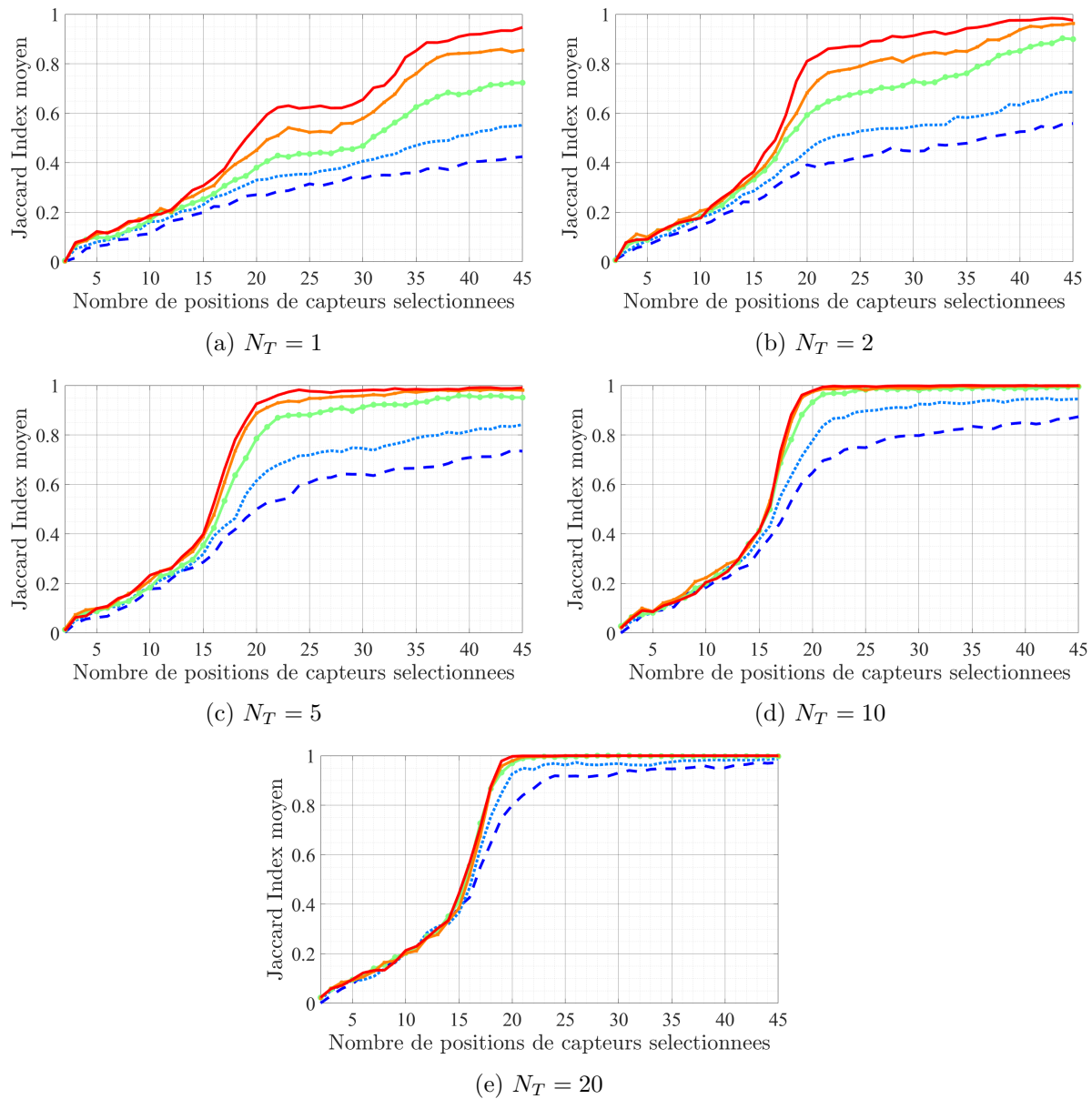


FIGURE 3.17 – Indice de Jaccard moyen en fonction du nombre de positions de capteurs sélectionnées pour l’antenne optimale à différents niveaux de RSB : 12 dB (rouge), 9 dB (orange), 6 dB (vert), 3 dB (bleu clair) et 0 dB (bleu foncé). Chaque sous-figure correspond à un nombre de snapshots différent.

temps, on a pu observer et interpréter l'allure du critère de sélection de positions de capteurs en fonction des différents paramètres physiques de la localisation de sources. Deux liens intéressants entre le critère de sélection et les fondamentaux en antennerie classique a été mis en évidence. Notamment, on a remarqué que le critère favorise naturellement :

- une amélioration de la géométrie d'antenne, dans le cas où le critère de Shannon Nyquist spatial n'est pas respecté, et
- une augmentation de la taille de l'antenne en basse fréquence.

Puis, ajoutant l'étape de seuillage en post-traitement, la pertinence de l'algorithme a pu être démontrée. Une étude paramétrique en fonction du RSB a été présentée. Ensuite, une comparaison avec différentes conceptions d'antennes a été présentée. On a démontré le niveau de performance supérieur de notre algorithme par rapport à l'état de l'art.

Enfin, nous avons étudié les extensions de notre algorithme, en modifiant la partie estimation des positions de sources, ou la partie sélection de positions de capteurs de l'algorithme. On a pu confirmer que l'algorithme avec SBI, couplé avec une sélection de capteurs itérative D-optimale, montre les meilleures performances. Une dernière étude sur l'implémentation de l'algorithme SBI avec plusieurs snapshots a permis de valider son intérêt dans des scénarios à faible RSB.

Pour finir sur ce chapitre, les résultats présentés ont fait l'objet de deux conférences : une à *EUSIPCO* [132] et une à *Forum Acousticum* [133].

TROISIÈME PARTIE

**Application en localisation de
sources aériennes en salle
réverbérante à partir
d'expérimentations**

MISE EN PLACE D'UNE EXPÉRIENCE

Sommaire

4.1	Introduction	85
4.2	Présentation de la salle MODO	86
4.2.1	Histoire du projet	86
4.2.2	Le microphone MEMS	87
4.2.3	Description de la salle	88
4.2.4	Géométrie de l'antenne	89
4.2.5	Système d'acquisition	89
4.2.6	Calibration géométrique de l'antenne	90
4.3	Expérience	91
4.3.1	Source monopolaire	91
4.3.2	Positionnement de la source	93
4.3.3	Les signaux émis	95
4.3.4	Température et célérité du son	96
4.3.5	Pré-traitements	96
4.4	Localisation de sources acoustiques en milieu réverbérant	97
4.4.1	Estimation de l'ordre de réflexion	97
4.4.2	Estimation des coefficients de réflexion	101
4.5	Conclusion	101

4.1 Introduction

Après avoir présenté la localisation de sources et la sélection de positions de capteurs afin d'arriver à l'algorithme proposé dans ce manuscrit, la sélection séquentielle bayésienne orientée par les données, nous avons utilisé ce même algorithme dans le contexte de la localisation de sources acoustiques sous-marines. Cependant, cette étude n'était basée que sur des données simulées. Nous souhaitons maintenant évaluer notre algorithme sur des données réelles. Malheureusement, aucune expérience en localisation de sources acoustiques sous-marines correspondant à notre situation n'a été trouvée. En effet elle doit faire état des sources multiples ou un nombre

de capteurs suffisant. Or mettre en place une expérience d'acoustique sous-marine respectant ces critères est chère et compliquée.

À défaut du domaine sous-marin, nous avons une solution dans le domaine de l'aérien via une salle située à Saint-Cyr L'école, appartenant à des collègues acousticiens de Sorbonne Université, au nom de "Les Murs Ont des Oreilles" ou MODO. Cette salle, de forme parallélépipédique, est couverte d'un nombre important de microphones situés sur 5 des six faces. Malheureusement, comme en acoustique sous-marine, aucune expérience déjà menée dans cette salle convenait pour évaluer notre algorithme. Mais grâce à l'accord et l'aide de collègues, nous avons pu mener des expériences adéquates. Nous avons donc acquis un nombre important de mesures pour lesquelles il a fallu appliquer un pré-traitement et vérifier si la localisation était possible avant d'utiliser l'algorithme de sélection de positions de capteurs. Cette expérience étant dans le domaine aérien, les paramètres physiques et le modèle de propagation de l'onde sont donc différents de ceux du chapitre 3.

Dans ce chapitre, nous décrivons la mise en place de l'expérience, l'acquisition des signaux et les premiers traitements. Nous étudions l'impact de la réverbération de la salle sur la localisation, un point qui ne pouvait pas être traité en acoustique sous-marine. La sélection de positions de capteurs ne sera étudiée que dans le chapitre suivant. Dans un premier temps, nous présentons dans la partie 4.2 la salle MODO et ses particularités grâce à des études menées auparavant. La mise en place de notre expérience, les contraintes, les paramètres physiques et le pré-traitement appliqué sur les mesures sont donnés dans la section 4.3. Avant de conclure en 4.5, nous étudions en partie 4.4 la nécessité d'introduire la réverbération dans le modèle de propagation.

4.2 Présentation de la salle MODO

Dans cette partie nous présentons l'histoire de la conception de cette salle et du projet qui lui a donné naissance. Ce projet existe et évolue autour d'un composant technologique précis : le microphone MEMS (pour *MicroElectroMechanical Systems*). Ceux-ci sont au nombre de 1024 et sont disposés sur les murs de la salle et au plafond de façon plus ou moins aléatoire. Une conception d'antenne avec 1024 microphones MEMS nécessite un système d'acquisition complexe afin que les mesures soient synchronisés. Il est donc important de présenter ce système d'acquisition.

4.2.1 Histoire du projet

La salle "Les Murs ont des Oreilles" ou MODO est le dernier produit du projet *Megamicros* de Sorbonne Université. En 2012, l'institut Jean le Rond d'Alembert (UMR 7190) développe et conçoit un système d'acquisition à grand nombre de voies en utilisant une technologie massivement démocratisée en 2011, le microphone *MEMS* [134]. Le MEMS, qui a reçu un développement

considérable grâce aux appareils mobiles connectés (par exemple les téléphones), a permis de réduire la contrainte matérielle de la conception d’antennes acoustiques. Cette technologie réunit toute la chaîne d’instrumentation dans un capteur de taille millimétrique. Sa faible consommation d’énergie permet de l’intégrer dans un réseau à grand nombre de capteurs. L’utilisation de ces microphones MEMS pour concevoir des antennes a commencé en 2012 par Hafizovic et al [135], et leur utilisation massive a permis d’arriver actuellement à des antennes de 4096 microphones [136].

Cependant le projet *Megamicros* n’a pas seulement pour but de déployer des antennes de très grandes dimensions mais aussi de proposer des géométries arbitraires afin de capter et d’étudier le champ acoustique dans un large spectre d’applications. Ces applications concernent essentiellement l’analyse des champs acoustiques audibles et la résolution de problèmes inverses pour la séparation, la localisation et l’identification de sources sonores [137].

Si le projet *Megamicros* a commencé avec un système de 128 microphones [138], il a évolué avec le temps en proposant un système à 256 microphones [139]. Puis l’augmentation du nombre de microphones à 1024 a permis le passage d’une antenne de 2 dimensions à 3 dimensions [140]. Tout au long de son évolution, le projet a permis la réalisation d’un grand nombre d’acquisitions de mesures expérimentales. On retrouve des mesures en environnement urbain devant une route et une piste d’essai [141], ou au sommet d’un immeuble [142], des études de bruit d’écoulement en soufflerie anéchoïque [143] et la création d’un holographe temps réel [144].

Dans notre cas, nous utilisons le système MODO. Les études réalisées dans MODO sont prometteuses et en lien avec les travaux de cette thèse. On retrouve notamment des travaux sur la résolution de problèmes inverses et de l’identification de sources complexes [32]. On trouve aussi l’utilisation d’approches parcimonieuses comme le Block Orthogonal Matching Pursuit [20]. Ces diverses études confirment que l’utilisation de MODO est envisageable pour évaluer notre algorithme à partir de données expérimentales.

4.2.2 Le microphone MEMS

Composant principal des antennes de *Megamicros*, le MEMS est un microphone de taille $5 \times 4 \times 1$ millimètres (voir figure 4.1). Ce composant, du fait de sa taille (de l’ordre d’une puce électronique) et de sa fabrication à l’échelle industrielle, a vu son prix unitaire réduire nettement. Sa construction industrielle amène aussi une forte homogénéité dans les caractéristiques entre différents MEMS. Sa principale caractéristique est qu’il assure une acquisition omnidirectionnelle entre 60 Hz et 15 kHz. De plus, le capteur permet en sortie de délivrer le signal de pression numérisé grâce à la présence d’un convertisseur analogique-numérique intégré. Celui-ci par son échantillonnage, réduit le bruit de quantification et, par son filtre numérique passe-bande, écarte les basses fréquences et évite le repliement spectral.

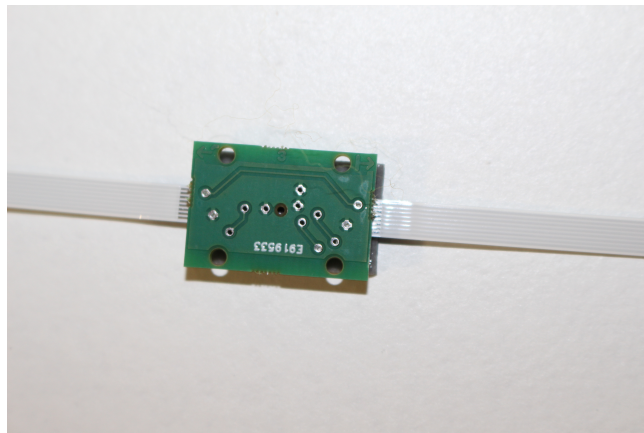


FIGURE 4.1 – Le microphone MEMS.

4.2.3 Description de la salle

La salle où sont disposés les microphones est un parallélépipède de taille $8,01 \times 3,75 \times 2,94$ mètres. Les murs et le plafond sont des plaques de plâtre sans revêtement. Un épais tapis est posé sur le sol en béton et apporte ainsi de l'absorption acoustique à la pièce. Les réflexions au sol seront donc moins importantes que sur les autres surfaces. Les parois de la salle sont utilisées comme support naturel de l'antenne, à l'exception du sol. Une hypothèse nécessaire mais approximative sera faite concernant la salle : la salle est parfaitement parallélépipédique avec des parois parfaitement parallèles deux à deux (voir figure 4.2). Cette hypothèse est importante dans la calibration géométrique ainsi que dans le modèle de propagation des réflexions de l'onde, qui seront présentés plus loin.

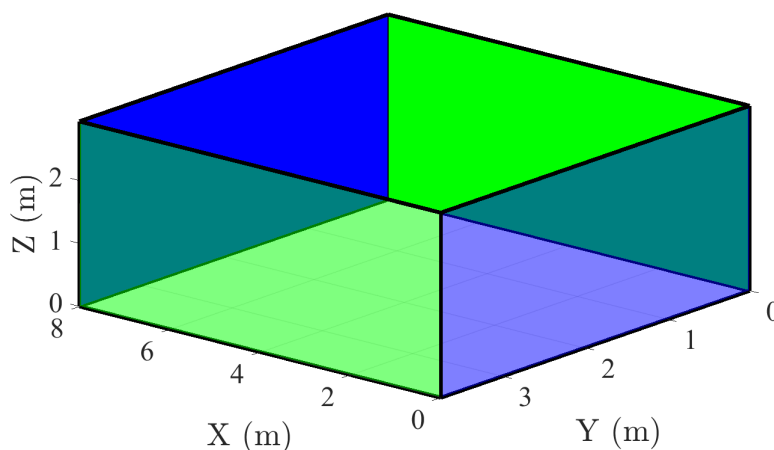


FIGURE 4.2 – Géométrie de la salle MODO.

4.2.4 Géométrie de l'antenne

De cette salle sont répartis 128 faisceaux sur les 4 murs et le plafond. Chaque faisceau (voir figure 4.3) correspond à une ligne composée de 8 MEMS. De par la légèreté et la taille réduite des MEMS, le faisceau est collé à la surface par une bande adhésive double face. Cela limite notamment les interactions acoustiques avec un éventuel portant métallique. Les MEMS sont connectés entre eux sur un même faisceau par des câbles plats FFC. La distance séparant 2 MEMS sur 1 faisceau est donc dictée par la longueur du câble FFC. 3 longueurs sont utilisables : 15, 30 et 45 centimètres. En jouant sur ces 3 longueurs et sur le positionnement des faisceaux sur les différentes surfaces, l'antenne adopte une géométrie pseudo-aléatoire (voir figure 4.4). Ce choix a été initialement fait afin de supprimer les lobes de réseaux. Une contrainte est cependant appliquée dans la disposition des faisceaux. Afin de garder une discrétisation homogène, une densité surfacique d'environ 8 MEMS par mètre carré est maintenue sur toute l'antenne.

Les premières études ont montré que [140] :

- la résolution spatiale est constante dans le milieu et est égale à $\frac{\lambda}{2}$ et est indépendante de la forme du réseau,
- le niveau de la dynamique de reconstruction varie avec le nombre de microphones,
- l'antenne possède une forte robustesse au repliement spectral.

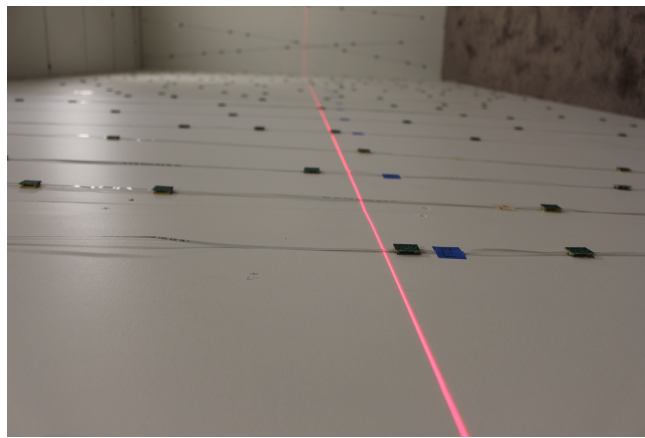


FIGURE 4.3 – Ensemble de faisceaux.

La figure 4.5 montre les signaux temporels mesurés par chaque microphone. On remarque rapidement que 4 MEMS ne fonctionnent pas. On ne considérera donc pas ces 4 microphones par la suite. Le nombre de capteurs total sera donc de 1020.

4.2.5 Système d'acquisition

Chaque faisceau est connecté à un buffer où les échantillons numériques sont agrégés par paquets. Un câble ethernet RJ45 assure la connexion, l'alimentation du faisceau ainsi que la

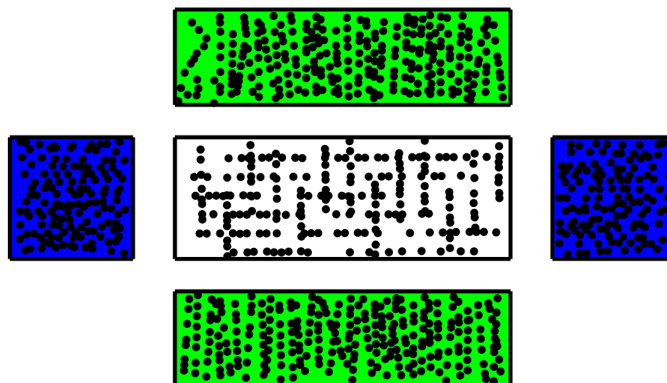


FIGURE 4.4 – Patron de la salle MODO et géométrie de l’antenne.

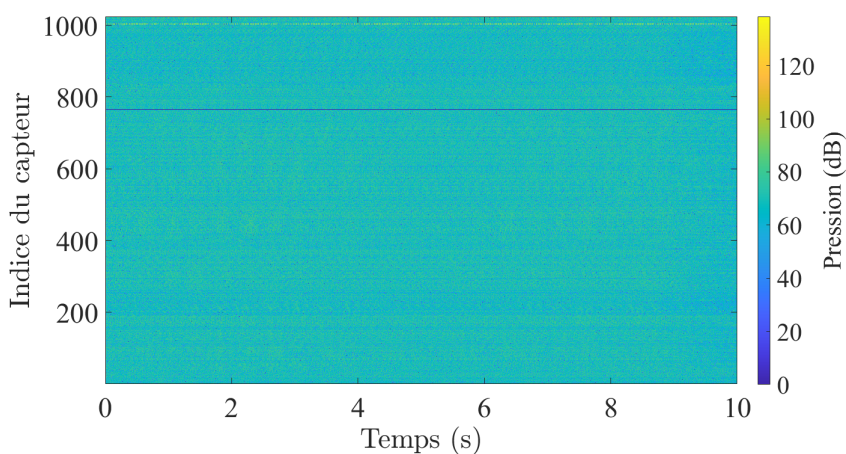


FIGURE 4.5 – Signaux temporels mesurés par chaque microphone.

transmission des données vers l’interface d’acquisition. Cette interface d’acquisition impose une horloge commune à tous les capteurs et permet le multiplexage temporel des signaux. Le grand nombre de voies est ainsi géré de façon synchrone. Les données récoltées sont ensuite transmises à l’ordinateur via un protocole standard USB 3.0 à débit élevé. Le système d’acquisition est en photo dans la figure 4.6.

4.2.6 Calibration géométrique de l’antenne

Une fois que les capteurs MEMS sont répartis dans la salle et le système d’acquisition est fonctionnel, il est nécessaire de connaître le plus précisément possible les positions des capteurs pour la localisation et la sélection, puisque le modèle de propagation est défini à partir des

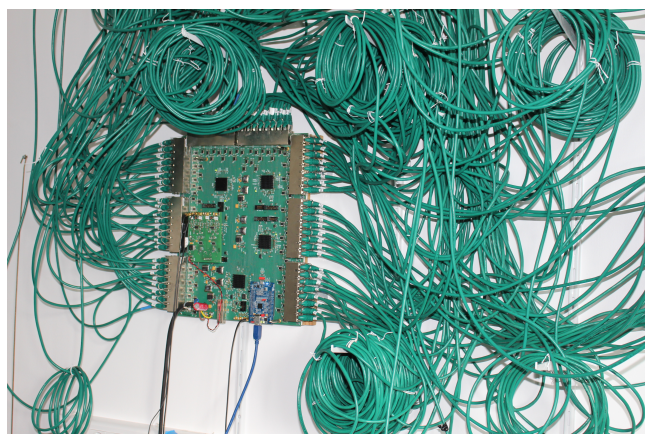


FIGURE 4.6 – Système d’acquisition.

positions. Malheureusement, en expérimental il existe toujours des imprécisions qui entraînent des erreurs de modèle qui ne pourront être connues ou déterminées. Lors de la mise en place de la géométrie de l’antenne, pour mesurer les positions de capteurs une procédure en 2 étapes a été suivie lors de la thèse de Hugo Demontis [140]. Cette procédure de calibration géométrique a permis d’obtenir les positions des capteurs avec une précision satisfaisante.

4.3 Expérience

Dans les études précédentes menées dans MODO [140], aucune acquisition ne correspondait exactement au problème que nous considérons. Nous avons cependant eu l’opportunité de mener nos propres expérimentations. Cette section résume la mise en place et la prise de mesures qui a pu être effectuée.

Cette prise de mesures consiste à placer une source monopolaire à différentes positions dans la salle, à la même hauteur, pour émettre trois types de signaux temporels qui seront captés par les 1020 capteurs. La température sera mesurée afin de connaître la célérité du son dans la salle. Dans la suite nous présentons : la source monopolaire, le positionnement (en incluant la hauteur) de la source, l’émission des signaux et la mesure de température.

Notons que d’autres acquisitions ont pu être menées, avec différents types de sources et sont présentées dans l’Annexe 5.7. Elles n’ont pas pu être exploitées lors de la thèse.

4.3.1 Source monopolaire

La source est un haut-parleur de 4 centimètres de diamètre, bafflé à un cube de 7 centimètres de côté (voir figure 4.7). Dans la suite, notamment pour l’utilisation de notre algorithme dans le chapitre 5, nous considérons la source comme ponctuelle et omnidirectionnelle. Ce modèle est

approximatif et engendrera des erreurs de modèle non mesurables. Afin de les réduire, il faudrait prendre en compte un diagramme de directivité plus réaliste :

- Demontis s'appuie sur un modèle piston bafflé infini [140],
- Aart et al. [145] suggèrent de modéliser une telle source par un dôme sphérique,
- faire une simulation par éléments finis ou des mesures du diagramme de directivité en chambre anéchoïque.

Mais ce travail sur la directivité est hors du sujet d'étude de ma thèse. Il pourra cependant être étudié dans des travaux futurs.

Nous souhaitons pouvoir considérer plusieurs configurations, dont certaines nécessitent de positionner plusieurs sources dans la salle. Cependant, nous n'avons à notre disposition qu'une seule source. Nous avons donc du procéder différemment en faisant des acquisitions pour différentes positions de source. Pour cela, nous allons déplacer la source à différentes positions, et émettre les signaux à chacune d'elles. Pour concevoir les différentes configurations, nous sélectionnons aléatoirement N_S sources, et additionnons les vecteurs de mesures correspondant à chacune de ces N_S sources.

Faire ainsi implique de considérer une hypothèse de linéarité entre les sources et amène que chaque source soit indépendante et identique. Cette procédure permet de créer différentes configurations avec différentes positions de sources sans avoir à faire de nouvelles manipulations. De plus, considérer que les sources présentes, lors de l'écriture du problème de localisation de sources, émettent toutes le même signal et avec la même directivité permet de simplifier l'estimation des positions de sources.

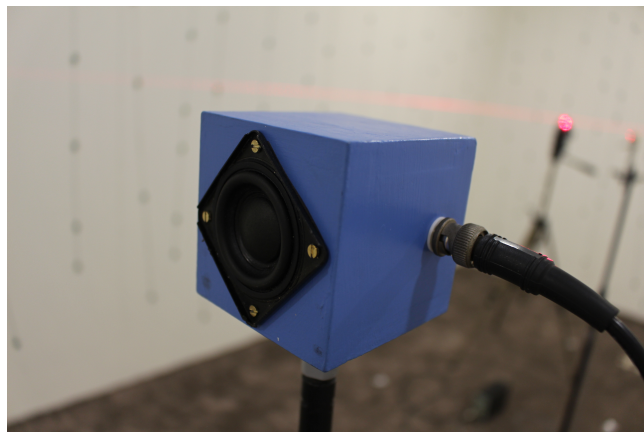


FIGURE 4.7 – Source monopolaire.

4.3.2 Positionnement de la source

Nous nous limitons au cas où les positions de la source sont toutes dans le même plan. Une hauteur $z_S = 1,35$ mètres est choisie arbitrairement (figure 4.8). Cependant le sol de la salle n'est pas parfaitement plan, même si nous faisons cette hypothèse par la suite avec le problème que nous considérons. Afin de s'assurer de placer la source toujours à la même hauteur, nous avons utilisé un télémètre laser (figure 4.9). Nous plaçons le milieu de la source à la hauteur du laser, comme illustré dans la figure 4.10.

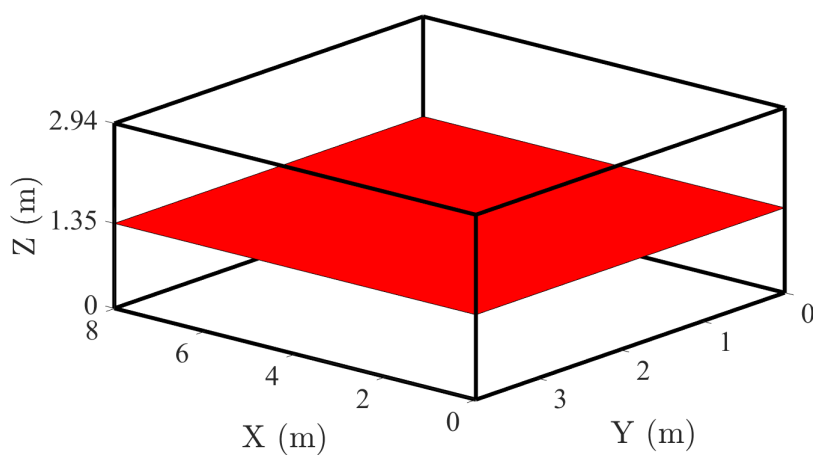


FIGURE 4.8 – Hauteur des sources dans la salle.



FIGURE 4.9 – Télémètre-laser.

On définit arbitrairement 25 positions de sources, comme le montre la figure 4.11 et donc à la hauteur z_S . Ces positions sont choisies aléatoirement. Une condition a été considérée : aucune source ne peut être placée à moins de 0,5 mètres d'un mur afin d'éviter les trop fortes réflexions sur les murs et la possible saturation des microphones. De plus cela permettra par la suite de

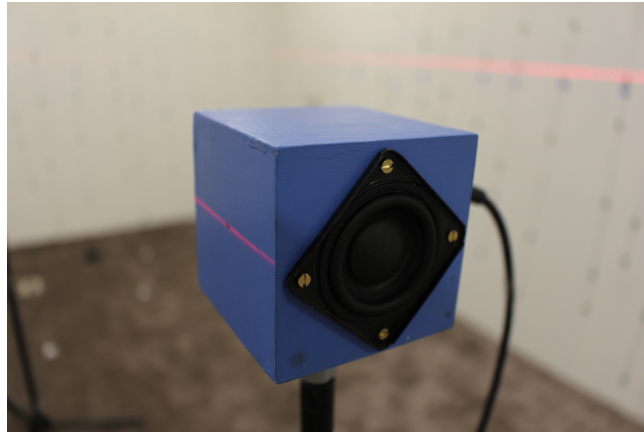


FIGURE 4.10 – Positionnement du centre de la source en z_S via le télémètre laser.

réduire l'espace de recherche des sources et donc de limiter les temps de calcul et la taille des matrices.

Par laser, on mesure la distance de la source par rapport à 2 murs : distance qui nous donne la position de la source x_S selon l'axe X et y_S selon l'axe Y . On considère que l'origine de l'onde est au centre de la membrane.

Enfin la source n'étant pas omnidirectionnelle, son orientation peut jouer sur la puissance reçue par les capteurs. Nous avons donc décidé de placer la source toujours dans le sens des x croissants (voir figure 4.11).

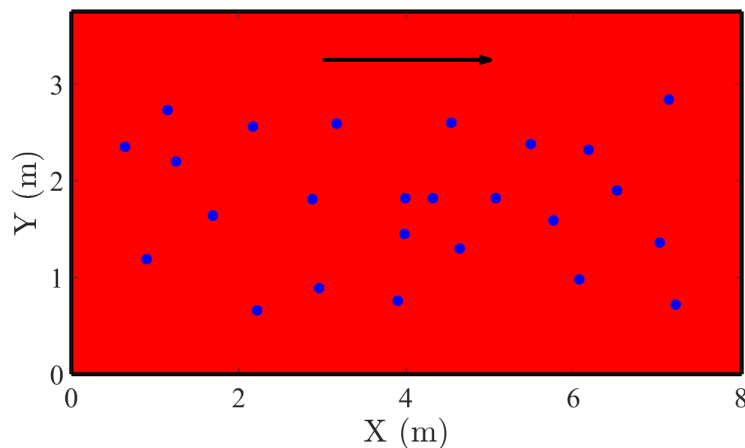


FIGURE 4.11 – Positions dans le plan XY (à la hauteur z_S) de la source. L'orientation de la source est donnée par la flèche noire.

4.3.3 Les signaux émis

À chaque position, la source émet trois types de signaux : un bruit blanc, un chirp et un peigne fréquentiel. Les signaux sont réels lors de l'émission et durent 10 secondes. La figure 4.12 présente les spectrogrammes des trois signaux.

Le bruit blanc est un signal temporel aléatoire dont sa densité spectrale de puissance est constante en fonction de la fréquence. Le chirp commence à 125 Hz et finit à 8000 Hz. La fréquence du chirp évolue de manière exponentielle au cours du temps. Le peigne fréquentiel est la somme de 5 sinus. Chaque sinus correspond à une fréquence. Les fréquences choisies pour le peigne fréquentiel sont les suivantes : 50, 500, 1000, 2000 et 4000 Hz.

Si pour le peigne fréquentiel et le bruit blanc, les fréquences souhaitées apparaîtront sur un grand nombre d'échantillons temporels, cela n'est pas le cas pour le chirp où le nombre d'échantillons contenant une fréquence donnée est faible. Afin de posséder un nombre important d'échantillons à chaque fréquence, nous émettons pour chaque position, non pas 1 fois comme pour le peigne fréquentiel et bruit blanc, mais 10 fois le chirp. Pour chaque position, nous avons donc 12 prises de mesures.

Rappelons que l'algorithme proposé dans ce manuscrit est mono-fréquentiel. Nous verrons lors du pré-traitement (partie 4.3.5), que la transformée de Fourier nous permettra de travailler avec une seule fréquence, quelque soit le type de signal émis. Pour la suite du manuscrit, nous avons choisi de travailler uniquement sur le peigne fréquentiel. Nous avons considéré que le chirp et le bruit blanc étaient moins adaptés pour une étude mono-fréquentielle que le peigne fréquentiel. Cependant, l'étude de Shoichi Koyama [118] montre qu'il est possible de concevoir une antenne à partir d'une minimisation du déterminant de la matrice de covariance pour un signal large bande. Dans des travaux futurs, il serait alors intéressant de travailler sur des signaux large bande comme le chirp ou bruit blanc, en prenant en compte la modification apportée par cet article.

Le peigne fréquentiel contient 5 fréquences différentes, or comme écrit précédemment, l'algorithme proposé est mono-fréquentiel. Il est donc nécessaire de choisir une des fréquences du signal lors de l'utilisation de l'algorithme. Afin d'éviter les comportements basses et hautes fréquences, vus dans les chapitres 1 et 3, nous choisissons la fréquence centrale : 1000 Hz. À la suite de cette thèse, il serait intéressant de comparer la conception d'antenne en fonction de la fréquence du signal émis par les sources.

On précise que les signaux émis par les enceintes sont des sons enregistrés avec une fréquence d'échantillonnage de 44100 Hz et que la fréquence d'échantillonnage pour le système de réception est de 25000 Hz. Par la suite, dans le pré-traitement (partie 4.3.5), nous utiliserons une transformée de Fourier. Afin de respecter le théorème d'échantillonnage, la fréquence maximale en sortie de la transformée de Fourier est la fréquence d'échantillonnage en réception (ici 25 kHz) divisée par deux.

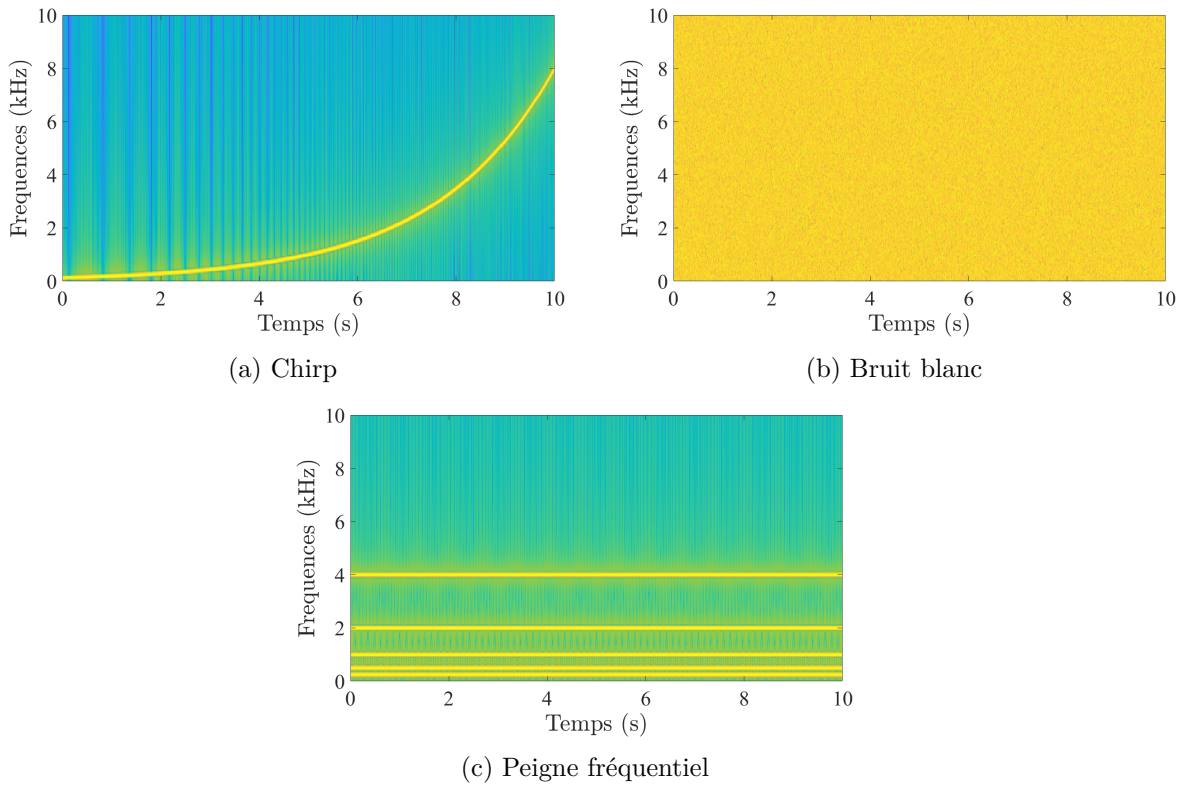


FIGURE 4.12 – Spectrogrammes des signaux.

4.3.4 Température et célérité du son

Les prises de mesures pour les 25 positions de sources prenant du temps (quelques heures), nous avons effectué des mesures de température à chaque position. En effet la célérité c du son varie en fonction de la température $Temp$ (en ° Celsius) selon l'équation :

$$c = 331.5 + 0.607 Temp. \quad (4.1)$$

Le capteur de température a été placé sur le support cubique de la source, comme le montre la figure 4.13.

Les mesures de température menées lors de l'expérience sont en moyenne de $21,1 \pm 0,1^\circ$. La célérité du son, selon l'équation (4.1), est en moyenne de $344,3 \pm 0,1$ mètres par secondes. On peut en déduire que les variations de célérité lors de l'expérience sont faibles.

4.3.5 Pré-traitements

On présente les pré-traitements qui ont été réalisés afin de pouvoir appliquer un algorithme de localisation de sources ou notre algorithme de sélection de positions de capteurs.

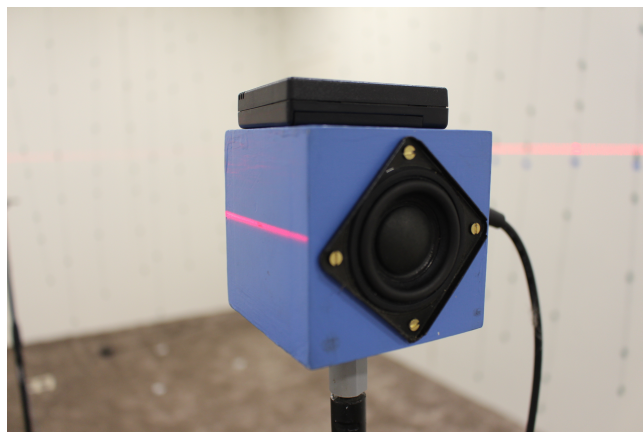


FIGURE 4.13 – Capteur de température positionné sur le support cubique de la source.

Les données enregistrées sont des entiers codés sur 32 bits. Elles sont transformées en pression en appliquant un gain donné par le constructeur des MEMS. On applique ensuite une transformée de Fourier sur chaque ligne de la matrice des pressions temporelles. Pour finir, pour chaque ligne de la matrice spectrale, on ne conserve que l'amplitude fréquentielle correspondant à la fréquence émise par la source.

Dans le cas où le signal émis est le peigne fréquentiel, il est possible de diminuer la durée du signal considérée lors de la transformée de Fourier. Ce choix peut permettre d'obtenir plusieurs snapshots. L'hypothèse émise dans ce contexte est que chaque partie du signal utilisée pour les snapshots sont indépendantes entre elles. Ayant pour but de fournir dans ce manuscrit une étude sur l'utilisation du multi-snapshots, nous décidons d'utiliser seulement 0,25 secondes de signal lors de la transformée de Fourier. Cette durée nous permettra dans la partie 5.5.1, d'aller jusqu'à $N_T = 40$ snapshots.

4.4 Localisation de sources acoustiques en milieu réverbérant

Cette section a pour but de montrer la nécessité de prendre en compte la réverbération dans le modèle de propagation pour l'estimation des positions de sources. Dans un premier temps, nous proposons de visualiser la différence dans le champ acoustique estimé par 4 algorithmes de localisation de sources pour différents ordres de réflexions. Le critère utilisé pour définir l'ordre de réflexion est la corrélation normalisée entre les mesures synthétiques et les mesures expérimentales. Ce même critère nous permettra en un second temps, de déterminer les coefficients de réflexion de chaque surface (murs, plafond et sol).

4.4.1 Estimation de l'ordre de réflexion

On décide de décomposer cette étude en trois parties :

- visualiser le champ acoustique estimé par 4 algorithmes de localisation de sources pour différents ordres de réflexions,
- visualiser le champ acoustique estimé par formation de voies en considérant un champ libre sur un espace étendu,
- calcul de la corrélation normalisée pour différents ordres de réflexions.

Le premier point nous permet de comparer l'estimation fournie par les algorithmes de localisation de sources selon l'ordre de réflexion considéré, et donc de voir l'impact de la réverbération sur l'estimation du champ acoustique. Le deuxième point permet de visualiser de manière élégante l'énergie des réflexions. Pour finir, le dernier point permet par un critère mathématique de définir l'ordre de réflexion correspondant au mieux aux données expérimentales.

Visualisation du champ acoustique selon l'algorithme de localisation de sources

Dans cette partie, on compare la localisation de sources si le modèle de propagation est défini en champ libre ou en considérant un ordre de réflexion égal à 3. Pour fournir une visualisation interprétable et intéressante, on sélectionne aléatoirement 2 sources parmi les 25. Nous avons remarqué que le champ acoustique avec un nombre plus faible ou plus grand de sources était plus difficile à visualiser et à interpréter. Les positions des sources selon les axes ($X; Y; Z$) en mètres sont les suivantes : (4, 54; 7, 14; 1, 35) et (2, 6; 2, 8; 1, 35). Pour le problème inverse de localisation de sources, on discrétise notre domaine de recherche avec un pas de 0,25 mètres selon l'axe X et l'axe Y . L'axe Z est fixé à la hauteur des sources, c'est-à-dire, à 1,35 mètres.

Dans la figure 4.14, on compare les champs acoustiques estimés par 4 algorithmes de localisation de sources : l'inférence bayésienne parcimonieuse, l'orthogonal matching pursuit, le basis pursuit de-noising et la formation de voies.

On remarque que quelque soit l'algorithme de localisation de sources choisi, à l'exception de l'orthogonal matching pursuit, la prise en compte des réverbérations améliore les performances d'estimation des positions de sources. Pour l'inférence bayésienne parcimonieuse, la prise en compte des réverbérations permet de diminuer les amplitudes estimées aux points où les sources ne sont pas présentes. On peut supposer que ces amplitudes sont liées aux réflexions sur les murs et que ne pas les prendre en compte dans le modèle de propagation les font apparaître dans l'estimation du champ acoustique.

Visualisation du champ acoustique en champ libre

Pour connaître le niveau de l'énergie contenue dans les réflexions, nous proposons de visualiser le champ acoustique estimé par formation de voies. Cependant nous considérons deux points importants : le modèle de propagation est considéré comme du champ libre et l'espace considéré n'est plus limité à la salle. Pour ce second point, le milieu de propagation est étendu afin de

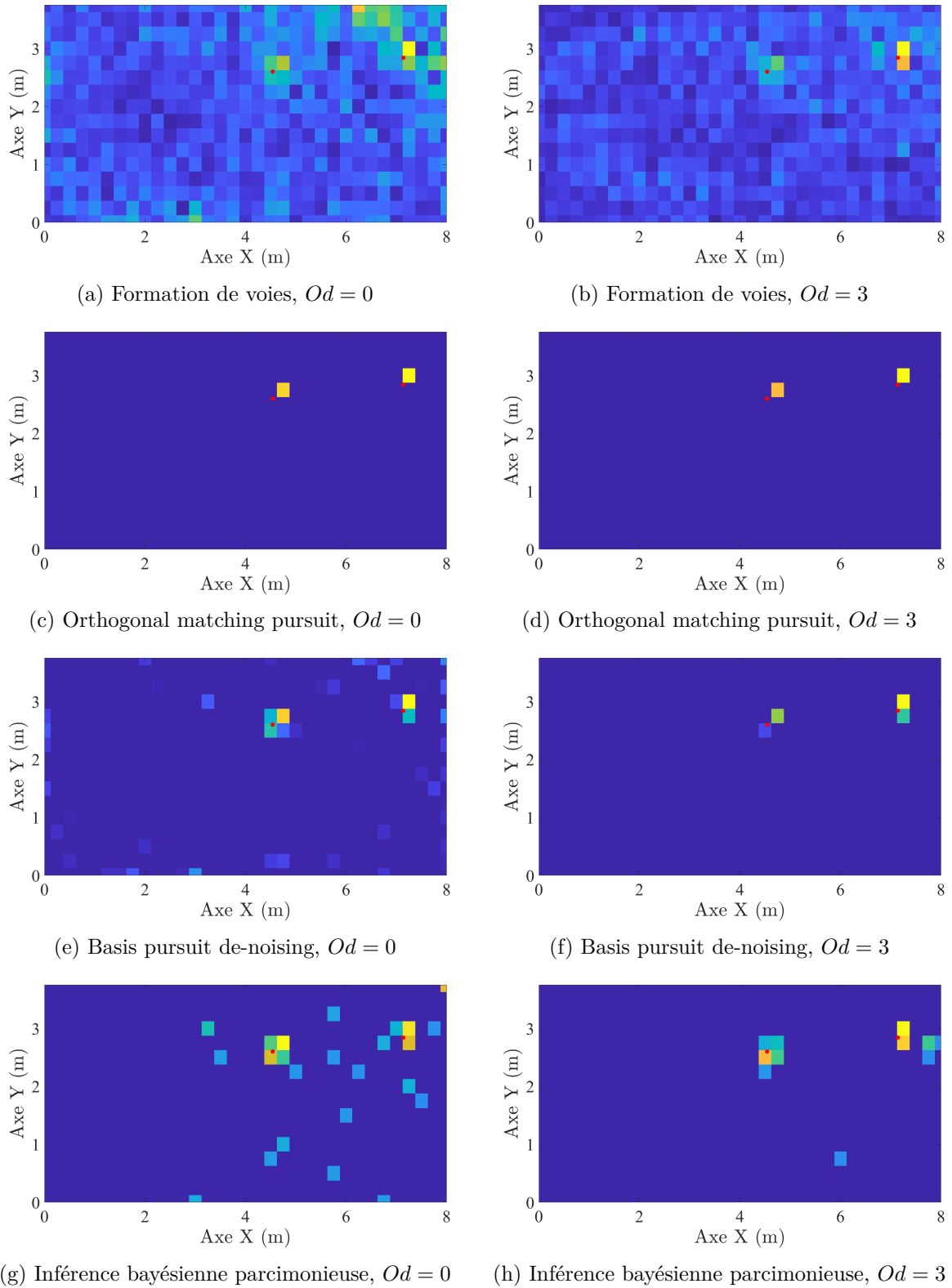


FIGURE 4.14 – Champs acoustiques estimés à partir des mesures expérimentales par différents algorithmes de localisation de sources en considérant un ordre de réflexion (Od) égal à 0 ou 3.

pouvoir observer les amplitudes coorespondant à un ordre de réflexion égal à 4, c’est-à-dire, que l’axe X est délimité entre $-16,02$ et $24,03$ mètres et que l’axe Y entre $-7,5$ et $11,25$ mètres. Dans la figure 4.15, les ordres de réflexion sont visibles par les délimitations de différentes couleurs. Notons que cette représentation visuelle ne permet d’observer que l’énergie liée aux réflexions par rapport aux murs. L’énergie liée aux réflexions par rapport au plafond et au sol ne seront pas visibles.

Par la figure 4.15, on constate que les amplitudes estimées pour un ordre de réflexion de 1 à 3 restent fortes par rapport aux amplitudes estimées pour un ordre de réflexion égal à 0. Les amplitudes estimées pour un ordre de réflexion à 4 sont quant à elles faibles par rapport aux autres amplitudes. Par cette visualisation, on confirme que les réverbérations sont trop importantes et ne peuvent être négligées par la suite dans le modèle de propagation.

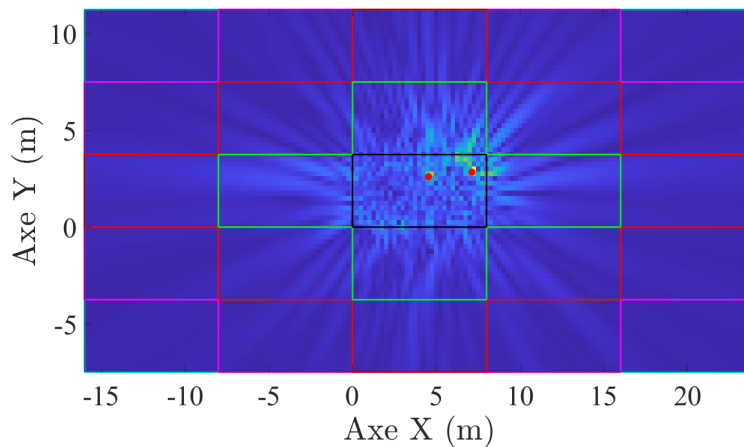


FIGURE 4.15 – Champ acoustique estimé par formation de voies en champ libre sur un espace étendu. L’ordre de réflexion est illustré par les délimitations de différentes couleurs : 0 (noire), 1 (vert), 2 (rouge), 3 (magenta) et 4 (cyan).

Calcul de la corrélation normalisée

Nous cherchons à estimer l’ordre de réflexion correspondant aux données expérimentales. Pour cela nous allons synthétiser des données que nous comparons par la suite aux données expérimentales par la corrélation normalisée. On cherche donc l’ordre de réflexion qui maximise la corrélation normalisée entre les données synthétiques et les données expérimentales. Les valeurs sont données dans le tableau 4.1. On remarque une augmentation de la corrélation normalisée avec l’ordre de réflexion entre 0 et 2 puis une diminution de la corrélation normalisée pour les ordres 3 et 4. Remarquons que la différence de la corrélation entre l’ordre 2 et 3 est faible. L’ordre 2 est donc l’ordre qui correspond au mieux aux données expérimentales, puisque c’est l’ordre qui maximise la corrélation normalisée.

Ordre de réflexion	0	1	2	3	4
Corrélation normalisée	0,4320	0,5840	0,6919	0,6819	0,6107

TABLE 4.1 – Corrélation normalisée entre données expérimentales et données simulées avec différents ordres de réflexion considérés.

4.4.2 Estimation des coefficients de réflexion

On cherche à déterminer les coefficients de réflexion pour chaque surface de la salle, c'est-à-dire, les murs, le plafond et le sol. On utilise une nouvelle fois la corrélation normalisée pour comparer les données expérimentales avec des données synthétiques. Le but est de trouver le jeu de coefficients de réflexion qui maximise la corrélation normalisée entre les 2 jeux de données. Par une recherche exhaustive, nous avons pu estimer les coefficients de réflexions. On donne les valeurs utilisées par la suite du manuscrit dans le tableau 4.2.

Surfaces considérées	Murs	Plafond	Sol
Coefficient de réflexion	0.74	0.5	0.73

TABLE 4.2 – Coefficients de réflexion estimés pour chaque surface.

4.5 Conclusion

Ce chapitre décrit l'expérience menée lors du manuscrit. Cette expérience, rendue possible grâce aux collègues de Sorbonne Université, se sert de la salle MODO. Après une présentation de la salle, des microphones, de la géométrie de l'antenne et de son système d'acquisition, nous avons entrepris l'acquisition de nombreuses mesures. Nous avons décrit par la suite la source et les signaux émis par celle-ci, les différentes positions de la source lors de l'expérience ainsi que les mesures de température nous donnant la célérité du son dans le milieu. Pour finir, nous avons explicité la transformation des mesures du domaine temporel au domaine fréquentiel qui nous permet d'utiliser les algorithmes de localisation de source et de conception d'antenne. Ces algorithmes sont mono-fréquentiel et seront étudiés dans le chapitre suivant.

Ce chapitre propose aussi une étude, principalement visuelle, sur la réverbération du milieu. Cette recherche permet, par l'utilisation de corrélation normalisée, de fixer les paramètres concernant la réverbération du milieu. Ces paramètres sont nécessaires pour la suite du manuscrit.

APPLICATION DE LA SÉLECTION DE CAPTEURS AUX DONNÉES EXPÉRIMENTALES

Sommaire

5.1	Introduction	103
5.2	Évaluation des performances en fonction du nombre de sources	104
5.3	Impact de l'ordre de réflexion	106
5.3.1	Données simulées	107
5.3.2	Données expérimentales	109
5.4	Comparaison avec d'autres méthodes	111
5.5	Etudes complémentaires	115
5.5.1	Extension multi-snapshots	115
5.5.2	Choix du critère d'optimalité	116
5.6	Applications en conception d'antenne	118
5.7	Conclusion	121

5.1 Introduction

Dans le chapitre précédent, nous avons décrit l'expérimentation dans la salle MODO qui va nous permettre d'évaluer notre algorithme de sélection de capteurs et de localisation de sources en milieu aérien et réverbérant. Le but est de montrer l'intérêt de notre approche dans le cadre d'une autre application et surtout de l'évaluer à partir de données réelles. En effet dans le chapitre 3, qui porte sur de l'acoustique sous-marine, nous n'avons pu utiliser que des données simulées. Comme les erreurs de modèle causées par les imprécisions de mesures sont plus nombreuses avec des données expérimentales, nous souhaitons aussi vérifier que notre algorithme reste performant. D'autre part, nous nous plaçons maintenant dans un contexte de champ proche et en milieu réverbérant. Quel impact cela va-t-il avoir sur le résultat des différentes méthodes et notamment celle proposée ?

Le nombre de sources et l'ordre de réflexion sont deux paramètres qui impactent les performances des algorithmes de localisation de sources et de sélection de capteurs. Pour cela, nous présentons une étude sur l'évolution des performances de notre algorithme en fonction du nombre de sources présentes dans le problème de localisation en section 5.2, puis en fonction de l'ordre de réflexion en données simulées et expérimentales dans la section 5.3. Cette étude nous permet par la suite de fixer ces paramètres pour la suite des études présentées dans ce chapitre.

Ensuite, similairement au chapitre 3, nous proposons une étude comparative avec des approches de l'état de l'art dans la section 5.4 et quelques études annexes dans la section 5.5 avec l'utilisation de multi-snapshots et la comparaison entre les différents critères d'optimalité de la conception optimale.

Avant de conclure en section 5.7, nous proposons une étude dans la section 5.6 sur des conceptions d'antennes provenant de données simulées. Ces antennes sont ensuite appliquées aux données expérimentales. Faire ceci permet de concevoir une antenne aux performances satisfaisantes. On montre alors que la sélection orientée par des données peut se faire à partir de données simulées pour ensuite l'utiliser avec des données expérimentales.

5.2 Évaluation des performances en fonction du nombre de sources

Présentation de l'étude

Nous proposons dans cette partie d'évaluer les performances de la sélection de capteur en fonction du nombre de sources. Nous allons considérer pour cela notre algorithme et une sélection aléatoire. À notre connaissance, ce sont les deux seules approches de conception d'antenne séquentielle qui peuvent être utilisées avec un grand nombre de capteurs.

L'étude consiste à appliquer les 2 algorithmes pour différents nombres de sources N_S : 1 à 6 et évaluer les performances de localisation de sources en fonction du nombre de capteurs sélectionnés. L'évaluation de la performance de localisation ne se fera pas par l'indice de Jaccard comme dans le chapitre 3. En effet, celui-ci nécessite de fixer un seuil qui est compliqué à trouver pour les mesures expérimentales, puisque celui-ci dépend de l'amplitude des signaux mesurés. Le seuil est alors dépendant de la position de la source. Afin de ne pas biaiser l'évaluation du fait de ce seuil, nous avons préféré considérer connaître le nombre de sources à estimer et leurs positions. Nous utilisons alors le pourcentage de vrais positifs donné par l'Algorithme 4. Ce dernier nécessite le rayon d'admission r_a qui par la suite sera fixé à $r_a = 0.25$ mètres.

On va donc moyenner le pourcentage de vrais positifs sur 300 réalisations. À chaque réalisation, on initialise l'algorithme avec $N_{K_0} = 2$ capteurs sélectionnés aléatoirement. Puis l'algorithme de sélection aléatoire et le notre sont appliqués jusqu'à avoir sélectionné les $N_M = 1020$ capteurs disponibles. On va alors regarder comment évolue les performances en fonction du nombre de capteurs sélectionnés et ainsi voir s'il est possible d'atteindre des limites hautes de

performances avec un nombre de capteurs réduit.

Notons que le domaine de recherche, c'est-à-dire, la salle MODO, est discrétisé par pas de 0,25 mètres. En considérant que les sources ne peuvent pas être dans un espace de 0,5 mètres autour des murs, on discrétise donc entre 0,5 et 7,5 mètres selon l'axe X et entre 0,5 et 3,25 mètres selon l'axe Y . On considère aussi que la hauteur du problème est connue (1,35 m), permettant de se restreindre à un problème à 2 dimensions. Par cette discrétisation, le nombre d'inconnues dans le problème inverse est de $N_G = 348$ inconnues. La discrétisation est illustrée dans la figure 5.1.

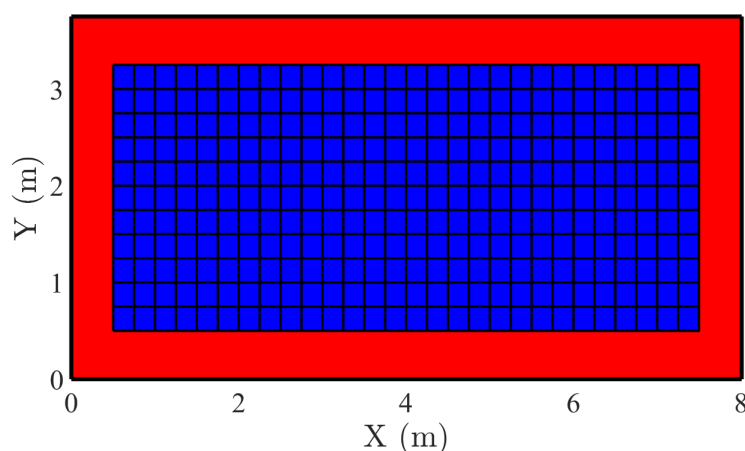


FIGURE 5.1 – Discretisation du domaine de recherche. Chaque intersection de la grille en noir correspond à une position de l'ensemble \mathcal{G} . La zone rouge correspond à l'espace de 0,5 mètres autour des murs qui n'est pas pris en compte.

Le résumé des paramètres de l'étude est donné dans le tableau 5.1.

Résultats

La figure 5.2 représente l'évolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées et du nombre de sources obtenu avec notre algorithme. On observe que le nombre de positions de capteurs sélectionnées nécessaire pour atteindre le niveau de performance maximal augmente avec le nombre de sources. Par exemple, avec 1 source, le nombre de positions de capteurs nécessaire pour atteindre ce niveau est inférieur à 50, alors qu'avec 6 sources ce nombre est supérieur à 200. Nous remarquons aussi que le niveau de performance maximal diminue avec le nombre de sources. Selon ce critère d'évaluation et le rayon d'admission choisi, il n'est pas possible de localiser 6 sources à toutes les réalisations.

Dans la figure 5.3, nous présentons une comparaison entre notre sélection optimisée et la sélection aléatoire pour différents nombres de sources. On observe que quelque soit le nombre de sources présentes dans le problème de localisation, la sélection optimisée a des meilleures

Paramètres	Valeurs
Nombre de sources N_S	[1; 2; 3; 4; 5; 6]
Positions des sources \mathbf{p}_S	Positions illustrées dans la figure 4.11
Hauteur des sources et du domaine de recherche	1,35 mètres
Type de signal émis	Peigne fréquentiel
Temps du signal émis	10 secondes
Temps du signal utilisé pour le snapshot	0,25 secondes
Fréquence considérée F	1000 Hz
Célérité de l'onde c	$\simeq 344,3$ mètres/secondes
Ordre de réflexion	3
Coefficient de réflexion pour les murs Cr_{murs}	0,74
Coefficient de réflexion pour le plafond Cr_{plafond}	0,5
Coefficient de réflexion pour le sol Cr_{sol}	0,73
Nombre d'inconnues N_G	348
Pas de discrétisation	0,25 mètres
Positions de la grille discrétisée \mathbf{p}_G	Positions illustrées dans la figure 5.1
Critère d'évaluation	Pourcentage de vrais positifs
Rayon d'admission du critère d'évaluation r_a	0,25 mètres
Nombre de réalisations	300
Nombre de capteurs disponible N_M	1020
Positions des capteurs \mathbf{p}_M	Positions illustrées dans la figure 4.4
Nombre de capteurs sélectionnés aléatoirement à l'initialisation des algorithmes N_{K_0}	2
Nombre de capteurs sélectionnés en sortie des algorithmes N_K	1020

TABLE 5.1 – Paramètres de l'étude.

performances que la sélection aléatoire. En effet, le nombre de positions de capteurs sélectionnées nécessaire pour atteindre le niveau de performance maximal est plus petit pour la sélection optimisée que la sélection aléatoire. Notons aussi que ce gain est de plus en plus important quand le nombre de sources augmente. Il est d'une dizaine de capteurs quand la source est unique, et une centaine quand le nombre de sources est supérieur ou égal à 4. On peut donc conclure que plus le nombre de sources est important, plus il est intéressant d'utiliser notre sélection optimisée par rapport à une sélection aléatoire afin de réduire le nombre de capteurs.

5.3 Impact de l'ordre de réflexion

Dans cette section, nous allons évaluer l'impact de l'ordre de réflexion, puisque nous sommes en milieu réverbérant, à partir des données simulées puis réelles. Les données simulées vont permettre d'évaluer l'impact des erreurs de modèles liées au nombre de réflexions considérées

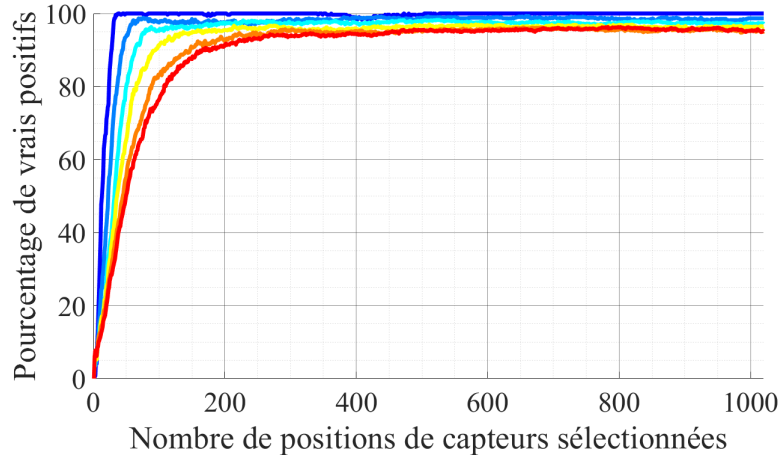


FIGURE 5.2 – Évolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées par notre algorithme pour différents nombres de sources présentes : 1 (bleu foncé), 2 (bleu), 3 (bleu clair), 4 (jaune), 5 (orange) et 6 sources (rouge).

pour simuler les données et donc le modèle. Puis nous évaluerons ces effets à partir des données expérimentales.

5.3.1 Données simulées

Présentation de l'étude

Nous cherchons donc à évaluer l'impact des erreurs de modèle sur les performances des différentes approches. Pour cela, on génère un ensemble à partir du modèle de signal incluant des réflexions à l'ordre 3. Puis dans les algorithmes, nous considérerons un modèle à travers la matrice $\mathbf{A}_{\mathcal{M},\mathcal{G}}$ (équation (1.4)) utilisant différents ordres de réflexions : 0, 1, 2 et 3. De cette manière, on contrôle les erreurs de modèle et on peut évaluer leurs impacts sur la localisation et la sélection des positions de capteurs. Lorsque nous considérons l'ordre 3, forcément il n'y a pas d'erreur de modèle. La plus forte présence d'erreurs de modèle correspond à l'ordre 0 (propagation en champ libre). Les paramètres utilisés, à l'exception du nombre de sources et de l'ordre de réflexion, sont identiques à ceux décrits dans le tableau 5.1. Dans cette étude, le nombre de sources est fixé à $N_S = 4$.

Résultats

En comparant les performances de la sélection optimisée en fonction de l'ordre de réflexion dans la figure 5.4, on observe le fort impact que les erreurs de modèles ont sur la localisation et la sélection. En effet, plus l'ordre de réflexion pris en compte dans $\mathbf{A}_{\mathcal{M},\mathcal{G}}$ diminue et plus les performances se dégradent.

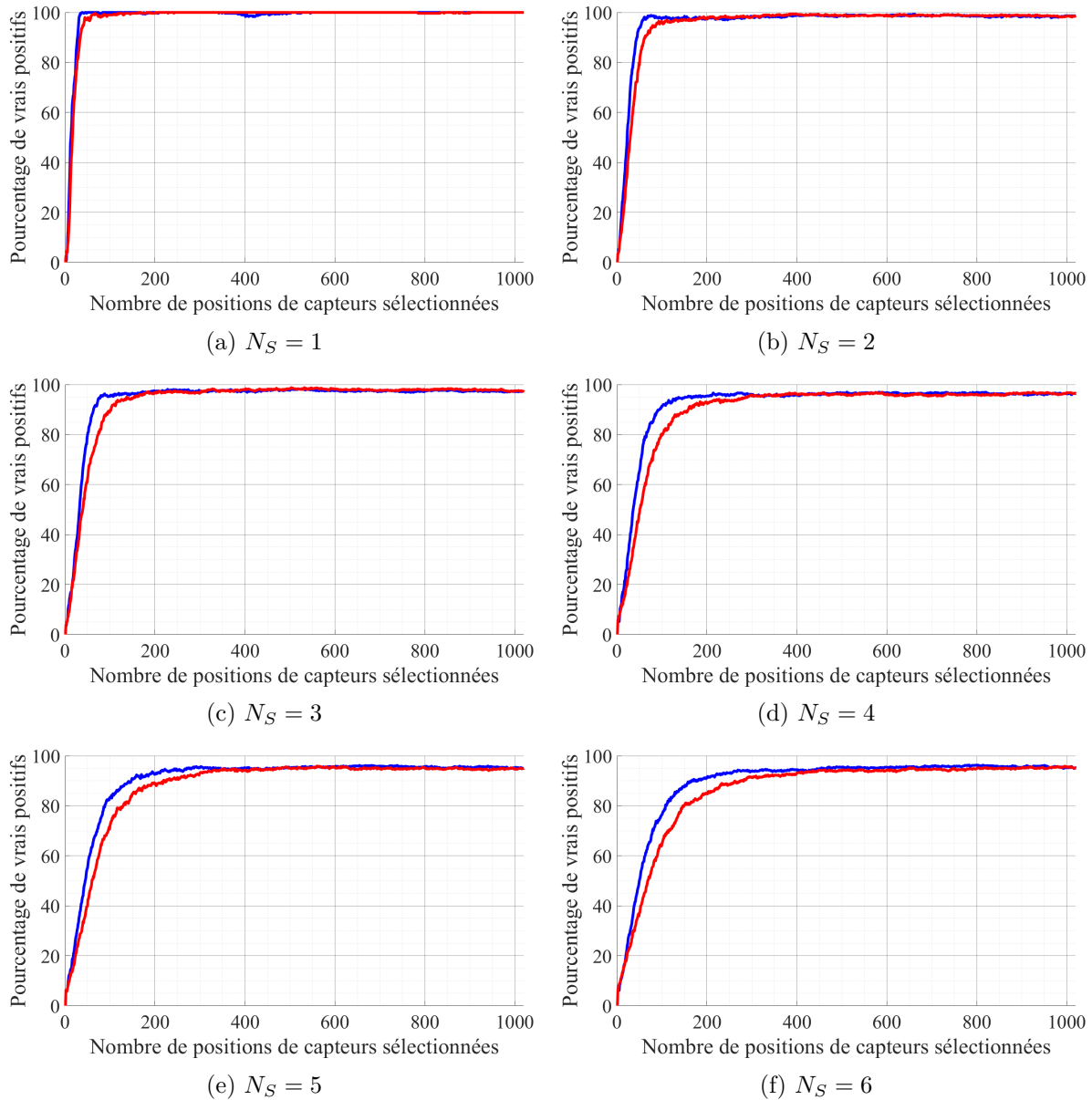


FIGURE 5.3 – Evolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées par notre algorithme (bleu) et par une sélection aléatoire (rouge) pour différents nombres de sources présentes dans le problème.

Si on compare, pour chaque ordre de réflexion, la sélection optimisée et la sélection aléatoire, on obtient la figure 5.5. On note que quelque soit l'ordre de réflexion, la sélection optimisée a de meilleures performances que la sélection aléatoire, même si ce gain est très faible quand l'ordre de réflexion est égal à 3. Cependant quand la présence d'erreurs de modèle devient plus importante, la différence entre les performances de la sélection optimisée et aléatoire est plus

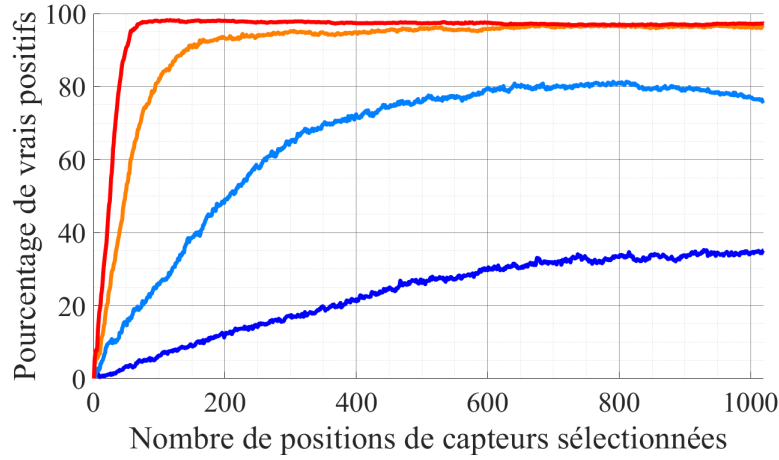


FIGURE 5.4 – Évolution du pourcentage de vrais positifs sur données simulées en fonction du nombre de positions de capteurs sélectionnées par notre algorithme pour différents ordres de réflexion considérés : 0 (bleu foncé), 1 (bleu clair), 2 (orange), et 3 (rouge).

visible. Cette différence est la plus visible dans le pire des cas, c'est-à-dire, quand l'ordre de réflexion est égal à 0.

5.3.2 Données expérimentales

Présentation de l'étude

Nous proposons ici la même étude que la précédente mais avec des données expérimentales. La principale différence est que nous ne connaissons pas l'ordre de réflexion qu'il faudrait considérer, ni les coefficients de réflexions des murs, sol et plafond. Dans le chapitre 4, on a pu faire une première estimation de ces paramètres. Nous proposons d'analyser ici les performances d'une sélection suivant notre algorithme par rapport à une sélection aléatoire en fonction de l'ordre de réflexion pris en compte dans $\mathbf{A}_{\mathcal{M},\mathcal{G}}$.

Les paramètres utilisés, à l'exception du nombre de sources et de l'ordre de réflexion, sont identiques à ceux décrits dans le tableau 5.1. Dans cette étude, le nombre de sources est fixé à $N_S = 4$.

Résultats

Comme pour l'étude avec les données simulées, on observe dans un premier temps l'impact de la prise en compte de la réverbération sur les performances de notre sélection optimisée (voir figure 5.6). Dans le cas des données expérimentales, comme dans le cas simulé, la non prise en compte des réverbérations dans $\mathbf{A}_{\mathcal{M},\mathcal{G}}$ dégradent les performances de localisation. Cependant, contrairement aux résultats obtenus avec les données simulées, la dégradation des performances

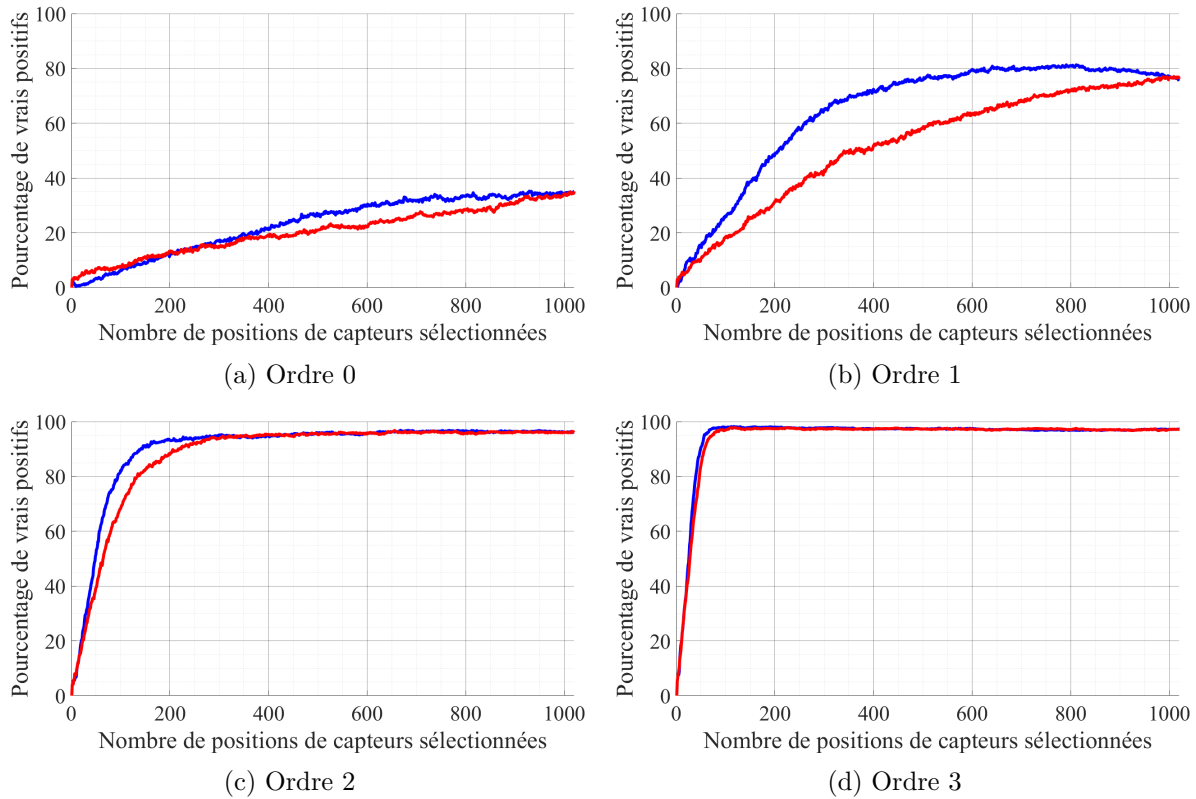


FIGURE 5.5 – Evolution du pourcentage de vrais positifs sur données simulées en fonction du nombre de positions de capteurs sélectionnées par notre algorithme (bleu) et par une sélection aléatoire (rouge) pour différents ordres de réflexions considérés.

est plus faible avec les données réelles. On observe qu’entre l’ordre 2 et 3 de réflexion, les performances ne sont pas dégradées. On note aussi qu’en ne prenant pas en compte les réverbérations, le pourcentage de vrais positifs dépasse 80 %. Cette différence entre les performances avec les données expérimentales et avec les données simulées est surprenante. Trois hypothèses peuvent être données pour expliquer cela :

- soit les autres erreurs de modèles présentes dans les données expérimentales contrecarrent l’impact de la non prise en compte des réverbérations,
- soit dans les données expérimentales, les réflexions avec un ordre de réflexion supérieur à 2 ne sont pas contributives,
- soit la puissance d’émission a un impact sur la réflexion et son impact.

Personnellement, je trouve que la troisième hypothèse est plus vraisemblable. Notamment car l’amplitude des sources ne sont pas connues dans les données expérimentales.

On compare à la figure 5.7, les performances de la sélection optimisée et de la sélection aléatoire en fonction de l’ordre des réflexions considéré dans $\mathbf{A}_{\mathcal{M},\mathcal{G}}$. Si aucune réflexion n’est

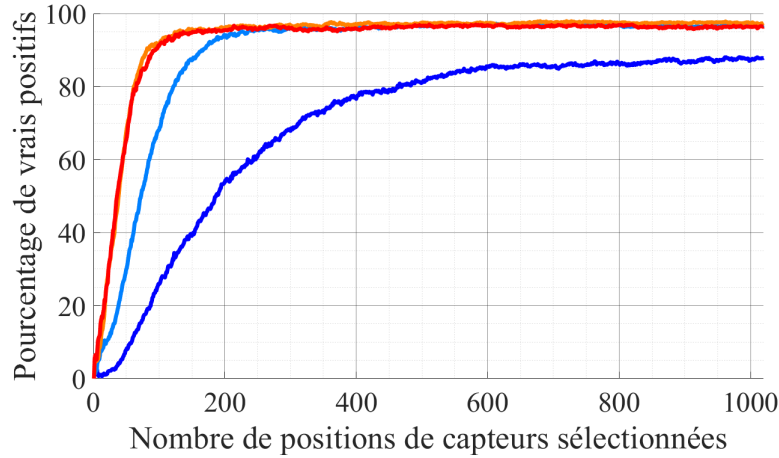


FIGURE 5.6 – Evolution du pourcentage de vrais positifs sur données expérimentales en fonction du nombre de positions de capteurs sélectionnées par notre algorithme pour différents ordres de réflexion considérés : 0 (bleu foncé), 1 (bleu clair), 2 (orange), et 3 (rouge).

prise en compte, utiliser notre algorithme ne permet pas de réduire le nombre de positions de capteurs sélectionnées par rapport à une sélection aléatoire. Par contre, en utilisant un modèle de propagation avec un ordre de réflexion supérieur ou égal à 1, on obtient un gain d'environ 200 positions de capteurs sélectionnées.

5.4 Comparaison avec d'autres méthodes

Présentation et paramètres de l'étude

Comme au chapitre 3, nous proposons une étude comparative de différentes méthodes de sélection de positions de capteurs. Les performances de notre algorithme sont comparées à trois autres approches de la littérature : la sélection séquentielle aléatoire [128, 129], la sélection non séquentielle suivant un algorithme génétique [95] (équation (2.14)) et la sélection non séquentielle maximisant la plus petite valeur propre de la matrice de Fisher par approche relaxée [121] (équation (2.16)).

Les paramètres utilisés, à l'exception du nombre de sources, sont identiques à ceux décrits dans le tableau 5.1. Dans cette étude, le nombre de sources est fixé à $N_S = 4$.

Présentation des méthodes de sélection de positions de capteurs

L'antenne aléatoire généralement utilisée dans l'acquisition comprimée permet de faire une sélection séquentielle, ce qui correspond à notre approche.

Les deux méthodes non séquentielles sont proches de notre contexte et sont donc d'intérêt

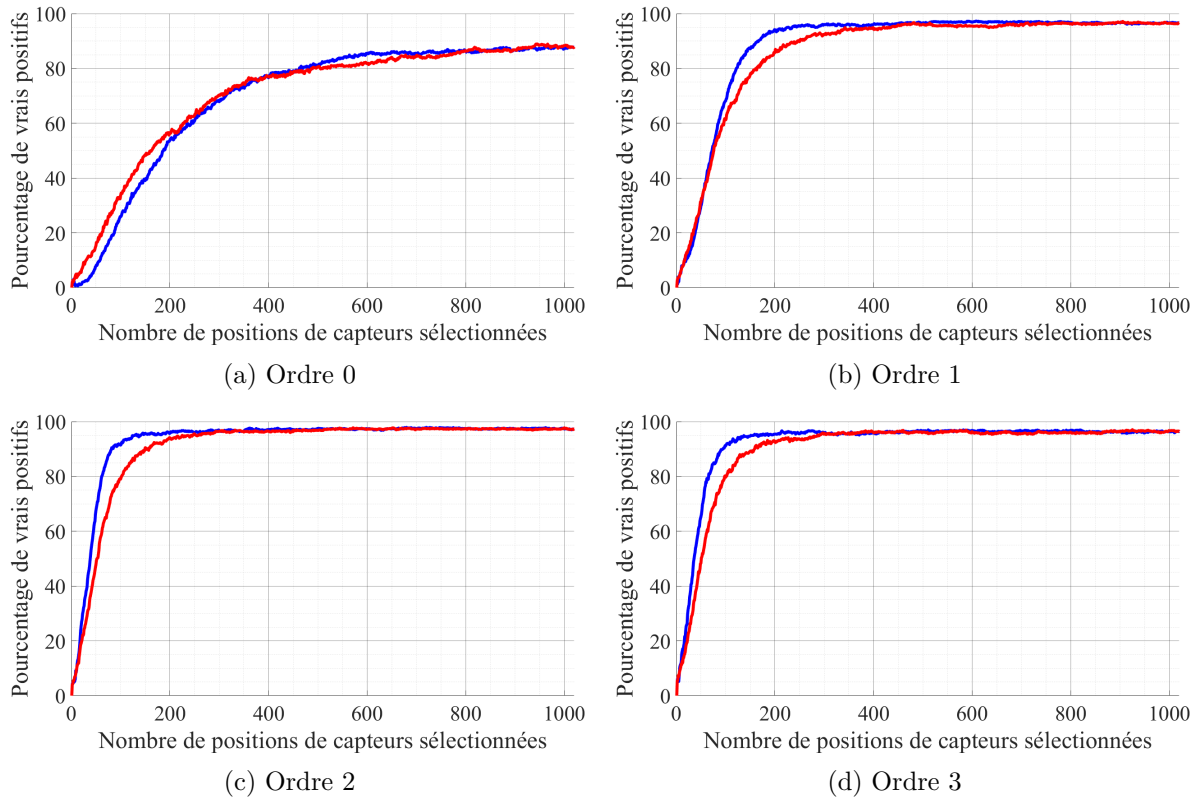


FIGURE 5.7 – Evolution du pourcentage de vrais positifs sur données expérimentales en fonction du nombre de positions de capteurs sélectionnées par notre algorithme (bleu) et par une sélection aléatoire (rouge) pour différents ordres de réflexions considérés.

ici pour une comparaison. La sélection génétique a été utilisée pour de la sélection de microphones dans un contexte de localisation de source aérienne [95]. Dans l'article, l'algorithme de localisation de sources est la formation de voies. Or nous utilisons comme algorithme l'inférence bayésienne parcimonieuse. On peut supposer alors que cette différence majeure diminue les performances de la sélection génétique. L'algorithme génétique nécessite des paramètres d'entrées qui sont définis arbitrairement de la manière suivante : le nombre d'antennes aléatoires est de $N_{ant} = 100$, le nombre de générations est de $N_{gen} = 250$ et le coefficient de mutation est de $\chi = 0,0001$. De plus, cette méthode est une sélection non séquentielle. Il est donc nécessaire de concevoir l'antenne pour un nombre de capteurs fixé. Ce processus s'avère être long. Nous avons donc décidé de calculer les performances pour des nombres de capteurs allant de 50 à 1000 par pas de 50.

L'autre méthode non séquentielle est la sélection maximisant la petite valeur propre de la matrice de Fisher. Elle a été utilisée dans la littérature pour de la localisation de sources par temps de vol avec un grand nombre de capteurs [121]. Similairement à la sélection génétique,

nous évaluons les performances de cette sélection non séquentielle pour des nombres de capteurs allant de 50 à 1000 par pas de 50. Pour implémenter cette méthode, nous avons utilisé la toolbox CVX. Mais la dimension de la matrice de Fisher dépend (voir équation (2.7)) :

- du nombre de sources $N_S = 4$,
- du nombre d'inconnues $N_G = 348$ (les positions discrétisées dans la grille),
- du nombre de paramètres qui définissent une source, au nombre de 5 :
 1. la position selon l'axe X ,
 2. la position selon l'axe Y ,
 3. la position selon l'axe Z ,
 4. le module de l'amplitude complexe émise par la source,
 5. la phase de l'amplitude complexe émise par la source.

De fait, il ne va pas être possible de résoudre le problème avec CVX si la dimension résultante du problème est trop grande. Pour pouvoir utiliser cette approche dans notre comparaison, il est nécessaire de réduire la dimension de la matrice de Fisher et nous avons donc modifier certains éléments. Pour le calcul de la matrice de Fisher, on considère :

- une seule source $N_S = 1$,
- $N_G = 25$ qui correspondent aux 25 positions de sources placées dans l'expérience (illustrées dans la figure 4.11),
- les 5 paramètres.

Avec ces modifications, il est possible de réduire la dimension de la matrice de Fisher permettant ainsi l'utilisation de CVX.

Résultats

On compare dans la figure 5.8 l'évolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées. Les performances de l'antenne génétique et relaxée suivent les performances de la sélection aléatoire. Elles sont inférieures à celles de notre algorithme, notamment pour un faible nombre de capteurs (< 300).

Afin de compléter l'analyse des résultats nous proposons, un peu comme au chapitre 3, de voir si notre algorithme conduit à sélectionner préférentiellement telles ou telles positions de capteurs. Pour cela, nous prenons en compte l'ordre dans lesquelles sont sélectionnées les positions de capteurs. Pour chaque réalisation, on attribue un score allant de 1 à 1018 à chaque position de capteur, dans l'ordre de sélection. Dit autrement, si la position de capteur a un score de 1, celle-ci a donc été sélectionnée en premier. Notons que les deux positions initiales de capteurs sélectionnées aléatoirement n'ont donc pas de score pour la réalisation donnée. Mais

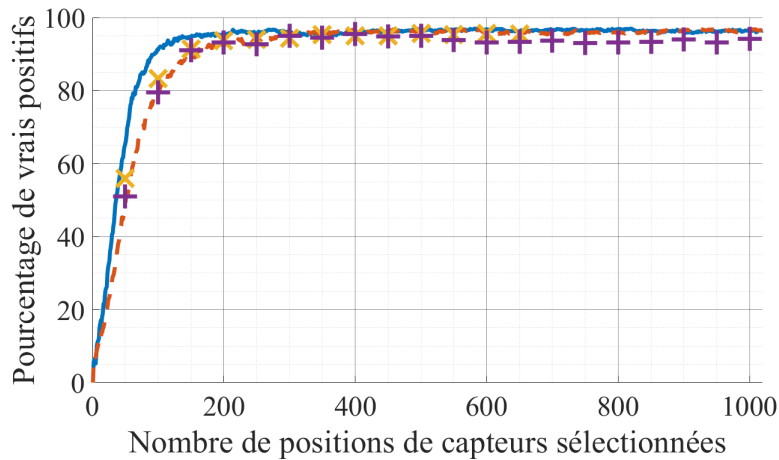


FIGURE 5.8 – Evolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées pour différentes méthodes de sélection de positions de capteurs : sélection optimisée (trait bleu), sélection aléatoire (tirets oranges), sélection génétique (plus violets) et sélection relaxée (croix jaunes).

comme l'initialisation est modifiée à chaque réalisation, cela n'impactera que très peu la moyenne du score des positions de capteurs. Ce score, calculé pour chaque position de capteur, est ensuite moyenné sur toutes les réalisations. On montre dans la figure 5.9 les 50 positions de capteurs ayant le score le plus faible et donc les 50 positions de capteurs sélectionnées généralement en premier.

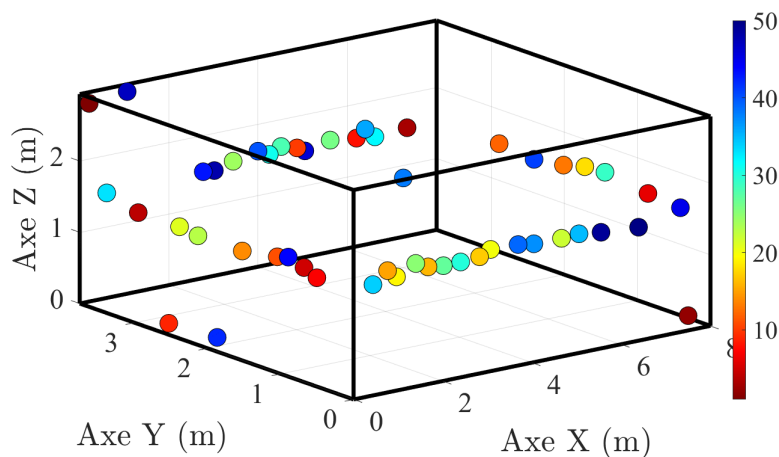


FIGURE 5.9 – Les 50 positions de capteurs sélectionnées généralement en premier par notre algorithme. Le score est donné par la couleur du capteur.

On remarque deux comportements dans la sélection optimisée : la sélection de deux positions de capteurs aux coins formant la grande diagonale de la salle, et une sélection de positions de

capteurs essentiellement au niveau du plan XY où sont les sources.

Pour le premier comportement, les deux capteurs sélectionnés sont les capteurs les plus éloignés. Ils forment donc l'antenne avec la plus grande ouverture d'antenne. Ce comportement est surtout intéressant en connaissant le score de chacune de ces positions de capteurs. Elles ont respectivement 1 et 2 comme score moyen. C'est-à-dire que ces deux positions de capteurs sont sélectionnées respectivement en premier et en deuxième pour toutes les réalisations. Quelque soit le problème de localisation de sources donné, l'algorithme cherche à créer l'antenne avec la plus grande ouverture d'antenne possible. C'est un phénomène connu de ce manuscrit qu'on retrouve quand la fréquence du signal est basse (voir partie 3.2.3).

L'autre comportement montre préférentiellement des positions de capteurs à la même hauteur que les sources. Rappelons que le problème de localisation est en 2 dimensions et que toutes les sources sont à la même hauteur. L'hypothèse pour expliquer ce comportement est que l'algorithme, après avoir maximisé l'ouverture d'antenne, cherche à sélectionner les positions de capteurs les plus proches des sources, qui sont donc les capteurs à la même hauteur que les sources. Cette hypothèse repose sur le fait que le modèle de propagation de l'onde en champ proche possède une atténuation géométrique. L'amplitude de l'onde sonore diminue avec la distance parcourue, donc les positions de capteurs majoritairement retenues mesurent l'amplitude de source la plus élevée.

5.5 Etudes complémentaires

Nous présentons dans cette partie des résultats en lien avec ceux vus en chapitre 3. Dans un premier temps, une étude avec une mesure de plusieurs snapshots est réalisée. Celle-ci est possible grâce au fait que nous avons fait des acquisitions sur une durée de 10 secondes et que nous effectuons les traitements sur des durées de 0,25 secondes. Dans un second temps, nous proposons une comparaison entre les différents critères d'optimalité de la conception optimale. Le but de cette étude est de voir si un critère d'optimalité est plus robuste aux erreurs de modèle.

5.5.1 Extension multi-snapshots

Présentation et paramètres de l'étude

Dans cette étude, on a pour but de démontrer que l'utilisation de plusieurs snapshots permet dans un premier temps une amélioration de l'estimation des positions des sources, mais aussi une augmentation des performances de la sélection de positions de capteurs. Dans le chapitre 3, nous avons démontré l'intérêt d'exploiter l'extension multi-snapshots avec des données simulées en acoustique sous-marine. Ici, nous tentons de le démontrer avec des données réelles en milieu aérien réverbérant.

Rappelons que chaque mesure considère la même scène acoustique (hypothèse de stationnarité temporelle). On remplace l'inférence bayésienne parcimonieuse dans l'algorithme par l'inférence bayésienne parcimonieuse à plusieurs snapshots (voir Algorithme 3). Rappelons que l'ajout de plusieurs snapshots ne modifie pas les dimensions de la matrice de covariance, le calcul du critère de sélection de positions de capteurs est donc identique.

Pour cela chaque mesure va utiliser une partie du signal temporel différente. En utilisant 0,25 secondes de signal, il est possible d'utiliser 40 mesures (puisque le signal utile est de 10 secondes). On compare la sélection suivant notre algorithme et la sélection aléatoire en fonction du nombre de mesures considérées. Le nombre de mesures N_T vont de 1 à 40.

Résultats

Commençons par comparer l'évolution du pourcentage de vrais positifs de notre sélection optimisée et la sélection aléatoire en fonction du nombre de capteurs sélectionnés et pour différents nombre de snapshots considérés. Dans la figure 5.10, on note que quelque soit le nombre de snapshots considérés, les performances de la sélection optimisée sont meilleures que la sélection aléatoire. Cependant il apparait un décrochage autour de 100 à 150 capteurs sélectionnés l'augmentation de nombre de snapshots. On peut aussi noter que les performances semblent légèrement diminuer avec l'augmentation du nombre de snapshots, ce qui n'est pas le résultat qui était attendu ! Malheureusement, nous n'avons pas pu trouver d'explications dans la littérature ou que nous aurons pu vérifiées par des simulations. Néanmoins, on peut supposer que l'hypothèse de stationnarité entre les snapshots n'est pas respectée.

Remarquons aussi que l'ajout du nombre de snapshots a souvent pour but de pallier au problème d'un rapport signal à bruit trop faible. Or, il s'avère que nos expérimentations sont effectuées en milieu contrôlé et que donc le bruit est très faible. En comparant l'amplitude des mesures expérimentales et l'amplitude des mesures de bruit (du scénario 4, voir Annexe 5.7), il nous est possible d'estimer un rapport signal à bruit d'environ 68 dB. Le bruit est donc presque inexistant et l'intérêt d'utiliser le multi-snapshots devient moindre.

5.5.2 Choix du critère d'optimalité

Présentation de l'étude

Similairement à la section 3.5.2, nous cherchons à comparer les performances supposées identiques des différents critères d'optimalité de la conception optimale. Comme auparavant, le critère A-Optimal et le critère D-Optimal sont deux approches pouvant être accélérées dans la sélection séquentielle (parties 2.5.2 et 2.5.1), ce qui n'est pas le cas pour le critère E-Optimal.

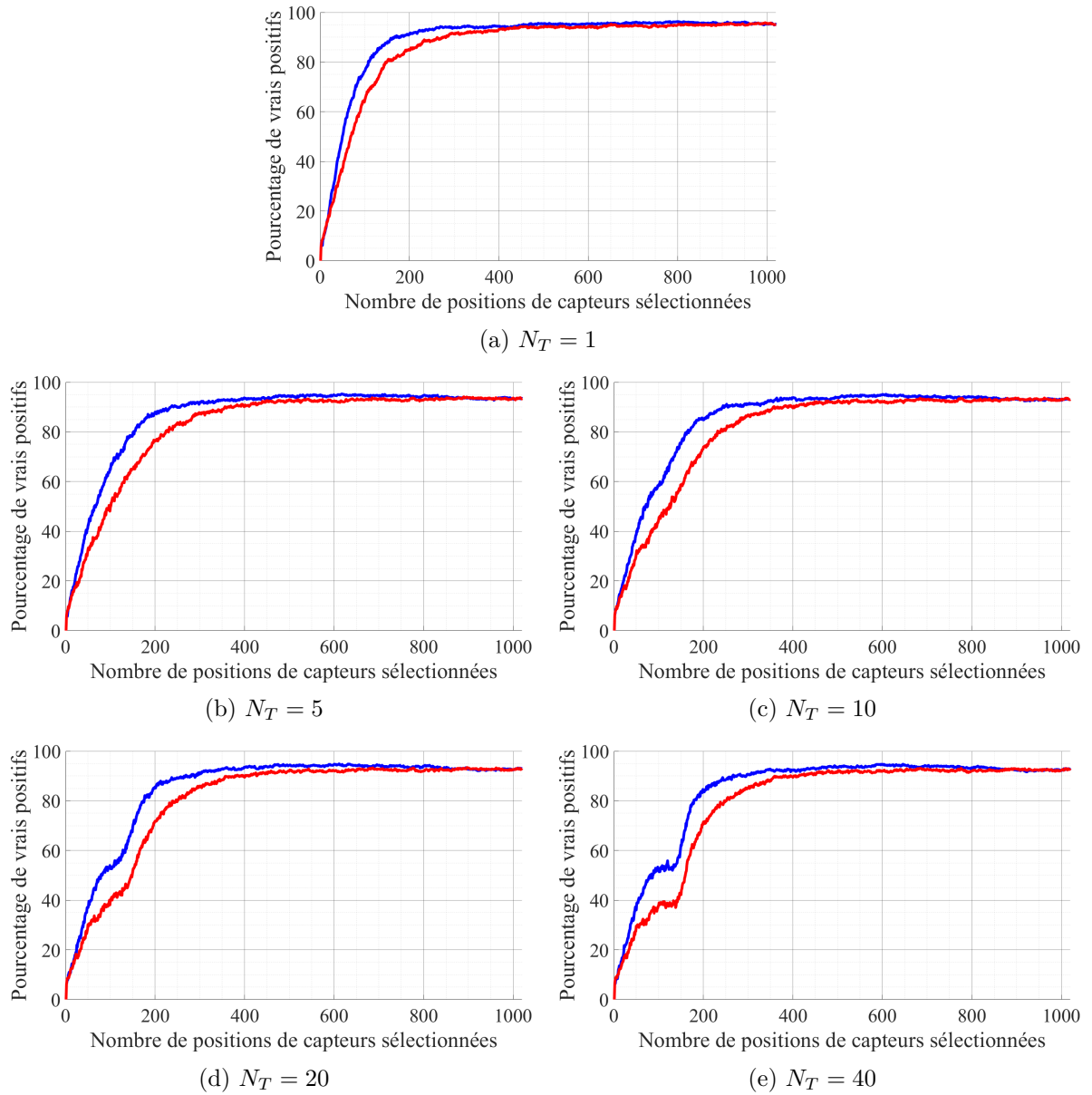


FIGURE 5.10 – Evolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées par notre algorithme (bleu) et par une sélection aléatoire (rouge) pour différents le nombre de snapshots utilisés.

Résultats

On présente l'évolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées pour chaque critère d'optimalité dans la figure 5.11. Similairement à l'étude dans la partie 3.5.2, on peut remarquer que :

1. l'ajout de capteurs permet une augmentation des performances en localisation de sources

quelque soit le critère optimisé,

2. les performances obtenues par les critères D-optimal et A-optimal sont similaires,
3. les performances de l'antenne conçue par le critère E-optimal sont plus faibles que celles des deux autres critères.

Le plan d'expériences (A-optimal, D-optimal et E-optimal) permet de réduire significativement le nombre de capteurs sans dégrader les performances de localisation. En effet, pour atteindre le seuil maximal de performances (ici un pourcentage de vrais positifs de 95%), il faut en moyenne 150 capteurs aux 3 méthodes, contrairement à 300 pour les 3 antennes de l'état de l'art vues précédemment (voir figure 5.8).

Les performances des critères D-optimal et A-optimal sont similaires mais le temps de calcul est très différent selon le critère d'optimalité choisi. La sélection par optimisation du critère D-optimal est la plus rapide, suivi du critère A-optimal, pour finir avec celui du critère E-optimal. Il est donc plus intéressant d'utiliser le critère D-optimal dans ce contexte. On rappelle que le critère E-optimal fait intervenir des calculs potentiellement instables numériquement, que ce soit dans l'inversion de la matrice de covariance, ou dans le calcul des valeurs propres.

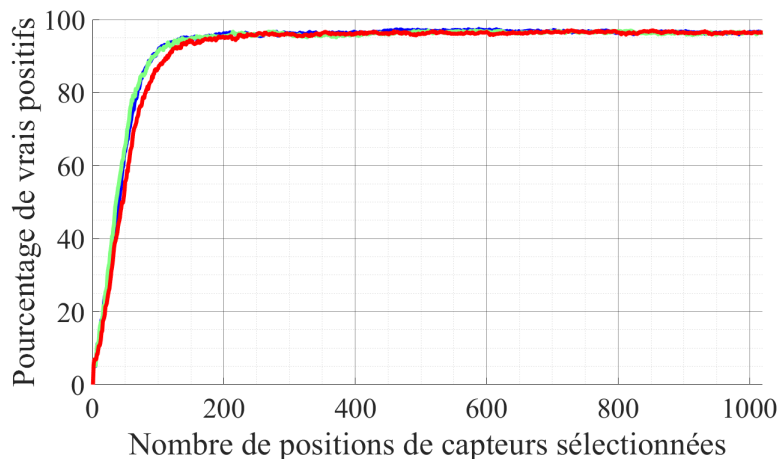


FIGURE 5.11 – Evolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées selon le critère d'optimalité choisi : A-Optimal (bleu), D-Optimal (vert) et E-Optimal (rouge).

5.6 Applications en conception d'antenne

Présentation et paramètres de l'étude

La dernière étude de ce manuscrit a pour but d'exploiter notre algorithme via une autre stratégie. En effet, un reproche commun aux méthodes de sélection de positions de capteurs

orientées par les données est un manque d'applications possibles dû à la nécessité de la prise de mesures pour la sélection. Nous proposons donc ici d'aborder ce problème en exploitant des données simulées pour concevoir l'antenne et pour ensuite l'utiliser sur des données réelles.

Dans un premier temps nous présentons les différentes méthodes de sélection de capteurs considérées. Puis nous les évaluons sur des données réelles caractérisées par les paramètres suivants :

- 4 sources émettent un peigne fréquentiel de 10 secondes,
- le pré-traitement utilise 0,25 secondes de signal temporel et la fréquence considérée est de 1000 Hz,
- une discrétisation de la salle en $N_G = 348$ points,
- le milieu de propagation est considéré réverbérant à l'ordre 3 avec des coefficients de réflexions $[Cr_{\text{murs}}; Cr_{\text{plafond}}; Cr_{\text{sol}}]$ égaux à $[0,74; 0,5; 0,73]$,
- 2 capteurs sont sélectionnés aléatoirement pour l'initialisation des méthodes,
- les performances sont évaluées en moyennant 300 réalisations, chacune sélectionnant aléatoirement 4 sources parmi les 25 disponibles.

Présentation des méthodes

Dans cette étude on compare quatre méthodes de conceptions d'antennes :

- L'antenne *aléatoire*. L'antenne est différente à chaque réalisation et est identique à l'antenne aléatoire utilisée dans la section 5.4.
- L'antenne *optimisée sur données réelles*. L'antenne est différente à chaque réalisation et est identique à la sélection optimisée utilisée dans la section 5.4.
- L'antenne *optimisée à partir d'un jeu de données simulées*. Les positions des sources dans les données simulées sont identiques à celles des données réelles qui seront utilisées dans l'évaluation des performances. Donc les données simulées considèrent les positions des sources a priori connues. L'antenne est différente à chaque réalisation. Remarquons que la principale différence entre l'antenne *optimisée à partir d'un jeu de données simulées* et l'antenne *optimisée sur données réelles* est donc la présence des erreurs de modèles (erreurs sur les positions de capteurs, sur les positions de sources, sur la géométrie de la salle...) dans les mesures pour l'antenne *optimisée sur données réelles*.
- L'antenne *optimisée à partir de l'ensemble des jeux de données simulées*. De cette manière, on s'affranchit de la connaissance des positions de sources. Notons que l'antenne est unique. Elle est identique pour toutes les réalisations.

Pour l'antenne *optimisée à partir de l'ensemble des jeux de données simulées*, le score est calculé puis moyenné à partir d'un grand nombre d'antennes déjà construites. Les antennes utilisées pour le score correspondent aux antennes *optimisées à partir d'un jeu de données simulées*

(ce sont les seules antennes qui utilisent les données simulées lors de leurs conceptions). Pour chaque position de capteur de chaque antenne *optimisée à partir d'un jeu de données simulées*, on attribue une valeur allant de 1, si celle-ci a été sélectionnée en premier, à 1018, si celle-ci a été sélectionnée en dernier. Après que le score soit calculé pour les 300 antennes *optimisées à partir d'un jeu de données simulées*, on moyenne le score. Le score permet d'ordonner les positions de capteurs et permet de concevoir l'antenne *optimisée à partir de l'ensemble des jeux de données simulées*. Notons que les positions de capteurs sélectionnées aléatoirement pour l'initialisation n'ont pas de scores, mais cela n'a peu d'influence sur le score des positions de capteurs.

Résultats

On observe à la figure 5.12 l'évolution des performances obtenues sur données expérimentales pour les quatre conceptions d'antennes présentées précédemment. On observe dans un premier temps que les performances de l'antenne *optimisée sur données réelles* et l'antenne *optimisée à partir d'un jeu de données simulées* sont très proches. La présence d'erreurs de modèles dans les mesures pour la conception de l'antenne ne permet pas d'augmenter les performances de l'antenne.

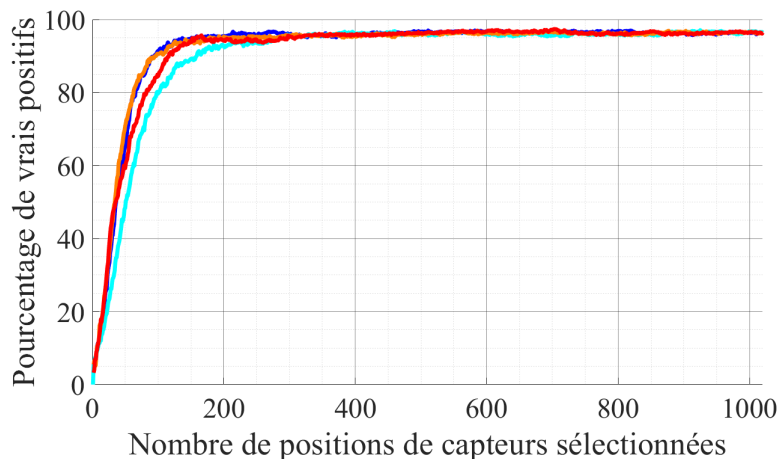


FIGURE 5.12 – Evolution du pourcentage de vrais positifs en fonction du nombre de positions de capteurs sélectionnées selon différentes conceptions d'antennes : l'antenne *optimisée sur données réelles* (bleu foncé), l'antenne *aléatoire* (bleu clair), l'antenne *optimisée à partir d'un jeu de données simulées* (orange) et l'antenne *optimisée à partir de l'ensemble des jeux de données simulées* (rouge).

Les performances de l'antenne *optimisée à partir de l'ensemble des jeux de données simulées* sont en dessous des antennes optimisées mais au dessus de l'antenne *aléatoire*. Si ses performances par rapport aux antennes optimisées semblent logiques, vu que l'antenne *optimisée à partir de l'ensemble des jeux de données simulées* n'est pas optimale, ses performances par rapport à celles

de l'antenne *aléatoire* sont prometteuses. En effet, l'antenne *optimisée à partir de l'ensemble des jeux de données simulées* est unique et convient à un grand nombre de scènes acoustiques différentes. De plus, comme pour l'antenne *optimisée à partir d'un jeu de données simulées*, l'antenne *optimisée à partir de l'ensemble des jeux de données simulées* permet de connaître en amont de l'expérience, le nombre de capteurs minimal pour atteindre les performances maximales de localisation avec peu de connaissances sur la scène acoustique. Or connaître le nombre nécessaire de positions de capteurs est utile afin de prévoir le budget financier (le nombre de capteurs) pour mener à bien son expérience. Cependant, il est important de noter que actuellement, l'antenne *optimisée à partir de l'ensemble des jeux de données simulées* nécessite un grand nombre de données simulées et donc un temps de calcul en amont de l'expérience important.

5.7 Conclusion

Ce chapitre a montré l'utilisation de notre algorithme, la sélection séquentielle bayésienne orientée par les données sur données expérimentales, dans un contexte acoustique difficile : champ proche et salle réverbérante. Les évaluations ont été réalisées à partir des données expérimentales décrites au chapitre 4. Les résultats ont permis de confirmer l'importance d'une sélection optimisée par rapport à une sélection aléatoire dans le contexte fourni par l'expérience.

Nous avons fini ce chapitre en proposant d'utiliser notre algorithme différemment : l'idée est de concevoir, en amont d'une expérience, l'antenne à partir de données simulées pour l'exploiter sur des données réelles. Nous avons montré la faisabilité de cette approche en évaluant les performances de l'antenne ainsi conçue à partir de nos expérimentations dans la salle MODO. On peut particulièrement noter que l'antenne *optimisée à partir de l'ensemble des jeux de données simulées* (qui n'est pas optimale) convient à un grand nombre de scènes acoustiques. Cette conception permet notamment de connaître en amont le nombre de capteurs minimal avec peu de connaissances sur la scène acoustique (la fréquence et le nombre de sources sont connus).

On pourrait poursuivre cette étude en évaluant de façon plus complète les performances quand on modifie le nombre de sources, leurs fréquences ou directivités, en les rendant cohérentes entre elles, ou même en utilisant cette antenne générale dans d'autres applications.

CONCLUSION ET PERSPECTIVES

Ce chapitre dresse le bilan des travaux menés et en propose quelques perspectives.

Contexte et objectifs

Dans un premier temps, nous rappelons le contexte et les objectifs de la thèse en quelques phrases. Les travaux proposés dans ce manuscrit concernent la localisation passive de sources acoustiques sous-marines ou aériennes. Dans ce domaine, l'utilisation d'antennes de capteurs est nécessaire pour l'estimation des positions des sources. La géométrie de l'antenne a un impact sur les performances de localisation. La problématique de cette thèse est de trouver un critère qui nous permet de concevoir l'antenne avec les plus hautes performances de localisation de sources. Pour résoudre cette problématique, les travaux se sont divisés en trois parties (correspondant aux trois objectifs de la thèse) :

1. proposer une méthode de sélection de capteurs pour de la localisation de sources. L'objectif est réalisé dans la partie I.
2. l'application et l'évaluation des performances de l'algorithme sur données simulées. L'objectif est atteint dans la partie II.
3. l'application et l'évaluation des performances de l'algorithme sur données réelles. L'objectif est atteint dans la partie III.

Conclusion

Ces travaux de thèse ont pu mettre en lumière l'importance de l'optimisation des positions de capteurs pour l'estimation des positions de sources acoustiques, sous-marines ou aériennes. Pour un problème de localisation de sources donné, ce manuscrit a proposé une approche de conception d'antenne permettant de réduire le nombre de capteurs nécessaire à une performance satisfaisante.

Le chapitre 1 a dressé un spectre d'algorithmes d'estimations de positions de sources, allant de la classique formation de voies aux méthodes parcimonieuses en passant par les approches bayésiennes. Dans ce manuscrit, nous avons choisi l'Inférence Bayésienne Parcimonieuse. Cette approche considère un nombre de sources faible par rapport aux positions observés. Cette condition est souvent respectée en localisation de sources acoustiques. De plus dans ce chapitre, nous

avons pu voir l'impact des paramètres physiques du problème, comme la fréquence émise par la source, sur le dimensionnement d'antenne.

Sur ce dernier point, le chapitre 2 a permis d'explicitier quelques méthodes de conception d'antenne afin de résoudre les problèmes de dimensionnement d'antenne. Mais ce chapitre est allé plus loin en donnant des algorithmes avec des performances maximales en localisation de sources tout en ayant le plus petit nombre de capteurs utilisés. La sélection de positions de capteurs, c'est-à-dire, la conception d'antenne, a été résolue par des algorithmes génétiques ou des méthodes provenant du plan d'expériences ("*Experimental Design*"). Le plan d'expériences propose une antenne minimisant suivant une métrique de la matrice de covariance. De plus, il peut proposer des sélections séquentielles et orientées par les données. C'est en mettant en commun ces caractéristiques que nous arrivons à l'algorithme proposé lors de ce manuscrit : la sélection séquentielle bayésienne orientée par les données.

Dans le chapitre 3, cet algorithme a été évalué dans un contexte de localisation de sources acoustiques sous-marines. L'utilisation de données synthétiques a permis, à travers différentes simulations, de caractériser et comparer les performances, de fixer les limites et de spécifier ses atouts. Une comparaison avec l'état de l'art nous a montré le gain de capteurs envisageable en utilisant notre algorithme par rapport à une autre approche. Ce chapitre s'est cependant concentré sur de localisation de sources discrètes à 1 dimension et de la sélection de positions de capteurs à 1 dimension de l'espace.

Le chapitre 4 a introduit la mise en place de l'expérience ainsi que les premiers essais sur les données réelles. L'expérience s'est déroulée dans une salle du nom de "les Murs Ont Des Oreilles" à Saint-Cyr L'École, au sein de l'Institut Jean Le Rond d'Alembert. La localisation de sources et donc la conception d'antenne se fait ainsi dans un milieu aérien, fermé et réverbérant avec à disposition, un grand nombre de capteurs. Ce chapitre a pu mettre en évidence l'impact non négligeable de la réverbération sur les estimateurs de positions de sources.

L'application des algorithmes de conception d'antenne a été présentée lors du chapitre 5. Les performances de notre algorithme surpasse les autres méthodes testées dans la minimisation du nombre de capteurs nécessaires. En évaluant les performances sur les données réelles d'une antenne conçue par des données simulées, nous avons pu justifier l'utilisation d'une conception orientée par les données dans un contexte de localisation de sources acoustiques.

Perspectives

Ces travaux nous ont permis de dégager plusieurs perspectives en conception d'antenne. Nous les avons structurées en quatre grands axes, qui pourront être considérés à plus ou moins long terme.

Conception d'antenne prenant en compte la directivité des sources

Dans la thèse de Hugo Demontis [140], la salle MODO est utilisée pour estimer la directivité de la source. Il est donc possible de prendre en compte la directivité des sources dans le modèle de propagation. En créant plusieurs jeux de données avec une directivité différente de la source, il doit être possible de mettre en évidence l'impact de la directivité sur la localisation et sur la conception d'antenne. Rappelons que dans les mesures utilisées dans le manuscrit, la source est supposée monopolaire et donc omnidirectionnelle. Mais, nous avons réalisé des mesures avec une source dipolaire (voir Annexe 5.7). Hypothétiquement, on pourrait s'attendre à ce que la directivité de la source induise une sélection des capteurs situés dans les lobes principaux de directivité.

Conception d'antenne prenant en compte la fluctuation du milieu de propagation

L'océan est un milieu qui fluctue dans l'espace et le temps. Mais ces fluctuations peuvent être prises en compte dans le modèle de propagation par un modèle probabiliste dans un contexte de localisation de sources [33]. Or les estimateurs utilisés sont bayésiens et peuvent en sortie d'algorithme restituer la matrice de covariance a posteriori, cette même matrice qui nous permet ensuite de sélectionner les positions de capteurs. L'idée serait de conserver la structure de notre algorithme mais de remplacer l'algorithme SBI par l'algorithme proposé dans la thèse de Guillaume Beaumont [33], *phase-aware Swept Approximate Message Passing* ou PASAMP. Il alternerait donc entre l'algorithme de localisation de sources bayésien PASAMP et l'algorithme D-Optimal de sélection de position de capteur. Cette étude permettrait de considérer un milieu de propagation sous-marin plus réaliste et il serait notamment possible d'observer l'impact de la fluctuation du milieu sur la conception d'antenne.

Conception d'antenne avec des signaux large bande

Dans les travaux présentés dans ce manuscrit, les signaux émis par les sources ne contiennent qu'une seule fréquence. Notons que dans les chapitres 4 et 5, le signal émis contenait 5 fréquences différentes mais le traitement était effectué que sur une seule de ces fréquences. L'étude de Shoichi Koyama [118] montre qu'il est possible de concevoir une antenne à partir d'une minimisation du déterminant de la matrice de covariance pour un signal large bande. La solution donnée est de calculer puis de moyenner le critère d'optimalité sur toutes les fréquences du signal émis. Cette opération est possible sur notre algorithme. Cependant, cette modification ne concerne que la partie sélection de positions de capteurs. Il est important que l'estimateur de positions de sources, et donc l'Inférence Bayésienne Parcimonieuse, prenne en compte toutes ces fréquences. Il faudrait donc trouver une manière adéquate de prendre en compte les covariances issues de SBI, calculée sur un ensemble de fréquences distinctes.

Conception d'antenne pour de la reconstruction modale

Si notre algorithme peut être appliqué sur des signaux large bande, alors il est possible d'appliquer l'algorithme de conception d'antenne non plus pour de la localisation de sources mais pour de la reconstruction modale. La reconstruction modale, telle qu'elle est employée dans la thèse de Thomas Paviet-Salomon [31], permet de caractériser le milieu de propagation, en déterminant la profondeur du milieu ou la célérité du son dans ce milieu. Pour cela, on cherche à estimer les modes de propagation dans un guide d'onde. Le faible nombre de modes de propagation permet l'utilisation de méthodes parcimonieuses. Un algorithme bayésien permettant la reconstruction modale et permettant l'estimation de la matrice de covariance a posteriori est le *Soft Bayesian Pursuit* ou SOBAP [146]. On pourrait envisager d'utiliser notre algorithme pour réaliser de la reconstruction modale en remplaçant SBI par SOBAP. La nouvelle version de l'algorithme alternerait donc entre SOBAP et l'algorithme D-Optimal de sélection de position de capteur.

Annexes

LIENS ENTRE LES CRITÈRES D'OPTIMALITÉ

Sommaire

Divergence de Kullback-Leibler	b
L'entropie de Shannon	c

Cette annexe se concentre à démontrer des liens mathématiques entre les critères d'optimisation par conception optimale développés dans le manuscrit et quelques critères les plus connus de l'état de l'art pour la sélection de capteurs : la Divergence de Kullback-Leibler [83] et l'entropie de Shannon [102]. Cette annexe fait donc suite à la section 2.4.

Divergence de Kullback-Leibler

La divergence de Kullback-Leibler est une mesure de dissimilarité entre deux distributions de probabilités. Soit q_1 et q_2 ces deux distributions de probabilités, la divergence de Kullback-Leibler est définie pour [116] :

$$KL(q_1 : q_2) = \int_{\Omega(x)} q_1(x) \log \left(\frac{q_1(x)}{q_2(x)} \right) dx, \quad (5.1)$$

avec x une variable aléatoire et avec $\Omega(x)$ le domaine d'intégration de la variable aléatoire. Il est possible de faire le lien entre la divergence de Kullback-Leibler et l'information de Fisher [147, 148]. Pour cela il faut redéfinir les deux distributions de probabilités q_1 et q_2 . q_1 est définie comme la distribution de vraisemblance $f(\mathbf{v}; \psi)$ d'une variable multidimensionnelle \mathbf{v} paramétrée par ψ , qui est doublement différentiable par rapport à ψ pour tout \mathbf{v} . q_2 est une distribution de vraisemblance $f(\mathbf{v}; \psi + \delta\psi)$ mais avec un paramètre ψ subissant une petite différence $\delta\psi$. On a ainsi :

$$\begin{aligned} KL(f(\mathbf{v}; \psi) : f(\mathbf{v}; \psi + \delta\psi)) &= \int_{\Omega(\mathbf{v})} f(\mathbf{v}; \psi) \log \left(\frac{f(\mathbf{v}; \psi)}{f(\mathbf{v}; \psi + \delta\psi)} \right) d\mathbf{v} \\ &= \int_{\Omega(\mathbf{v})} f(\mathbf{v}; \psi) \log(f(\mathbf{v}; \psi)) d\mathbf{v} - \int_{\Omega(\mathbf{v})} f(\mathbf{v}; \psi) \log(f(\mathbf{v}; \psi + \delta\psi)) d\mathbf{v}. \end{aligned} \quad (5.2)$$

En considérant que la variation $\delta\psi$ est faible, il est possible de remplacer la distribution de vraisemblance $f(\mathbf{v}; \psi + \delta\psi)$ par un développement de Taylor à l'ordre 2 :

$$\begin{aligned} KL(f(\mathbf{v}; \psi) : f(\mathbf{v}; \psi + \delta\psi)) &= \int_{\Omega(\mathbf{v})} f(\mathbf{v}; \psi) \log(f(\mathbf{v}; \psi)) d\mathbf{v} \\ &\quad - \int_{\Omega(\mathbf{v})} f(\mathbf{v}; \psi) \log(f(\mathbf{v}; \psi)) d\mathbf{v} \\ &\quad - \delta\psi \int_{\Omega(\mathbf{v})} f(\mathbf{v}; \psi) \frac{\partial}{\partial \psi} \log(f(\mathbf{v}; \psi)) d\mathbf{v} \\ &\quad - \frac{1}{2} \delta\psi^2 \int_{\Omega(\mathbf{v})} f(\mathbf{v}; \psi) \frac{\partial^2}{\partial \psi^2} \log(f(\mathbf{v}; \psi)) d\mathbf{v}. \end{aligned} \quad (5.3)$$

On remarque que les deux premiers termes s'annulent et donne la forme suivante, avec seulement le terme à l'ordre 1 et le terme à l'ordre 2 du développement de Taylor :

$$\begin{aligned} KL(f(\mathbf{v}; \psi) : f(\mathbf{v}; \psi + \delta\psi)) &= -\delta\psi \int_{\Omega(\mathbf{v})} f(\mathbf{v}; \psi) \frac{\partial}{\partial \psi} \log(f(\mathbf{v}; \psi)) d\mathbf{v} \\ &\quad - \frac{1}{2} \delta\psi^2 \int_{\Omega(\mathbf{v})} f(\mathbf{v}; \psi) \frac{\partial^2}{\partial \psi^2} \log(f(\mathbf{v}; \psi)) d\mathbf{v}. \end{aligned} \quad (5.4)$$

D'après l'équation (2.5), le terme à l'ordre 1 du développement de Taylor est nul, simplifiant une nouvelle fois la formulation de la divergence de Kullback-Leibler. Il est ainsi possible de transformer le terme à l'ordre 2 du développement de Taylor par une espérance mathématique et ainsi retrouver l'équation (2.6), prouvant le lien entre la divergence de Kullback-Leibler et l'information de Fisher :

$$\begin{aligned} KL(f(\mathbf{v}; \psi) : f(\mathbf{v}; \psi + \delta\psi)) &= -\frac{1}{2} \delta\psi^2 \int_{\Omega(\mathbf{v})} f(\mathbf{v}; \psi) \frac{\partial^2}{\partial \psi^2} \log(f(\mathbf{v}; \psi)) d\mathbf{v} \\ &= -\frac{1}{2} \delta\psi^2 \mathbb{E} \left[\frac{\partial^2}{\partial \psi^2} \log f(\mathbf{v}; \psi) | \psi \right] \\ &= \frac{1}{2} \delta\psi^2 Fi(\psi). \end{aligned} \quad (5.5)$$

La divergence de Kullback-Leibler peut être utilisée comme critère d'optimisation dans la sélection de positions de capteurs en considérant que le vecteur de la variable aléatoire est le vecteur d'amplitude des sources et que les paramètres soit les positions de capteurs [83, 149]. La variation du paramètre est l'ajout d'une position de capteur dans l'ensemble, comme dans les approches itératives pour la matrice de covariance dans l'équation (2.23). Le problème d'optimisation devient alors :

$$\mathbf{p}_{\mathcal{K}_{i+1}} = \operatorname{argmax}_{\mathbf{p}_{\mathcal{K}_{i+1}} \in \mathcal{M} \setminus \mathcal{K}} KL(f(\tilde{\mathbf{s}}_{\mathcal{K}, \mathcal{G}}; \mathcal{K}) : f(\tilde{\mathbf{s}}_{\mathcal{K} \cup \mathbf{p}_{\mathcal{K}_{i+1}}, \mathcal{G}}; \mathcal{K} \cup \mathbf{p}_{\mathcal{K}_{i+1}})) \quad (5.6)$$

L'entropie de Shannon

L'entropie de Shannon ou l'entropie différentielle définie en fonction d'une distribution de vraisemblance $q(x)$ [150] est donnée par :

$$h(\mathbf{x}) = - \int q(x) \log(q(x)) dx = -\mathbb{E}[\log(q(x))]. \quad (5.7)$$

En considérant les notations utilisées pour la matrice de Fisher ou la divergence de Kullback-Leibler on a :

$$h(\mathbf{v}) = -\mathbb{E}[\log(f(\mathbf{v}; \boldsymbol{\psi}))]. \quad (5.8)$$

Si on considère que la distribution de vraisemblance suit une loi gaussienne complexe multivariée, comme dans l'équation (2.10), l'entropie différentielle devient :

$$\begin{aligned} h(\mathbf{v}) &= -\mathbb{E}[\log(\pi^{-N_G} \det(\boldsymbol{\Sigma})^{-1} \exp^{-(\mathbf{v}-\mathbb{E}[\mathbf{v}])^H \boldsymbol{\Sigma}^{-1} (\mathbf{v}-\mathbb{E}[\mathbf{v}])})] \\ &= \log(\det(\boldsymbol{\Sigma})) + N_G \log(\pi) + \mathbb{E}[(\mathbf{v} - \mathbb{E}[\mathbf{v}])^H \boldsymbol{\Sigma}^{-1} (\mathbf{v} - \mathbb{E}[\mathbf{v}])]. \end{aligned} \quad (5.9)$$

En utilisant les propriétés de la trace, on modifie la formulation de l'entropie différentielle :

$$\begin{aligned} h(\mathbf{v}) &= \log(\det(\boldsymbol{\Sigma})) + N_G \log(\pi) + \mathbb{E}[\text{tr}[\boldsymbol{\Sigma}^{-1} (\mathbf{v} - \mathbb{E}[\mathbf{v}]) (\mathbf{v} - \mathbb{E}[\mathbf{v}])^H]] \\ &= \log(\det(\boldsymbol{\Sigma})) + N_G \log(\pi) + \text{tr}[\boldsymbol{\Sigma}^{-1} \mathbb{E}[(\mathbf{v} - \mathbb{E}[\mathbf{v}]) (\mathbf{v} - \mathbb{E}[\mathbf{v}])^H]]. \end{aligned} \quad (5.10)$$

Pour finir on remarque que le dernier terme, c'est-à-dire, l'espérance mathématique correspond à la matrice de covariance et permet de transformer une dernière fois la formulation de l'entropie différentielle :

$$\begin{aligned} h(\mathbf{v}) &= \log(\det(\boldsymbol{\Sigma})) + N_G \log(\pi) + \text{tr}[\boldsymbol{\Sigma}^{-1} \boldsymbol{\Sigma}] \\ &= \log(\det(\boldsymbol{\Sigma})) + \epsilon. \end{aligned} \quad (5.11)$$

avec ϵ une constante. On remarque que l'entropie différentielle est égal au logarithme du déterminant de la matrice de covariance et donc au logarithme du D-optimal. Il est donc possible d'écrire le problème d'optimisation en fonction de l'entropie différentielle [151, 152] :

$$\mathcal{K} = \underset{\mathcal{K}}{\text{argmin}} h(\tilde{\mathbf{s}}_{\mathcal{K}, \mathcal{G}}). \quad (5.12)$$

ACQUISITIONS RÉALISÉES DANS LA SALLE MODO

Sommaire

Source dipolaire	e
Sources monopolaires cohérentes	f
Calibration géométrique	f
Aspirateur	g
Cocktail party	g

Plusieurs acquisitions ont été menées en plus de celle présentée dans le chapitre 4. On présente ces acquisitions, qui pourront dans l'avenir amener à des travaux. Ces acquisitions diffèrent par la source utilisée. Les 5 acquisitions que nous présentons dans cette annexe concernent les sources suivantes :

- une source dipolaire
- quatre sources monopolaires cohérentes
- une source monopolaire à différentes hauteurs pour de la calibration géométrique
- un aspirateur
- une configuration de type "cocktail party" (des personnes lisant des textes différents simultanément)

Source dipolaire

Le but de cette acquisition est de voir l'impact de la directivité dans la localisation et dans la sélection des capteurs. Par rapport à la prise de mesures présentées dans le manuscrit (section 4.3), peu de choses sont différentes. La mesure de température, les signaux émis et la hauteur de la source sont identiques. Par contre la source est dipolaire (voir figure 5.13) et les positions en X et Y sont différentes : une seule position de source est testée, autour du point $[4, 1; 1, 8; 1, 35]$ mètres. Quatre mesures sont réalisées qui diffèrent par l'orientation de la source. Modifier l'orientation de la source implique des imprécisions sur la position de la source, mais on considère que la position de la source est identique pour toutes les mesures. L'orientation de la source pour chaque mesure est la suivante :

- vers le plafond,
- dans l'axe Y , dans le sens négatif
- dans l'axe X , dans le sens positif
- dans la grande diagonale, c'est-à-dire, vers le plafond, dans le sens négatif de l'axe Y et dans le sens positif de l'axe X .



FIGURE 5.13 – Source dipolaire.

Sources monopolaires cohérentes

Cette acquisition peut être utilisée afin de voir si la cohérence entre les sources influe sur la localisation et la sélection. 4 sources monopolaires cohérentes sont placées sur un support métallique (voir figure 5.14) et émettent simultanément. Les ondes sonores émises interfèrent entre elles. Trois mesures sont réalisées où différentes positions de sources sont testées, comme illustrées dans la figure 5.15. La mesure des positions de chacune des sources varie selon la mesure. Pour la première mesure, vu dans la figure 5.15a, les positions des sources ne varient que selon l'axe Y . La position de la source en vert est mesurée au mètre laser. Les positions des trois autres sources (en bleu) sont estimées par les mesures des distances entre les sources à la règle. Pour la seconde mesure (voir figure 5.15b), la procédure est similaire. La différence réside dans le fait que les positions des sources ne varient que selon l'axe X . Pour la troisième mesure (voir figure 5.15c), comme les positions des sources varient selon les deux axes, il n'est pas possible d'utiliser la règle. Les positions des sources sont mesurées au laser.

Calibration géométrique

Cette acquisition a pour but de faire de la calibration géométrique et de comparer les positions de capteurs mesurées par les collègues et celles estimées par d'autres approches de calibration géométrique. Pour cela il est nécessaire de positionner une source à différentes hauteurs. On

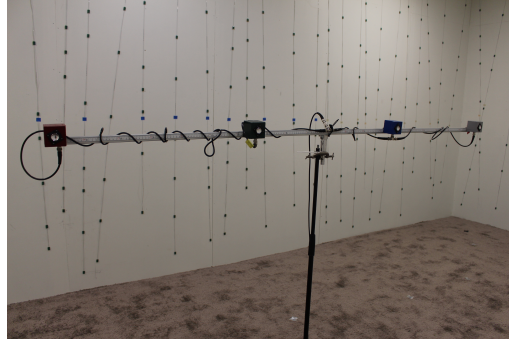


FIGURE 5.14 – Sources monopolaires cohérentes.

définit une position dans le plan XY arbitrairement à $[3,99; 1,82]$ mètres. Trois hauteurs sont choisies :

- au sol,
- à 0,92 mètres,
- à 2,29 mètres.

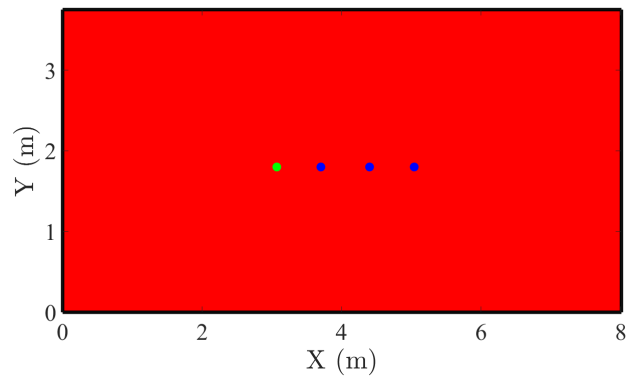
Pour mesurer les hauteurs de la source, à l'exception de celle au sol, nous utilisons le laser pour pointer le mur, puis de mesurer la distance du pointage laser avec le laser plan du chapitre 4 (fixé à une hauteur de 1,35 mètres) par une règle.

Aspirateur

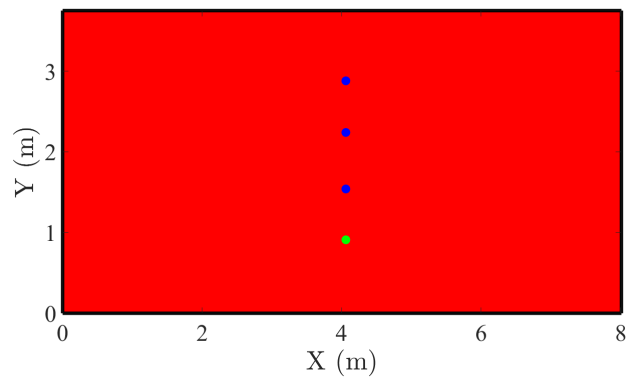
Cette acquisition cherche à localiser un aspirateur dans la salle. Le son émis par un aspirateur comporte plusieurs fréquences émises à différents endroits de l'aspirateur. On pourrait utiliser cette étude pour localiser des sources proches à des fréquences différentes. L'algorithme en l'état ne peut pas être utilisé pour des sources émettant plusieurs fréquences. Une solution envisageable serait de moyenner, par rapport aux fréquences, le critère de sélection de positions de capteurs et l'estimation des positions de source (voir la perspective sur la conception d'antenne avec des signaux large bande de la conclusion) [118].

Cocktail party

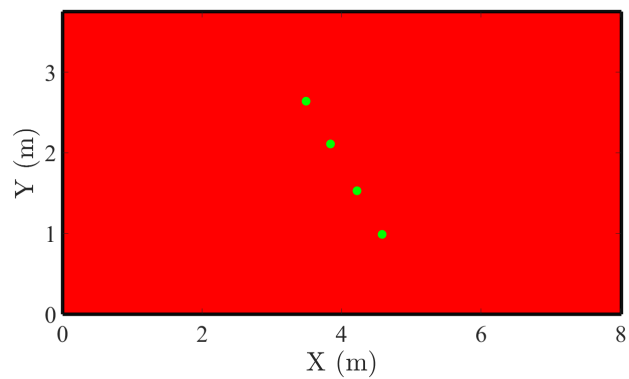
La dernière acquisition est celle d'un "cocktail party". Cela correspond à une salle dans laquelle différentes personnes parlent simultanément. Les personnes sont à des positions différentes et sont orientées dans des directions différentes. Ce point est important puisqu'une personne n'est pas une source omnidirectionnelle. L'autre point à souligner est que chaque personne n'est pas à la même hauteur, et ne parle pas à la même fréquence. Grâce à cette acquisition, on pourrait localiser chacune des personnes, les identifier, puis tenter la sélection de capteurs.



(a) Première mesure



(b) Deuxième mesure



(c) Troisième mesure

FIGURE 5.15 – Positions dans le plan XY des sources mesurées : au laser (vertes) ou à la règle (bleue).

LISTE DES PUBLICATIONS

Articles de journal international

- Courcoux-Caro, M., Vanwynsberghe, C., Herzet, C., & Baussard, A. (2021, December). Sequential sensor selection for sparse source localization. In 2022 The Journal of the Acoustical Society of America (JASA) (submitted).

Articles de conférence international

- Courcoux-Caro, M., Vanwynsberghe, C., Herzet, C., & Baussard, A. (2021, January). Sequential Sensor Placement using Bayesian Compressed Sensing for Source Localization. In 2020 28th European Signal Processing Conference (EUSIPCO) (p. 241-245). IEEE. [\[132\]](#)
- Courcoux-Caro, M., Vanwynsberghe, C., Baussard, A., & Herzet, C. (2020, December). Sequential sensor placement using Bayesian compressive sensing for direction of arrival estimation. In Forum Acusticum (p. 303-304). [\[133\]](#)

Participations Scientifiques

- Courcoux-Caro, M., Vanwynsberghe, C., Herzet, C., & Baussard, A. (2019, Juin). Conception optimisée d'antenne pour la localisation passive de sources acoustiques sous-marines. In 2019 Journées Jeunes Chercheurs en Acoustique Physique (JJCAP).
- Courcoux-Caro, M., & Vanwynsberghe, C. (2019, Juillet). Conception optimisée d'antenne pour la localisation passive de sources acoustiques sous-marines. In 2019 Groupe de travail à Roscoff.

RÉFÉRENCES

- [1] Harry L Van Trees. *Optimum array processing : Part IV of detection, estimation, and modulation theory*. John Wiley & Sons, 2004.
- [2] Whitlow WL Au. *The sonar of dolphins*. Springer Science & Business Media, 2012.
- [3] R Busnel. *Animal sonar systems*, volume 28. Springer Science & Business Media, 2013.
- [4] VM Janik, SM Van Parijs, and PM Thompson. A two-dimensional acoustic localization system for marine mammals. *Marine Mammal Science*, 16(2) :437–447, 2000.
- [5] Robert J Urick. *Principles of underwater sound-2*. 1975.
- [6] MS Babbitt. Smoothing periodograms from time-series with continuous spectra. *Nature*, 161(4096) :686–687, 1948.
- [7] Hamid Krim and Mats Viberg. Two decades of array signal processing research : the parametric approach. *IEEE signal processing magazine*, 13(4) :67–94, 1996.
- [8] Joe C Chen, Kung Yao, and Ralph E Hudson. Source localization and beamforming. *IEEE Signal Processing Magazine*, 19(2) :30–39, 2002.
- [9] Barry D Van Veen and Kevin M Buckley. Beamforming : A versatile approach to spatial filtering. *IEEE assp magazine*, 5(2) :4–24, 1988.
- [10] Don H Johnson and Dan E Dudgeon. *Array signal processing : concepts and techniques*. Simon & Schuster, Inc., 1992.
- [11] Claude Elwood Shannon. Communication in the presence of noise. *Proceedings of the IRE*, 37(1) :10–21, 1949.
- [12] Jack Capon. High-resolution frequency-wavenumber spectrum analysis. *Proceedings of the IEEE*, 57(8) :1408–1418, 1969.
- [13] Ralph Schmidt. Multiple emitter location and signal parameter estimation. *IEEE transactions on antennas and propagation*, 34(3) :276–280, 1986.
- [14] Richard Roy and Thomas Kailath. Esprit-estimation of signal parameters via rotational invariance techniques. *IEEE Transactions on acoustics, speech, and signal processing*, 37(7) :984–995, 1989.

- [15] David L Donoho. Compressed sensing. *IEEE Transactions on information theory*, 52(4) :1289–1306, 2006.
- [16] Emmanuel J Candès, Justin Romberg, and Terence Tao. Robust uncertainty principles : Exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on information theory*, 52(2) :489–509, 2006.
- [17] Angeliki Xenaki and Peter Gerstoft. Grid-free compressive beamforming. *The Journal of the Acoustical Society of America*, 137(4) :1923–1935, 2015.
- [18] Geoffrey F Edelmann and Charles F Gaumond. Beamforming using compressive sensing. *The Journal of the Acoustical Society of America*, 130(4) :EL232–EL237, 2011.
- [19] Francois Ollivier, Antoine Peillot, Gilles Chardon, and Laurent Daudet. Acoustic sources joint localization and characterization using compressed sensing. *The Journal of the Acoustical Society of America*, 131(4) :3257–3257, 2012.
- [20] Hugo Demontis, François Ollivier, and Jacques Marchal. Block-sparse approach for the identification of complex sound sources in a room. 2019.
- [21] Emmanuel J Candès and Carlos Fernandez-Granda. Towards a mathematical theory of super-resolution. *Communications on pure and applied Mathematics*, 67(6) :906–956, 2014.
- [22] Gongguo Tang, Badri Narayan Bhaskar, Parikshit Shah, and Benjamin Recht. Compressed sensing off the grid. *IEEE transactions on information theory*, 59(11) :7465–7490, 2013.
- [23] Chunwei Jethro Lam and Andrew C Singer. Bayesian beamforming for doa uncertainty : theory and implementation. *IEEE Transactions on Signal Processing*, 54(11) :4435–4445, 2006.
- [24] Kristine L Bell, Yariv Ephraim, and Harry L Van Trees. A bayesian approach to robust adaptive beamforming. *IEEE Transactions on Signal Processing*, 48(2) :386–398, 2000.
- [25] Michael E Tipping. Sparse bayesian learning and the relevance vector machine. *Journal of machine learning research*, 1(Jun) :211–244, 2001.
- [26] Zai Yang, Lihua Xie, and Cishen Zhang. Off-grid direction of arrival estimation using sparse bayesian inference. *IEEE Transactions on Signal Processing*, 61(1) :38–43, 2012.
- [27] Quentin Denoyelle. *Theoretical and numerical analysis of super-resolution without grid*. PhD thesis, 2018.

- [28] Joe C Chen, Kung Yao, and Ralph E Hudson. Acoustic source localization and beamforming : theory and practice. *EURASIP Journal on Advances in Signal Processing*, 2003(4) :926837, 2003.
- [29] Finn B Jensen, William A Kuperman, Michael B Porter, and Henrik Schmidt. *Computational ocean acoustics*. Springer Science & Business Media, 2011.
- [30] Florent Le Courtois and Julien Bonnel. Compressed sensing for wideband wavenumber tracking in dispersive shallow water. *The Journal of the Acoustical Society of America*, 138(2) :575–583, 2015.
- [31] Thomas Paviet-Salomon, Clément Dorffer, Julien Bonnel, Barbara Nicolas, Thierry Chonavel, and Angélique Drémeau. Dispersive grid-free orthogonal matching pursuit for modal estimation in ocean acoustics. In *ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4602–4606. IEEE, 2020.
- [32] Hugo Demontis, François Ollivier, and Jacques Marchal. 3d identification of acoustic sources in rooms using a large-scale microphone array. In *2018 16th International Workshop on Acoustic Signal Enhancement (IWAENC)*, pages 506–509. IEEE, 2018.
- [33] Guillaume Beaumont, Ronan Fablet, and Angélique Drémeau. *DOA estimation in fluctuating environments : an approximate message-passing approach*. PhD thesis, ENSTA Bretagne ; Lab-STICC, 2017.
- [34] Riwal Lefort, Gaultier Real, and Angélique Drémeau. Direct regressions for underwater acoustic source localization in fluctuating oceans. *Applied Acoustics*, 116 :303–310, 2017.
- [35] Rui Duan, Kunde Yang, Yuanliang Ma, Qiulong Yang, and Hui Li. Moving source localization with a single hydrophone using multipath time delays in the deep ocean. *The Journal of the Acoustical Society of America*, 136(2) :EL159–EL165, 2014.
- [36] Jean-Marc Valin, François Michaud, and Jean Rouat. Robust localization and tracking of simultaneous moving sound sources using beamforming and particle filtering. *Robotics and Autonomous Systems*, 55(3) :216–228, 2007.
- [37] Eric A Lehmann and Anders M Johansson. Prediction of energy decay in room impulse responses simulated with an image-source model. *The Journal of the Acoustical Society of America*, 124(1) :269–277, 2008.
- [38] Lauri Savioja and U Peter Svensson. Overview of geometrical room acoustic modeling techniques. *The Journal of the Acoustical Society of America*, 138(2) :708–730, 2015.

- [39] Jont B Allen and David A Berkley. Image method for efficiently simulating small-room acoustics. *The Journal of the Acoustical Society of America*, 65(4) :943–950, 1979.
- [40] Georgios I Koutsouris, Jonas Brunskog, Cheol-Ho Jeong, and Finn Jacobsen. Combination of acoustical radiosity and the image source method. *The Journal of the Acoustical Society of America*, 133(6) :3963–3974, 2013.
- [41] Marc Aretz, Pascal Dietrich, and Michael Vorländer. Application of the mirror source method for low frequency sound prediction in rectangular rooms. *Acta Acustica united with Acustica*, 100(2) :306–319, 2014.
- [42] Bernard D Steinberg. Principles of aperture and array system design : Including random and adaptive arrays. *New York*, 1976.
- [43] Tukaram Baburao Lavate, VK Kokate, and AM Sapkal. Performance analysis of music and esprit doa estimation algorithms for adaptive array smart antenna in mobile communication. In *2010 Second International Conference on Computer and Network Technology*, pages 308–311. IEEE, 2010.
- [44] Peter Gerstoft, Angeliki Xenaki, and Christoph F Mecklenbräuer. Multiple and single snapshot compressive beamforming. *The Journal of the Acoustical Society of America*, 138(4) :2003–2014, 2015.
- [45] Quentin Denoyelle, Vincent Duval, and Gabriel Peyré. Support recovery for sparse super-resolution of positive measures. *Journal of Fourier Analysis and Applications*, 23(5) :1153–1194, 2017.
- [46] Zhao Tan and Arye Nehorai. Sparse direction of arrival estimation using co-prime arrays with off-grid targets. *IEEE Signal Processing Letters*, 21(1) :26–29, 2013.
- [47] Richard C Aster, Brian Borchers, and Clifford H Thurber. *Parameter estimation and inverse problems*. Elsevier, 2018.
- [48] Iain McCowan. Microphone arrays : A tutorial. *Queensland University, Australia*, pages 1–38, 2001.
- [49] Constantine A Balanis. *Antenna theory : analysis and design*. John wiley & sons, 2015.
- [50] Chandra Vaidyanathan and Kevin M Buckley. Performance analysis of the mvdr spatial spectrum estimator. *IEEE Transactions on Signal Processing*, 43(6) :1427–1437, 1995.
- [51] Youssef Khmou, Said Safi, and Miloud Frikel. Comparative study between several direction of arrival estimation methods. *Journal of Telecommunications and Information Technology*, 2014.

- [52] Lal C Godara. Application of antenna arrays to mobile communications. ii. beam-forming and direction-of-arrival considerations. *Proceedings of the IEEE*, 85(8) :1195–1245, 1997.
- [53] Mark A Davenport, Marco F Duarte, Yonina C Eldar, and Gitta Kutyniok. Introduction to compressed sensing. *preprint*, 93(1) :2, 2011.
- [54] Angeliki Xenaki, Peter Gerstoft, and Klaus Mosegaard. Compressive beamforming. *The Journal of the Acoustical Society of America*, 136(1) :260–271, 2014.
- [55] Stéphane G Mallat and Zhifeng Zhang. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on signal processing*, 41(12) :3397–3415, 1993.
- [56] Rémi Gribonval, Philippe Depalle, Xavier Rodet, Emmanuel Bacry, and Stéphane Mallat. Sound signals decomposition using a high resolution matching pursuit. In *Proceedings of the International Computer Music Conference*, pages 293–296. Citeseer, 1996.
- [57] Yagyensh Chandra Pati, Ramin Rezaifar, and Perinkulam Sambamurthy Krishnaprasad. Orthogonal matching pursuit : Recursive function approximation with applications to wavelet decomposition. In *Proceedings of 27th Asilomar conference on signals, systems and computers*, pages 40–44. IEEE, 1993.
- [58] Joel A Tropp and Anna C Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on information theory*, 53(12) :4655–4666, 2007.
- [59] Scott Shaobing Chen, David L Donoho, and Michael A Saunders. Atomic decomposition by basis pursuit. *SIAM review*, 43(1) :129–159, 2001.
- [60] Ewout Van Den Berg and Michael P Friedlander. Probing the pareto frontier for basis pursuit solutions. *SIAM Journal on Scientific Computing*, 31(2) :890–912, 2009.
- [61] Amir Beck and Marc Teboulle. A fast iterative shrinkage-thresholding algorithm for linear inverse problems. *SIAM journal on imaging sciences*, 2(1) :183–202, 2009.
- [62] Michael Grant and Stephen Boyd. Cvx : Matlab software for disciplined convex programming, version 2.1, 2014.
- [63] In Jae Myung. Tutorial on maximum likelihood estimation. *Journal of mathematical Psychology*, 47(1) :90–100, 2003.
- [64] Manohar Shamaiah, Siddhartha Banerjee, and Haris Vikalo. Greedy sensor selection : Leveraging submodularity. In *49th IEEE conference on decision and control (CDC)*, pages 2572–2577. IEEE, 2010.

- [65] Karin C Knudson, Jacob L Yates, Alexander C Huk, and JonathanW Pillow. Inferring sparse representations of continuous signals with continuous orthogonal matching pursuit. *Advances in neural information processing systems*, 27, 2014.
- [66] Quentin Denoyelle, Vincent Duval, Gabriel Peyré, and Emmanuel Soubies. The sliding frank–wolfe algorithm and its application to super-resolution microscopy. *Inverse Problems*, 36(1) :014001, 2019.
- [67] Anup Das. Theoretical and experimental comparison of off-grid sparse bayesian direction-of-arrival estimation algorithms. *IEEE access*, 5 :18075–18087, 2017.
- [68] Y-D Huang and Mourad Barkat. Near-field multiple source localization by passive sensor array. *IEEE Transactions on antennas and propagation*, 39(7) :968–975, 1991.
- [69] Dmitry Malioutov, Müjdat Cetin, and Alan S Willsky. A sparse signal reconstruction perspective for source localization with sensor arrays. *IEEE transactions on signal processing*, 53(8) :3010–3022, 2005.
- [70] Tarik Yardibi, Jian Li, Petre Stoica, Ming Xue, and Arthur B Baggeroer. Source localization and sensing : A nonparametric iterative adaptive approach based on weighted least squares. *IEEE Transactions on Aerospace and Electronic Systems*, 46(1) :425–443, 2010.
- [71] Peter Gerstoft, Christoph F Mecklenbräuker, Angeliki Xenaki, and Santosh Nannuru. Multisnapshot sparse bayesian learning for doa. *IEEE Signal Processing Letters*, 23(10) :1469–1473, 2016.
- [72] Paul Jaccard. Distribution de la flore alpine dans le bassin des dranses et dans quelques régions voisines. *Bull Soc Vaudoise Sci Nat*, 37 :241–272, 1901.
- [73] Roger E Kirk. Experimental design. *Handbook of Psychology, Second Edition*, 2, 2012.
- [74] Allen Edwards. Experimental design in psychological research, rev. 1960.
- [75] Everet Franklin Lindquist. Design and analysis of experiments in psychology and education. 1953.
- [76] Jerry P Queen, Gerry P Quinn, and Michael J Keough. *Experimental design and data analysis for biologists*. Cambridge university press, 2002.
- [77] Michelle Jackson and David R Cox. The principles of experimental design and their application in sociology. *Annual Review of Sociology*, 39 :27–49, 2013.
- [78] Vinayak Govind Panse, Pandurang Vasudeo Sukhatme, et al. Statistical methods for agricultural workers. *Statistical methods for agricultural workers.*, 1954.

- [79] Thomas Morton Little and Frederick Jackson Hills. Agricultural experimentation : design and analysis. 1978.
- [80] Charles E Metz. Some practical issues of experimental design and data analysis in radiological roc studies. *Investigative radiology*, 24(3) :234–245, 1989.
- [81] Michael Papadopoulos and Ephraim Garcia. Sensor placement methodologies for dynamic testing. *AIAA journal*, 36(2) :256–263, 1998.
- [82] Andreas Krause, Carlos Guestrin, Anupam Gupta, and Jon Kleinberg. Near-optimal sensor placements : Maximizing information while minimizing communication cost. In *Proceedings of the 5th international conference on Information processing in sensor networks*, pages 2–10, 2006.
- [83] Laurent Gilquin, Pierre Lecomte, Jérôme Antoni, Thibaut Le Magueresse, and Clément Marteau. Iterative positioning of microphone arrays for acoustic imaging. *Journal of Sound and Vibration*, 469 :115116, 2020.
- [84] David Moreno-Salinas, Antonio Pascoal, and Joaquin Aranda. Optimal sensor placement for acoustic underwater target positioning with range-only measurements. *IEEE Journal of Oceanic Engineering*, 41(3) :620–643, 2016.
- [85] Zebb Prime and Con Doolan. A comparison of popular beamforming arrays. *Proceedings of the Australian Acoustical Society AAS2013 Victor Harbor*, 1 :5, 2013.
- [86] Ennes Sarradj. A generic approach to synthesize optimal array microphone arrangements. In *Proceedings of the 6th Berlin Beamforming Conference*, page 4, 2016.
- [87] A Nordborg, J Wedemann, and L Willenbrink. Optimum array microphone configuration. *Proceedings of Inter-Noise 2000*, 2000.
- [88] Lawrence Carin. On the relationship between compressive sensing and random sensor arrays. *IEEE Antennas and Propagation Magazine*, 51(5) :72–81, 2009.
- [89] Lawrence Carin, Dehong Liu, and Bin Guo. Coherence, compressive sensing, and random sensor arrays. *IEEE Antennas and Propagation Magazine*, 53(4) :28–39, 2011.
- [90] J Hald and JJ Christensen. A class of optimal broadband phased array geometries designed for easy construction. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 2002, pages 1993–1998. Institute of Noise Control Engineering, 2002.
- [91] Haohai Sun, Shefeng Yan, and U Peter Svensson. Robust minimum sidelobe beamforming for spherical microphone arrays. *IEEE transactions on audio, speech, and language processing*, 19(4) :1045–1051, 2010.

- [92] Miloš Bjelić, Miodrag Stanojević, Dragana Šumarac Pavlović, and Miomir Mijić. Microphone array geometry optimization for traffic noise analysis. *The Journal of the Acoustical Society of America*, 141(5) :3101–3104, 2017.
- [93] Melanie Mitchell. *An introduction to genetic algorithms*. MIT press, 1998.
- [94] Lawrence Davis. *Handbook of genetic algorithms*. 1991.
- [95] Florent Le Courtois, Jean-Hugh Thomas, Franck Poisson, and Jean-Claude Pascal. Genetic optimisation of a plane array geometry for beamforming. application to source localisation in a high speed train. *Journal of Sound and Vibration*, 371 :78–93, 2016.
- [96] Friedrich Pukelsheim. *Optimal design of experiments*. SIAM, 2006.
- [97] Valerii Vadimovich Fedorov. *Theory of optimal experiments*. Elsevier, 2013.
- [98] Kirstine Smith. On the standard deviations of adjusted and interpolated values of an observed polynomial function and its constants and the guidance they give towards a proper choice of the distribution of observations. *Biometrika*, 12(1/2) :1–85, 1918.
- [99] Sundeep Prabhakar Chepuri and Geert Leus. Continuous sensor placement. *IEEE Signal Processing Letters*, 22(5) :544–548, 2014.
- [100] Siddharth Joshi and Stephen Boyd. Sensor selection via convex optimization. *IEEE Transactions on Signal Processing*, 57(2) :451–462, 2008.
- [101] Kathryn Chaloner and Isabella Verdinelli. Bayesian experimental design : A review. *Statistical Science*, pages 273–304, 1995.
- [102] Shihao Ji, Ya Xue, and Lawrence Carin. Bayesian compressive sensing. *IEEE Transactions on signal processing*, 56(6) :2346–2356, 2008.
- [103] Matthias W Seeger and Hannes Nickisch. Compressed sensing and bayesian experimental design. In *Proceedings of the 25th international conference on Machine learning*, pages 912–919, 2008.
- [104] Krithika Manohar, Bingni W Brunton, J Nathan Kutz, and Steven L Brunton. Data-driven sparse sensor placement for reconstruction : Demonstrating the benefits of exploiting known patterns. *IEEE Control Systems Magazine*, 38(3) :63–86, 2018.
- [105] Prabhakar S Naidu. *Sensor array signal processing*. CRC press, 2009.
- [106] K Yao, C Reed, RE Hudson, and F Lorenzelli. Beamforming performance of a randomly distributed sensor array system. In *1997 IEEE Workshop on Signal Processing Systems*.

-
- SiPS 97 Design and Implementation formerly VLSI Signal Processing*, pages 438–447. IEEE, 1997.
- [107] Kit-Sang Tang, Kim-Fung Man, Sam Kwong, and Qun He. Genetic algorithms and their applications. *IEEE signal processing magazine*, 13(6) :22–37, 1996.
- [108] Kim-Fung Man, Kit-Sang Tang, and Sam Kwong. Genetic algorithms : concepts and applications [in engineering design]. *IEEE transactions on Industrial Electronics*, 43(5) :519–534, 1996.
- [109] Leehter Yao, William A Sethares, and Daniel C Kammer. Sensor placement for on-orbit modal identification via a genetic algorithm. *AIAA journal*, 31(10) :1922–1928, 1993.
- [110] Wei Liu, Wei-cheng Gao, Yi Sun, and Min-jian Xu. Optimal sensor placement for spatial lattice structure based on genetic algorithms. *Journal of Sound and Vibration*, 317(1-2) :175–189, 2008.
- [111] Sébastien Debert and François Ollivier. Large array optimization using a genetic algorithm and application to wide band moving sources imaging. *The Journal of the Acoustical Society of America*, 123(5) :3388–3388, 2008.
- [112] Tomasz Dominik Gwiazda. *Crossover for single-objective numerical optimization problems*, volume 1. Tomasz Gwiazda, 2006.
- [113] Costas Papadimitriou. Optimal sensor placement methodology for parametric identification of structural systems. *Journal of sound and vibration*, 278(4-5) :923–947, 2004.
- [114] Harald Cramér. *Mathematical Methods of Statistics (PMS-9), Volume 9*. Princeton university press, 2016.
- [115] Calyampudi Radhakrishna Rao. *Linear statistical inference and its applications*, volume 2. Wiley New York, 1973.
- [116] Solomon Kullback and Richard A Leibler. On information and sufficiency. *The annals of mathematical statistics*, 22(1) :79–86, 1951.
- [117] Claude Elwood Shannon. A mathematical theory of communication. *ACM SIGMOBILE mobile computing and communications review*, 5(1) :3–55, 2001.
- [118] Shoichi Koyama, Gilles Chardon, and Laurent Daudet. Optimizing source and sensor placement for sound field control : An overview. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 28 :696–714, 2020.

- [119] Nathaniel R Goodman. Statistical analysis based on a certain multivariate complex gaussian distribution (an introduction). *The Annals of mathematical statistics*, 34(1) :152–177, 1963.
- [120] Robin A Wooding. The multivariate distribution of complex normal variables. *Biometrika*, 43(1/2) :212–215, 1956.
- [121] Sundeep Prabhakar Chepuri and Geert Leus. Sparsity-promoting sensor selection for non-linear measurement models. *IEEE Transactions on Signal Processing*, 63(3) :684–698, 2014.
- [122] David A Harville. Matrix algebra from a statistician’s perspective, 1998.
- [123] Nicholas J Higham. *Accuracy and stability of numerical algorithms*. SIAM, 2002.
- [124] Mark A Davenport, Andrew K Massimino, Deanna Needell, and Tina Woolf. Constrained adaptive sensing. *IEEE Transactions on Signal Processing*, 64(20) :5437–5449, 2016.
- [125] Ery Arias-Castro, Emmanuel J Candes, and Mark A Davenport. On the fundamental limits of adaptive sensing. *IEEE Transactions on Information Theory*, 59(1) :472–481, 2012.
- [126] Shihao Ji, David Dunson, and Lawrence Carin. Multitask compressive sensing. *IEEE Transactions on Signal Processing*, 57(1) :92–106, 2008.
- [127] Claude Elwood Shannon. A mathematical theory of communication. *The Bell system technical journal*, 27(3) :379–423, 1948.
- [128] Kuoye Han, Yanping Wang, Bo Kou, and Wen Hong. Parameters estimation using a random linear array and compressed sensing. In *2010 3rd International Congress on Image and Signal Processing*, volume 8, pages 3950–3954. IEEE, 2010.
- [129] Hongmei Zhang, Zhenguogao Gao, and Huixuan Fu. A high resolution random linear sonar array based music method for underwater doa estimation. In *Proceedings of the 32nd Chinese Control Conference*, pages 4592–4595. IEEE, 2013.
- [130] Charles F Gaumond and Geoffrey F Edelmann. Sparse array design using statistical restricted isometry property. *The Journal of the Acoustical Society of America*, 134(2) :EL191–EL197, 2013.
- [131] Hongqing Liu, Luming Zhao, Yong Li, Xiaorong Jing, and Trieu-Kien Truong. A sparse-based approach for doa estimation and array calibration in uniform linear array. *IEEE sensors journal*, 16(15) :6018–6027, 2016.

- [132] Milan Courcoux-Caro, Charles Vanwynsberghe, Cédric Herzet, and Alexandre Baussard. Sequential sensor placement using bayesian compressed sensing for source localization. In *2020 28th European Signal Processing Conference (EUSIPCO)*, pages 241–245. IEEE, 2021.
- [133] Milan Courcoux-Caro, Charles Vanwynsberghe, Alexandre Baussard, and Cédric Herzet. Sequential sensor placement using bayesian compressive sensing for direction of arrival estimation. In *Forum Acousticum*, pages 303–304, 2020.
- [134] Analog Devices. Omnidirectional microphone with bottom port and i2s digital output. *ADMP421, Data Sheet*, 2012.
- [135] Ines Hafizovic, Carl-Inge Colombo Nilsen, Morgan Kjølervbakken, and Vibeke Jahr. Design and implementation of a mems microphone array system for real-time speech acquisition. *Applied acoustics*, 73(2) :132–143, 2012.
- [136] Guinness World Records. Largest microphone array, 2014.
- [137] Charles Vanwynsberghe. *Réseaux à grand nombre de microphones : applicabilité et mise en œuvre*. PhD thesis, Paris 6, 2016.
- [138] Charles Vanwynsberghe, Régis Marchiano, François Ollivier, Pascal Challande, Hélène Moingeon, and Jacques Marchal. Design and implementation of a multi-octave-band audio camera for realtime diagnosis. *Applied Acoustics*, 89 :281–287, 2015.
- [139] Raphaël Leiba. *Conception d’un outil de diagnostic de la gêne sonore en milieu urbain*. PhD thesis, Université Pierre et Marie Curie-Paris VI, 2017.
- [140] Hugo Demontis. *Identification de sources acoustiques complexes en milieu réverbérant par grands réseaux de microphones*. PhD thesis, Sorbonne université, 2019.
- [141] Raphaël Leiba, Régis Marchiano, Nicolas Misdariis, and Jacques Marchal. Urban acoustic imaging : from measurement to the soundscape perception evaluation. In *INTER-NOISE and NOISE-CON Congress and Conference Proceedings*, volume 253, pages 6803–6809. Institute of Noise Control Engineering, 2016.
- [142] Nicolas Misdariis, Régis Marchiano, Patrick Susini, François Ollivier, Raphael Leiba, and Jacques Marchal. Mobility and life quality relationships—measurement and perception of noise in urban context. In *Internoise-International Congress on Noise Control Engineering*, page 21, 2014.
- [143] Régis Marchiano, Philippe Druault, Raphael Leiba, Jacques Marchal, François Ollivier, Vincent Valeau, and Charles Vanwynsberghe. Localisation de sources aéroacoustiques

- par une méthode de retournement temporel tri-dimensionnelle. In *Congrès Français d'Acoustique*, 2016.
- [144] Thibaut Meurisse, François Ollivier, Jacques Marchal, and Charles Vanwynsberghe. Analyse modale en temps réel par holographie de champ proche et calculs parallèles. In *Congrès Français d'Acoustique*, 2016.
- [145] Ronald M Aarts and Augustus JEM Janssen. Sound radiation from a resilient spherical cap on a rigid sphere. *The Journal of the Acoustical Society of America*, 127(4) :2262–2273, 2010.
- [146] Angélique Drémeau, Florent Le Courtois, and Julien Bonnel. Reconstruction of dispersion curves in the frequency-wavenumber domain using compressed sensing on a random array. *IEEE Journal of Oceanic Engineering*, 42(4) :914–922, 2017.
- [147] Dongning Guo. Relative entropy and score function : New information-estimation relationships through arbitrary additive perturbation. In *2009 IEEE International Symposium on Information Theory*, pages 814–818. IEEE, 2009.
- [148] Solomon Kullback. *Information theory and statistics*. Courier Corporation, 1997.
- [149] Alen Alexanderian and Arvind K Saibaba. Efficient d-optimal design of experiments for infinite-dimensional bayesian linear inverse problems. *SIAM Journal on Scientific Computing*, 40(5) :A2956–A2985, 2018.
- [150] Thomas M Cover. *Elements of information theory*. John Wiley & Sons, 1999.
- [151] Michael C Shewry and Henry P Wynn. Maximum entropy sampling. *Journal of applied statistics*, 14(2) :165–170, 1987.
- [152] Hanbiao Wang, Kung Yao, Greg Pottie, and Deborah Estrin. Entropy-based sensor selection heuristic for target localization. In *Proceedings of the 3rd international symposium on Information processing in sensor networks*, pages 36–45, 2004.
- [153] Roger Dashen, Walter H Munk, Kenneth M Watson, and Frederik Zachariasen. *Sound transmission through a fluctuating ocean*. Cambridge University Press, 2010.
- [154] WH Munk and F Zachariasen. Sound propagation through a fluctuating stratified ocean : Theory and observation. *The Journal of the Acoustical Society of America*, 59(4) :818–838, 1976.
- [155] Jeffrey Borish. Extension of the image model to arbitrary polyhedra. *The Journal of the Acoustical Society of America*, 75(6) :1827–1836, 1984.

- [156] Roger Penrose. A generalized inverse for matrices. In *Mathematical proceedings of the Cambridge philosophical society*, volume 51, pages 406–413. Cambridge University Press, 1955.
- [157] Taesang Yoo and Andrea Goldsmith. On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming. *IEEE Journal on selected areas in communications*, 24(3) :528–541, 2006.
- [158] Raymond Knopp and Pierre A Humblet. Information capacity and power control in single-cell multiuser communications. In *Proceedings IEEE International Conference on Communications ICC'95*, volume 1, pages 331–335. IEEE, 1995.
- [159] Xi-Ren Cao and Ruey-wen Liu. General approach to blind source separation. *IEEE Transactions on signal Processing*, 44(3) :562–571, 1996.
- [160] J-F Cardoso. Source separation using higher order moments. In *International Conference on Acoustics, Speech, and Signal Processing*,, pages 2109–2112. IEEE, 1989.
- [161] Emmanuel J Candès and Michael B Wakin. An introduction to compressive sampling. *IEEE signal processing magazine*, 25(2) :21–30, 2008.
- [162] Dragana Bajovic, Bruno Sinopoli, and Joao Xavier. Sensor selection for event detection in wireless sensor networks. *IEEE Transactions on Signal Processing*, 59(10) :4938–4953, 2011.
- [163] Futoshi Asano, Yôiti Suzuki, and David C Swanson. Optimization of control source configuration in active control systems using gram-schmidt orthogonalization. *IEEE transactions on speech and audio processing*, 7(2) :213–220, 1999.
- [164] Andreas Krause, Ajit Singh, and Carlos Guestrin. Near-optimal sensor placements in gaussian processes : Theory, efficient algorithms and empirical studies. *Journal of Machine Learning Research*, 9(Feb) :235–284, 2008.
- [165] Juri Ranieri, Amina Chebira, and Martin Vetterli. Near-optimal sensor placement for linear inverse problems. *IEEE Transactions on signal processing*, 62(5) :1135–1146, 2014.
- [166] Gilles Chardon, Wolfgang Kreuzer, and Markus Noisternig. Design of spatial microphone arrays for sound field interpolation. *IEEE Journal of Selected Topics in Signal Processing*, 9(5) :780–790, 2015.
- [167] Dongsheng Li. *Sensor placement methods and evaluation criteria in structural health monitoring*. PhD thesis, Universitätsbibliothek der Universität Siegen, 2011.

- [168] Shilpa Rao, Sundeep Prabhakar Chepuri, and Geert Leus. Greedy sensor selection for non-linear models. In *2015 IEEE 6th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 241–244. IEEE, 2015.
- [169] Sijia Liu, Sundeep Prabhakar Chepuri, Makan Fardad, Engin Maşazade, Geert Leus, and Pramod K Varshney. Sensor selection for estimation with correlated measurement noise. *IEEE Transactions on Signal Processing*, 64(13) :3509–3522, 2016.
- [170] Wei Zhang, Wei Liu, Ju Wang, and Siliang Wu. Computationally efficient 2-d doa estimation for uniform rectangular arrays. *Multidimensional Systems and Signal Processing*, 25(4) :847–857, 2014.
- [171] Xionghu Zhong and A Benjamin Premkumar. Particle filtering approaches for multiple acoustic source detection and 2-d direction of arrival estimation using a single acoustic vector sensor. *IEEE Transactions on Signal Processing*, 60(9) :4719–4733, 2012.
- [172] Qinghua Huang, Longfei Xiang, and Kai Liu. Off-grid doa estimation in real spherical harmonics domain using sparse bayesian inference. *Signal Processing*, 137 :124–134, 2017.
- [173] Kay L Gemba, Santosh Nannuru, Peter Gerstoft, and William S Hodgkiss. Multi-frequency sparse bayesian learning for robust matched field processing. *The Journal of the Acoustical Society of America*, 141(5) :3411–3420, 2017.
- [174] Biao Wu, Yong Huang, Xiang Chen, Sridhar Krishnaswamy, and Hui Li. Guided-wave signal processing by the sparse bayesian learning approach employing gabor pulse model. *Structural Health Monitoring*, 16(3) :347–362, 2017.
- [175] Haiqiang Niu, Peter Gerstoft, Emma Ozanich, Zhenglin Li, Renhe Zhang, Zaixiao Gong, and Haibin Wang. Block sparse bayesian learning for broadband mode extraction in shallow water from a vertical array. *The Journal of the Acoustical Society of America*, 147(6) :3729–3739, 2020.
- [176] Sylvie Marcos and Jacques Munier. Source localization using a distorted antenna. In *International Conference on Acoustics, Speech, and Signal Processing*,, pages 2756–2759. IEEE, 1989.
- [177] Charles Vanwynsberghe, Pascal Challande, François Ollivier, Jacques Marchal, and Régis Marchiano. Geometric calibration of very large microphone arrays in mismatched free field. *The Journal of the Acoustical Society of America*, 145(1) :215–227, 2019.

Titre : Conception optimisée d'antenne pour de la localisation passive de sources acoustiques

Mot clés : localisation de sources acoustiques, aérien, sous-marin, conception d'antenne, données expérimentales, sélection séquentielle, approches bayésiennes.

Résumé : L'objectif de ces travaux de thèse est de proposer une approche de conception d'antenne pour de la localisation de sources acoustiques aériennes ou sous-marines. Dans un premier temps, nous verrons comment décrire la propagation des ondes émises par les sources. Puis nous dresserons les différentes méthodes permettant d'estimer les positions des sources à partir des mesures acquises par les capteurs et du modèle de propagation. Le travail bibliographique sur le domaine de la localisation de sources acoustiques mettra en évidence l'importance des positions des capteurs dans les performances des estimateurs de positions de sources. Cela amènera donc à un second travail bibliographique sur la conception d'antenne, et plus précisément sur la sélection de positions de capteurs.

L'état de l'art sur les méthodes de conception d'antenne nous permettra de proposer la méthode suivante : une sélection séquentielle bayésienne orientée par les données. Cette approche prend en compte l'information contenue dans les mesures acquises précédemment dans l'antenne pré-conçue, afin de sélectionner les futurs capteurs. L'application de cette

approche est innovante dans un contexte de localisation de sources acoustiques.

Dans un deuxième temps, des premiers résultats réalisés sur données synthétiques s'avèrent prometteurs quant à l'application de cette approche dans un cadre de localisation de sources acoustiques sous-marines. Les analyses fournies dans ce manuscrit permettront de juger la pertinence de cette approche dans le contexte testé, ainsi que d'évaluer et de comparer la performance de cette conception d'antenne par rapport à la littérature.

Pour finir, nous appliquerons cette approche sur des données réelles provenant d'une expérience mise en place lors de la thèse. Le milieu de propagation sera aérien dans une salle fermée réverbérante avec un grand nombre de capteurs disponibles. Acquérir des données réelles permettront, en plus de l'évaluation des performances de notre approche, de fournir une nouvelle conception d'antenne. Celle-ci est conçue à partir de données synthétiques et est appliquée ensuite sur les données réelles.

Title: Optimized array design for passive acoustic source location

Keywords: acoustic source location, airborne, underwater, antenna design, real data, sequential selection, Bayesian approaches.

Abstract: The objective of this thesis work is to propose an approach of array design for the localization of aerial or underwater acoustic sources. First, we will see how to describe the propagation of waves emitted by the sources. Then we will draw up the different methods allowing to estimate the positions of the sources from the measurements acquired by the sensors and from the propagation model. The bibliographic work on the field of acoustic source localization will highlight the importance of the sensors' positions in the performance of the source position estimators. This will lead to a second bibliographic work on antenna design, and more precisely on the selection of sensor positions.

The state of the art on antenna design methods will allow us to propose the following method: a sequential Bayesian data-driven selection. This approach takes into account the information contained in the measurements previously acquired in the pre-designed antenna, in order to select the future sen-

sors. The application of this approach is innovative in the context of acoustic source localization.

In a second step, first results realized on synthetic data prove to be promising for the application of this approach in an underwater acoustic source localization context. The analyses provided in this manuscript will allow to judge the relevance of this approach in the tested context, as well as to evaluate and compare the performance of this antenna design with the literature.

Finally, we will apply this approach on real data from an experiment set up during the thesis. The propagation environment will be aerial in a closed reverberant room with a large number of available sensors. Acquiring real data will allow, in addition to evaluating the performance of our approach, to provide a new antenna design. This is designed from synthetic data and then applied to real data.

