



Estimation de la direction d'arrivée d'échos sonores à large bande noyés dans le signal direct

Mémoire

Marc-André Guérard

Maîtrise en génie électrique
Maître ès sciences (M.Sc.)

Québec, Canada

© Marc-André Guérard, 2016

Estimation de la direction d'arrivée d'échos sonores à large bande noyés dans le signal direct

Mémoire

Marc-André Guérard

Sous la direction de:

Directeur de recherche Dominic Grenier

Résumé

Ce mémoire présente deux algorithmes qui ont pour but d'améliorer la précision de l'estimation de la direction d'arrivée de sources sonores et de leurs échos. Le premier algorithme, qui s'appelle la méthode par élimination des sources, permet d'améliorer l'estimation de la direction d'arrivée d'échos qui sont noyés dans le bruit. Le second, qui s'appelle Multiple Signal Classification à focalisation de phase, utilise l'information dans la phase à chaque fréquence pour déterminer la direction d'arrivée de sources à large bande. La combinaison de ces deux algorithmes permet de localiser des échos dont la puissance est de -17 dB par rapport à la source principale, jusqu'à un rapport écho-à-bruit de -15 dB. Ce mémoire présente aussi des mesures expérimentales qui viennent confirmer les résultats obtenus lors de simulations.

Abstract

This memoir presents two algorithms which goal is to determine the direction of arrival of wideband acoustic sources. The first algorithm, called Source Elimination Method, eliminate the contribution of every source from the covariance matrix to improve the estimation of the direction of arrival estimate of echoes in a low echo-to-noise ratio. The second algorithm, called Multiple Signal Classification with Phase Focalisation, uses the information contained in the phase of every frequency to find the direction of arrival of wideband sources. The combination of the two algorithm permits to find the direction of arrival of two echoes, which power is - 17 dB compared to the main source, up to an echo to noise ratio of -15 dB. Experimental results confirming the precision of these algorithm are presented.

Table des matières

Résumé	iii
Abstract	iv
Table des matières	v
Liste des tableaux	vii
Liste des figures	viii
Liste des symboles	xi
Liste des acronymes	xv
1 Introduction	1
1.1 Plan du mémoire	2
2 Estimation de directions d'arrivée	4
2.1 Modèle du signal	4
2.2 Méthodes à bande étroite	8
2.3 Méthodes à large bande	13
2.4 Corrélation de sources	17
2.5 Détermination du nombre de sources	22
2.6 Conclusion	22
3 Méthode par élimination des sources	23
3.1 Principe de base	24
3.2 Génération des données	26
3.3 Estimation des paramètres	30
3.4 Simulations	34
3.5 Conclusion	48
4 MUSIC à focalisation de phase	51
4.1 Algorithme	52
4.2 Détermination de la fréquence de référence optimale	54
4.3 Essais Monte-Carlo	56

4.4	Conclusion	62
5	Estimation de DOA de signaux acoustiques réfléchis	67
5.1	Sources de puissances différentes	67
5.2	Test de la puissance des sources	68
5.3	Simulations Monte-Carlo	71
5.4	Conclusion	83
6	Mesures expérimentales	86
6.1	Montage expérimental	87
6.2	Mesures de signaux à bande étroite	89
6.3	Signaux à large bande	95
6.4	Conclusion	100
7	Conclusion	103
7.1	Travaux futurs	105
	Bibliographie	107
A	Article soumis	110

Liste des tableaux

4.1	Fréquences de références maximales pour différentes distances inter éléments.	55
-----	---	----

Liste des figures

2.1	Réseau de capteurs à espacement linéaire. Les points représentent des microphones.	5
2.2	Illustration du filtrage spatial. Chaque sortie du réseau est multipliée par un coefficient du filtre.	8
2.3	Illustration montrant la décomposition du réseau initial en sous-réseaux pour la méthode de lissage spatial.	18
3.1	Exemple de signal généré avec la méthode présentée dans cette section, avec les différentes étapes illustrées.	27
3.2	Pseudospectre MUSIC en assumant une source pour un signal composé d'une source principale et de deux sources secondaires (échos) sans bruit pour différents EDR.	32
3.3	Comportement de la première et de la dernière valeur propre de la matrice de covariance partiellement réduite en fonction de l'erreur sur le puissance posée.	33
3.4	Comparaison des résultats obtenus avec MUSIC-FBSS et la méthode de réduction de la matrice de covariance pour deux cas différents, l'un fait à l'aveuglette et l'autre en utilisant les positions exactes de départ. Les données proviennent de deux sources différentes non corrélées, dont l'une est 13 dB plus puissance que l'autre.	36
3.5	Comparaison des résultats obtenus avec MUSIC avec FBSS et SEM pour deux cas différents, l'un fait à l'aveuglette et l'un en utilisant les positions exactes de départ. Les données proviennent de deux sources différentes non corrélées, dont l'une est 3 dB plus puissance que l'autre.	37
3.6	Erreur moyenne de la DOA estimée d'une source faible d'une série de 200 essais Monte-Carlo en fonction du SNR pour des signaux non corrélés.	39
3.7	Écart-type d'une série de 200 essais Monte-Carlo en fonction du SNR pour des signaux non corrélés.	40
3.8	Probabilité de fausse alarme en fonction du SNR obtenu lors de 200 essais Monte-Carlo avec des signaux non corrélés.	41
3.9	Erreur moyenne de la DOA estimée d'une source faible d'une série de 200 essais Monte-Carlo en fonction du ENR pour un signal composé d'une source en trajet direct puissant et d'un écho beaucoup plus faible.	42

3.10	Écart-type d'une série de 200 essais Monte-Carlo en fonction du ENR pour un signal composé d'une source en trajet direct puissant et d'un écho beaucoup plus faible.	43
3.11	Probabilité de fausse alarme en fonction du ENR obtenu avec 200 essais Monte-Carlo pour un signal composé d'une source en trajet direct puissant et d'un écho beaucoup plus faible.	44
3.12	Probabilité de fausse alarme en fonction du ENR obtenu lors de 200 essais Monte-Carlo pour un signal composé d'une source et de deux échos.	47
3.13	Moyenne de l'erreur des bonnes détections de 200 essais Monte-Carlo en fonction du ENR pour un signal composé d'une source et de deux échos après retrait des cas de fausse alarme.	48
3.14	Écart-type de l'erreur des bonnes détections de 200 essais Monte-Carlo en fonction du ENR pour un signal composé d'une source et de deux échos après retrait des cas de fausse alarme.	49
4.1	Impact du choix de la fréquence de référence sur le pseudospectre obtenu à l'aide de MUSIC-FP lorsque le signal mesuré est composé d'une seule source.	56
4.2	Impact du choix de la fréquence de référence sur le pseudospectre obtenu à l'aide de MUSIC-FP lorsque le signal mesuré est composé de deux sources.	57
4.3	Moyenne de l'erreur de 200 essais de Monte-Carlo en fonction du rapport signal à bruit pour des signaux non corrélés.	59
4.4	Écart-type de l'erreur de 200 essais de Monte-Carlo en fonction du rapport signal à bruit pour des signaux non corrélés.	60
4.5	Probabilité de fausse alarme en fonction du SNR de 200 essais de Monte-Carlo avec deux sources non corrélées.	61
4.6	Moyenne de l'erreur de 200 essais Monte-Carlo en fonction du rapport signal à bruit pour des signaux corrélés.	63
4.7	Écart-type de 200 essais Monte-Carlo en fonction du rapport signal à bruit pour des signaux corrélés.	64
4.8	Probabilité de fausse alarme en fonction du SNR de 200 essais Monte-Carlo avec deux sources corrélées	65
5.1	Effet de la variation du rapport de puissance entre deux sources à large bande et non corrélées pour deux méthodes différentes, MUSIC-FP et CSSM, dans un cas sans bruit.	69
5.2	Effet de la variation du rapport de puissance entre deux sources indépendantes à large bande et non corrélées pour deux méthodes différentes, MUSIC-FP et CSSM, dans un cas sans bruit.	70
5.3	Valeurs propres de la matrice de covariance réduite en fonction de l'erreur sur la puissance enlevée pour différentes valeurs de SNR.	72
5.4	Moyenne de l'erreur de l'estimation de la DOA de la source faible en fonction du SNR de 200 essais de Monte-Carlo dans un cas sans corrélation entre les sources.	74

5.5	Écart-type de l'estimation de la DOA de la source faible en fonction du SNR de 200 essais de Monte-Carlo dans un cas sans corrélation entre les sources.	75
5.6	Probabilité de fausse alarme de la source faible en fonction du SNR obtenue avec 200 essais de Monte-Carlo pour des sources non corrélées.	76
5.7	Moyenne de l'erreur de l'estimation de la DOA de l'écho en fonction du SNR calculée à partir de 200 essais de Monte-Carlo dans un cas où les sources sont complètement corrélées.	77
5.8	Écart-type de l'estimation de DOA pour la source faible en fonction du ENR obtenue à l'aide de 200 essais de Monte-Carlo dans un cas où les sources sont complètement corrélées.	79
5.9	Estimation de la probabilité de fausse alarme en fonction du ENR calculée à partir de 200 essais Monte-Carlo pour des sources corrélées.	80
5.10	Probabilité de fausse alarme en fonction du ENR pour un signal composé d'une source à large bande et de deux échos.	81
5.11	Moyenne de l'erreur des bonnes détections de 200 essais de Monte-Carlo en fonction du ENR pour un signal composé d'une source à large bande et de deux échos.	81
5.12	Écart-type des bonnes détections de 200 essais de Monte-Carlo en fonction du ENR pour un signal composé d'une source à large bande et de deux échos.	82
6.1	Schéma du parcours du signal provenant d'un microphone.	87
6.2	Schéma du montage expérimental.	89
6.3	Résultat expérimental du test de puissance pour deux sources corrélées dont le rapport de puissance est de -10 dB.	90
6.4	Comparaison des résultats expérimentaux de MUSIC avec FBSS et avec SEM pour deux mesures différentes, chacune avec une seule source.	92
6.5	Exemple des signaux reçus par les capteurs 1 et 2 en fonction du temps lorsqu'une seule source est présente dans le signal.	93
6.6	Comparaison des résultats expérimentaux de MUSIC avec FBSS et SEM pour différents rapports de puissance entre deux sources corrélées.	94
6.7	Variation du pseudospectre de MUSIC-FP seul pour différentes valeurs de fréquences de référence.	96
6.8	Résultats expérimentaux de MUSIC avec et sans SEM pour deux mesures différentes, chacune avec une seule source à large bande.	98
6.9	Résultats expérimentaux de MUSIC-FP avec FBSS et avec SEM pour différents rapports de puissance entre les sources, dans le cas où les signaux à large bande sortant des haut-parleurs sont différents	99
6.10	Résultats expérimentaux de MUSIC-FP avec FBSS et avec SEM pour différents rapports de puissance entre deux sources à large bande.	101

Liste des symboles

Symbole	Définition
$\mathbf{a}(\theta_m)$	Vecteur directeur de la source m
\mathbf{A}	Matrice directrice des sources
$\mathbf{A}(\omega_j, \theta_m)$	Matrice directrice de la source m à la fréquence ω_j
\mathbf{A}_0	Matrice directrice à la fréquence de référence
α	Angles de focalisation pour CSSM
α_b	Facteur d'ajustement du SNR.
BW	Largeur de faisceau du réseau de capteurs
β	Constante de propagation de l'onde
d	Distance inter éléments d'un réseau de capteurs
$d_s(\theta)$	Distance de MUSIC pour la position θ
$\hat{d}(\theta, \omega_j)$	Distance de MUSIC pour l'angle θ à la fréquence ω_j
D	Matrice directrice des sous-réseaux pour le lissage spatial
f_0	Fréquence de référence de CSSM et MUSIC-FP
f_{0max}	Fréquence maximale du réseau
f_{em}	Fréquence d'échantillonnage du signal modulant
f_p	Fréquence de la porteuse
f_s	Fréquence d'échantillonnage des signaux générés
γ	Rapport de fréquences pour la focalisation des phases
G	Nombre de groupes
\mathbf{I}	Matrice identité
j	Nombre imaginaire
k	Indice de l'échantillon de temps
K	Nombre d'épreuves
λ_i	i^{eme} valeur propre de R_{xx}
Λ_{yy}	Matrice diagonale des valeurs propres de \mathbf{R}_{yy}
Λ	Matrice diagonale contenant les valeurs propres de \mathbf{R}_{xx}

m	Indice de la source
M	Nombre de sources
\tilde{M}	Nombre de sources estimées par AICe
M_c	Nombre de sources corrélées dans le signal
M_u	Nombre de sources non corrélées dans le signal
μ	Coefficient de corrélation entre les deux sources
n	Indice du capteur
$\mathbf{n}(t_k)$	Vecteur du bruit additif capté au temps t_k
n_{iter}	Nombre d'itérations pour CSSM
N	Nombre de capteurs dans le réseau
\mathbf{N}	Matrice du bruit additif
$\mathbf{N}_n(\omega_j)$	Matrice de la transformée de Fourier du bruit additif du capteur n à la fréquence ω_j
\mathbf{N}_0	Matrice du bruit additif focalisée à la fréquence de référence
N_r	Nombre d'éléments dans chaque sous-réseaux pour le lissage spatial
ν_n	Retard caractéristique du réseau
P	Nombre de série de données utilisées pour estimer les matrices de covariance à chaque fréquence
\mathbf{P}	Spectre spatial de MVDR
$\hat{\mathbf{P}}(\theta)$	Estimée de $\mathbf{P}(\theta)$
Ψ	Nombre d'angles de focalisation
ϕ_m	Déphasage entre deux capteurs pour la source m
ϕ_0	Déphasage entre les capteurs à la fréquence de référence
Q	Nombre de données dans chaque série utilisées pour estimer les matrices de covariance à chaque fréquence
r	Indice du sous-réseau du lissage spatial
R	Nombre de sous-réseaux pour le lissage spatial
\mathbf{R}_m	Matrice de covariance de la source m reçue par les capteurs
$\hat{\mathbf{R}}_m$	Estimée de la matrice de covariance de la source m reçue par les capteurs
\mathbf{R}_{nn}	Matrice de covariance du bruit
\mathbf{R}_{redm}	Matrice de covariance réduite d'ordre m
$\hat{\mathbf{R}}_{redm}$	Estimée de la matrice de covariance réduite d'ordre m
\mathbf{R}_{ss}	Matrice de covariance des sources
$\bar{\mathbf{R}}_{ss}$	Matrice de covariance des sources lissée
\mathbf{R}_{ss0}	Matrice de covariance des sources focalisée à f_0
$\mathbf{R}_{ss}(\omega_j)$	Matrice de covariance de la transformée de Fourier des sources à ω_j

\mathbf{R}_{xx}	Matrice de covariance des signaux reçus
\mathbf{R}_{xx-}	Matrice de covariance des signaux reçus sans bruit
$\hat{\mathbf{R}}_{xx}$	Estimée de la matrice de covariance des signaux reçus
$\bar{\mathbf{R}}_{xx}$	Matrice de covariance des signaux reçus lissée
\mathbf{R}_{xx_0}	Matrice de covariance des signaux reçus focalisée à f_0
\mathbf{R}_{xx_r}	Matrice de covariance reçue par le sous-réseau r
$\mathbf{R}_{xx_r}^f$	Matrice de covariance reçue par le sous-réseau r dans la direction avant
$\mathbf{R}_{xx_r}^b$	Matrice de covariance reçue par le sous-réseau r dans la direction arrière
$\mathbf{R}(\omega_j)$	Matrice de covariance reçue à la fréquence ω_j
\mathbf{R}_{yy}	Matrice de covariance théorique pour la génération de signaux corrélés
$s_m(t_k)$	Enveloppe complexe envoyée par la source m au temps t_k
\mathbf{S}	Matrice de l'enveloppe complexe des sources
$\mathbf{S}_m(\omega_j)$	Matrice de la transformée de Fourier de l'enveloppe complexe de la source m à la fréquence ω_j
$\hat{\mathbf{S}}(\theta, \omega_j)$	Estimée du spectre spatial à l'angle θ à la fréquence ω_j
snr	Rapport signal-à-bruit
σ_m	Puissance de la source m
σ_{pb}	Moyenne de la diagonale de la matrice de covariance de bruit, donc une indication de sa puissance
t_k	k^{ieme} valeur de temps
$\mathbf{T}(\omega_j)$	Matrice de transformation de CSSM
\mathbf{T}_y	Matrice de transformation pour générer des données corrélées
τ	Décalage en temps entre d'une source entre les capteurs
\mathbf{V}	Matrice des vecteurs propres colonnes de \mathbf{R}_{xx}
\mathbf{V}_n	Matrice du sous-espace bruit de \mathbf{R}_{xx}
\mathbf{V}_s	Matrice du sous-espace source de \mathbf{R}_{xx}
\mathbf{V}_{yy}	Matrice des vecteurs propres colonnes de \mathbf{R}_{yy}
ω_{0max}	Fréquence angulaire maximale du réseau
ω_L	Fréquence limite basse de la bande utile du signal
ω_H	Fréquence limite haute de la bande utile du signal
$\hat{\mathbf{X}}$	Matrice de la mesure faite avec les capteurs
\mathbf{X}_0	Matrice des signaux reçus focalisée à f_0
$\mathbf{X}_n(\omega_j)$	Matrice de la transformée de Fourier des signaux reçus par le capteur n à la fréquence ω_j
θ_m	Angle d'arrivée de la source m , mesuré par rapport à l'axe du réseau
v	Vitesse de propagation de l'onde

\mathbf{w}	Coefficients du filtre spatial
$x_n(t_k)$	Signal reçu par le capteur n au temps t_k
$\mathbf{x}(t_k)$	Vecteur des signaux reçus à t_k
\mathbf{X}	Matrice des signaux reçus
ξ	Indice de la fréquence
Ξ	Nombre de bandes de fréquences utilisés pour estimer la matrice de covariance
\mathbf{Z}	Matrice de nombres aléatoires utilisée pour générer des signaux corrélés

Liste des acronymes

Abréviation	Définition
AICe	Akaike Information Criterion Estimate
CSSM	Coherent Signal Subspace Method
DOA	Direction of Arrival
EDR	Echo-to-Direct Path Ratio
ENR	Echo-to-Noise Ratio
FBSS	Forward-Backward Spatial Smoothing
FFT	Fast Fourier Transform
ISSM	Incoherent Signal Subspace Method
MUSIC	Multiple Signal Classification
MUSIC-FP	Multiple Signal Classification à focalisation de phase
MUSIC-S	Multiple Signal Classification Summation
MVDR	Minimum Variance Distortionless Response
MVDR-S	Minimum Variance Distortionless Response Summation
SEM	Source Elimination Method
SNR	Signal-to-Noise Ratio

À Steeve

This semester, I teach the same
thing I teach every semester.
The Mathematics of Quantum
Neutrino Fields. I made up the
title so that no student would
dare take it.

Professor Hubert Farnsworth,
Futurama, épisode *Mars*
University, 1999

Remerciements

Je tiens tout d'abord à remercier ma merveilleuse compagne Evelyne qui m'a écouté et encouragé pendant les moments les plus difficiles de ma maîtrise, et m'a félicité lors des meilleurs moments. Je tiens également à remercier ma famille et spécialement mes parents qui m'ont aidé tout au long de mon parcours scolaire et m'ont permis de me rendre là où je suis.

Je veux aussi remercier tout ce que j'ai côtoyé dans le LRTS. Je parle ici de mes collègues et amis de bureau Ahmed Kebe, Carl Poirier et Alexia Blanchard-Lapierre qui ont fait passer ces deux années beaucoup plus rapidement. Je remercie Mathieu Gallichand pour avoir fabriqué mon montage expérimental et d'avoir pris le temps de me montrer comment le faire fonctionner. Un merci aux professeurs du LRTS qui étaient présents aux pauses et aux dîners, vous m'avez appris beaucoup avec toutes vos discussions.

Finalement, je tiens à remercier mon directeur de recherche, Dominic Grenier, de m'avoir donné conseils et support à de nombreuses reprises pendant toute ma maîtrise. Nos discussions m'ont beaucoup éclairé et m'ont amené à avoir une approche rigoureuse par rapport à mon projet. Nos rapports harmonieux ont fait que j'ai eu beaucoup de plaisir pendant mes études.

Avant-propos

Ce mémoire contient, à l'annexe A, une première version de l'article qui a été rédigé lors des travaux faits pendant cette maîtrise. Cet article a été soumis au Journal canadien de génie électrique et informatique et est présentement en cours d'études. Il n'est donc pas dans son état final. L'étudiant est l'auteur principal de l'article, ayant créé l'algorithme qui y est présenté, produit les résultats de simulations et expérimentaux en plus d'en avoir fait l'analyse. Le deuxième auteur, Dominic Grenier, est le directeur de recherche de l'étudiant et a eu un rôle de soutien tout au long du projet et a permis l'avancement du projet à de nombreuses reprises. Cet article porte sur la même matière que le chapitre 3 de ce mémoire.

Chapitre 1

Introduction

La localisation de sources sonores est un sujet étudié depuis de nombreuses années. Plusieurs méthodes pour faire de telles estimations utilisent un réseau de microphones pour déterminer l'origine des signaux reçus. Ces algorithmes fonctionnent généralement bien lorsque les sources sont de même puissance et lorsqu'elles proviennent de sources complètement indépendantes, ce qui pose problème lorsque l'objectif est de déterminer la direction d'arrivée (DOA) de sources sonores et de leurs échos, en particulier lorsque ces derniers sont de faibles puissances. Très peu d'articles scientifiques portent sur ce sujet puisque les échos sont généralement une nuisance pour la détection du signal en trajet direct et beaucoup d'articles portent sur la suppression des échos pour diminuer le bruit. Pourtant, la localisation d'échos sonores a plusieurs applications pratiques telles que l'analyse de l'acoustique d'une salle de spectacle ou bien dans le domaine de la bathymétrie [1]. Une autre application pratique est l'écho-détection passive cohérente, qui utilise les signaux sonores produits par diverses sources, telles que des voitures, pour déterminer la position d'obstacles.

Les principaux problèmes de la détection d'échos acoustique sont les suivants :

1. Les sources sont toutes corrélées, ce qui limite la capacité des algorithmes à déterminer la DOA des sources.
2. Le coefficient de réflexion des matériaux est petit, signifiant que le rapport écho à trajet direct (EDR) est petit, ce qui a pour effet de noyer l'écho dans le signal en trajet direct et dans le bruit.
3. Les réflexions multiples dans une salle causent un bruit de fond corrélé supplémentaire au bruit déjà présent.
4. La largeur de bande des sources sonores est assez grande pour que les méthodes

de DOA à bande étroite ne fonctionnent pas, ce qui complexifie les algorithmes utilisables et augmente par le fait même la quantité de calculs nécessaires pour trouver leur DOA.

L'objectif de ce mémoire est de créer un algorithme complet qui permet de contourner ces problèmes et ainsi d'obtenir de meilleurs résultats que les méthodes les plus présentes dans la littérature.

1.1 Plan du mémoire

Ce mémoire commence en présentant, au chapitre 2, de nombreuses méthodes qui sont liées à l'estimation de la direction d'arrivée de sources sonores. En particulier, on développe dans ce chapitre des modèles pour les signaux reçus par un réseau de microphones pour le cas où les sources sont à bande étroite et pour un cas où les sources sont à large bande. Par la suite, on présente différentes méthodes pour estimer les DOA de sources sonores à bande étroite et à large bande à partir des modèles bâtis. Finalement, on voit des algorithmes qui servent à contourner le problème de la corrélation entre les sources et un algorithme qui sert à déterminer le nombre de sources à partir des signaux reçus.

Au chapitre 3, on développe un algorithme appelé méthode par élimination des sources, dont l'objectif est d'améliorer la précision des méthodes à bande étroite lorsque le signal est composé de sources très puissantes comparativement à leurs échos. On commence alors le chapitre en expliquant l'idée à la base de l'algorithme, qui est d'éliminer la contribution des sources une par une pour ensuite estimer la contribution du bruit dans le signal capté. Par la suite, on développe deux méthodes permettant de déterminer la DOA et la puissance des sources une par une, les deux paramètres essentiels pour calculer la contribution de chacune des sources dans le signal capté. Finalement, on effectue des simulations Monte-Carlo dans un cas sans et un cas avec corrélation pour vérifier la robustesse de la méthode comparativement à Multiple Signal Classification (MUSIC), qui est la méthode de référence dans le domaine.

Dans le chapitre 4, on développe un nouvel algorithme d'estimation de direction d'arrivée à large bande, que l'on appelle MUSIC à focalisation de phase, dont le principe de base est d'ajuster la phase de chacune des fréquences pour se rapporter à un cas à bande étroite. Pour ce faire, on commence par déterminer les paramètres optimaux de la méthode pour ensuite vérifier ses limites en faisant des essais Monte-Carlo pour des signaux sans et avec corrélation.

Au chapitre 5, on combine les algorithmes des chapitres 3 et 4 dans le but de finaliser les objectifs du mémoire. On commence alors par vérifier que la méthode par élimination des sources fonctionne toujours pour le cas à large bande pour ensuite effectuer des essais Monte-Carlo pour plusieurs cas différents, ce qui permet de vérifier les limites de fonctionnement de la fusion des algorithmes d'élimination des sources et à focalisation des phases.

Finalement, au chapitre 6, on vérifie les algorithmes développés dans le mémoire en effectuant des mesures expérimentales. On commence alors par présenter le montage expérimental, pour ensuite vérifier le comportement des algorithmes sur des signaux mesurés expérimentalement.

Chapitre 2

Estimation de directions d'arrivée

Depuis l'apparition de la méthode par *Minimum Variance Distortionless Response* (MVDR) en 1969, le domaine de l'estimation de directions d'arrivée s'est énormément développé. Bien qu'au départ ce fut pour des signaux à bande étroite telles que MUSIC, des méthodes pour des signaux à large bande ont aussi été développées par la suite. Dans ce chapitre, on commencera par bâtir un modèle du signal dans le cas à bande étroite pour ensuite généraliser en construisant un nouveau modèle pour le cas à large bande. Ensuite, on va présenter quelques algorithmes très présents dans la littérature pour faire de l'estimation des DOA pour le cas à bande étroite et le cas à large bande. On verra finalement comment réduire l'effet néfaste d'une corrélation entre les sources et l'on présentera un algorithme qui permet d'estimer le nombre de sources présentes dans le signal.

2.1 Modèle du signal

2.1.1 Bande étroite

Avant de se lancer dans la présentation d'algorithmes effectuant l'estimation de DOA, il est nécessaire de se bâtir un modèle théorique du signal. Pour commencer, on va considérer un réseau composé de plusieurs capteurs qui, dans le cas de sources sonores, sont des microphones. On peut utiliser de nombreuses configurations de capteurs, la plus simple étant un réseau linéaire uniforme pour l'estimation de l'angle dans un plan, mais il est aussi possible d'utiliser un réseau planaire [2] ou sphérique [3] si l'on veut faire une estimation en deux dimensions. Dans le cas de ce mémoire, on se concentrera sur un cas à une seule dimension avec un réseau linéaire uniforme avec des capteurs séparés

d'une distance d , tel qu'illustré à la figure 2.1. Un tel réseau permet de déterminer la direction d'arrivée sur une plage de 180° , tandis qu'un réseau distribué sur un cercle permet de le faire sur 360° .

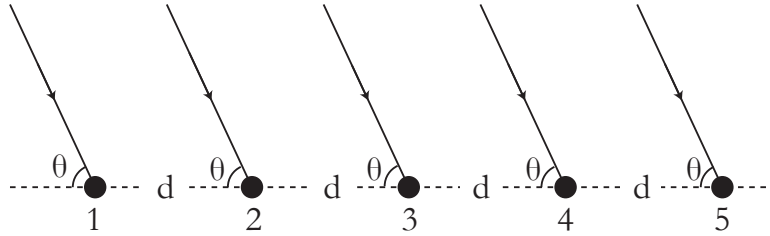


FIGURE 2.1 – Réseau de capteurs à espacement linéaire. Les points représentent des microphones.

On peut alors représenter le signal reçu par les capteurs au temps t_k comme étant :

$$\mathbf{x}_k = [x_1(t_k) \ x_2(t_k) \ \dots \ x_N(t_k)]^T \quad (2.1)$$

où $x_n(t_k)$ correspond au signal reçu par le capteur n au temps t_k , et T représente l'opération de transposition sur le vecteur. Si l'on effectue K échantillons à une fréquence f_s , on peut représenter les signaux enregistrés dans une matrice, que l'on appelle matrice des signaux reçus \mathbf{X} :

$$\underset{(N \times K)}{\mathbf{X}} = [\mathbf{x}_1 \ \mathbf{x}_2 \ \dots \ \mathbf{x}_K] \quad (2.2)$$

Maintenant, considérons que chaque source produit une enveloppe complexe $s_m(t_k)$ au temps t_k . Considérons aussi un bruit gaussien additif, qui est causé par des parasites captés et par du bruit électronique sur les capteurs. On peut modéliser une écoute comme étant une somme de chaque source avec un déphasage différent pour chacun des capteurs :

$$\mathbf{x}_k = \sum_{m=1}^M s_m(t_k) \mathbf{a}(\theta_m) + \mathbf{n}(t_k) \quad (2.3)$$

Dans cette équation, les paramètres sont donnés par les équations suivantes :

$$\underset{(Nx1)}{\mathbf{a}(\theta_m)} = [1 \ e^{-j\phi_m} \ e^{-j2\phi_m} \ \dots \ e^{-j(N-1)\phi_m}]^T \quad (2.4)$$

$$\phi_m = \beta d \cos(\theta_m) = \frac{2\pi f d \cos(\theta_m)}{v} \quad (2.5)$$

$$\underset{(M \times 1)}{\mathbf{s}_k} = [s_1(t_k) \ s_2(t_k) \ \dots \ s_M(t_k)]^T \quad (2.6)$$

$$\underset{(N \times 1)}{\mathbf{n}_k} = [n_1(t_k) \ s_2(t_k) \ \dots \ s_M(t_k)]^T \quad (2.7)$$

où ϕ_m est le déphasage de la modulation de la source m causé par la distance que l'onde doit parcourir entre les capteurs, v est la vitesse de l'onde dans son milieu, $\mathbf{a}(\theta_m)$ est le vecteur directeur pour la source m , \mathbf{s}_k est l'enveloppe complexe de la source m à l'instant t_k et, finalement, \mathbf{n}_k est le bruit additif qui s'ajoute à chacun des capteurs. Tel que mentionné précédemment, on utilise un réseau linéaire uniforme, ce qui a pour effet que ϕ_m est constant. Si l'on veut adapter le cas à un réseau circulaire, il suffit de modifier la valeur de ϕ_m pour chaque capteur.

On peut aussi simplifier l'écriture de l'équation (2.3) en utilisant des matrices. Ainsi, si l'on pose

$$\underset{(N \times M)}{\mathbf{A}} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_M] \quad (2.8)$$

$$\underset{(M \times K)}{\mathbf{S}} = [\mathbf{s}_1 \ \mathbf{s}_2 \ \dots \ \mathbf{s}_K] \quad (2.9)$$

$$\underset{(N \times K)}{\mathbf{N}} = [\mathbf{n}_1 \ \mathbf{n}_2 \ \dots \ \mathbf{n}_K] \quad (2.10)$$

L'équation (2.3) devient alors :

$$\mathbf{X} = \mathbf{A}\mathbf{S} + \mathbf{N} \quad (2.11)$$

Le principe de base des méthodes d'estimation de direction d'arrivée est de déterminer l'espérance mathématique entre les signaux reçus sous la forme d'une matrice de covariance, qui est définie comme :

$$\mathbf{R}_{xx} = E \left\{ \mathbf{x}_k \mathbf{x}_k^\dagger \right\} \quad (2.12)$$

$$= \underbrace{\mathbf{A} E \left\{ \mathbf{s}_k \mathbf{s}_k^\dagger \right\} \mathbf{A}^\dagger}_{\mathbf{R}_{ss}} + \underbrace{E \left\{ \mathbf{n}_k \mathbf{n}_k^\dagger \right\}}_{\mathbf{R}_{nn}} \quad (2.13)$$

$$= \underbrace{\mathbf{A} \mathbf{R}_{ss} \mathbf{A}^\dagger}_{\mathbf{R}_{xx-}} + \mathbf{R}_{nn} \quad (2.14)$$

$$(2.15)$$

Le terme \dagger correspond à la combinaison transposition-conjugaison de la matrice en question. La matrice \mathbf{R}_{xx} correspond à la matrice des signaux reçus incluant le bruit et c'est elle qui est estimée par les K épreuves, tandis que la matrice \mathbf{R}_{xx-} est la matrice de covariance reçue par les capteurs sans tenir compte du bruit additif \mathbf{R}_{nn} . Finalement, \mathbf{R}_{ss} est la matrice de covariance des sources. Cette dernière compte trois particularités :

1. Les éléments de sa diagonale principale correspondent à la variance du signal de la source m , qui elle est liée à la puissance de cette source.
2. Les éléments extra diagonaux correspondent à la covariance entre deux signaux. On sait d'ailleurs que la covariance entre deux signaux p et q est donnée par $\mu\sigma_{s_p}\sigma_{s_q}$, où μ est le coefficient de corrélation entre les sources et est toujours situé entre 0 et 1 alors que σ_{s_p} est la variance du signal p . Ainsi, si les sources ne sont pas corrélées, les éléments extra diagonaux sont nuls.
3. Comme toute matrice de covariance, elle est à symétrie hermitienne.

En pratique, alors que l'on a mesuré une série de K épreuves, on peut approximer \mathbf{R}_{xx} par la formule suivante :

$$\mathbf{R}_{xx} \simeq \hat{\mathbf{R}}_{xx} = \frac{1}{K} \mathbf{X}\mathbf{X}^\dagger \quad (2.16)$$

2.1.2 Large bande

Toutes les équations présentées à la section précédente ne sont valides que lorsque le signal est à bande étroite. On a plutôt considéré que le retard de propagation correspondait à un déphasage, ce qui n'est vrai qu'à une fréquence donnée. Pour adapter notre modèle à des signaux à large bande, on va plutôt décrire le signal reçu par un capteur comme étant un décalage temporel de signaux provenant de différentes sources :

$$\mathbf{x}_n(t) = \sum_{m=1}^M \mathbf{s}_m(t - \nu_n \cos(\theta_m)) + \mathbf{n}_n(t) \quad (2.17)$$

où $\mathbf{s}_m(t)$ est le signal de la source m , $\mathbf{n}_n(t)$ est le bruit du capteur n et $\nu_n = (n-1)d/v$. La variable v est la vitesse de propagation de l'onde dans son milieu, l'angle θ_m est la direction d'arrivée à estimer. On peut présumer que le bruit n'est pas corrélé avec le signal des sources, autant temporellement que spatialement. La transformée de Fourier discrète du capteur n est :

$$\mathbf{X}_n(\omega_\xi) = \sum_{m=1}^M \mathbf{S}_m(\omega_\xi) \exp(-j\omega\nu_n \sin \theta_m) + \mathbf{N}_n(\omega_\xi) \quad (2.18)$$

En forme vectorielle, on peut réécrire cette équation sous la forme suivante :

$$\mathbf{X}(\omega_\xi) = \mathbf{A}(\omega_\xi, \theta_m)\mathbf{S}(\omega_\xi) + \mathbf{N}(\omega_\xi), \quad \xi = 1..\Xi \quad (2.19)$$

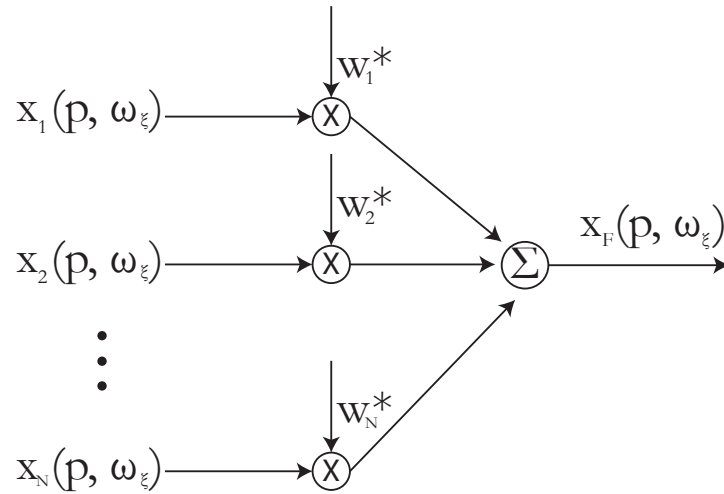


FIGURE 2.2 – Illustration du filtrage spatial. Chaque sortie du réseau est multipliée par un coefficient du filtre.

Où $\omega_L \leq \omega_\xi \leq \omega_H$ et $\mathbf{A}(\omega_\xi, \theta)$ est la matrice directrice dont chacune des colonnes correspond au vecteur directeur de l'une des sources.

On considère que le premier capteur sert de référence. On peut alors faire la matrice de covariance pour une fréquence :

$$\mathbf{R}(\omega_\xi) = E [\mathbf{X}(\omega_\xi) \mathbf{X}^\dagger(\omega_\xi)] = \mathbf{A}(\omega_\xi, \theta_m) \mathbf{R}_{ss}(\omega_\xi) \mathbf{A}(\omega_\xi, \theta_m)^\dagger \quad (2.20)$$

2.2 Méthodes à bande étroite

2.2.1 MVDR

Une première méthode permettant de déterminer les directions d'arrivée est appelée *minimum variance distortionless response* (MVDR) fut développée par Capon [4] en 1969 et on lui donne parfois le nom de son auteur. Cette méthode se base sur un filtrage spatial dans le but de filtrer le signal provenant de toutes les directions sauf une, située précisément à l'angle θ . Pour faire un filtrage spatial, il est simplement nécessaire de multiplier la sortie de chacun des capteurs par un coefficient de filtre, tel qu'illustré à la figure 2.2.

Dans le domaine des fréquences, cela revient à :

$$\mathbf{x}_f(p, \omega_k) = \mathbf{w}^\dagger \mathbf{x}(p, \omega_k) \quad (2.21)$$

Combinant l'équation (2.3) et l'équation (2.21), on obtient l'équation suivante :

$$\mathbf{x}_F = \mathbf{w}^H (\mathbf{A}\mathbf{x} + \mathbf{N}) = \mathbf{w}^H \mathbf{a}(\theta_1) s_1 + \mathbf{w}^H \mathbf{a}(\theta_2) s_2 + \dots + \mathbf{w}^H \mathbf{a}(\theta_M) s_M + \mathbf{w}^H \mathbf{e} \quad (2.22)$$

Il devient alors évident que pour éviter d'atténuer les signaux provenant d'un angle donné et de causer de la distorsion, il est nécessaire que :

$$\mathbf{w}^\dagger \mathbf{a}(\theta = \theta_m) = 1 \quad (2.23)$$

De plus, puisque l'idée de base de cette méthode est d'éliminer le signal provenant de toutes les autres directions, il faut que :

$$\mathbf{w}^H \mathbf{a}(\theta = \theta_k \forall k \neq m) = 0 \quad (2.24)$$

En pratique, il est impossible de créer un filtre qui respecte toutes ces conditions puisque l'on a une infinité de contraintes. Toutefois, il est possible de créer un filtre qui permet de filtrer le signal de toutes les directions, sauf une direction désirée. Dans le cas de la méthode par MVDR, le but est de minimiser l'énergie du signal, que l'on note $E[|y_f|^2]$. Ce filtre ne doit pas causer de distorsion dans la direction θ à la fréquence ω . Ainsi, on obtient le problème de minimisation suivant :

$$\min E[|x_f|^2] \quad (2.25)$$

$$\text{sujet à : } \mathbf{w}^\dagger \mathbf{a}(\theta) = 1 \quad (2.26)$$

La solution de ce problème peut se démontrer avec des multiplicateurs de Lagrange. Cette solution est :

$$\mathbf{w} = \frac{\hat{\mathbf{R}}^{-1} \mathbf{a}(\theta)}{\mathbf{a}^\dagger(\theta) \hat{\mathbf{R}}^{-1} \mathbf{a}(\theta)} \quad (2.27)$$

Ainsi, on peut obtenir une estimation du spectre spatial par la relation suivante :

$$\mathbf{P}(\theta) = \frac{1}{\mathbf{a}(\theta)^\dagger \hat{\mathbf{R}}^{-1} \mathbf{a}(\theta)} \quad (2.28)$$

La position angulaire des pics de $\mathbf{P}(\theta)$ correspond à la position des sources. Cette méthode est vue comme une méthode à haute résolution puisqu'elle possède une meilleure résolution que la méthode par formation des faisceaux (Beamforming), qui se base aussi sur un filtrage spatial. Toutefois, cette méthode n'est pas aussi efficace que d'autres qui ont été développés par la suite. La liste des étapes de la méthode MVDR est présentée à l'algorithme 1.

Algorithme 1 : MVDR

- 1 Estimer la matrice de covariance des signaux reçus $\hat{\mathbf{R}} = \mathbf{X}\mathbf{X}^\dagger$
 - 2 **for** $\theta = 0^\circ$ **to** 180° **do**
 - 3 Former le vecteur d'analyse $\mathbf{a}(\theta) = [1 \ e^{-j\phi} \ e^{-j2\phi} \ \dots e^{-j(N-1)\phi}]^T$, avec $\phi = \beta d \cos(\theta)$
 - 4 Calculer $\mathbf{P}(\theta) = \frac{1}{\mathbf{a}(\theta)^\dagger \hat{\mathbf{R}}^{-1} \mathbf{a}(\theta)}$
 - 5 **end**
 - 6 Déterminer la position des M plus hauts pics
-

2.2.2 MUSIC

Bien que les méthodes utilisant un filtrage spatial permettent de déterminer la bonne direction d'arrivée d'une source, elles ont généralement une variance assez élevée et sont fortement affectées par le bruit. C'est ainsi qu'une nouvelle classe de méthodes d'estimation de direction d'arrivée fut développée indépendamment par G. Bienvenu [5] en 1979 et par R.O. Schmidt [6] en 1986. L'idée de base est de faire une projection orthogonale pour obtenir la direction d'arrivée. Pour commencer, on se rappelle que la matrice de covariance reçue par les capteurs est :

$$\mathbf{R}_{xx} = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^\dagger + \mathbf{R}_{nn} \quad (2.29)$$

En premier lieu, on va considérer un cas sans bruit et sans corrélation entre les sources, ce qui implique que les éléments extras diagonaux de \mathbf{R}_{ss} sont nuls. On peut alors établir un lien entre cette matrice et une décomposition en vecteurs propres :

$$\mathbf{R}_{xx-} = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^\dagger = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\dagger \quad (2.30)$$

Ainsi, dans ce cas, on sait que la matrice \mathbf{A} a un rang complet selon ses colonnes, tant que le nombre de capteurs est plus élevé que le nombre de sources et que les sources soient situées à des angles différents. Les vecteurs colonnes de \mathbf{A} forment alors un sous-espace vectoriel de dimension M que l'on appelle sous-espace sources. On sait qu'une matrice de covariance est à symétrie hermitienne, ce qui implique que ses vecteurs propres sont orthogonaux entre eux. L'équation (2.30) a pour conséquence que les M vecteurs propres correspondant aux M valeurs propres les plus élevées forment aussi un sous-espace vectoriel de dimension M , que l'on note \mathbf{V}_s . Ce sous-espace sera le même que celui de la matrice directrice \mathbf{A} . L'idée à la base de MUSIC est d'utiliser ce principe en créant une série de vecteurs d'analyse $\mathbf{a}(\theta)$ et de vérifier, pour chacun d'entre eux, s'il est contenu dans le sous-espace source. En pratique, on préfère utiliser le sous-espace complémentaire au sous-espace source, que l'on appelle sous-espace bruit et que l'on note \mathbf{V}_n pour réaliser un goniomètre. Le choix de l'utilisation du sous-espace bruit s'est faite parce que les pics obtenus dans un pseudospectre du sous-espace bruit sont beaucoup plus étroits que ceux du sous-espace sources, ce qui amène une meilleure résolution. On définit la distance de MUSIC par la relation suivante :

$$\mathbf{d}_s^2(\theta) = \frac{1}{N} \mathbf{a}(\theta)^\dagger \underbrace{\mathbf{V}_n \mathbf{V}_n^\dagger}_{\mathbf{P}_n} \mathbf{a}(\theta) \quad (2.31)$$

Cette distance correspond à la norme du vecteur projeté dans le sous-espace bruit soit $\|\mathbf{V}_n^\dagger \mathbf{a}(\theta)\|^2$. On utilise généralement l'inverse de cette distance pour trouver la position des sources puisque les zéros de la distance de MUSIC donnent les positions pour lesquels le sous-espace bruit est orthogonal au vecteur directeur de cette direction, ce qui implique qu'il y a une source à cet endroit.

Bien que la méthode soit développée dans un cas sans bruit, on peut noter qu'en cas de présence de bruit idéal, c'est-à-dire d'un bruit non corrélé, la matrice de covariance du bruit tendra vers $\sigma^2 \mathbf{I}$ et n'affectera pas les vecteurs propres du sous-espace source, le bruit affectera seulement les valeurs propres, ce qui n'a pas d'importance pour la méthode. Si le bruit est corrélé avec un faible rapport signal à bruit (SNR), la présence d'éléments extra diagonaux dans la matrice de covariance de bruit viendra ajouter un biais sur la position estimée des sources et ainsi être beaucoup plus problématique qu'un bruit non idéal. Dans le cas des signaux acoustiques, on est souvent en présence de bruit non idéal parce que dans une pièce fermée, si la source sonore émet en continu des ondes, le son va se réfléchir dans la salle à de multiples reprises pour finalement devenir un bruit de fond corrélé qui vient d'un peu partout dans la pièce. C'est alors dans ce cas que MUSIC régulier perd beaucoup en efficacité.

Algorithme 2 : MUSIC

- 1 Estimer la matrice de covariance des signaux reçus $\hat{\mathbf{R}} = \mathbf{X}\mathbf{X}^\dagger$
 - 2 Effectuer la décomposition en vecteurs propres de $\hat{\mathbf{R}}$
 - 3 Déterminer le sous-espace bruit \mathbf{P}_n à partir des vecteurs propres
 - 4 **for** $\theta = 0^\circ$ **to** 180° **do**
 - 5 Former le vecteur d'analyse $\mathbf{a}(\theta) = [1 \ e^{-j\phi} \ e^{-j2\phi} \ \dots \ e^{-j(N-1)\phi}]^T$ avec $\phi = \beta d \cos(\theta)$
 - 6 Calculer $\mathbf{d}^{-2}(\theta) = \frac{N}{\mathbf{a}(\theta)^\dagger \mathbf{P}_n \mathbf{a}(\theta)}$
 - 7 **end**
 - 8 Déterminer la position des M plus hauts pics
-

2.2.3 Méthode par analyse propre

Une méthode plus récente, développée par Colone [7] et Gill [8], que l'on appelle méthode par analyse propre, utilise les propriétés de la matrice de covariance des signaux reçus afin d'éliminer la contribution du signal en trajet direct. On a déjà vu, dans la section 2.2.2, que la décomposition en vecteurs propres de la matrice de covariance des sources \mathbf{R}_{xx} donne N vecteurs propres \mathbf{v}_n formant une base orthogonale d'un espace vectoriel de dimension N . La méthode de suppression des sources par analyse propre consiste à projeter le signal capté dans le sous-espace complémentaire à celui des sources directes. On l'exprime comme :

$$\mathbf{x}_c(t_k) = \mathbf{V}_{th}^\dagger \mathbf{x}(t_k) \quad (2.32)$$

$$\mathbf{V}_{th} = [\mathbf{v}_{M+1} \ \mathbf{v}_{M+2} \ \dots \ \mathbf{v}_N] \quad (2.33)$$

Dans l'article de Gill et coll. [8], les auteurs utilisent cette méthode pour faire de la détection radar de type passive cohérente pour une cible en mouvement. Ce type de détection radar utilise des signaux électromagnétiques générés pour d'autres fins que du radar, par exemple des signaux de télévision, afin de déterminer la position d'avions présents dans le ciel.

Ils peuvent alors supposer que les sources ne sont pas corrélées pour deux raisons :

1. L'effet Doppler est important.
2. Le décalage en temps entre la détection de la source directe et l'écho est grand.

On peut alors faire l'approximation que seul le premier vecteur propre contient complètement le signal direct. Ces simplifications sont réalistes pour la détection radar passive cohérente, mais elles ne le sont pas dans le cas de sources sonores puisqu'elles ne sont pas en mouvement et parce qu'elles ne sont pas nécessairement très distantes.

2.3 Méthodes à large bande

Les méthodes vues dans la section précédente ont été bâties pour des signaux à bande étroite, ce qui implique que le retard temporel entre les capteurs cause un déphasage constant. Dans le cas à large bande, le fait que l'expression de la constante de phase soit une fonction de la fréquence a pour effet que le retard de phase n'est constant entre les capteurs que si l'on considère chaque fréquence individuellement. Dans la littérature, on retrouve principalement deux types d'algorithmes dont le but est de considérer la largeur de bande du signal. Le premier, que l'on désigne comme incohérent, consiste à traiter individuellement chaque bande de fréquence pour faire une estimation de la direction d'arrivée, pour finalement les combiner et obtenir un résultat final. Le second, que l'on appelle cohérente, utilise différentes techniques pour se ramener à un cas à bande étroite, pour ensuite faire une seule estimation de la DOA. Le but de cette section est de présenter brièvement quelques méthodes très présentes dans la littérature pour déterminer les directions d'arrivée, dont certaines sont incohérentes et d'autres cohérentes.

2.3.1 Prétraitement

Avant de commencer la présentation des méthodes incohérentes, il est nécessaire de faire un pré traitement des données dans le but d'obtenir une matrice de covariance pour chaque bande de fréquence. En faisant une transformée de Fourier rapide (FFT) des signaux reçus par les capteurs, il n'est pas possible d'avoir une matrice de covariance pour chaque fréquence puisqu'on a une seule valeur pour chaque fréquence. Le principe du pré traitement est de diviser les échantillons mesurés en P séries de données, contenant Q données chacun. Il est alors possible de faire la FFT pour chacune des séries de données et de finalement calculer une matrice de covariance pour chaque fréquence utile du signal. Dans ce mémoire, on considère que la largeur de bande des sources est connue *à priori*. L'algorithme 3 résume toutes les étapes du pré traitement.

Algorithme 3 : Estimation de matrices de covariance pour plusieurs fréquences

- 1 Séparer les données en P séries de Q données chacune. Plus Q est grand, plus le nombre de composantes spectrales à la sortie de la FFT sera grand.
 - 2 Faire la FFT sur chacune des séries de données
 - 3 **for** $\omega_\xi = \omega_L$ **to** ω_H **do**
 - 4 | Évaluer $\mathbf{R}_{xx}(\omega_\xi)$
 - 5 **end**
-

2.3.2 Méthodes incohérentes

Les méthodes incohérentes sont les premières méthodes qui sont apparues dans la littérature, avec l'article de Su et coll. [9] en 1983. Le principe présenté dans cet article est d'utiliser un algorithme de DOA pour bande étroite avec la matrice de covariance de chaque bande de fréquence et de moyenner les résultats obtenus. Ces méthodes sont généralement moins précises que les méthodes cohérentes, mais sont simples à implanter et demandent peu de puissance de calcul. Il arrive d'ailleurs que des méthodes cohérentes demandent une pré estimation de la position des sources, alors il devient intéressant de s'attarder à ces méthodes puisqu'elles peuvent servir à faire cette première estimation.

MVDR-S

Cette première méthode est simplement une version sommée de la méthode par MVDR, que l'on a présenté à la section 2.2.1. La méthode est brièvement présentée à l'article [10] et dans le mémoire [11]. L'algorithme est le suivant :

Algorithme 4 : MVDR-S

- 1 Estimer les matrices de covariance avec l'algorithme 3.
 - 2 **for** $\omega_\xi = \omega_L$ **to** ω_H **do**
 - 3 **for** $\theta = 0^\circ$ **to** 180° **do**
 - 4 Former le vecteur d'analyse $\mathbf{a}(\theta, \omega_\xi) = [1 \ e^{-j\phi_\xi} \ e^{-j2\phi_\xi} \ \dots \ e^{-j(N-1)\phi_\xi}]^T$
 - 5 Calculer $\mathbf{S}(\theta, \omega_\xi) = \frac{1}{\mathbf{a}(\theta, \omega_\xi)^\dagger \widehat{\mathbf{R}}_{xx}(\omega_\xi)^{-1} \mathbf{a}(\theta, \omega_\xi)}$
 - 6 **end**
 - 7 **end**
 - 8 Calculer $\hat{\mathbf{P}}(\theta) = \sum_{\omega_\xi} \mathbf{S}(\theta, \omega_\xi)$
 - 9 Déterminer la position des M plus hauts pics
-

Cette méthode n'est pas la meilleure puisque sa variance est beaucoup trop élevée, surtout en présence de bruit. Toutefois, elle n'est pas biaisée et ne demande pas trop de calculs. Elle est donc parfaite pour les méthodes qui demandent une pré estimation de la direction d'arrivée des sources. Il existe d'autres méthodes pour combiner les résultats des MVDR à bande étroite, telles qu'une moyenne géométrique [12], ou bien l'utilisation de la méthode de Capon robuste [13], mais on ne s'attardera pas à ces méthodes dans ce mémoire.

MUSIC-S

De la même façon que l'on a adapté la méthode de Capon au cas à large bande, on peut évaluer MUSIC pour plusieurs bandes spectrales et sommer les résultats, ce qui donne la méthode MUSIC-S, présentée à l'algorithme 5.

Algorithme 5 : MUSIC-S

- 1 Estimer les matrices de covariance avec l'algorithme 3.
 - 2 **for** $\omega_\xi = \omega_L$ **to** ω_H **do**
 - 3 Effectuer la décomposition en vecteurs propres de $\mathbf{R}_{\mathbf{xx}}(\omega_\xi)$
 - 4 Déterminer le sous-espace bruit \mathbf{P}_n à partir des vecteurs propres
 - 5 **for** $\theta = 0^\circ$ **to** 180° **do**
 - 6 Former le vecteur d'analyse $\mathbf{a}(\theta, \omega_\xi) = [1 \ e^{-j\phi_\xi} \ e^{-j2\phi_\xi} \ \dots \ e^{-j(N-1)\phi_\xi}]^T$
 - 7 Calculer $\hat{\mathbf{d}}^{-2}(\theta, \omega_\xi) = \frac{N}{\mathbf{a}(\theta, \omega_\xi)^H \mathbf{P}_n \mathbf{a}(\theta, \omega_\xi)}$
 - 8 **end**
 - 9 **end**
 - 10 Calculer $\hat{\mathbf{d}}^{-2}(\theta) = \sum_{\omega} \hat{\mathbf{d}}^{-2}(\theta, \omega_\xi)$
 - 11 Déterminer la position des M plus hauts pics
-

2.3.3 Méthodes cohérentes

Coherent signal subspace method (CSSM)

Une première méthode cohérente a été développée en 1988 par Wang [14] et est devenue la méthode de référence pour le traitement d'antenne pour des sources à large bande. La méthode se base principalement sur le lemme suivant :

Lemme

Si la matrice $\mathbf{A}(\omega_\xi)$, $\xi = 1, 2, \dots, \Xi$ à un rang d , il existe des matrices de transformation $\mathbf{T}(\omega_\xi)$, $\xi = 1, 2, \dots, \Xi$ ($M \times M$) telles que :

$$\mathbf{T}(\omega_\xi) \mathbf{A}(\omega_\xi) = \mathbf{A}(\omega_0) \quad (2.34)$$

Le choix de la matrice de transformation $\mathbf{T}(\omega_\xi)$ pour chaque fréquence n'est pas unique, et une très grande quantité ont été présentées par différents auteurs. Ces matrices demandent d'ailleurs une pré estimation des angles d'arrivée, ce qui implique alors d'utiliser une méthode telle que MVDR-S au préalable. Une classe de matrices très présentes dans la littérature est *Rotational Signal-Subspace Matrices* (RSSM) [15]. Le

principe derrière ce type de matrice est de minimiser l'équation suivante :

$$\min \|\mathbf{A}(\omega_0, \alpha) - \mathbf{T}(\omega_i)\mathbf{A}(\omega_i, \alpha)\|_F \quad (2.35)$$

où $\|\cdot\|_F$ signifie l'opérateur de la norme de Frobenius. De plus, il faut que :

$$\mathbf{T}^\dagger(\omega_\xi)\mathbf{T}(\omega_\xi) = \mathbf{I} \quad (2.36)$$

Dans ces équations, α est un vecteur formé des angles de focalisation, qui contiennent les angles pré estimés par MVDR-S et des angles auxiliaires, que l'on verra comment déterminer un peu plus bas. Pour bâtir la matrice, considérons premièrement deux matrices, $\mathbf{A}(\omega_0, \alpha)$ et $\mathbf{A}(\omega_i, \alpha)$. La multiplication matricielle des deux matrices $\mathbf{A}(\omega_0, \alpha)\mathbf{A}^\dagger(\omega_i, \alpha)$, résulte en une matrice $N \times N$, mais de rang $\Psi < N$, où Ψ est le nombre d'angles de focalisation. En faisant la décomposition en valeurs singulières de $\mathbf{A}(\omega_0, \alpha)\mathbf{A}^\dagger(\omega_i, \alpha)$, on obtient les matrices $\mathbf{U}(\omega_\xi)$ et $\mathbf{V}(\omega_\xi)$ dont les colonnes sont, respectivement, les vecteurs singuliers gauches et droits de la multiplication des deux matrices.

Une solution à l'équation 2.35 est de faire la multiplication suivante :

$$\mathbf{T}(\omega_\xi) = \mathbf{V}(\omega_\xi)\mathbf{U}^\dagger(\omega_\xi) \quad (2.37)$$

Pour déterminer les angles auxiliaires, il faut déterminer le nombre de groupes de sources G . Un groupe comprend toutes les sources dont les positions angulaires sont situées à moins d'une largeur de faisceau les unes des autres. On utilise la définition au sens de Rayleigh pour largeur de faisceau, qui est donnée par la relation suivante :

$$BW = \frac{v}{l\omega_0} \quad (2.38)$$

Dans cette équation, v est la vitesse de l'onde dans le milieu, soit 340,29 m/s pour le son, l est la longueur du réseau et ω_0 est la fréquence de focalisation qui correspond à la fréquence moyenne des fréquences que l'on a conservées pour l'analyse.

Pour que la méthode fonctionne bien, on utilise alors trois angles auxiliaires pour chacun des groupes. Le premier est la position estimée de la première source du groupe, que l'on note α_1 . Les deux autres angles sont situés à $\alpha_1 \pm 0.25BW$. On peut ensuite calculer les matrices de transformation et utiliser CSSM pour déterminer les angles d'arrivée de

toutes les sources. Si toutes les sources sont contenues dans un seul groupe, il devient possible d'utiliser la matrice de transformation suivante :

$$\mathbf{T}(\omega_\xi) = \text{diag}(a_n(\omega_0)/a_n(\omega_\xi)) \quad (2.39)$$

Pour améliorer la précision des angles de focalisation α , on utilise une boucle de rétroaction de n_{iter} itérations, permettant d'utiliser les résultats obtenus avec CSSM comme nouveaux angles de direction d'arrivée pour la pré estimation. Cette boucle permet de raffiner les résultats à chaque itération. Les étapes à suivre pour cette méthode sont présentées à l'algorithme 6.

Algorithme 6 : CSSM

```

1 Estimer les matrices de covariance avec l'algorithme 3.
2 Première estimation de la position des sources  $\hat{\theta}_m$  avec MVDR-S.
3 Calcul de  $BW = \frac{v}{l\omega_0}$ .
4 Estimer le nombre de groupe  $G$ . Un groupe est un ensemble de sources qui sont à une
   largeur de faisceau les unes des autres.
5 for  $i = 1$  to  $n_{iter}$  do
6   if  $G=1$  then
7     for  $\omega_\xi = \omega_L$  to  $\omega_H$  do
8       Calculer  $\mathbf{T}(\omega_\xi) = \text{diag}(a_n(\omega_0)/a_n(\omega_\xi))$ 
9       et  $\mathbf{A}(\omega_0) = \mathbf{A}(f_0) + \mathbf{T}(\omega_\xi)\mathbf{R}(\omega_\xi)$ 
10    end
11   else
12     Calculer les angles de projection  $\alpha$ 
13     for  $\omega_\xi = \omega_L$  to  $\omega_H$  do
14       Évaluer  $\mathbf{V}(\omega_\xi)$  et  $\mathbf{U}(\omega_\xi)$ 
15       Trouver  $\mathbf{T}(\omega_\xi) = \mathbf{V}(\omega_\xi)\mathbf{U}^\dagger(\omega_\xi)$ 
16       Calculer  $\mathbf{A}(\omega_0) = \mathbf{A}(f_0) + \mathbf{T}(\omega_\xi)\mathbf{R}(\omega_\xi)$ 
17     end
18   end
19   Utiliser MUSIC pour trouver les nouveaux  $\alpha$ 
20 end

```

2.4 Corrélation de sources

2.4.1 Lissage spatial

On a déjà mentionné dans la section sur MUSIC que les sources ne devaient pas être corrélées pour que l'on puisse estimer leur direction d'arrivée. Cela est dû au fait que

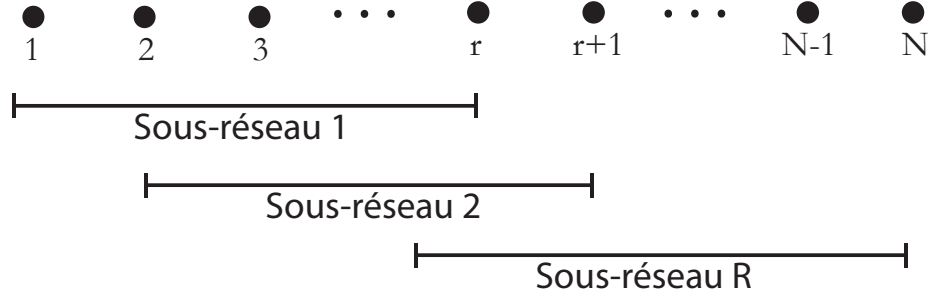


FIGURE 2.3 – Illustration montrant la décomposition du réseau initial en sous-réseaux pour la méthode de lissage spatial.

s'il y a une corrélation entre les sources, la matrice de covariance des sources \mathbf{R}_{ss} n'est plus diagonale. Ainsi, l'analogie faite à l'équation 2.30 n'est plus vraie dans ce cas. L'effet de la corrélation sera l'apparition d'une source imaginaire, positionnée dans le pseudospectre à l'angle moyen pondéré par les puissances des signaux corrélés, avec un pic très large par rapport à un cas sans corrélation. Par conséquent, le groupe des sources corrélées se comporte comme une seule source du point de vue de MUSIC. On définit alors la notion de groupe comme étant l'ensemble des sources corrélées entre elles, mais qui ne sont pas corrélées avec les autres sources.

Pour contrer l'effet de la corrélation des sources, la méthode la plus utilisée dans la littérature est celle du lissage spatial, qui fut développée par Shan et coll. [16]. La première étape de cette méthode est de séparer le réseau de microphones en R sous-réseaux, contenant chacun N_r capteurs (figure 2.3). Dans ce cas, on peut remarquer que la relation entre le nombre de capteurs, le nombre d'éléments dans le sous-réseau et le nombre de sous-réseaux est :

$$R = N - N_r + 1 \quad (2.40)$$

Le signal reçu par le r -ième sous réseau est donné par la formule suivante :

$$x_{r_k} = \mathbf{A}_1 \mathbf{D}^{(r-1)} x_k \quad (2.41)$$

$$\mathbf{D} = \text{diag}(e^{j\phi_1}, e^{j\phi_2}, e^{j\phi_3}, \dots, e^{j\phi_M}) \quad (2.42)$$

La matrice \mathbf{A}_1 est la matrice des vecteurs directeurs pour le premier sous-réseau, tandis que la matrice \mathbf{D} contient l'information de déphasage entre les éléments de chacune des sources. La matrice de covariance reçue pour le r -ième sous-réseau est donnée par :

$$\mathbf{R}_{xx_r} = \mathbf{A}_1 \mathbf{D}^{r-1} \mathbf{R}_{ss} (\mathbf{D}^{r-1})^\dagger \mathbf{A}_1^\dagger \quad (2.43)$$

On obtient alors la matrice de covariance lissée comme suit :

$$\hat{\mathbf{R}}_{xx-} = \frac{1}{R} \sum_{r=1}^R \mathbf{R}_{xx_r-} \quad (2.44)$$

Maintenant, si on définit $\hat{\mathbf{R}}_{ss}$ comme étant :

$$\hat{\mathbf{R}}_{ss} = \frac{1}{R} \sum_{r=1}^R \mathbf{D}^{r-1} \mathbf{R}_{ss} (\mathbf{D}^{r-1})^t \quad (2.45)$$

On remarque que l'on obtient finalement :

$$\hat{\mathbf{R}}_{xx} = \mathbf{A}_1 \hat{\mathbf{R}}_{ss} \mathbf{A}_1^t \quad (2.46)$$

La matrice de covariance lissée des sources est régulière si $R \geq M$.

Algorithme 7 : Lissage spatial

- 1 **for** $r=1$ **to** R **do**
 - 2 | Évaluer $\hat{\mathbf{R}}_{xx_r}$
 - 3 **end**
 - 4 Évaluer la matrice de covariance lissée $\hat{\mathbf{R}}_{xx-} = \frac{1}{R} \sum_{r=1}^R \mathbf{R}_{xx_r}$
 - 5 Utiliser MUSIC sur $\hat{\mathbf{R}}_{xx-}$
-

Pillai et Kwon [17] ont développé une amélioration de l'algorithme du lissage spatial, qui s'appelle Forward-Backward Spatial Smoothing Technique (FBSS). Pour cette méthode, on commence par évaluer une première matrice de covariance lissée $\hat{\mathbf{R}}_{xx-}^f$ dans la direction avant, c'est-à-dire en calculant exactement la même matrice de covariance lissée que pour le lissage spatial régulier. Par la suite, on calcule une seconde matrice de covariance lissée $\hat{\mathbf{R}}_{xx-}^b$, cette fois-ci qualifiée de direction arrière, que l'on calcule de la même façon que celle de direction avant, sauf que l'on intervertit la position de tous les éléments du réseau et que l'on fait une transposition-conjugaison de la matrice obtenue. L'avantage principal de cette méthode par rapport à lissage spatial direct est

la réduction du nombre d'éléments nécessaires pour être capable de faire une bonne détection de DOA. Si l'on considère que M_c est le nombre de sources corrélées et M_u est le nombre de sources non corrélées, le lissage spatial régulier a de besoin de $N = 2M_c + M_u$ capteurs pour être capable de détecter toutes les sources, tandis que ce nombre est de $N = 2M_c$ pour le lissage spatial forward-backward. L'algorithme 8 présente le lissage spatial forward-backward.

Algorithme 8 : Lissage spatial forward-backward

- 1 **for** $r=1$ **to** R **do**
 - 2 | Évaluer $\hat{\mathbf{R}}_{xx_r}^f$
 - 3 **end**
 - 4 Évaluer la matrice de covariance lissée de direction avant $\hat{\mathbf{R}}_{xx-}^f = \frac{1}{R} \sum_{r=1}^R \mathbf{R}_{xx_r}^f$
 - 5 Inverser l'ordre de tous les capteurs, ce qui revient à inverser l'ordre de tous les vecteurs colonnes de \mathbf{X} .
 - 6 **for** $r=1$ **to** R **do**
 - 7 | Évaluer $\hat{\mathbf{R}}_{xx_r}^b$
 - 8 **end**
 - 9 Évaluer la matrice de covariance lissée de direction avant $\hat{\mathbf{R}}_{xx-}^b = \frac{1}{R} (\sum_{r=1}^R \mathbf{R}_{xx_r}^b)^\dagger$
 - 10 Calculer $\hat{\mathbf{R}}_{xx-}^{fb} = \hat{\mathbf{R}}_{xx-}^f + \hat{\mathbf{R}}_{xx-}^b$
 - 11 Utiliser MUSIC sur $\hat{\mathbf{R}}_{xx-}^{fb}$
-

2.4.2 Méthode par Covariance Differencing and Iterative Spatial Smoothing

Une autre méthode, développée par Al-Ardi [18], appelée *Covariance Differencing and Iterative Spatial Smoothing*, permet de faire de la détection de DOA avec des sources corrélées. L'idée derrière cet algorithme est de faire un traitement en deux étapes, la première ayant pour but de déterminer la DOA de toutes les M_u sources non corrélées et la seconde de faire un processus itératif en considérant de plus de plus de capteurs dans le but de déterminer la DOA des M_c sources corrélées. L'algorithme 16 présente la liste des étapes pour la méthode.

La boucle présente dans la deuxième étape de l'algorithme sert à utiliser un réseau de plus en plus large. On a déjà vu que le lissage spatial direct utilisé seul a de besoin de $N = 2M_c + M_u$ pour être capable de détecter. Si le nombre de capteurs utilisé à une itération donnée est plus petit que ce nombre, on ne verra pas de pics bien définis et on sera capable de déterminer qu'il y a au moins une source de plus que le nombre assumé pendant l'itération, ce qui implique la nécessité de faire une autre itération. L'intérêt de

Algorithme 9 : Covariance Differencing and Iterative Spatial Smoothing

```
1 Estimer la matrice de covariance des signaux reçus  $\widehat{\mathbf{R}}^{-1} = \mathbf{X}\mathbf{X}^\dagger$ 
2 Calculer le pseudospectre de MUSIC de  $\widehat{\mathbf{R}}$ 
3 Estimer  $\widehat{\mathbf{R}}_u$ , la contribution des sources non corrélées à  $\widehat{\mathbf{R}}$ 
4 Calculer  $\widehat{\mathbf{R}}_c = \widehat{\mathbf{R}} - \widehat{\mathbf{R}}_u$ , la matrice de covariance des sources corrélées.
5 Poser  $N = 3$ ,  $R = 2$  et  $M_c = 0$ 
6 while  $M_c = 0$  do
7   | Faire du lissage spatial avec  $N$  capteurs et  $R$  sous-réseaux sur  $\widehat{\mathbf{R}}_c$ 
8   | Faire le pseudospectre MUSIC
9   | Évaluer  $M_p$ , le nombre de pics dans le pseudospectre.
10  | if  $M_p > 0$  then
11  |   |  $M_c = M_p$ 
12  |   | Déterminer la position des sources corrélées avec les pics du pseudospectre
13  | else
14  |   |  $M_c = 0$ 
15  | end
16 end
```

cette méthode est ainsi d'être plus efficace en calculs que la méthode par lissage spatial puisque l'on peut utiliser un nombre réduit de capteurs pour pouvoir utiliser la méthode.

Ray et al. [19] a repris cet algorithme et a décidé d'utiliser le lissage spatial Forward-Backward au lieu du lissage spatial régulier. L'avantage de cette méthode est qu'au lieu d'avoir de besoin de $N = 2M_c + M_u$ capteurs, on a seulement de besoin de $N = 2M_c$, ce qui demande moins de temps de calcul.

Les articles de Al-Ardi [18] et de Ray et coll. [19] présentent cet algorithme comme une alternative efficace de l'algorithme de lissage spatial, mais il manque quelques points importants dans ces deux articles :

1. L'auteur n'explique pas comment calculer \mathbf{R}_u , ce qui n'est pas aussi évident puisqu'on n'a pas seulement besoin de la DOA de chacune des sources, mais aussi de leur puissance.
2. Aucun résultat de simulations n'a été présenté dans les deux articles. On ne connaît donc pas les limites de l'algorithme et on ne sait même pas s'ils ont programmé l'algorithme pour vérifier son fonctionnement.

2.5 Détermination du nombre de sources

Pour toutes les méthodes présentées précédemment, on considérait que le nombre de sources est connu, ce qui est rarement le cas en pratique. Si l'on est en présence de SNR élevé, il est généralement facile de déterminer le nombre de sources puisque les valeurs propres correspondant au sous-espace bruit sont beaucoup plus petites que celles du sous-espace source. Toutefois, en présence de bruit corrélé, ce n'est plus le cas. Il existe un critère qui est très utilisé dans la littérature pour déterminer le nombre de source, le critère de Akaike [20], souvent noté AICe pour *Akaike Information Criterion Estimate*. Le nombre de sources correspond au nombre \widetilde{M} qui maximise la fonction de vraisemblance suivante :

$$AICe(\widetilde{M}) = K(N - \widetilde{M}) \log \left(\frac{a_0}{g_0} \right) \widetilde{M}(2N - \widetilde{M}) \quad (2.47)$$

où K est le nombre d'épreuves et

$$a_0 = \frac{1}{N - \widetilde{M}} \sum_{i=\widetilde{M}+1}^N \lambda_i \quad (2.48)$$

$$g_0 = \left(\prod_{i=\widetilde{M}+1}^N \lambda_i \right)^{\frac{1}{N-\widetilde{M}}} \quad (2.49)$$

et λ_i est la $i^{\text{ème}}$ valeur propre de la matrice de covariance.

2.6 Conclusion

Dans ce chapitre, on fait une revue des méthodes existantes dans le domaine de l'estimation de DOA pour des sources à bande étroite et des sources à large bande. Pour commencer, on a bâti un modèle du signal pour un cas à bande étroite et à large bande. Par la suite, on a utilisé ces modèles pour présenter différentes méthodes pour estimer la DOA de sources, qu'elles soient à bande étroite ou à large bande. Finalement, on a vu quelques algorithmes donc le but est de réduire l'impact de la corrélation entre les sources sur l'efficacité des algorithmes. Ces algorithmes sont très efficaces lorsque les sources sont de même puissance et lorsque le bruit est faible, mais perdent en efficacité lorsque les conditions idéales ne sont plus présentes. Les prochains chapitres auront pour but d'obtenir des algorithmes qui repoussent la limite de ceux présentés dans ce chapitre.

Chapitre 3

Méthode par élimination des sources

Les algorithmes d'estimation de directions d'arrivée pour un signal à bande étroite que l'on a présenté au chapitre précédent sont très efficaces lorsque les conditions sont idéales, c'est-à-dire pour des sources non corrélées entre elles et pour un bruit indépendant entre les capteurs. L'utilisation de lissage spatial permet d'améliorer les résultats, mais cette technique a ses limites, en particulier lorsqu'on est en présence d'un signal principal très puissant par rapport aux échos. Une solution à ce problème est d'éliminer la contribution du signal en trajet direct du signal total reçu par les capteurs. De nombreuses méthodes dont l'objectif est d'exploiter cette idée ont été développées au fil du temps, dont notamment l'utilisation de filtrage adaptatif par Howland et coll. [21], de rétroaction et soustraction par Saini et coll. [22], Kulpa et coll. [23] ou bien avec l'utilisation de l'algorithme CLEAN [24]. Ces méthodes demandent toutes de déterminer l'amplitude du signal, ce qui peut poser problème. La méthode d'analyse propre, tel que discuté à la section 2.2.3, n'est pas utilisable puisque les deux approximations faites pendant l'élaboration de l'algorithme ne peuvent pas être faites dans le cas où les sources sont stationnaires et relativement près les unes des autres. Dans ce chapitre, on commencera par présenter l'idée derrière un nouvel algorithme, ainsi que deux paramètres inconnus que l'on doit estimer pour pouvoir utiliser l'algorithme, soit la DOA et la puissance de chaque source. Par la suite, on va présenter des méthodes de génération de données et de bruit utilisées pendant tout ce mémoire. Finalement, on fera deux types de simulations, les premières ayant pour but de déterminer les meilleurs tests pour estimer les deux paramètres inconnus et les suivantes seront des simulations Monte Carlo qui ont pour but de tester les limites du nouvel algorithme par rapport à

des algorithmes bien connus.

3.1 Principe de base

Bien que la méthode développée par Gill et coll. [8] n'est pas utilisable pour des sources sonores et que la méthode par *Covariance Differencing and Iterative Spatial Smoothing* soit incomplète, il est possible d'utiliser les idées derrière les deux méthodes pour améliorer l'estimation de réverbérations sonores à faible puissance, c'est-à-dire de tenter d'éliminer la contribution du signal en trajet direct de la matrice de covariance des signaux reçus. On appelle donc cette méthode la méthode par élimination des sources, ou *Source Elimination Method* (SEM) en anglais.

Pour commencer le développement de la méthode, on rappelle l'équation (2.30), qui fait le lien entre la matrice de covariance des signaux reçus et sa décomposition en valeurs propres :

$$\mathbf{R}_{xx-} = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^\dagger = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\dagger \quad (3.1)$$

Il est aussi possible de réécrire cette équation comme étant la somme de la contribution de toutes les sources :

$$\mathbf{R}_{xx-} = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^\dagger = \mathbf{a}_1\sigma_1^2\mathbf{a}_1^\dagger + \mathbf{a}_2\sigma_2^2\mathbf{a}_2^\dagger + \dots + \mathbf{a}_M\sigma_M^2\mathbf{a}_M^\dagger \quad (3.2)$$

L'équation (3.2) n'est vraie que s'il n'y a pas de corrélation entre les sources, ce qui n'est pas le cas en présence d'échos. Toutefois, puisque l'on considère un cas où l'on est en présence de quelques sources directes non corrélées et d'échos beaucoup moins puissants, les termes de puissances σ_m^2 des échos sont tellement faibles que les éléments extra diagonaux de la matrice de covariance seront petits par rapport aux éléments de la diagonale principale des sources directes. Dans ce cas, l'équation (3.2) devient une bonne approximation de la matrice de covariance des signaux reçus.

Avant d'aller plus loin dans la description de la méthode, on définit cinq éléments importants :

1. **Matrice de contribution de la source m (\mathbf{R}_m)** : Matrice de covariance de la source m si elle était la seule présente dans le signal.
2. **Matrice de covariance réduite d'ordre 1 (\mathbf{R}_{red_1})** : Matrice de covariance obtenue lorsque l'on enlève la matrice de contribution d'une source à la matrice de covariance des signaux reçus.

3. **Matrice de covariance partiellement réduite** : Matrice obtenue lorsque la contribution d'une source n'est pas complètement enlevée de la matrice de covariance originale.
4. **Vecteur propre résiduel** : Vecteur propre de la matrice de covariance partiellement réduite qui est parallèle au vecteur propre de la source dont on essaie d'éliminer la contribution. Ce vecteur est calculé à partir de la matrice de covariance des signaux reçus pour la source 1 et à partir de la matrice de covariance réduite d'ordre $m - 1$ pour la source m .
5. **Valeur propre résiduelle** : Valeur propre de la matrice de covariance partiellement réduite correspondant au vecteur propre résiduel.

Pour le reste de ce mémoire, on ordonne les sources en fonction de leur puissance. Ainsi, la première source est la source la plus puissante captée par le réseau de microphones.

Mathématiquement, la matrice de covariance réduite d'ordre 1 s'exprime de la manière suivante :

$$\mathbf{R}_{red_1} = \mathbf{R}_{xx} - \mathbf{R}_1 = \mathbf{R}_{xx} - \mathbf{a}_1 \sigma_1^2 \mathbf{a}_1^\dagger \quad (3.3)$$

On qualifie cette matrice de réduite d'ordre 1 parce que son sous-espace source a été réduit d'une dimension. On peut généraliser avec la matrice réduite d'ordre m , que l'on exprime ainsi :

$$\mathbf{R}_{red_m} = \mathbf{R}_{xx} - \mathbf{R}_1 - \mathbf{R}_2 - \dots - \mathbf{R}_m \quad (3.4)$$

En pratique, on doit estimer toutes ces matrices, ce qui donne l'expression suivante :

$$\hat{\mathbf{R}}_{red_m} = \hat{\mathbf{R}}_{xx} - \hat{\mathbf{R}}_1 - \hat{\mathbf{R}}_2 - \dots - \hat{\mathbf{R}}_m \quad (3.5)$$

L'idée derrière SEM est de déterminer la contribution de chaque source dans la matrice de covariance des signaux reçus. Au lieu d'enlever la matrice de covariance pour toutes les sources non corrélées en même temps, tel que fait dans la méthode *Covariance Differencing and Iterative Spatial Smoothing* [18], on peut estimer, une par une, la contribution de chaque source et de les enlever de la matrice de covariance des sources jusqu'à ce qu'il ne reste que la matrice de covariance du bruit. Sans définir l'algorithme dès maintenant, on peut faire une liste d'étapes de l'algorithme pour déterminer les différents problèmes à résoudre avant de pouvoir l'utiliser :

1. Calculer $\hat{\mathbf{R}}_1$, la version estimée de \mathbf{R}_1
2. Calculer $\hat{\mathbf{R}}_{red_1} = \hat{\mathbf{R}}_{xx} - \hat{\mathbf{R}}_1$
3. Recommencer les étapes 1-2 pour toutes les sources.
4. Calculer $\hat{\mathbf{R}}_{red_m}$ à l'aide de l'équation 3.5.
5. Recommencer les étapes 1-4 avec $\hat{\mathbf{R}}_{xx0-} = \hat{\mathbf{R}}_{xx} - \hat{\mathbf{R}}_{red_m}$ au lieu de $\hat{\mathbf{R}}_{xx}$ pour améliorer les résultats.

Ainsi, pour être capable d'utiliser SEM, il est nécessaire de pouvoir estimer avec un bon degré de précision la contribution de chaque source à la matrice de covariance des signaux reçus. À l'équation (3.2), on a vu qu'il y avait deux paramètres dans l'expression de chaque matrice de contribution, le premier étant le vecteur directeur et le second étant la puissance du signal. Dans le cas du vecteur directeur, puisque l'on connaît tous les paramètres du réseau que l'on utilise, la seule variable inconnue est l'angle d'arrivée de la source.

3.2 Génération des données

Dans les prochaines sections, on teste différentes parties de SEM en faisant plusieurs simulations. Il est très important de s'attarder à la génération des données de simulations puisqu'il est nécessaire d'être capable de bien contrôler les différents paramètres des signaux tels que leur puissance, leur corrélation et leur largeur de bande.

3.2.1 Génération de données corrélées

Pour générer des sources corrélées avec une puissance différente, il faut premièrement créer les signaux émis par chaque source. Si l'on considère un cas d'une source en trajet direct et de ses échos, cela implique de produire un signal avec tous les paramètres désirés pour ce signal pour ensuite ajuster la puissance du signal (via l'amplitude) pour obtenir les signaux des échos. Voici donc les étapes pour générer le signal en trajet direct :

1. À l'aide d'un générateur de nombres aléatoires uniformément distribués, produire une série de données. Il est alors possible de postuler que cette série de nombres aléatoires est l'enveloppe du signal modulant produit par la source, qui a été échantillonné avec une fréquence f_{em} .

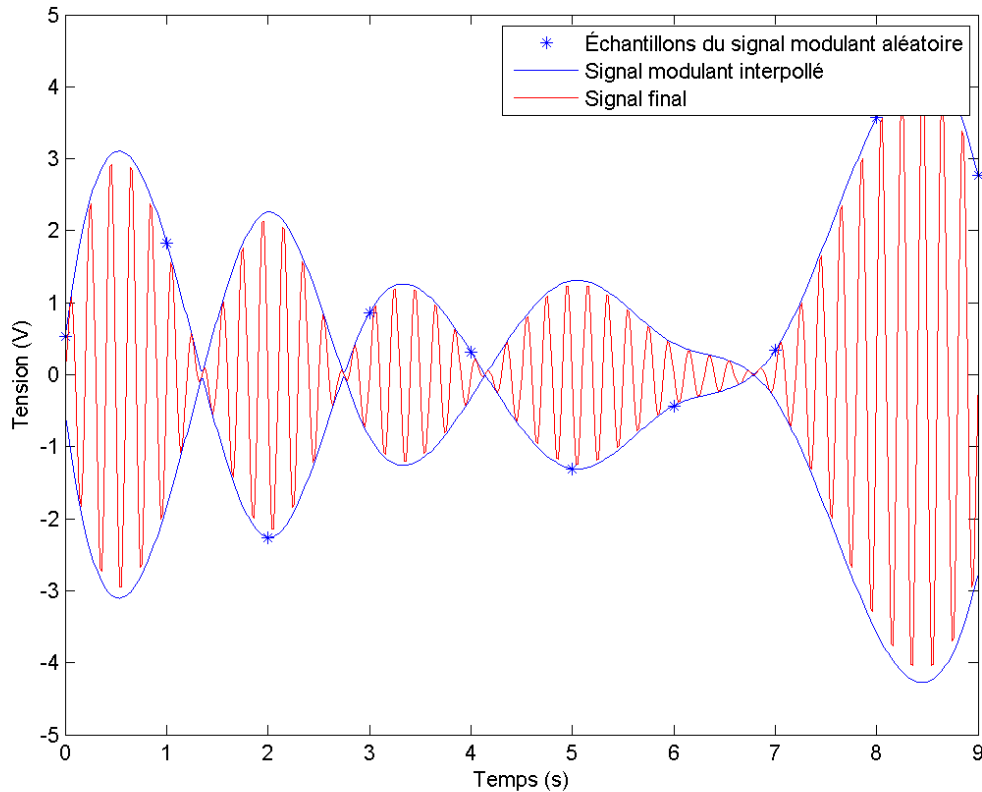


FIGURE 3.1 – Exemple de signal généré avec la méthode présentée dans cette section, avec les différentes étapes illustrées.

2. Utiliser une spline pour interpoler les points générés par le signal modulant aléatoire pour ainsi se ramener à une fréquence d'échantillonnage f_s assez élevée pour dépasser la fréquence de Nyquist.
3. Multiplier l'enveloppe extrapolée avec une porteuse, dont la fréquence f_p est beaucoup plus grande que la fréquence d'échantillonnage de la modulation de l'enveloppe de l'étape 1. On obtient alors le signal généré par la source.

Ces différentes étapes sont illustrées à la figure 3.1.

Le signal modulant aléatoire générée à l'étape 1 a une largeur de bande f_{em} et conserve cette largeur de bande après avoir été interpolée par la spline. Lorsque l'on multiplie par la porteuse, il se produit un décalage de la bande utile autour de la fréquence de la porteuse. Ainsi, pour que le signal soit à bande étroite, il faut que $f_{em} \ll f_p$. De plus, pour respecter le critère de Nyquist, il faut que $2f_p \leq f_s$.

Le niveau de corrélation entre deux signaux générés de cette façon est donné par le niveau de corrélation entre les séries de données générées à l'étape 1. Ainsi, plus les séries de données sont longues, plus les signaux sont indépendants. Lorsque l'on veut des signaux à bande étroite, cela pose problème puisque la faible fréquence d'échantillonnage de modulation f_{em} implique d'avoir des signaux d'une très longue durée. Dans un cas où l'on veut des signaux à large bande, la longueur des séries aléatoires générées n'est plus un problème parce que f_{em} est élevée. Donc pour un signal de quelques secondes, on génère des milliers d'échantillons pour les signaux modulateurs, ce qui est amplement suffisant pour obtenir des signaux indépendants à la fin de l'algorithme.

Une fois que l'on a généré des signaux ayant les propriétés voulues pour chaque source corrélée, il devient nécessaire de simuler ce que chaque capteur reçoit. Pour ce faire, la meilleure technique est de déterminer quel capteur va détecter le signal en premier, soit le capteur 1 si la DOA est située entre 0° et 90° , ou le capteur N si la DOA est située entre 90° et 180° . Ensuite, il faut calculer le décalage en temps entre les capteurs selon :

$$\tau = \frac{d}{v} \quad (3.6)$$

on peut ensuite décaler de τ le signal reçu par le premier capteur en interpolant avec une spline pour obtenir le signal capté par le deuxième capteur. On fait de même pour d'autres multiples de τ pour obtenir le signal de la source pour chaque capteur. Finalement, on fait les mêmes étapes pour chacune des sources et l'on somme les résultats pour chacun des capteurs pour obtenir le signal reçu à chacun des capteurs.

Dans le cas où l'on veut deux signaux complètement corrélés, par exemple pour simuler un écho, le plus simple est de générer seulement un signal ayant les propriétés voulues et d'utiliser un décalage temporel différent pour la position de chaque source. De cette façon, on simule bien un signal en trajet direct et un écho à des positions différentes. Si l'on veut diminuer la puissance de l'écho, il suffit simplement de multiplier le signal par un facteur plus petit que 1 avant de le décaler pour calculer ce qui est reçu par chacun des capteurs pour l'écho.

3.2.2 Génération de données non corrélées

La génération de signaux acoustiques non corrélés est assez difficile puisque les fréquences peu élevées font qu'on doit le faire avec un petit nombre de données, ce qui a pour effet qu'il reste toujours une certaine corrélation dans les données, à moins de générer un signal d'une durée infinie. Dans ce cas-ci, on va utiliser une autre approche,

au lieu de générer des signaux, on génère directement une matrice de covariance. Pour ce faire, il suffit de créer la matrice diagonale des signaux \mathbf{R}_{ss} et ensuite produire la matrice directrice \mathbf{A} . Ensuite, on utilise l'équation (2.14) :

$$\mathbf{R}_{xx} = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^\dagger + \mathbf{R}_{nn} \quad (3.7)$$

Dans un cas sans bruit, il suffit de ne pas additionner le terme \mathbf{R}_{nn} dont la génération suit la procédure décrite dans la section suivante.

3.2.3 Génération du bruit

Les deux sections précédentes avaient pour but de présenter les méthodes qui seront utilisées dans ce mémoire pour générer les signaux produits par des sources ponctuelles. Toutefois, il est important d'être capable de générer du bruit qui possède des propriétés ajustables. Pour ce faire, on présente dans cette section deux méthodes, la première générant une matrice de covariance de bruit qui peut être ajoutée à la matrice de covariance sans bruit générée avec les méthodes vues précédemment. Cette méthode est présentée à l'algorithme 10.

Algorithme 10 : Génération d'une matrice de covariance de bruit

- 1 Générer N vecteurs de nombres aléatoires uniformément distribués de longueur l_b .
 - 2 Calculer \mathbf{R}_{yy} , la matrice de covariance des vecteurs générées à l'étape 1.
 - 3 Calculer σ_{pb} , la moyenne de la diagonale de \mathbf{R}_{yy}
 - 4 Soit SNR , le rapport signal à bruit désiré. Calculer $\alpha_b = \frac{SNR}{\sigma_{pb}}$.
 - 5 Faire la multiplication $\mathbf{R}_{nn} = \alpha_b \mathbf{R}_{yy}$ pour obtenir la matrice de covariance du bruit désirée.
-

Cette méthode permet donc d'ajuster le SNR au niveau désiré. Il est aussi possible de contrôler la corrélation du bruit en faisant varier l_b , plus les vecteurs sont longs et plus ils sont indépendants.

L'algorithme 10 permet ainsi de générer une matrice de covariance avec les paramètres désirés. Toutefois, il est parfois intéressant d'être capable de générer du bruit partiellement corrélé avec des paramètres ajustables pour que l'on puisse l'ajouter avec des signaux de simulations des sources au lieu d'ajouter une matrice de covariance. La méthode la plus simple pour le faire est de simplement générer des séries de nombres aléatoires et d'ajuster leur puissance au niveau désiré pour ensuite les additionner aux signaux des sources. Cette méthode est toutefois peu pratique puisque la corrélation

du bruit est donnée par la longueur des séries aléatoires, qui elle-même doit être égale à la longueur des données simulées des sources. Pour contrer ce problème, l'algorithme 11 présente une méthode permettant de générer des signaux suivant une matrice de covariance donnée. Une bonne explication des étapes de l'algorithme se retrouve en [25], mais n'est pas présentée ici pour alléger le texte.

Algorithme 11 : Génération de données corrélées

- 1 Utiliser l'algorithme 10 pour générer \mathbf{R}_{nn} , la matrice de covariance du bruit désirée.
 - 2 Générer \mathbf{Z} , une matrice composée de N vecteurs colonnes de nombres aléatoires uniformément distribués de la même longueur que celle du signal sans bruit.
 - 3 Calculer \mathbf{V}_{nn} et $\mathbf{\Lambda}_{nn}$, les matrices contenant respectivement les vecteurs propres colonnes et les valeurs propres de \mathbf{R}_{nn} .
 - 4 Calculer la racine carrée des valeurs propres, ce qui donne la matrice diagonale $\mathbf{\Lambda}_{nn}^{1/2}$
 - 5 Calculer la matrice de transformation à l'aide de la formule suivante :

$$\mathbf{T}_y = \left(\mathbf{V}_{nn} \mathbf{\Lambda}_{nn}^{1/2} \right)^T$$
 - 6 Générer la série de données voulue en utilisant la formule suivante $\mathbf{Y} = \mathbf{T}_y^T \mathbf{Z}$.
-

Le but de cet algorithme est de générer des séries de données ayant la matrice de covariance \mathbf{R}_{yy} . Dans notre cas, on désire avoir une bonne corrélation dans le bruit capté par chacun des microphones, alors on génère des séries de seulement 50 données à l'étape 1, ce qui permet de laisser une certaine corrélation entre les données. Il est ensuite possible d'ajuster la puissance du bruit en le multipliant par le facteur approprié.

3.3 Estimation des paramètres

3.3.1 Direction d'arrivée individuelle

Tel que mentionné à plusieurs reprises dans ce mémoire, il est nécessaire d'utiliser MUSIC en assumant le nombre exact de sources, sinon le pseudospectre peut avoir des pics à de mauvaises DOA. Il n'est donc pas possible de déterminer la DOA d'une seule source dans une mesure contenant de nombreuses sources, bien que ce soit ce dont on a besoin pour SEM.

Pour commencer la création d'une méthode permettant de le faire, on bâtit un modèle du signal qui considère la corrélation des sources. Pour ce faire, il est nécessaire de développer l'équation (3.1), qui est valide même s'il y a de la corrélation entre les sources. Dans un cas sans corrélation entre les sources, la matrice \mathbf{R}_{ss} est diagonale

et dans le cas avec corrélation des sources, des éléments extra diagonaux apparaissent. Puisque toutes les équations sont linéaires, il est possible de séparer \mathbf{R}_{ss} en une somme de deux matrices. La première, notée \mathbf{R}_{ssdiag} , est une matrice diagonale correspondant à \mathbf{R}_{ss} dans le cas où il y a absence de corrélation entre les sources. La seconde, notée $\mathbf{R}_{ssdiag0}$, est une matrice à diagonale nulle, dont tous les termes extra diagonaux pour l'emplacement (i, j) sont $\mu_{ij}\sigma_i\sigma_j$, μ_{ij} étant le coefficient de corrélation entre les sources i et j et est nécessairement situé entre 0 et 1, σ_i et σ_j sont, respectivement, la puissance de la source i et de la source j . Cette seconde matrice correspond donc à la partie nuisible pour MUSIC de \mathbf{R}_{ss} . On peut alors développer l'équation (3.1) :

$$\mathbf{R}_{xx} = \underbrace{\mathbf{A}\mathbf{R}_{ssdiag}\mathbf{A}^\dagger}_{\mathbf{R}_{xx0-}} + \mathbf{A}\mathbf{R}_{ssdiag0}\mathbf{A}^\dagger + \mathbf{R}_{nn} \quad (3.8)$$

Dans cette équation, \mathbf{R}_{xx0-} est la matrice de covariance reçue par le réseau sans le bruit et sans la corrélation entre les sources. C'est donc cette matrice qui contient l'information importante pour MUSIC. Si l'on considère que le rapport écho à direct (EDR) est petit, les termes non nuls de $\mathbf{R}_{ssdiag0}$ sont petits par rapport aux termes de la diagonale de \mathbf{R}_{ssdiag} , ce qui fait qu'il n'affecte pas significativement le résultat de la somme des deux matrices. Ainsi, dans un cas avec une source en trajet direct et d'échos beaucoup plus faibles, le premier vecteur propre sera colinéaire avec le vecteur directeur de la première source, ce qui permet d'utiliser MUSIC en assumant une seule source pour déterminer sa DOA.

Pour vérifier cette affirmation, il est nécessaire de faire quelques simulations. La figure 3.2 présente les résultats obtenus pour quatre simulations, chaque simulation ayant été faite pour différentes valeurs de EDR pour une source en trajet direct positionnée à $\theta_1 = 70^\circ$, et deux échos de même puissance, positionnés à $\theta_2 = 115^\circ$ et $\theta_3 = 135^\circ$, le tout capté par un réseau de 20 microphones distants de 5 cm entre eux. Cette figure montre bien qu'on est capable de déterminer la DOA de la première source assez facilement si le EDR est assez faible. Toutefois, on voit que plus le EDR est élevé, plus les trois sources sont confondues dans le premier vecteur propre, ce qui rend difficile la localisation de la première source seule. La problématique adressée dans ce mémoire tombe justement dans le cas d'une source en trajet direct très puissante par rapport aux échos, donc cette méthode peut être utilisée, mais il est important de comprendre qu'il n'est pas possible d'utiliser cette méthode dans un cas où l'ordre de grandeur de la puissance des sources devient semblable.

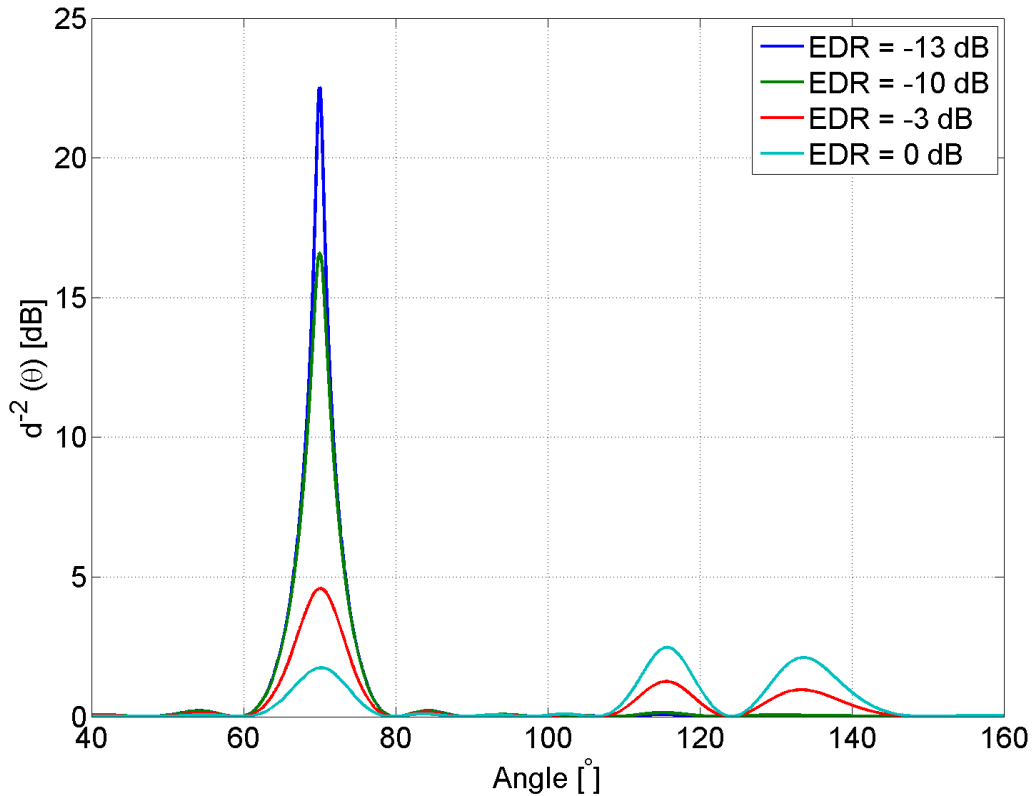


FIGURE 3.2 – Pseudospectre MUSIC en assumant une source pour un signal composé d’une source principale et de deux sources secondaires (échos) sans bruit pour différents EDR.

3.3.2 Estimation de la puissance des sources

Le fait que l’on soit capable de déterminer la DOA de chaque source individuellement ouvre la porte à de nombreuses techniques dans le but de déterminer la puissance des sources. L’une de ces méthodes est d’analyser le comportement des valeurs propres de la matrice de covariance partiellement réduite. En effet, on peut s’attendre à ce que les valeurs propres de cette matrice ne restent pas constantes en fonction de la puissance posée de la source principale. La figure 3.3 montre le comportement de la valeur absolue de la première et de la dernière valeur propre en fonction de l’erreur sur la puissance estimée de la source principale. Les courbes pour les autres valeurs propres n’ont pas été présentées sur la figure puisque leur variation est très minime et leur présence viendrait nuire à la visualisation des éléments du graphique. Puisque les signaux sont générés par ordinateur, ils n’ont pas de valeur physique réelle et c’est pour raison que l’on utilise des métriques comme unité. La valeur de cette unité pourrait aussi bien être des Watts que des chevaux-vapeur.

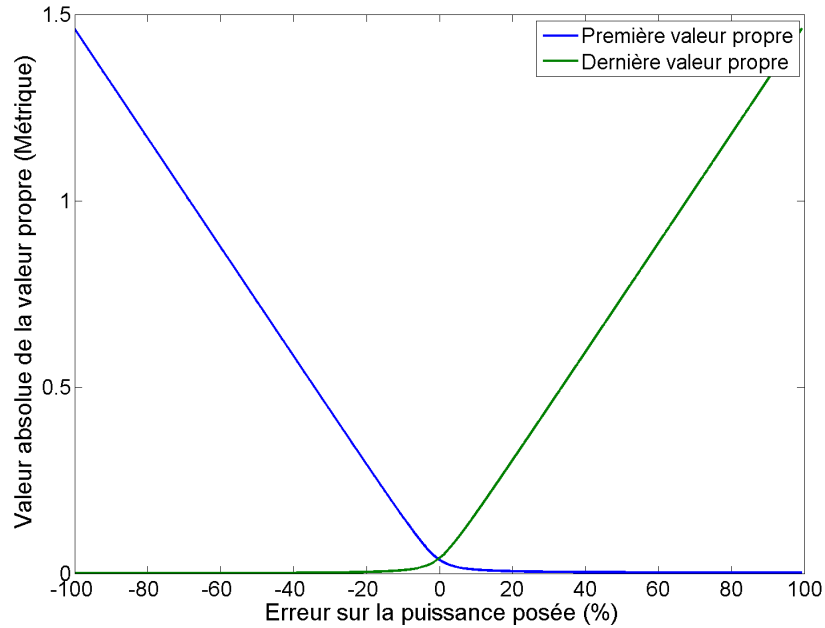


FIGURE 3.3 – Comportement de la première et de la dernière valeur propre de la matrice de covariance partiellement réduite en fonction de l’erreur sur le puissance posée.

On y voit donc que le point de rencontre entre la valeur absolue de la première et de la dernière valeur propre fournit une excellente estimation de la puissance de la source principale, avec une erreur de 1%. L’explication de ces résultats est qu’à chaque fois que l’on augmente la puissance retranchée, la valeur propre résiduelle associée à la source principale, qui est un indice de puissance, diminue de plus en plus jusqu’à être très près de celle des valeurs propres correspondant aux sources secondaires (échos). Ainsi, si l’on revient à l’équation (3.8), le terme des valeurs non-diagonales de la matrice de covariance des sources \mathbf{R}_{ssdiag_0} devient de plus en plus importante, ce qui nous ramène à un cas où toutes les sources sont de puissance semblables et corrélées. La première valeur propre de la matrice de covariance partiellement réduite correspond de plus en plus à celles des trois sources en même temps et le premier vecteur propre de cette matrice n’est plus parallèle au premier vecteur propre de la matrice originelle, donc celui correspondant à la source principale. En continuant d’enlever la contribution de la première source, la valeur propre résiduelle change de rang peu à peu pour devenir finalement la dernière valeur propre de la matrice de covariance partiellement réduite. À ce moment, plus l’on enlève la contribution de la première source, plus la dernière valeur propre devient négative, donc plus sa valeur absolue, telle que montrée à la figure 3.3, devient grande. Cette transition n’est toutefois pas drastique, elle se fait progressivement, ce qui a pour

effet qu'il n'y a pas un point précis où la dernière valeur propre commence à diminuer. Comme on peut le voir si la figure 3.3, le point d'intersection de la valeur absolue entre la première et la dernière valeur propre de la matrice de covariance partiellement réduite est toutefois une bonne approximation du point où la contribution de la source principale est complètement enlevée. C'est donc ce point que nous utiliserons dans le reste de ce mémoire.

Pour déterminer la position de ce point, il est possible d'utiliser une méthode numérique telle que la méthode de Newton-Raphson au lieu de balayer toutes les valeurs de puissance posée.

3.4 Simulations

Dans la section précédente, on a fait quelques simulations dans le but de déterminer comment estimer la DOA et la puissance de chaque source, qui sont des étapes intermédiaires de l'algorithme SEM. Toutefois, on n'a pas testé la méthode au complet pour vérifier si elle fonctionne bien. Le but de cette section est de faire une analyse de SEM en commençant par des tests individuels pour différents SNR pour valider que la méthode permet d'obtenir un pseudospectre avec les bonnes DOA. Par la suite, on présentera les résultats de deux séries de simulations Monte-Carlo, la première contenant 200 essais pour des sources non corrélées et la seconde pour 200 essais de sources corrélées. Pour toutes ces simulations, on utilise à nouveau un réseau composé de 20 microphones avec une distance inter capteurs de 5 cm.

3.4.1 Simulations individuelles

Pour commencer les simulations, il est important de vérifier si SEM permet bien de déterminer la DOA de toutes les sources. Utilisant la méthode vue à la section 3.2.2, on simule deux signaux non corrélés, situés aux positions $\theta_1 = 70^\circ$ et $\theta_2 = 120^\circ$ et mesurés avec un réseau linéaire de 20 capteurs distancés de 5 cm. Le rapport de puissance entre les sources est de 13 dB. Tel que l'on a vu à la section 3.2.3, lorsque l'on simule des sources non corrélées, il est plus facile de générer directement la matrice de covariance. Par la suite, on utilise l'algorithme 10 pour générer directement une matrice de covariance du bruit, que l'on ajoute à la matrice de covariance des sources pour obtenir la matrice de covariance des signaux reçus. Dans le but d'avoir du bruit assez nuisible à la détection des DOA, on génère la matrice de covariance du bruit avec des séries de longueurs de $l_b = 50$ échantillons. Puisque cette longueur est assez courte, le bruit est

corrélé et vient fortement nuire à la détection des DOA. Pour vérifier l'impact du bruit, on ajuste la matrice de covariance du bruit en la multipliant par différents facteurs α_b , ce qui permet d'obtenir des résultats pour différents SNR.

Une fois que la matrice de covariance avec bruit pour le réseau est générée, on peut alors utiliser différentes techniques pour estimer la DOA. Dans notre cas, on utilise premièrement MUSIC avec FBSS avec 8 sous-réseaux. L'utilisation de FBSS peut paraître inutile lorsque les sources ne sont pas corrélées, mais puisque cette méthode effectue un moyennage de plusieurs sous-réseaux, l'impact du bruit est diminué. On utilise ensuite SEM de deux façons différentes :

1. À l'aveuglette, c'est-à-dire en faisant une pré estimation à l'aide de MUSIC pour déterminer la première source et, pour la seconde source, en utilisant MUSIC sur la matrice de covariance réduite d'ordre 1.
2. En utilisant, comme estimation de DOA de départ, les positions théoriques des deux sources.

L'objectif d'utiliser SEM de ces deux manières est de déterminer d'où proviennent les erreurs. Lorsque l'on utilise les positions théoriques à la place des pré estimations, les erreurs provenant de cette méthode proviennent toutes de l'estimation de la puissance des sources. Ainsi, la différence entre les résultats obtenus avec les pré estimations théoriques et estimées devient l'erreur causée par la pré estimation.

Les résultats obtenus de cette manière se retrouvent à la figure 3.4 si l'on considère un rapport de puissance entre les sources de 13 dB, tandis que les résultats lorsque le rapport de puissance est de 3 dB se retrouvent à la figure 3.5. Sur ces figures, on calcule le SNR par rapport à la source la moins puissante, soit la source 2.

On remarque alors qu'à SNR très élevé, la nouvelle méthode fonctionne pratiquement aussi bien que MUSIC avec FBSS. De plus, on remarque qu'il est facile de déterminer la DOA des deux sources, les courbes faites avec et sans les valeurs de départ théoriques se confondent au point où il est impossible de les discerner. À -10 dB, on commence à voir une différence entre les courbes, ce qui indique qu'il devient plus difficile de déterminer la position de la deuxième source avec MUSIC-FBSS. À -15 dB, on voit que la méthode faite avec des positions estimées fonctionne de moins en moins bien pour la seconde source, tandis que celle faite avec des valeurs théoriques donne de très bons résultats. Il semble donc que la première estimation des sources soit très importante. On remarque aussi que MUSIC-FBSS fonctionne relativement bien, même lorsque le SNR est faible.

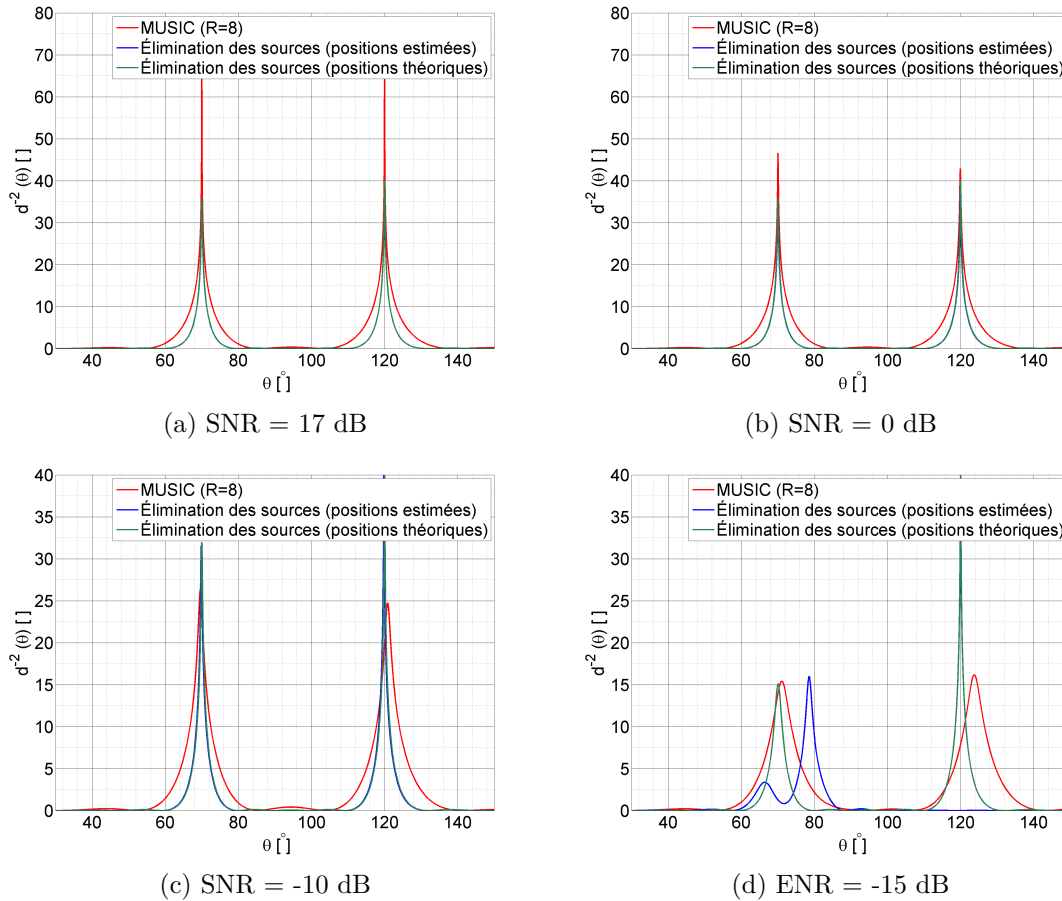


FIGURE 3.4 – Comparaison des résultats obtenus avec MUSIC-FBSS et la méthode de réduction de la matrice de covariance pour deux cas différents, l'un fait à l'aveuglette et l'autre en utilisant les positions exactes de départ. Les données proviennent de deux sources différentes non corrélées, dont l'une est 13 dB plus puissance que l'autre.

Finalement, comme on peut le voir sur la figure 3.5, SEM fonctionne bien même si la puissance des sources est assez près l'une de l'autre.

3.4.2 Simulations Monte-Carlo pour sources non corrélées

Les figures 3.4 et 3.5 sont intéressantes pour montrer que la méthode par réduction de la matrice de covariance fonctionne bien pour des sources non corrélées, mais il ne s'agit pas d'une analyse complète du sujet, mais bien d'un seul essai. Pour bien comparer la nouvelle méthode avec MUSIC, il est nécessaire de faire un grand nombre d'essais et d'évaluer la moyenne et l'écart-type des résultats. Réalisant 200 essais Monte-Carlo pour différents niveaux de bruit, on obtient la figure 3.6 pour la moyenne, la figure 3.7 pour l'écart-type et la figure 3.8 pour la probabilité de fausse alarme, que l'on définit comme

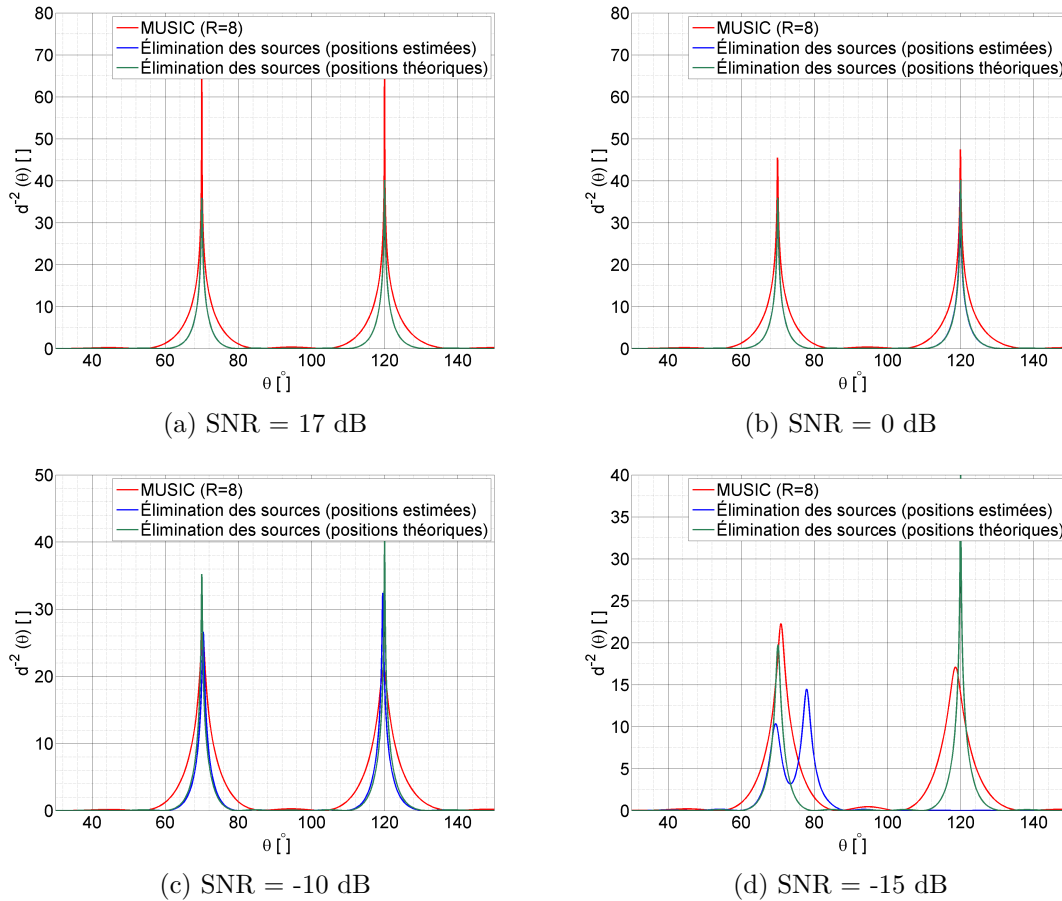


FIGURE 3.5 – Comparaison des résultats obtenus avec MUSIC avec FBSS et SEM pour deux cas différents, l’un fait à l’aveuglette et l’un en utilisant les positions exactes de départ. Les données proviennent de deux sources différentes non corrélées, dont l’une est 3 dB plus puissance que l’autre.

la probabilité qu’aucun pic ne soit détecté dans un écart de 10° autour de la position réelle de la source. Les SNR ont été calculés avec le rapport de puissance de la source la plus faible sur la puissance du bruit. On n’a pas présenté les courbes pour la position de la source la plus puissante dans ce graphique puisqu’autant son erreur moyenne, son écart-type et sa probabilité de fausse alarme étaient pratiquement parfaits pour toutes les valeurs de SNR simulées, ce qui est attendu puisque le rapport de puissance entre les sources est de 13 dB, ce qui implique que le bruit est faible comparativement à la source 1.

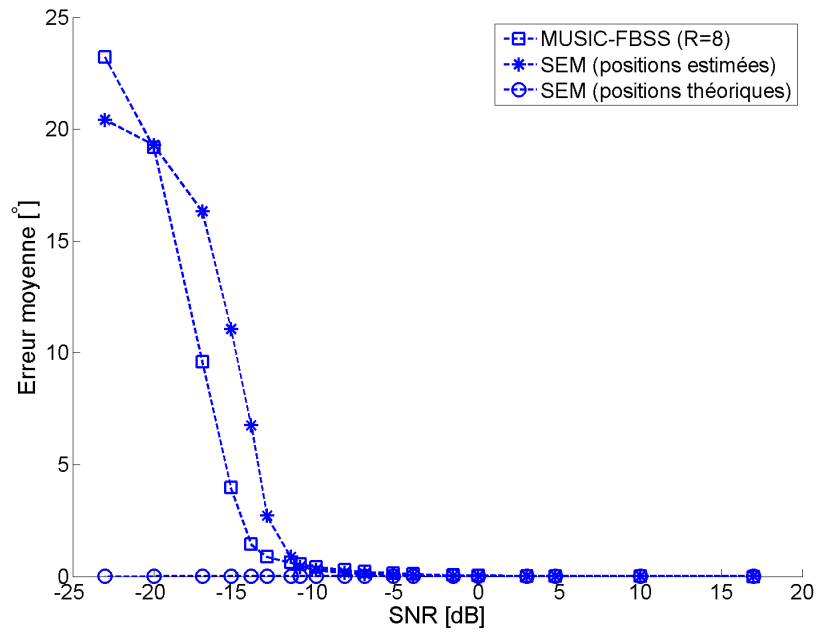
En observant la figure 3.6, on remarque qu’à SNR élevé, toutes les méthodes offrent un résultat assez semblable. Toutefois, on peut remarquer qu’à partir de 0 dB, l’utilisation de MUSIC-FBSS commence à perdre en efficacité et continue d’augmenter jusqu’aux

très faibles SNR. SEM avec positions estimées est légèrement plus précise que MUSIC entre 0 et -10 dB, mais son erreur augmente très rapidement par la suite pour devenir moins précise. Les courbes d'écart-type présentées à la figure 3.7 et de probabilité de fausse alarme à la figure 3.8 présentent la même tendance, mais avec une chute d'écart-type vers les plus petites valeurs de SNR. La courbe de SEM avec positions théoriques reste parfaite, peu importe le niveau de bruit. La courbe obtenue pour FBSS est bien ce que l'on s'attendait, c'est-à-dire que la méthode fonctionne bien lorsque le bruit est faible, mais perd en efficacité lorsque le bruit vient noyer le signal utile de la source la moins puissante. Les courbes obtenues pour les deux versions de SEM viennent montrer l'importance de la première estimation de la DOA. Lorsque cette première estimation comporte une certaine erreur, les étapes suivantes de l'algorithme viennent amplifier cette erreur, ce qui explique la forte augmentation que l'on observe à -12 dB. La diminution de l'écart-type que l'on observe vers -20 dB n'est pas causée par une amélioration des résultats, puisque l'erreur moyenne et la probabilité de fausse alarme continuent d'augmenter. Cela provient plutôt d'un pic fantôme qui apparaît dans le pseudospectre à une mauvaise position, ce qui diminue bien l'écart-type, mais qui augmente malgré tout l'erreur. Tous ces résultats montrent donc que SEM fonctionne bien, mais n'est pas significativement meilleure de MUSIC avec FBSS dans un cas sans corrélation entre les sources.

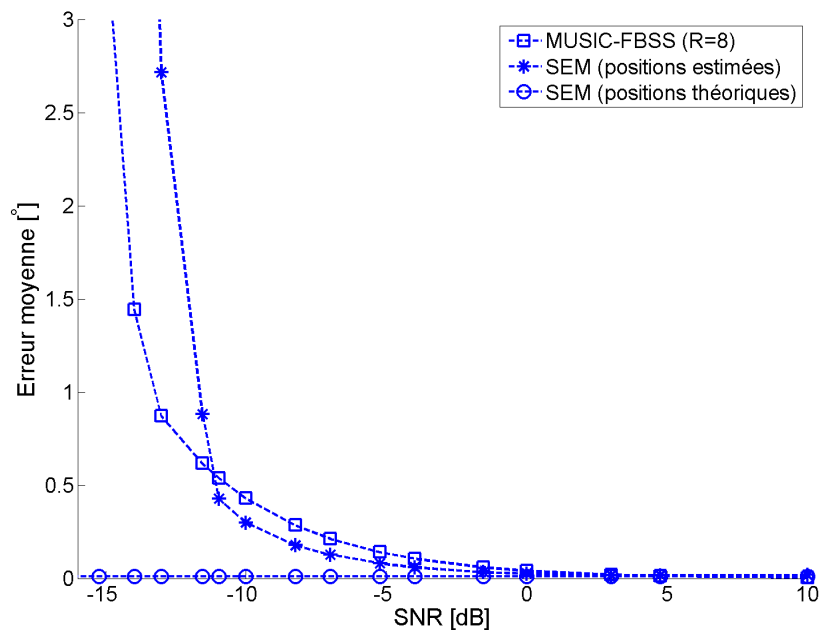
3.4.3 Simulations pour sources corrélées

Pour la continuation de l'analyse de SEM, on considère un nouveau cas, moins idéal pour les algorithmes, où les sources sont parfaitement corrélées. Pour commencer cette analyse, on débute avec une source puissante et d'un écho beaucoup plus faible. Dans ce contexte, on sait d'abord que MUSIC utilisé seul ne fonctionne pas et que même si l'on utilise du lissage spatial avant, il existe toujours un certain biais dans le résultat obtenu au final. Ce biais peut d'ailleurs devenir très important si le réseau de capteurs n'est pas assez grand pour le diviser en un grand nombre sous-réseau et complètement décorréler les sources. Pour le cas de SEM, on a déjà vu que la pré estimation de la première source n'est pas trop affectée par la corrélation entre les sources si le EDR est faible.

Pour commencer les simulations Monte-Carlo, on décide de placer deux sources, une première à $\theta_1 = 70^\circ$, et une deuxième, l'écho de la première, à $\theta_2 = 120^\circ$. On utilise alors une porteuse à $f_p = 2$ kHz, une fréquence d'échantillonnage de modulation de 50 Hz et une fréquence d'échantillonnage du signal de 10 kHz. Finalement, on ajuste l'amplitude

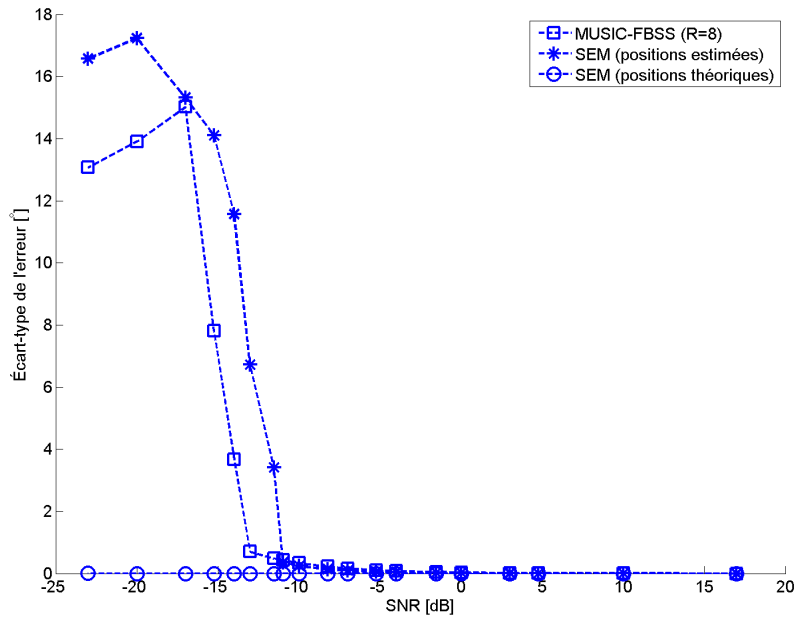


(a) Résultats complets

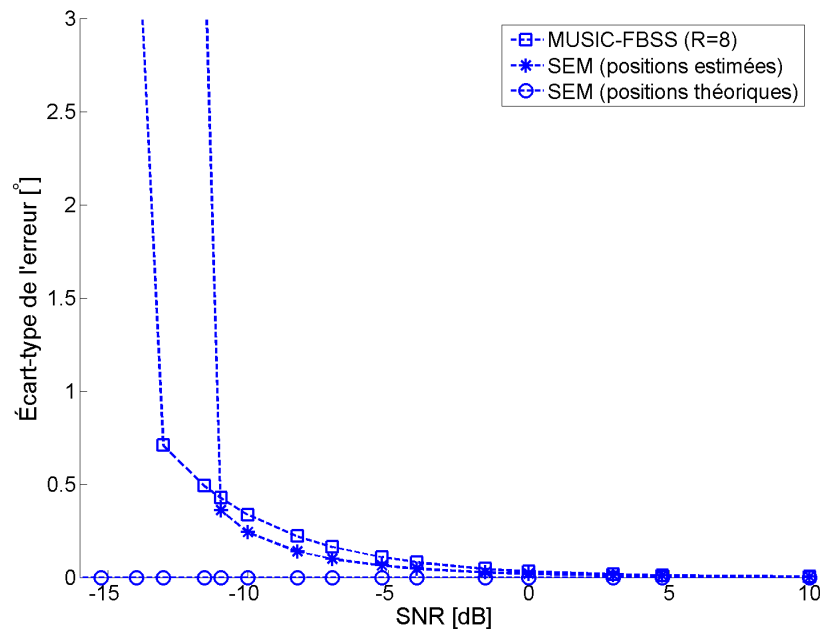


(b) Résultats agrandis pour montrer la différence pour les plus grandes valeurs de SNR

FIGURE 3.6 – Erreur moyenne de la DOA estimée d’une source faible d’une série de 200 essais Monte-Carlo en fonction du SNR pour des signaux non corrélés.



(a) Résultats complets



(b) Résultats agrandis pour montrer la différence pour les plus grandes valeurs de SNR pour des signaux non corrélés.

FIGURE 3.7 – Écart-type d'une série de 200 essais Monte-Carlo en fonction du SNR pour des signaux non corrélés.

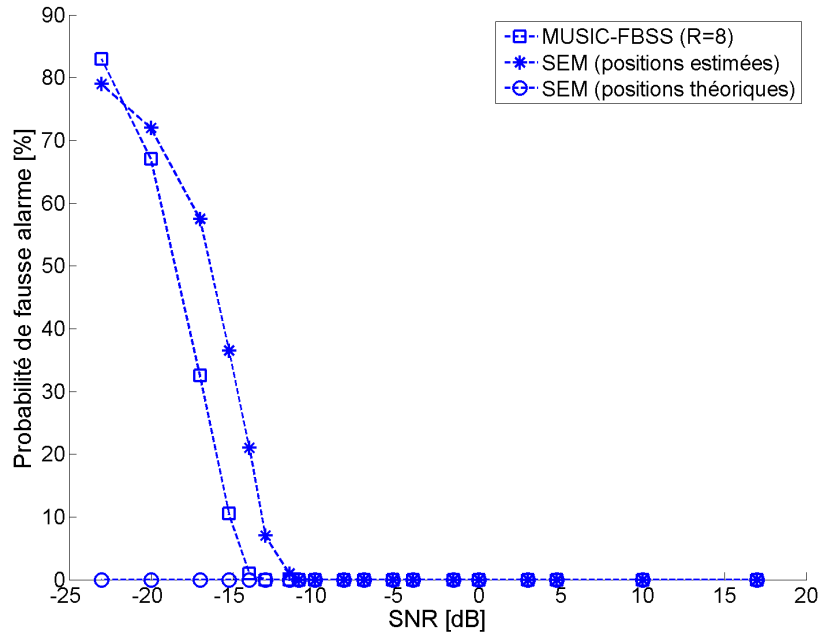
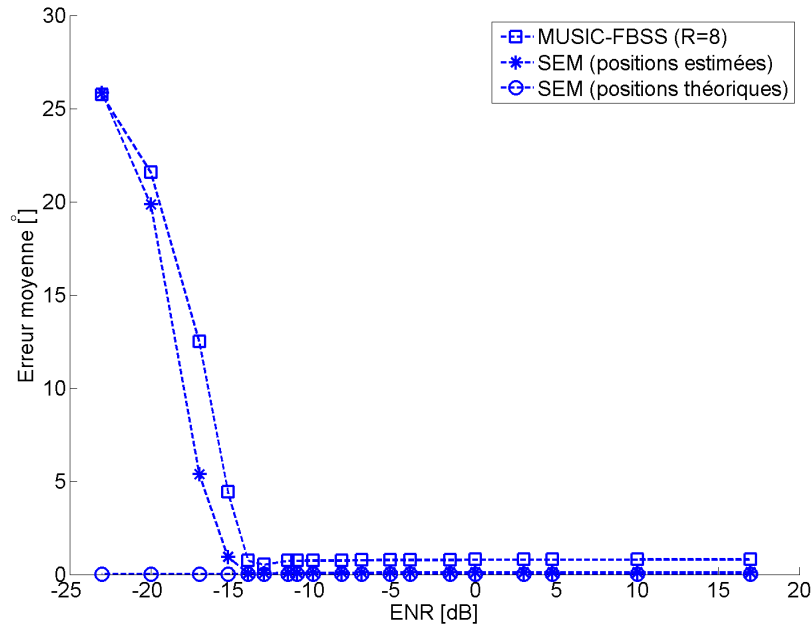


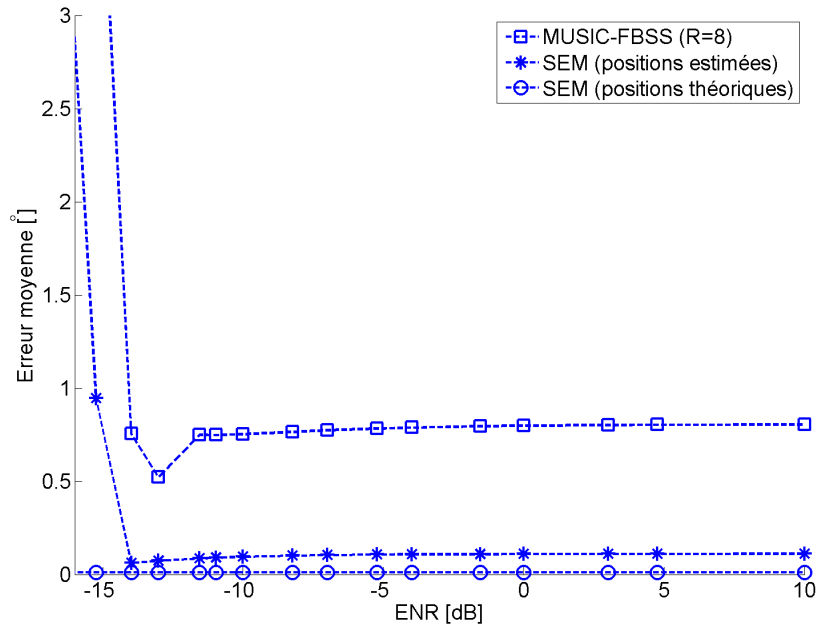
FIGURE 3.8 – Probabilité de fausse alarme en fonction du SNR obtenu lors de 200 essais Monte-Carlo avec des signaux non corrélés.

des signaux pour que le EDR soit de -34 dB. Faisant 200 essais Monte-Carlo pour différents ENR, on obtient les résultats présents sur les figures 3.9 (erreur moyenne), 3.10 (écart-type) et 3.11 (probabilité de fausse alarme). La matrice de covariance du bruit est générée à l'aide de l'algorithme 10 avec $N = 20$ séries contenant $l_b = 50$ échantillons chacune, que l'on ajuste pour obtenir différents ENR.

À nouveau, on commence par analyser les résultats de l'erreur moyenne en fonction du l'ENR. On remarque alors que MUSIC-FBSS a un biais constant de 0.75° pour les valeurs les plus hautes d'ENR, ce qui est possiblement un restant de corrélation entre les sources que FBSS n'a pas été capable de complètement enlever. Une fois rendu à -14 dB, le biais augmente très rapidement, jusqu'à atteindre 25° rendu à -23 dB. SEM, quant à elle, a un biais d'environ 0.1° pour les valeurs de haut ENR, et ce biais augmente rapidement à partir de -15 dB. La courbe de l'écart-type présente des résultats semblables, MUSIC-FBSS a un écart-type d'environ 0.4° pour les hauts ENR et il augmente exponentiellement à partir de -5 dB; SEM, quant à elle, possède un écart-type pratiquement nul et qui augmente tranquillement jusqu'à -15 dB, où il se met à augmenter très rapidement. On peut aussi remarquer, autant sur la figure 3.9 que sur la figure 3.10, que SEM avec les positions théoriques fournit des résultats très précis et peu variants. La figure 3.11 vient résumer les résultats présents sur les

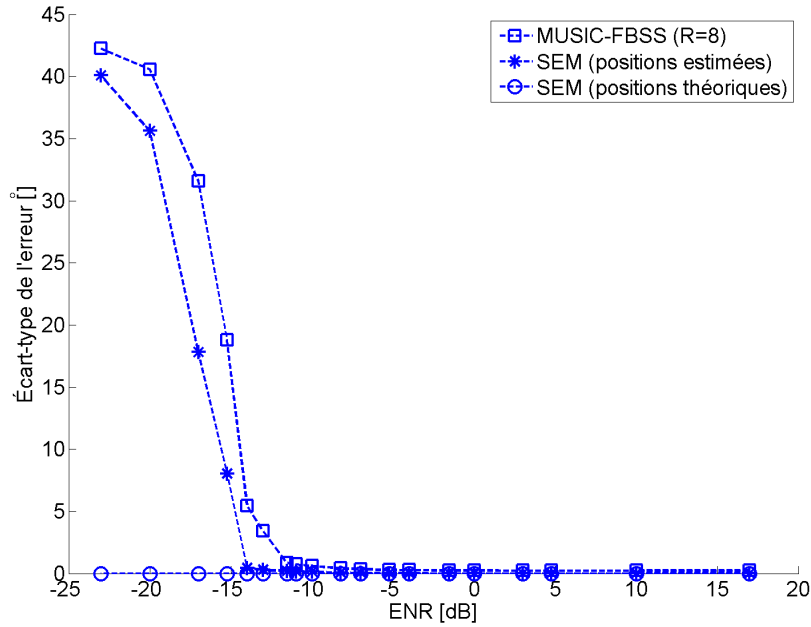


(a) Résultats complets

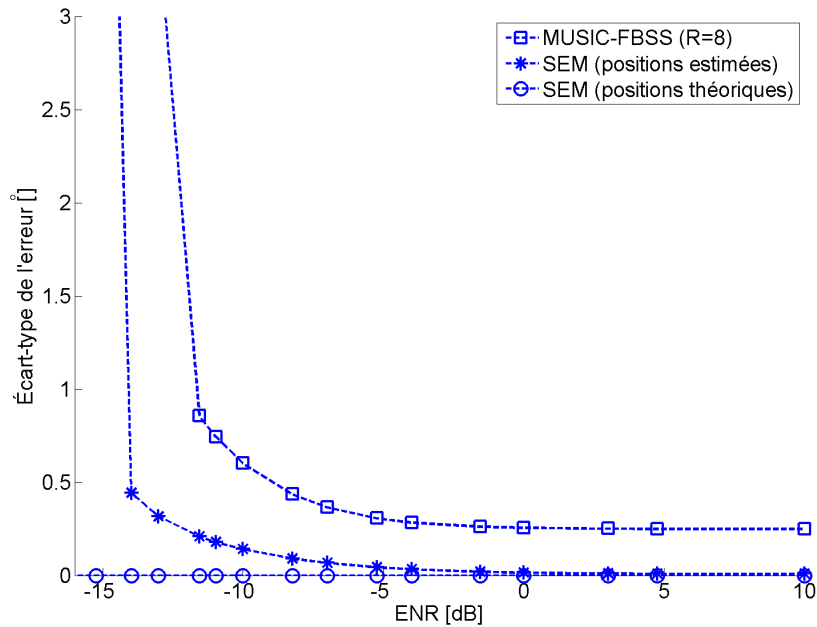


(b) Résultats agrandis pour montrer la différence pour les plus grandes valeurs de ENR

FIGURE 3.9 – Erreur moyenne de la DOA estimée d’une source faible d’une série de 200 essais Monte-Carlo en fonction du ENR pour un signal composé d’une source en trajet direct puissant et d’un écho beaucoup plus faible.



(a) Résultats complets



(b) Résultats agrandis pour montrer la différence pour les plus grandes valeurs de ENR pour des signaux non corrélés.

FIGURE 3.10 – Écart-type d'une série de 200 essais Monte-Carlo en fonction du ENR pour un signal composé d'une source en trajet direct puissant et d'un écho beaucoup plus faible.

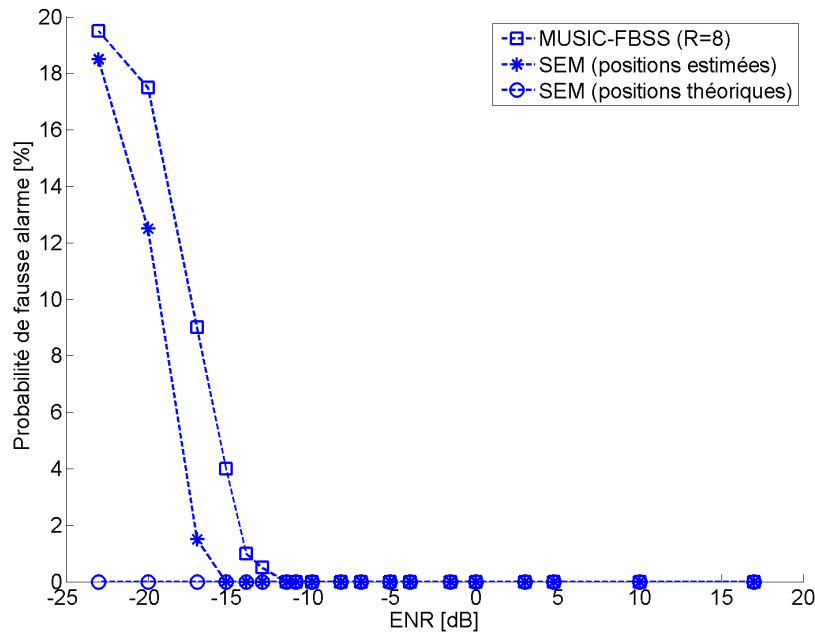


FIGURE 3.11 – Probabilité de fausse alarme en fonction du ENR obtenu avec 200 essais Monte-Carlo pour un signal composé d’une source en trajet direct puissant et d’un écho beaucoup plus faible.

deux autres figures, montrant qu’autant MUSIC-FBSS que SEM ne font pas de fausse alarme pour les ENR plus élevés. Toutefois, on y remarque que les cas de fausse alarme commencent à apparaître vers -12 dB et augmentent jusqu’à 12 % rendu à -23 dB pour MUSIC, tandis que pour SEM avec des positions estimées, les cas de fausse alarme apparaissent vers -17 dB et augmentent rapidement jusqu’à 18 % à -23 dB. Pour SEM avec positions théoriques, on remarque qu’il n’y a pas présence de fausse alarme, peu importe le niveau de bruit.

Ces résultats contrastent donc avec ceux obtenus lorsque les sources n’étaient pas corrélées, où les deux méthodes donnaient des résultats assez similaires. Ici, on peut remarquer jusqu’à quel point SEM permet d’obtenir de meilleurs résultats, peu importe le niveau de bruit, puisqu’il est capable de bien décorrélérer les sources. En comparant directement les graphiques des figures 3.6, 3.7 et 3.8 avec ceux des figures 3.9, 3.10 et 3.11, on pourrait penser que SEM fonctionne mieux dans le cas où les sources sont corrélées. Toutefois, il faut considérer que le rapport de puissance entre les sources n’est pas le même pour les deux simulations, il est de 13 dB pour les simulations de sources non corrélées, et de 34 dB pour les sources corrélées. Tel que l’on a vu en comparant les figures 3.5 et 3.4, SEM fonctionne mieux lorsque le rapport de puissance entre les

sources est plus élevé.

Le cas d'un seul écho est très intéressant pour étudier la qualité de la réduction de la matrice de covariance et pour analyser d'où proviennent les erreurs principales de l'algorithme. Toutefois, le cas où il y a une source directe et plusieurs échos est plus réaliste puisqu'en général, le son va réfléchir sur les parois d'une pièce de nombreuses fois. Ainsi, pour les simulations qui suivent, on simulera un signal composé d'une source en trajet direct, positionnée à $\theta_1 = 70^\circ$, et de deux échos, positionnés à $\theta_2 = 115^\circ$ et $\theta_3 = 135^\circ$. Pour ces simulations, le EDR est de -34 dB.

L'ajout d'une troisième source pose un problème qui n'était pas présent auparavant pour déterminer les statistiques d'estimation. En effet, il devient beaucoup plus difficile d'associer une DOA estimée avec une DOA théorique, surtout dans les cas où une source estimée est très loin de la position réelle de toutes les sources. Pour contourner le problème, on commence par identifier la DOA estimée qui est la plus près de la DOA de la source la plus puissante. Puisque le EDR est petit et que l'on ajuste le niveau de bruit par rapport à la puissance des échos, la DOA estimée pour la source en trajet direct est toujours très près de la valeur théorique. Par la suite, on détermine la DOA estimée qui est la plus près d'une DOA théorique des échos, et le second écho est associé à la dernière DOA estimée. Toutefois, il arrive que l'association d'une DOA estimée avec une DOA théorique soit faussée, en particulier à faible ENR, où il peut apparaître des pics qui ne correspondent pas du tout à une réelle DOA. Pour faire une bonne analyse des simulations, on effectue les graphiques pour les 3 éléments suivants :

1. La probabilité de fausse alarme en fonction du ENR (figure 3.12). À nouveau, on considère que l'on a une fausse alarme lorsqu'aucune source n'a été trouvée dans un rayon de 10° autour de la position réelle de la source. Considérant que la position des échos est de $\theta_2 = 115^\circ$ et $\theta_3 = 135^\circ$, le choix de la marge de 10° empêche les rayons des deux échos de se croiser.
2. La moyenne de l'erreur pour chacune des sources, mais en excluant de l'analyse les cas de fausse alarme (figure 3.13).
3. L'écart-type de l'erreur pour chacune des sources, mais en excluant de l'analyse les cas de fausse alarme (figure 3.14).

Les résultats présentés à la figure 3.12 démontrent à quel point il est difficile de déterminer la DOA avec MUSIC-FBSS. Bien qu'à fort ENR la probabilité de détection soit près de 100%, les résultats deviennent de moins en moins fiables à partir de -12

dB, jusqu'à atteindre une probabilité de fausse alarme de 65% à -23 dB. Ces résultats contrastent avec ceux de la figure 3.11 puisque la probabilité de fausse alarme est beaucoup plus élevée avec l'ajout d'un second écho. Il est donc assez clair que la présence d'un deuxième écho est très nuisible à cette méthode. Pour SEM, la probabilité de fausse alarme est quasi nulle jusqu'à -16 dB et atteint 60% à -23 dB. Ces résultats montrent que la nouvelle méthode diverge beaucoup moins que MUSIC-FBSS pour des ENR plus faibles que -12 dB alors qu'en dessous de -20 dB, les deux méthodes sont peu recommandables. De plus, on peut voir que la probabilité de fausse alarme est nulle si la première estimation de la position des sources est parfaite. Ces résultats montrent à nouveau l'importance de la première estimation de la position des sources. Dans le cas où l'on utilise les positions théoriques, on remarque que la probabilité de fausse alarme est nulle pour toutes les valeurs de ENR.

Pour faire l'analyse des figures 3.13 et 3.14, il faut toujours se rappeler que les cas de fausse alarme ont été enlevés puisqu'il était difficile d'associer une position estimée avec une source ; une estimation initiale impossible aurait faussé la précision des données de manière inutile. Il faut faire particulièrement attention en observant les courbes lorsque la probabilité de fausse alarme n'est plus nulle puisque l'on ne se base plus sur toutes les données.

Les résultats obtenus pour MUSIC avec FBSS montrent qu'à haut ENR, le biais qui était présent lorsqu'il y avait un seul écho disparaît. L'explication de ce phénomène vient du fait que lorsqu'il y a seulement un écho dans le signal, la position de ce dernier en cas de corrélation résiduelle est décalée vers la position de la première source de quelques degrés. En ajoutant un deuxième écho de l'autre côté du pseudospectre, le pic se décale légèrement de l'autre côté, ce qui vient annuler le biais. On remarque ensuite que pour les deux échos, l'erreur moyenne se met à augmenter significativement vers -6 dB et augmente rapidement jusqu'à se stabiliser vers -18 dB. L'écart-type, quant à lui, est déjà en augmentation à 0 dB et continue d'augmenter jusqu'à -16 dB, où les résultats se stabilisent à 2.75° . Malgré que les résultats semblent fiables à haut ENR, ils se mettent à diverger plus rapidement qu'avec un seul écho, ce qui est complémentaire à ce qu'on a obtenu dans la courbe de la probabilité de fausse alarme.

Dans le cas de SEM avec positions estimées, on remarque un biais constant d'environ 1° pour les valeurs de haut ENR et qui augmente à partir de -16 dB. Pour ce qui est de SEM avec des positions théoriques, on remarque aussi un biais d'environ 0.5° en moyenne et qui reste assez constant pour toutes les valeurs de ENR testées. Il n'est pas

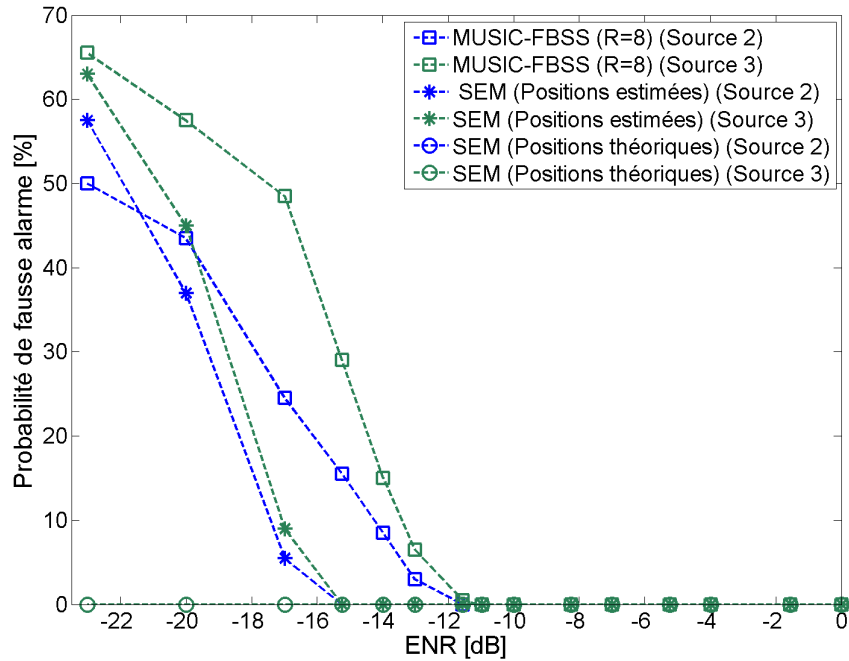


FIGURE 3.12 – Probabilité de fausse alarme en fonction du ENR obtenu lors de 200 essais Monte-Carlo pour un signal composé d’une source et de deux échos.

très surprenant d’avoir un biais avec deux échos puisque, une fois que l’on a totalement éliminé la contribution de la source en trajet direct, il reste deux sources corrélées et de même puissance dans la matrice de covariance. Ces deux conditions font que l’idée d’utiliser MUSIC seul en considérant une seule source sur la matrice de covariance réduite d’ordre 1 n’est pas optimale. Ce fait est aussi vrai pour le test de puissance qui sera faussé par le fait que le premier vecteur propre de la matrice de covariance réduite d’ordre 1 correspondra aux deux échos en même temps.

La figure 3.14 montre que l’écart-type de MUSIC-FBSS suit la même tendance que son erreur moyenne. Pour SEM, on remarque qu’il y a un écart-type significatif même pour les valeurs de haut ENR avec les positions estimées, et un écart-type nul si l’on utilise les valeurs théoriques. Ainsi, on peut voir que l’estimation de la position des échos induit un écart-type dans les résultats du à leur corrélation.

Tous les résultats obtenus dans le cas d’une source en trajet direct et de deux échos montrent donc que SEM est plus précis que MUSIC-FBSS pour des ENR allant de -11 dB à -18 dB. Dans le cas de haut ENR, SEM possède un biais et un écart-type non nul, ce qui rend la méthode moins précise que MUSIC-FBSS pour ces valeurs de ENR. Les résultats avec SEM utilisant les positions théoriques comme première estimation

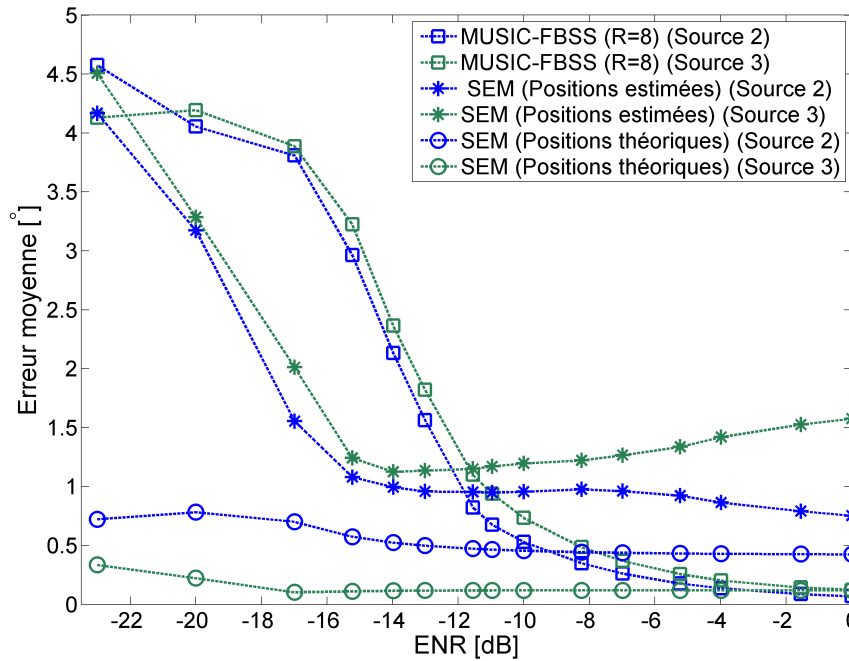


FIGURE 3.13 – Moyenne de l’erreur des bonnes détections de 200 essais Monte-Carlo en fonction du ENR pour un signal composé d’une source et de deux échos après retrait des cas de fausse alarme.

ont montré qu’une partie de l’erreur provient du test de puissance, qui n’est pas idéal lorsque l’on doit estimer la puissance de sources corrélées et de même puissance.

3.5 Conclusion

Dans ce chapitre, on a développé la méthode par élimination des sources, qui se base sur une estimation de la DOA et de la puissance de chacune des sources dans le but d’enlever, une par une, la contribution de chaque source de la matrice de covariance des signaux reçus pour ensuite estimer la matrice de covariance du bruit. Cette méthode donne toujours de bons résultats si la pré estimation des DOA est exacte. Toutefois, si la pré estimation est faite avec MUSIC, les résultats ne sont pas supérieurs à MUSIC avec FBSS lorsqu’il n’y a pas de corrélation entre les sources. L’ajout de corrélation entre les sources vient toutefois diminuer la précision de MUSIC-FBSS, mais n’affecte que peu SEM, qui devient alors beaucoup plus précise. En particulier, on a vu que dans le cas d’une source en trajet direct et de deux échos de même puissance, mais beaucoup plus faibles que la source principale, MUSIC-FBSS donne des résultats qui divergent des positions théoriques dès les ENR moyens, tandis que la méthode par élimination

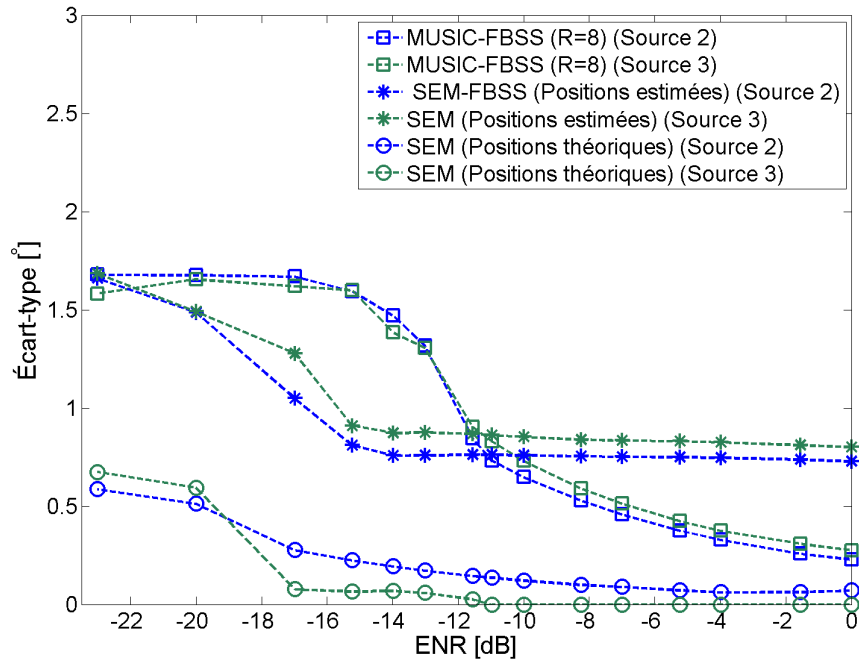


FIGURE 3.14 – Écart-type de l’erreur des bonnes détections de 200 essais Monte-Carlo en fonction du ENR pour un signal composé d’une source et de deux échos après retrait des cas de fausse alarme.

des sources donne des résultats fiables jusqu’à des ENR assez faibles. Ainsi, bien que SEM dépende beaucoup de la première estimation des sources, elle est plus fiable que MUSIC-FBSS à faible ENR lorsque le EDR est petit. On peut alors écrire l’algorithme 12, qui présente SEM et qui représente une des contributions du mémoire au domaine d’estimation de DOA. Il est à noter que dans cet algorithme, les étapes 8 à 13 peuvent être remplacées par la méthode de Newton-Raphson.

Algorithme 12 : Méthode par élimination des sources

```
1 Calculer  $\mathbf{R}_{xx}$ 
2 for  $w = 1$  to  $2$  do
3   for  $k=1$  to  $M$  do
4     Faire une décomposition des vecteurs propres pour obtenir  $\mathbf{V}$  et  $\lambda$ 
5     Définir le vecteur de référence  $\mathbf{V}_{ref_k}$ 
6     Utiliser MUSIC pour trouver  $\theta_k$ 
7     Calculer  $\mathbf{a}(\theta_k)$ 
8     for  $P = 0.5\lambda_k$  to  $1.5\lambda_k$  do
9       Calculer  $\hat{\mathbf{R}}_{kP} = P \mathbf{a}(\theta_k)\mathbf{a}(\theta_k)^\dagger$ 
10      Calculer  $\hat{\mathbf{R}}_{red_{kP}} = \hat{\mathbf{R}}_{red_{k-1}} - \hat{\mathbf{R}}_{kP}$ 
11      Calculer  $\lambda_1$  et  $\lambda_{N-k-1}$ , les valeurs propres de  $\hat{\mathbf{R}}_{red_{kP}}$ 
12    end
13    Déterminer  $P = \min(|\lambda_1| - |\lambda_{N-k-1}|)$  Calculer
14     $\hat{\mathbf{R}}_{red_k} = \hat{\mathbf{R}}_{red_{k-1}} - P\mathbf{a}(\theta_k)\mathbf{a}(\theta_k)^\dagger$ 
15  end
16  Calculer  $\hat{\mathbf{R}}_{nn} = \mathbf{R}_{xx} - \sum_k \hat{\mathbf{R}}_{red_k} = \hat{\mathbf{R}}_{xx} - \hat{\mathbf{R}}_{nn}$ 
17 end
```

Chapitre 4

MUSIC à focalisation de phase

Les méthodes d'estimation de directions d'arrivée pour des signaux à large bande que l'on a vu au chapitre 2 ne sont pas idéales. Les méthodes incohérentes, bien que simples à implanter, ont une faible résolution, tandis que les méthodes cohérentes demandent une grande quantité de calculs pour transformer les matrices de covariance, en plus de nécessiter une pré estimation de la DOA. Le but de ce chapitre est de présenter une nouvelle méthode permettant de déterminer la direction d'arrivée de sources à large bande pour alimenter l'algorithme SEM, qui ne demande pas beaucoup de puissance de calcul. Cette méthode est appelée MUSIC par focalisation de phase (MUSIC-FP).

4.1 Algorithme

On a vu dans le chapitre 2 que l'algorithme CSSM demande d'estimer la matrice de covariance pour chacune des fréquences pour ensuite les ramener à un cas à bande étroite en les multipliant par une matrice de focalisation. L'idée derrière MUSIC-FP est la même, mais au lieu de passer par des matrices de covariance, on utilise plutôt l'information contenue dans la phase pour réaliser cet objectif. On a déjà montré que le déphasage entre les capteurs du réseau pour une certaine bande spectrale est donné par la formule suivante :

$$\phi(f) = \frac{2\pi df \cos(\theta)}{v} \quad (4.1)$$

Dans le cas à bande étroite, le déphasage est constant dans tout le signal puisqu'on est en présence d'une seule fréquence. C'est pour cette raison que l'on peut simplement utiliser MUSIC sur l'enveloppe complexe du signal, dont le déphasage sera la même pour toutes les données en temps. Pour des signaux à large bande, ce ne sera plus le cas puisque la phase entre les capteurs dépend maintenant de la fréquence. Cette relation, que l'on peut voir à l'équation (4.1), est une droite avec une pente $\frac{2\pi d \cos(\theta)}{v}$ dans le cas où une seule source est présente dans le signal. Il est possible de retrouver l'information de la phase d'un signal à large bande en faisant les deux étapes suivantes :

1. Effectuer la transformée de Fourier rapide (FFT) du signal de sortie de chacun des capteurs.
2. Décomposer la FFT du signal en amplitude et en phase.
3. Effectuer un dépliage de la phase entre les capteurs

La dernière étape est très importante à effectuer puisque la phase calculée sur chacune des composantes spectrales sera contenue entre 0 et 2π , mais ce qui nous intéresse est la phase par rapport à celle du premier capteur. Cette phase fait un saut de $\phi(f)$ à chaque fois que l'on passe d'un capteur à l'autre, ce qui fait que le déphasage du capteur n par rapport au capteur 1 est :

$$\phi_n(f) = (n - 1) \frac{2\pi df \cos(\theta)}{v} \quad (4.2)$$

Si $\phi_n(f)$ dépasse 2π , la phase calculée au point 2 ne sera pas la vraie phase pour ce capteur puisque son déphasage avec le premier capteur ne sera pas $\phi_n(f)$. Le dépliage

de la phase entre les capteurs corrige ce problème.

Puisqu'il est possible de déterminer $\phi_n(f)$, il devient alors possible de modifier la phase pour se ramener à un cas à bande étroite en la multipliant par un terme γ qui ajuste la phase à une fréquence de focalisation f_0 . Ce terme est :

$$\gamma = \frac{f_0}{f} \quad (4.3)$$

En faisant la multiplication, la phase entre les capteurs devient :

$$\phi_0 = \phi(f) \cdot \gamma = \frac{2\pi f d \cos(\theta)}{v} \frac{f_0}{f} = \frac{2\pi f_0 d \cos(\theta)}{v} \quad (4.4)$$

La nouvelle phase obtenue ne dépend plus de la fréquence, à condition que la vitesse de l'onde soit constante pour toute la bande utile du signal. On peut alors construire un spectre avec la nouvelle phase focalisée, que l'on appelle le spectre focalisé. Mathématiquement, on écrit :

$$\mathbf{X}_0 = \mathbf{A}_0(\theta_m)\mathbf{S} + \mathbf{N}_0 \quad (4.5)$$

où \mathbf{A}_0 est la matrice directrice à la fréquence de référence f_0 , \mathbf{S} est la matrice de l'amplitude du spectre des sources et \mathbf{N}_0 est la matrice de covariance du bruit focalisée en phase. La matrice de covariance du spectre focalisé est alors :

$$\mathbf{R}_{xx_0} = \mathbf{A}_0(\theta_m)\mathbf{R}_{ss}\mathbf{A}_0(\theta_m)^\dagger + \mathbf{R}_{nn_0} \quad (4.6)$$

où \mathbf{R}_{ss} est la matrice de covariance des sources et \mathbf{R}_{nn_0} est la matrice de covariance du bruit focalisé. La forme de cette équation est exactement la même que celle de l'équation (2.14), ce qui implique qu'il devient possible d'utiliser MUSIC sur cette matrice pour en obtenir une estimation de la direction d'arrivée des sources.

En pratique, les signaux enregistrés sont discrets et il est nécessaire d'utiliser une FFT pour obtenir la transformée de Fourier discrète des signaux captés, ce qui implique de travailler en sous-bande. Il faut donc ajuster chaque phase $\phi_\xi(f)$ pour chaque ξ et chaque f telle que $f = \xi f_b$, telle que $f_b = \frac{f_e}{\Xi}$, où Ξ est le nombre d'échantillons utilisés pour faire la FFT. De plus, pour diminuer la quantité de calculs et limiter l'impact du bruit, on peut filtrer le spectre avant de focaliser la phase pour ne conserver que la bande utile du signal. Puisque l'on est déjà dans le domaine des fréquences rendu à ce point, le filtrage se résume à couper les fréquences en dehors de la bande utile.

4.2 Détermination de la fréquence de référence optimale

Dans la section précédente, on a défini f_0 comme étant la fréquence à laquelle on focalise la phase de toutes les autres fréquences. Le choix de cette fréquence de référence est donc très important à analyser, surtout qu'un mauvais choix pour sa valeur peut conduire à une erreur de l'estimation de la DOA.

Un premier critère à respecter pour le choix de f_0 est de s'assurer qu'il n'y ait pas présence de recouvrement spatial. Ce phénomène, qui est analogue à celui du recouvrement spectral, survient si la fréquence d'échantillonnage spatiale est trop petite, donc si la distance entre les capteurs est trop grande. En effet, si le déphasage focalisé entre deux capteurs adjacents dépasse 2π , un pic fantôme peut apparaître dans le pseudospectre, ce qui cause une fausse alarme. Pour éviter ce phénomène, il est nécessaire que :

$$-\pi \leq \phi_0 \leq \pi \quad (4.7)$$

Si l'on développe cette équation, on obtient :

$$-\pi \leq \frac{2\pi f_0 d \cos(\theta)}{v} \leq \pi \quad (4.8)$$

Sachant que $\cos(\theta)$ est compris entre -1 et 1, on obtient la condition suivante :

$$\frac{2\pi f_0 d}{\pi v} \leq 1 \Rightarrow \boxed{f_0 \leq \frac{v}{2d}} \quad (4.9)$$

Si l'on considère que le son voyage à 340,29 m/s, on peut calculer la fréquence f_{0max} pour différentes distances inter éléments. Le tableau 4.1 montre les valeurs pour quelques distances inter éléments.

Le développement de la condition sur f_0 permet aussi de voir une autre condition importante, mais cette fois-ci sur le signal. En effet, si la bande spectrale du signal sonore capté dépasse f_{0max} , il y aura du recouvrement spatial dans les signaux captés et cela peut induire des erreurs dans l'estimation d'arrivée. Il faut donc utiliser un réseau ayant une distance inter éléments suffisamment petite pour le signal mesuré.

Bien que la condition sur f_{0max} élimine une grande quantité de fréquences de référence, il reste une grande plage de fréquences de références possibles. Il est donc pertinent d'analyser quelle fréquence de référence fournit les meilleurs résultats. Pour ce faire, le plus simple est de considérer un cas sans bruit et de tester l'algorithme pour plusieurs

d	ω_{0max}	f_{0max}
[cm]	[rad/s]	[kHz]
1	106 940	17,02
5	21 382	3,40
7	15 273	2,43
10	10 690	1,70
15	7127	1,13
20	5345	0,851

TABLE 4.1 – Fréquences de références maximales pour différentes distances inter éléments.

fréquences de références. Pour générer des données, on utilise la méthode présentée à la section 3.2.1. Tel que l'on avait mentionné dans cette section, la largeur de bande du signal généré varie en fonction de la fréquence d'échantillonnage de la modulation f_{em} . Puisque, dans notre cas, on est intéressé à avoir une grande largeur de bande, on utilise $f_{em} = 3000$ Hz, que l'on multiplie avec une porteuse à une fréquence $f_p = 4000$ Hz. Le réseau utilisé est composé de 20 microphones dont la distance entre chaque microphone adjacent est de 2,5 cm. Dans ce cas, le signal est donc compris entre 2500 Hz et 5500 Hz, ce qui ne peut plus être considéré comme à bande étroite, et qui est plus faible de la fréquence limite du réseau, qui est d'environ 6500 Hz. En positionnant une seule source à $\theta = 70^\circ$, on obtient les résultats de la figure 4.1. Comme on peut voir sur cette figure, plus la fréquence de référence est élevée, plus le pic de la source est étroit, ce qui donne de meilleurs résultats. Cet effet vient du fait que plus la fréquence de référence est élevée, plus le déphasage ϕ entre les capteurs l'est aussi. Lors de la formation du vecteur d'analyse, une grande valeur de ϕ a fait que pour chaque incrément de θ , il y a une grande différence dans le vecteur d'analyse $\mathbf{a}(\theta)$, ce qui amincit le pic de la source

Bien que la figure 4.1 montre que MUSIC-FP fonctionne bien et que le choix idéal de f_0 semble être la fréquence limite du réseau, il est intéressant de vérifier le cas où l'on est en présence de plusieurs sources. Ainsi, on refait une simulation avec les mêmes paramètres, mais en ajoutant cette fois-ci une seconde source, indépendante de la première, à 120° . Les résultats obtenus se retrouvent à la figure 4.2.

Sur cette figure, on peut voir des résultats bien différents que ceux obtenus dans le cas d'une seule source. On remarque en effet que la fréquence de référence qui donne les meilleurs résultats est 4000 Hz, ce qui correspond exactement à la fréquence de la porteuse du signal. Si l'on augmente la fréquence de référence à 5500 Hz, les résultats sont beaucoup moins précis. Cette observation s'explique par le fait que l'ajout d'une

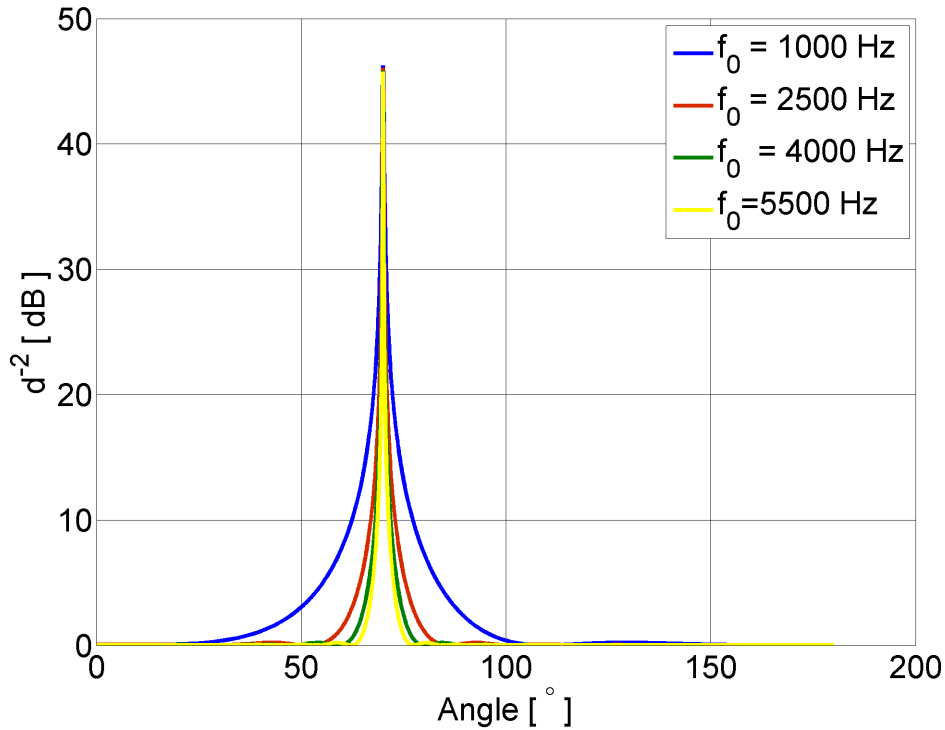


FIGURE 4.1 – Impact du choix de la fréquence de référence sur le pseudospectre obtenu à l’aide de MUSIC-FP lorsque le signal mesuré est composé d’une seule source.

deuxième source vient affecter d’une grande façon $\phi_0(f)$, qui n’est plus une droite selon l’indice des capteurs, mais devient plutôt, pour chaque fréquence, une valeur moyenne pondérée par l’amplitude des deux signaux pour chacune des fréquences. On sait toutefois que la porteuse représente exactement le centre du spectre du signal. Il est donc fort possible que la symétrie du spectre soit un élément très important pour le choix de la fréquence de référence. Pour cette raison, on utilisera toujours cette fréquence de référence pour MUSIC-FP.

4.3 Essais Monte-Carlo

La section précédente a servi à déterminer quelle était la meilleure fréquence de référence à utiliser, mais aussi à tester la méthode pour être certain qu’elle permet bien de déterminer la direction d’arrivée de différentes sources. Tout comme on a fait au chapitre 3, il est intéressant d’évaluer les limites de la méthode en vérifiant son erreur moyenne et son écart-type avec de nombreux essais Monte-Carlo. En plus de faire ces essais Monte-Carlo sur MUSIC-FP, il est aussi pertinent de comparer la performance de cette

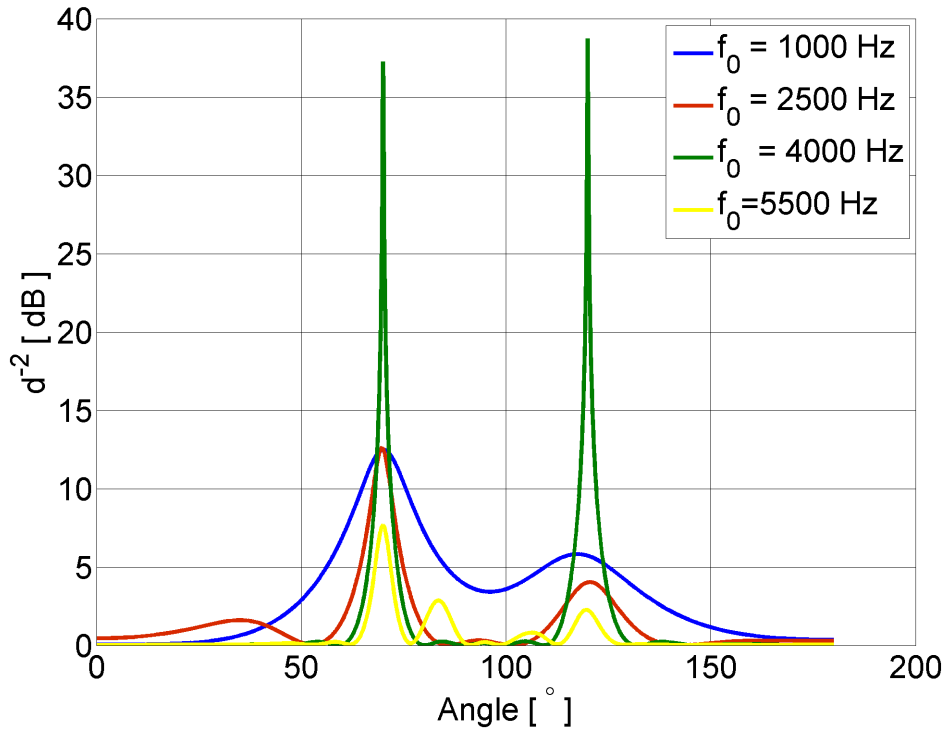


FIGURE 4.2 – Impact du choix de la fréquence de référence sur le pseudospectre obtenu à l’aide de MUSIC-FP lorsque le signal mesuré est composé de deux sources.

méthode avec celle d’une autre méthode connue, telle que CSSM. Puisque cette méthode est itérative, on va analyser deux cas différents ; un premier avec une seule itération en utilisant MVDR-S comme première estimation de la DOA et un second avec une seule itération, mais en utilisant les positions théoriques comme estimation préliminaire. Ce second cas permet alors d’avoir les meilleurs résultats possible avec CSSM, ce qui est intéressant pour voir la convergence de l’algorithme.

Tout comme on a fait au chapitre 3, on fait les simulations Monte-Carlo pour différents SNR, ce qui permet d’analyser jusqu’à quel niveau de bruit l’algorithme fonctionne convenablement. Dans ce chapitre, les algorithmes qui faisaient l’objet d’analyses nous permettaient de générer simplement une matrice de covariance du bruit et de l’ajouter à la matrice de covariance des signaux reçus sans bruit. Dans le cas de MUSIC-FP, il n’est pas possible d’utiliser cet algorithme puisque l’on doit ajuster les signaux eux-mêmes avant de calculer la matrice de covariance. Si l’on ajoutait le bruit après toutes ces transformations, on ne simulerait pas un cas réaliste de signaux bruités. Ainsi, il est nécessaire de générer des signaux de bruit corrélé, ce que l’on peut faire avec la méthode présentée à l’algorithme 11 de la section 3.2.3. Tout comme pour le chapitre 3, on utilise

des vecteurs de longueur $l_b = 50$, ce qui permet d'avoir des données corrélées et ainsi nuire aux algorithmes de DOA.

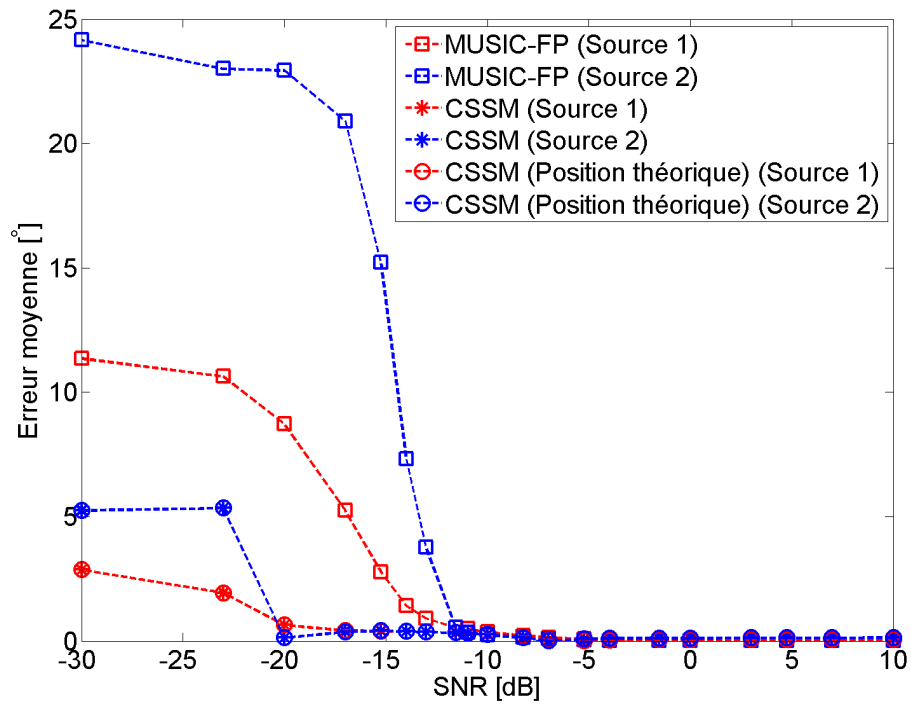
4.3.1 Signaux non corrélés

Comme première série de simulations, on fait 200 essais Monte-Carlo pour un cas composé de deux signaux. Ces signaux ont tous les deux une porteuse de 4000 Hz et leur signal modulant aléatoire est échantillonné à 3000 Hz, ce qui fait que la bande utile des deux signaux est située entre 2500 Hz et 5500 Hz. Puisque le signal modulant entre les deux signaux est aléatoire et est échantillonné à une fréquence élevée ils sont pratiquement, indépendants l'un de l'autre. Les résultats obtenus pour la moyenne de l'erreur se retrouvent à la figure 4.3, pour l'écart-type à la figure 4.4, tandis que les résultats pour la probabilité de fausse alarme, que l'on définit comme les cas où aucune source n'est positionnée à moins de 10° de la position réelle de la source, se retrouvent à la figure 4.5.

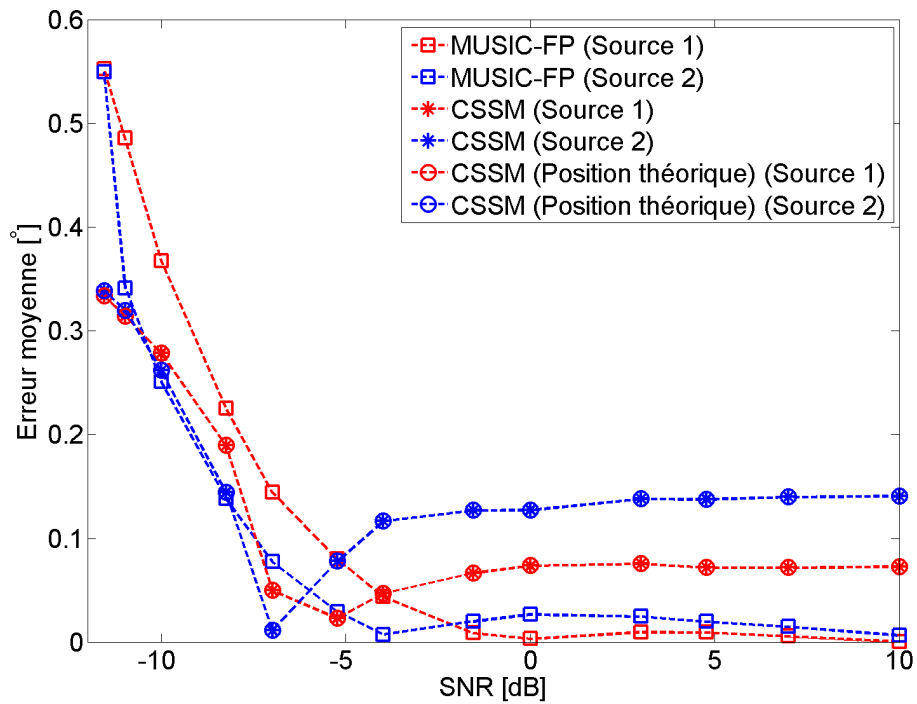
À haut SNR, on observe que l'erreur moyenne et l'écart-type sont très faibles, peu importe la méthode utilisée. Toutefois, si l'on observe la figure 4.4b, on remarque que MUSIC-FP a un biais plus petit que CSSM et un écart-type qui croît plus lentement. Dans de telles conditions, il est donc pratique d'utiliser MUSIC-FP au lieu de CSSM.

Lorsque l'on augmente le niveau de bruit à -10 dB, on remarque que MUSIC-FP perd beaucoup en efficacité et est, à ce point, aussi efficace que CSSM, que ce soit en biais qu'en écart-type. Passé ce point, le biais de MUSIC-FP devient significativement élevé, ce qui rend la méthode très difficile à utiliser. Si l'on observe la figure 4.5, on remarque qu'à partir de -12 dB, il y a une probabilité de 10% d'avoir une fausse alarme, ce qui est logique puisque l'écart-type à ce point est rendu significatif. On considère donc qu'à partir d'un SNR plus faible que -12 dB, il est difficile d'utiliser MUSIC-FP.

Les résultats obtenus à l'aide de l'algorithme CSSM peuvent porter à confusion puisque l'erreur moyenne reste relativement petite, même lorsque le SNR est très faible. Toutefois, si l'on observe les résultats pour l'écart-type, on remarque que cette dernière devient extrêmement élevée lorsque l'on dépasse -20 dB. La probabilité de fausse alarme suit le même comportement que l'écart-type, soit qu'il soit nul à haut SNR et augmente beaucoup passé -20 dB. Malgré cela, on peut conclure qu'à faible SNR, CSSM reste plus précise que MUSIC-FP, surtout considérant que si l'erreur moyenne est petite, il reste possible de faire plusieurs épreuves et les moyenner pour améliorer la précision des résultats obtenus. On remarque aussi que la différence entre la pré estimation avec MVDR-S

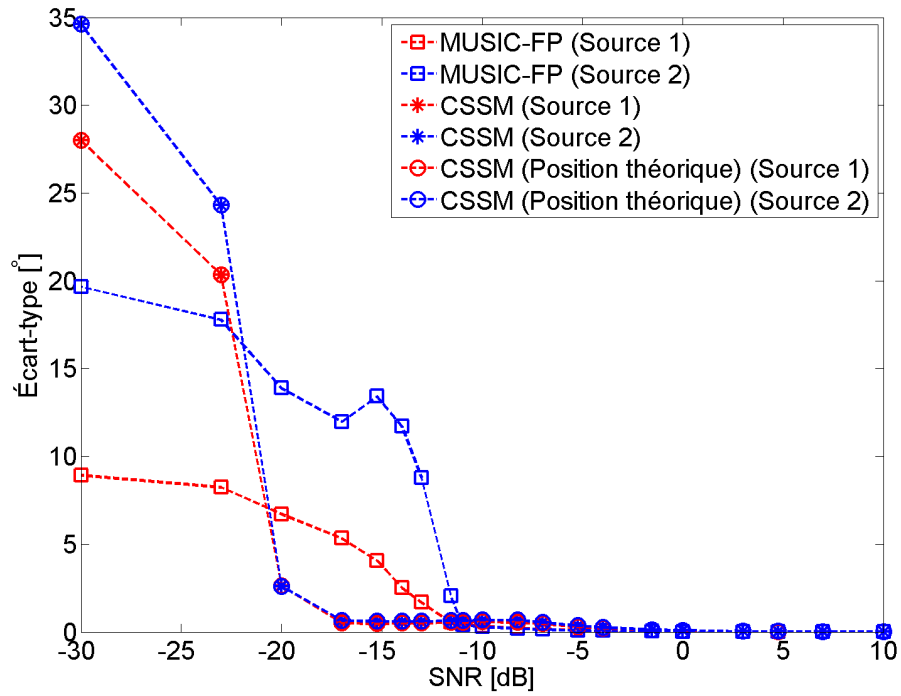


(a) Résultats complets

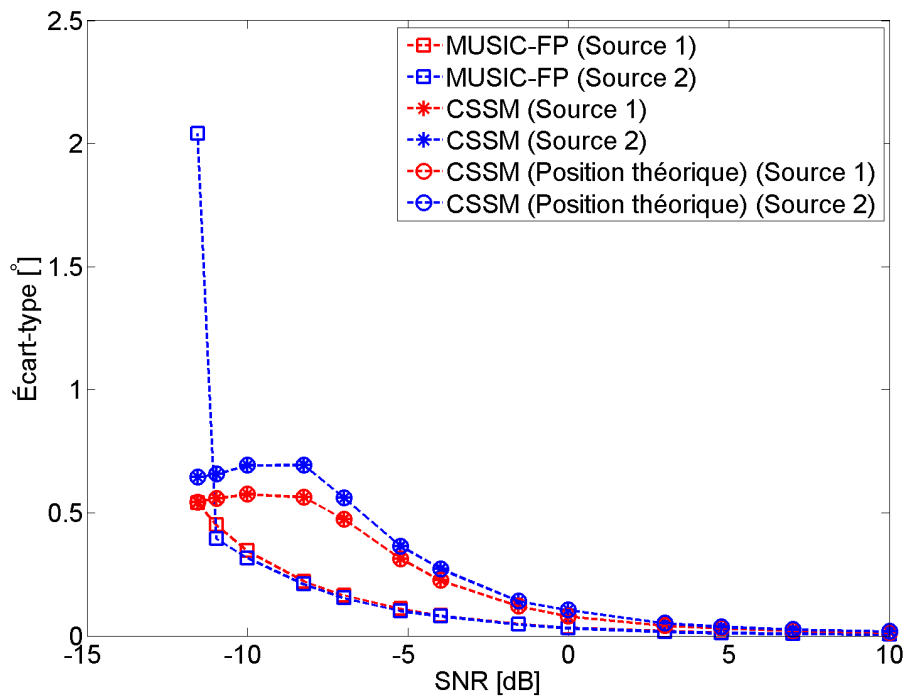


(b) Résultats agrandis pour montrer la différence pour les plus grandes valeurs de SNR.

FIGURE 4.3 – Moyenne de l'erreur de 200 essais de Monte-Carlo en fonction du rapport signal à bruit pour des signaux non corrélés.



(a) Résultats complets



(b) Résultats agrandis pour montrer la différence pour les plus grandes valeurs de SNR.

FIGURE 4.4 – Écart-type de l'erreur de 200 essais de Monte-Carlo en fonction du rapport signal à bruit pour des signaux non corrélés.

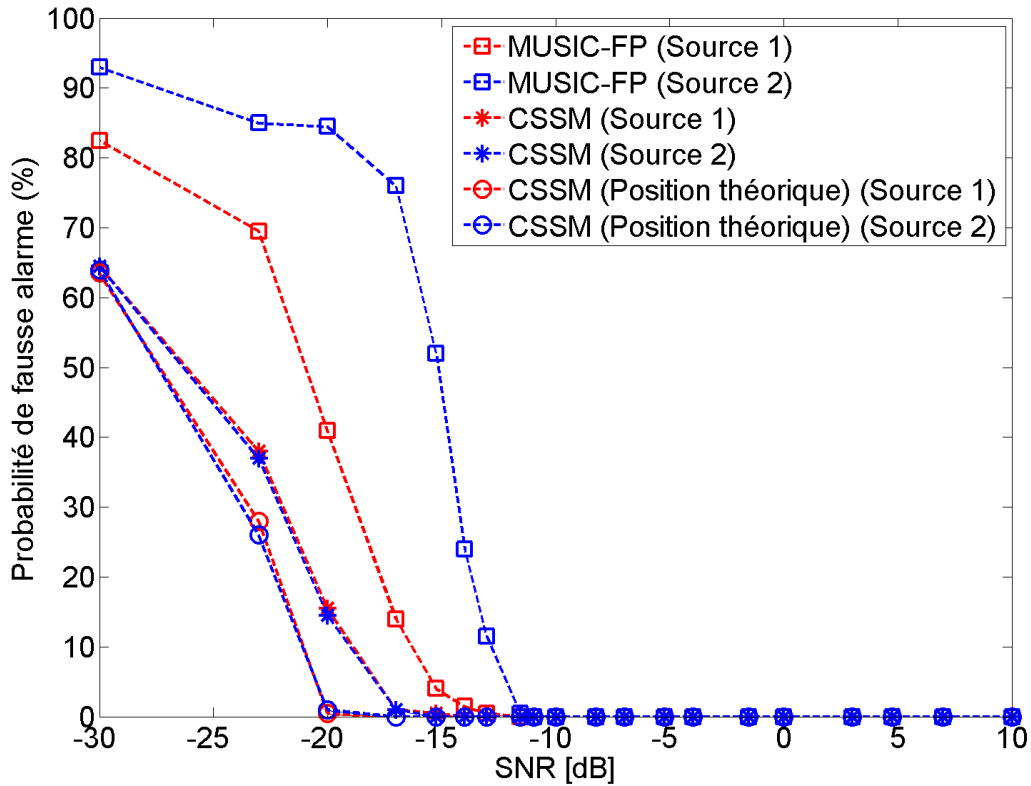


FIGURE 4.5 – Probabilité de fausse alarme en fonction du SNR de 200 essais de Monte-Carlo avec deux sources non corrélées.

ou avec l'utilisation des positions théoriques est très petite, démontrant qu'il n'est pas nécessaire de faire une boucle de rétroaction pour obtenir les meilleurs résultats possible avec CSSM.

Une dernière observation intéressante que l'on peut faire avec ces résultats est la très faible erreur moyenne à -7 dB pour l'algorithme CSSM. Il est difficile d'en déterminer la cause exacte, mais il est fort possible que cela provienne d'une erreur statistique. Il serait intéressant de faire un plus grand nombre d'essais Monte-Carlo pour le vérifier. On remarque toutefois que l'écart-type est assez élevé à ce niveau de bruit, ce qui montre que les résultats ne sont pas nécessairement meilleurs que ceux à plus faible SNR, seulement que leur biais est très petit.

4.3.2 Signaux corrélés

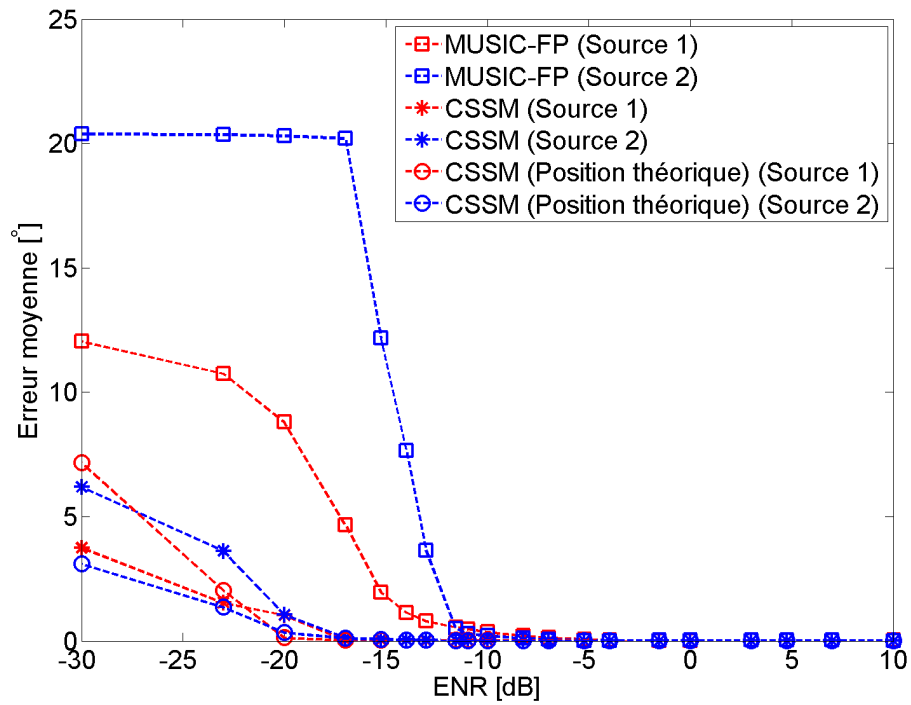
La forme de l'équation (4.6) montre qu'une corrélation entre les sources aura le même effet pour MUSIC-FP pour des signaux à large bande que pour MUSIC régulier pour

des signaux à bande étroite. Il devient alors nécessaire d'effectuer du lissage spatial dans le but de réduire l'impact de la corrélation entre les sources. On reprend alors les simulations de la section précédente, mais cette fois-ci en générant un seul signal pour les deux sources, ce qui implique une corrélation complète entre les deux sources. On utilise FBSS en divisant en 8 sous-réseaux la matrice de covariance focalisée générée avec MUSIC-FP. On fait aussi le même traitement pour la matrice de covariance focalisée de CSSM, puisque cette méthode est aussi sensible à la corrélation entre les sources. Les résultats pour l'erreur moyenne se retrouvent à la figure 4.6, pour l'écart-type de l'erreur à la figure 4.7 et pour la probabilité de fausse alarme à la figure 4.8.

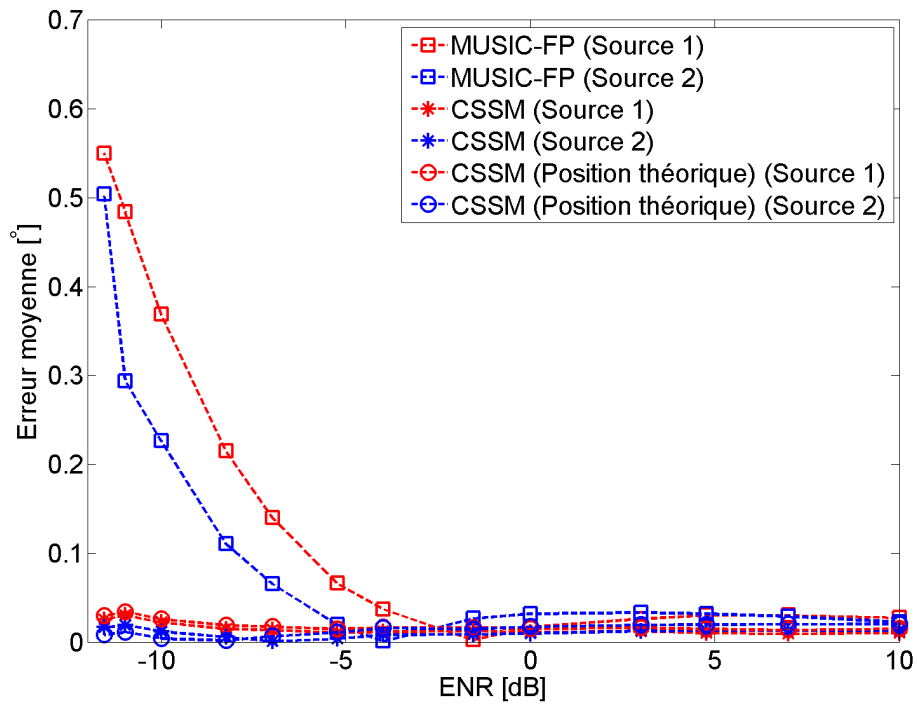
Les résultats présents sur la figure 4.6 montrent qu'à SNR élevé, il n'y a plus de biais. Cette amélioration vient du fait que FBSS fait une moyenne des matrices de covariances de différents sous-réseaux, ce qui améliore les résultats. Il se peut que d'ajouter le même lissage spatial au cas sans corrélation puisse enlever le biais qui était présent dans les simulations faites à la section précédente. On remarque aussi que si les angles de départ sont les positions exactes, la différence sur le biais est très faible. À partir de -12 dB, on remarque que le biais de MUSIC-FP augmente de façon significative et la méthode devient peu fiable, tandis que CSSM a une erreur assez faible pour toutes les valeurs de SNR calculées. À la figure 4.7, on peut remarquer que l'écart-type de CSSM augmente beaucoup plus lentement par rapport à celui de MUSIC-FP, surtout à haut SNR, ce qui rend la méthode CSSM plus attrayante que MUSIC-FP pour cette plage de SNR. À partir de -20 dB, on remarque toutefois que l'écart-type augmente très rapidement, peu importe si l'on fait CSSM à l'aveuglette ou pas. Il s'agit donc de la limite maximale de son utilisation. La figure 4.8 résume les résultats obtenus sur les deux autres figures, soit que MUSIC-FP commence à perdre en fiabilité à partir de -12 dB, tandis que CSSM est fiable jusqu'à environ -20 dB.

4.4 Conclusion

Dans ce chapitre, on a développé un nouvel algorithme qui permet d'estimer la DOA de sources à large bande qui s'appelle MUSIC-FP. Pour utiliser adéquatement cet algorithme, il est nécessaire de connaître la fréquence de la porteuse du signal capté et d'utiliser cette fréquence pour focaliser le déphasage entre les capteurs. Les résultats de 200 simulations Monte-Carlo pour des sources non corrélées montrent que sans l'utilisation de lissage spatial, MUSIC-FP est plus précise que CSSM à des SNR plus hauts que -10 dB, mais que pour des valeurs plus faibles que -12 dB, la méthode est ren-

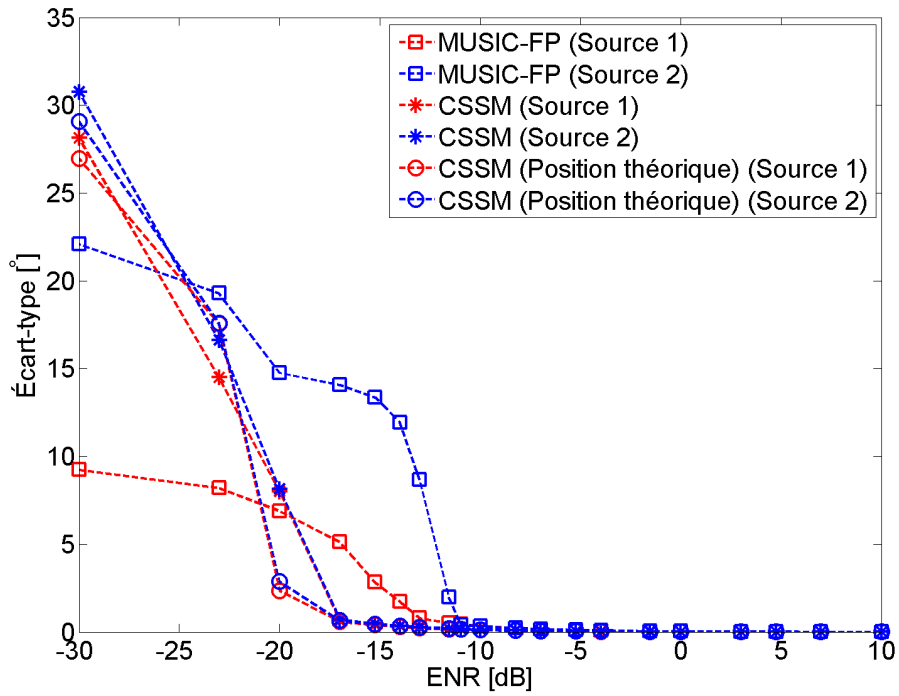


(a) Résultats complets

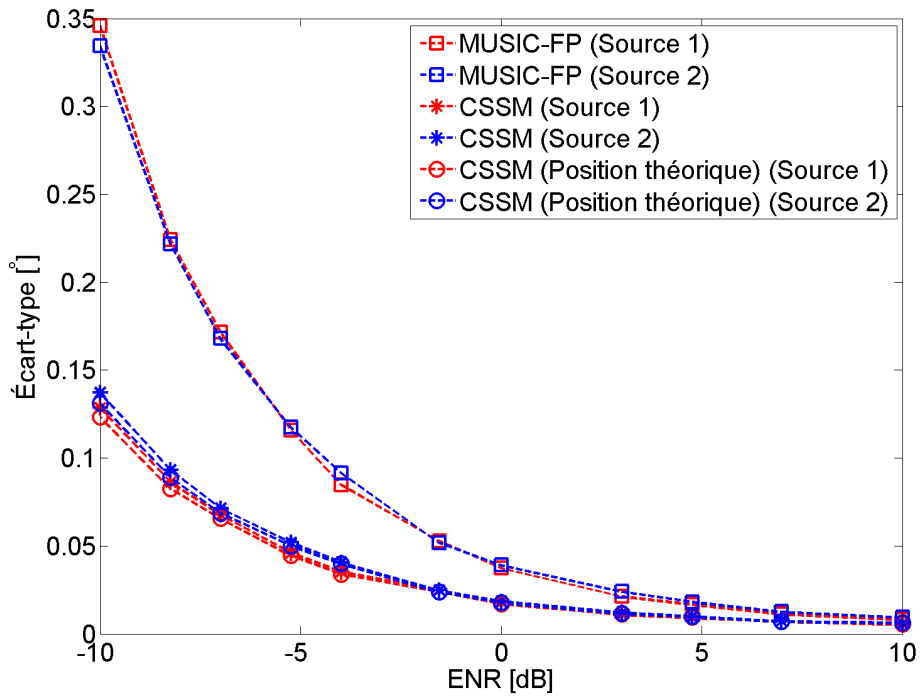


(b) Résultats agrandis pour montrer la différence pour les plus grandes valeurs de SNR.

FIGURE 4.6 – Moyenne de l'erreur de 200 essais Monte-Carlo en fonction du rapport signal à bruit pour des signaux corrélés.



(a) Résultats complets



(b) Résultats agrandis pour montrer la différence pour les plus grandes valeurs de SNR.

FIGURE 4.7 – Écart-type de 200 essais Monte-Carlo en fonction du rapport signal à bruit pour des signaux corrélés.

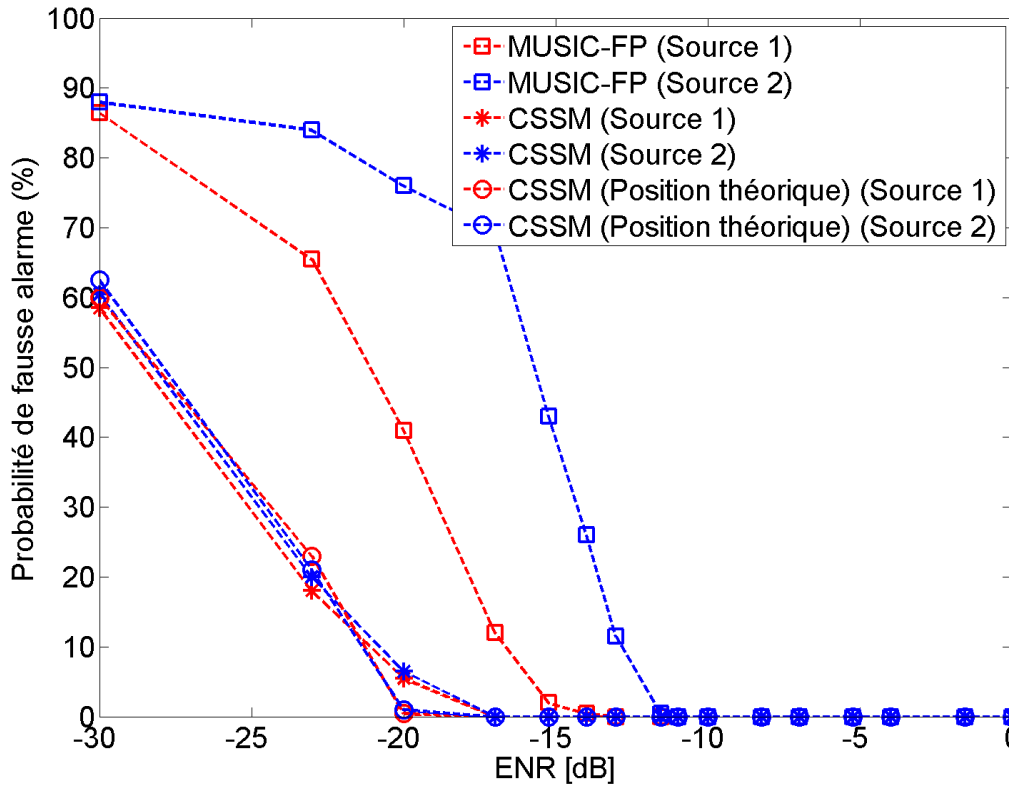


FIGURE 4.8 – Probabilité de fausse alarme en fonction du SNR de 200 essais Monte-Carlo avec deux sources corrélées

due inutilisable. La méthode CSSM, quant à elle, reste très efficace jusqu'à -20 dB, où l'écart-type devient très élevé.

Les 200 essais de Monte-Carlo pour des sources corrélées ont mis en lumière l'importance de l'utilisation de lissage spatial pour CSSM. En effet, le biais observé à haut SNR a disparu, tandis que l'écart-type et le biais pour de plus faibles valeurs de SNR sont restés semblables. La méthode MUSIC-FP, quant à elle, reste autant précise que dans le cas sans corrélation si l'on utilise du lissage spatial. Toutefois, on peut remarquer que CSSM est désormais la meilleure méthode des deux puisqu'elle est plus précise et moins variante pour presque toutes les valeurs de SNR. L'algorithme 13, présent ci-dessous, résume toutes les étapes de MUSIC-FP qui fut développé lors de ces travaux de maîtrise.

Algorithme 13 : MUSIC à focalisation de phases

- 1 Calculer la FFT de \mathbf{X} , la matrice des signaux reçus.
 - 2 Calculer \mathbf{f} , le vecteur des fréquences comprises dans la FFT.
 - 3 Couper les fréquences qui sont en dehors de l'intervalle $[f_L, f_H]$.
 - 4 Déterminer f_0 , la fréquence de la porteuse.
 - 5 Calculer $|\mathbf{X}(f)|$, l'amplitude de tous les points de la FFT qui n'ont pas été coupés.
 - 6 Calculer $\phi(f)$, la phase de la FFT pour chaque fréquence pour chaque capteur.
 - 7 Calculer $\phi_d(f)$, la phase dépliée de $\phi(f)$.
 - 8 Calculer $\gamma = \frac{f_0}{\mathbf{f}}$
 - 9 Focaliser la phase en faisant la multiplication $\phi_0 = \gamma \cdot \phi_d(f)$.
 - 10 Calculer les signaux reçus modifiés avec la formule suivante : $\mathbf{X}_0 = |\mathbf{X}(f)| \exp(\mathbf{j}\phi_0)$.
 - 11 Calculer $\mathbf{R}_{xx_0} = \mathbf{X}_0 \mathbf{X}_0^\dagger$.
 - 12 Utiliser MUSIC sur \mathbf{R}_{xx_0} pour déterminer les DOA.
-

Chapitre 5

Estimation de DOA de signaux acoustiques réfléchis

Dans les deux chapitres précédents, on a développé indépendamment deux algorithmes ayant chacun une fonction différente. Le premier, appelé la méthode par élimination des sources (SEM), a pour but de réduire l'influence d'une source très puissante et du bruit ambiant sur la matrice de covariance des signaux reçus pour améliorer la résolution de l'estimation de DOA d'échos. Le second, qui s'appelle MUSIC-FP, focalise la phase de signaux à large bande pour leur donner des propriétés de signaux à bande étroite, pour ensuite utiliser MUSIC pour déterminer leur DOA. Dans ce chapitre, l'objectif est de faire la combinaison de ces deux algorithmes pour améliorer les résultats dans le cas où le signal est composé d'une source à large bande très puissante et d'échos beaucoup moins puissants. Pour ce faire, on va commencer par analyser le comportement de CSSM et MUSIC-FP, dans le cas où l'on est en présence d'une source puissante et un écho. Par la suite, on fera les mêmes tests que dans le chapitre 3 pour vérifier si le critère retenu dans ce chapitre pour déterminer la puissance des sources se prête bien au cas à large bande. Finalement, on effectuera des essais Monte-Carlo pour la combinaison des deux méthodes.

5.1 Sources de puissances différentes

Avant de commencer l'intégration de SEM avec les autres méthodes à large bande, il est intéressant d'observer le comportement de chacune des méthodes au fur et à mesure que la puissance de la deuxième source diminue. Pour avoir des résultats comparables à ceux obtenus dans le chapitre 4, on utilise les mêmes paramètres, soit une fréquence

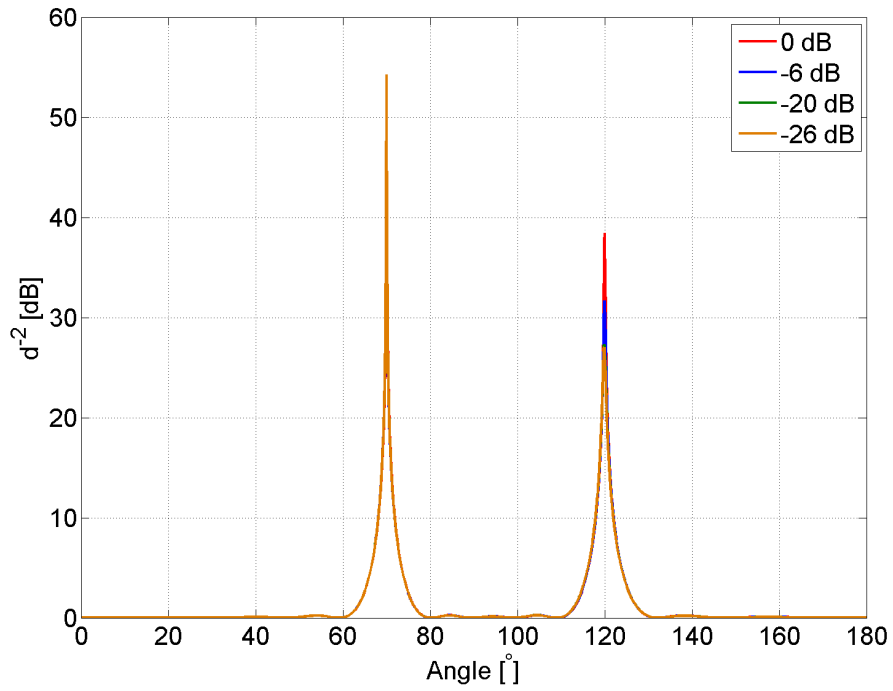
d'échantillonnage de signal modulant de 3 kHz, une porteuse de 4 kHz, une fréquence d'échantillonnage du signal de 50 kHz pour deux sources placés à $\theta_1 = 70^\circ$ et $\theta_2 = 120^\circ$. Les sources simulées sont indépendantes et leur rapport de puissance varie de 0 dB à 26 dB. Les résultats pour MUSIC-FP se retrouvent à la figure 5.1a et ceux pour CSSM se retrouvent à la figure 5.1b.

Les résultats pour MUSIC-FP montrent que cet algorithme fonctionne bien, peu importe le rapport de puissance entre les sources. On remarque toutefois que CSSM a plus de difficulté lorsque la puissance de la seconde source diminue puisque le pic pour cette source n'est pas situé à la bonne direction d'arrivée si le rapport de puissance entre les sources est de -20 dB ou moins. On voit donc que les résultats obtenus au chapitre précédent concernant CSSM ne sont valides que lorsque la puissance des sources est comparable. Bien que ces résultats semblent compromettre la possibilité d'utiliser SEM avec CSSM, il est nécessaire de faire d'autres tests pour s'en assurer. Pour ce qui est de MUSIC-FP, les résultats sont prometteurs et montrent que l'étape de focalisation de la phase conserve l'information de la source faible.

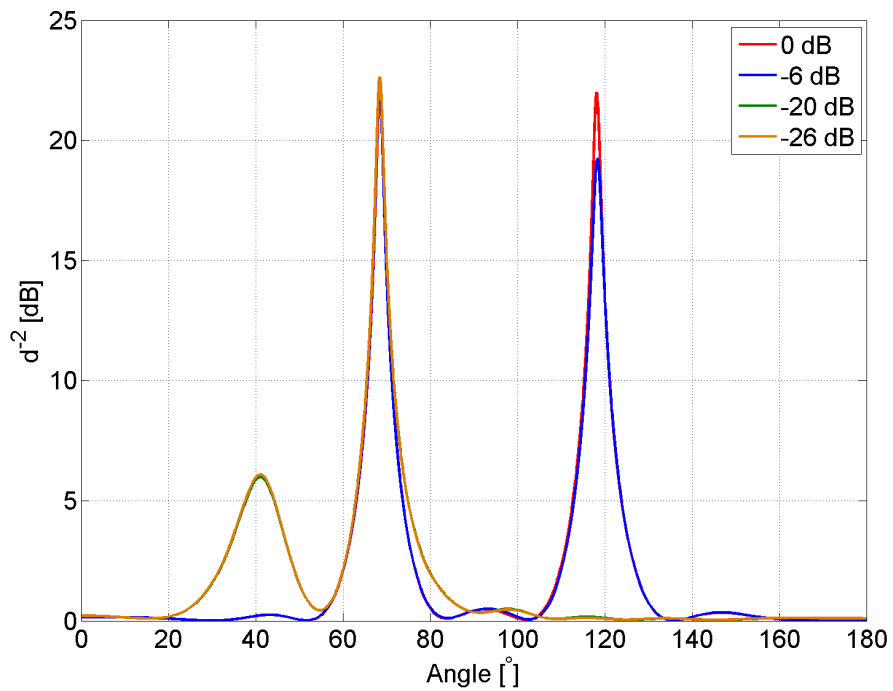
5.2 Test de la puissance des sources

Pour commencer l'ajout de SEM aux méthodes à large bande, il est important de vérifier que le test de la puissance des sources défini au chapitre 3 fonctionne pour le cas à large bande. Si tel est le cas, il sera facile d'intégrer les méthodes ensemble. Toutefois, si ce n'est pas le cas, il sera nécessaire de trouver un nouveau test de puissance ou bien abandonner l'idée de mélanger les deux méthodes. Si l'on utilise exactement les mêmes paramètres que dans la section précédente pour générer les signaux, on obtient les résultats présents à la figure 5.2a pour le test de puissance faite sur MUSIC-FP et, à la figure 5.2b, pour CSSM. Dans les deux cas, les signaux ne sont pas corrélés, ont un rapport de puissance de -13 dB. Pour SEM, on utilise les positions exactes et dans le but d'avoir le comportement idéal.

La courbe présente à la figure 5.2a montre que le test de puissance fonctionne bien pour l'algorithme MUSIC-FP, tandis que la figure 5.2b montre une courbe qui a une allure bien différente. Premièrement, l'axe des abscisses n'est pas l'erreur sur la puissance, mais bien la puissance enlevée. La raison est que la transformation faite pour CSSM empêche d'estimer la puissance théorique des sources. Deuxièmement, on peut aussi remarquer que seulement la dernière valeur propre varie en fonction de la puissance estimée. La raison de ce résultat est le fait que la transformation modifie la structure

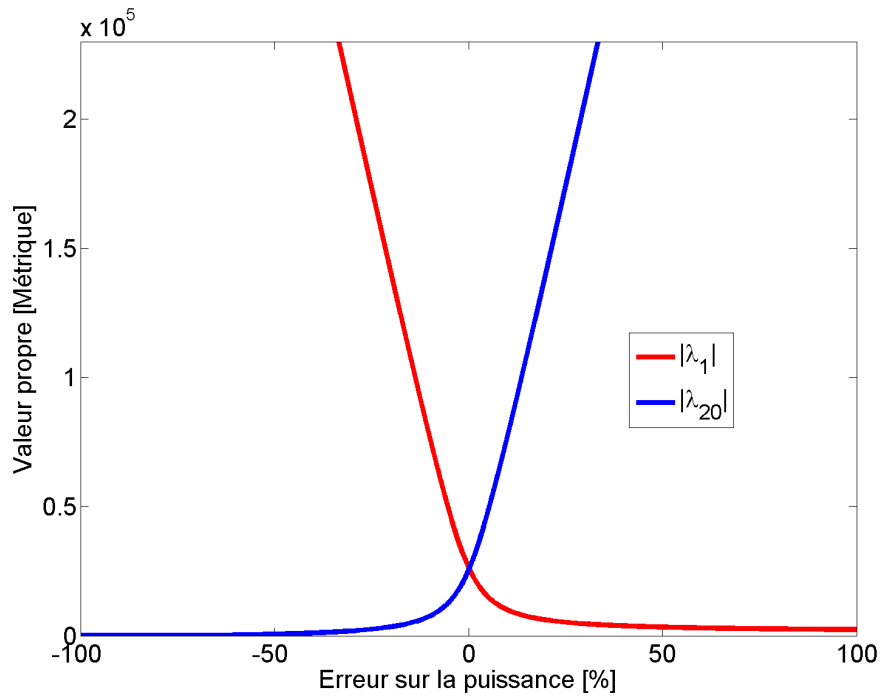


(a) MUSIC-FP

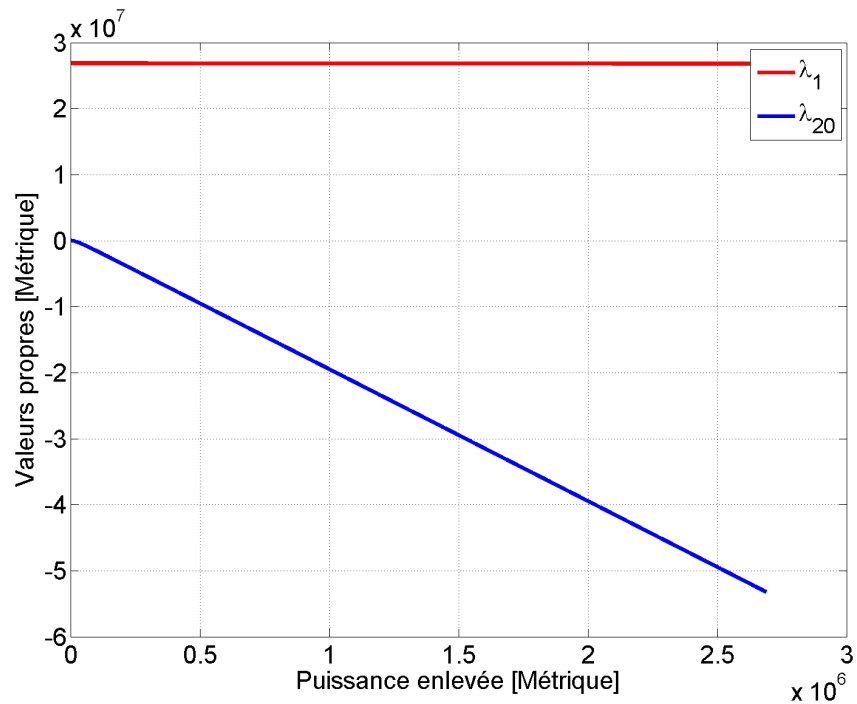


(b) CSSM

FIGURE 5.1 – Effet de la variation du rapport de puissance entre deux sources à large bande et non corrélées pour deux méthodes différentes, MUSIC-FP et CSSM, dans un cas sans bruit.



(a) MUSIC-FP



(b) CSSM

FIGURE 5.2 – Effet de la variation du rapport de puissance entre deux sources indépendantes à large bande et non corrélées pour deux méthodes différentes, MUSIC-FP et CSSM, dans un cas sans bruit.

de la matrice de covariance et que cette dernière ne suit pas exactement l'équation 3.1, donc la réduction de la matrice de covariance obtenue avec CSSM n'est pas possible avec la méthode utilisée avec SEM. Ainsi, lorsque l'on enlève la contribution de la source 1, on enlève une matrice de covariance qui n'est pas contenue dans la matrice de covariance que l'on analyse, ce qui a pour effet de faire diminuer la dernière valeur propre, qui devient alors négative. On peut aussi remarquer que l'ordre de grandeur des valeurs propres pour les deux algorithmes aux figures 5.2a et 5.2b est très différente. Cette grande différence peut sembler étrange puisque les signaux pour les simulations ont été générés de la même manière avec les mêmes paramètres. Toutefois, si l'on retourne voir l'algorithme 6 (CSSM), on peut voir qu'à l'étape 15, on additionne une grande quantité de matrices de covariance après avoir fait une transformation. Cette addition augmente la puissance perçue par MUSIC, donc les valeurs propres. Pour MUSIC-FP, puisque l'on ajuste directement les phases pour ensuite faire la matrice de covariance, on ne touche pas à la puissance du signal. Les résultats peu concluants pour la combinaison de SEM avec CSSM font qu'on va se contenter, pour le reste de ce mémoire, de combiner SEM avec MUSIC-FP.

Bien que la courbe présente à la figure 5.2a est très prometteuse, il est intéressant de voir le comportement des courbes si l'on augmente le niveau de bruit. La figure 5.3 présente le comportement de la première et de la dernière valeur propre pour différents SNR.

Sur cette figure, on remarque que les courbes de valeur absolue des deux valeurs propres se croisent à une erreur d'environ 0.1%, ce qui est très faible. On peut aussi remarquer, sur cette figure, qu'il faut que le SNR soit plus faible que -20 dB pour qu'il ait un impact important sur le point de rencontre des courbes. Évidemment, ces figures sont toutes basées sur une seule simulation et ne permettent pas de statuer sur la performance de l'algorithme en faible SNR.

5.3 Simulations Monte-Carlo

5.3.1 Sources non corrélées

Maintenant que l'on a un algorithme complet, il est important de vérifier ses limites. Pour ce faire, on fait, à l'instar des chapitres précédents, 200 essais Monte-Carlo. Pour commencer, on génère les données de deux sources non corrélées situées à $\theta_1 = 70^\circ$ et $\theta_2 = 120^\circ$. Le rapport de puissance entre les deux sources est de -26 dB, ce qui a pour

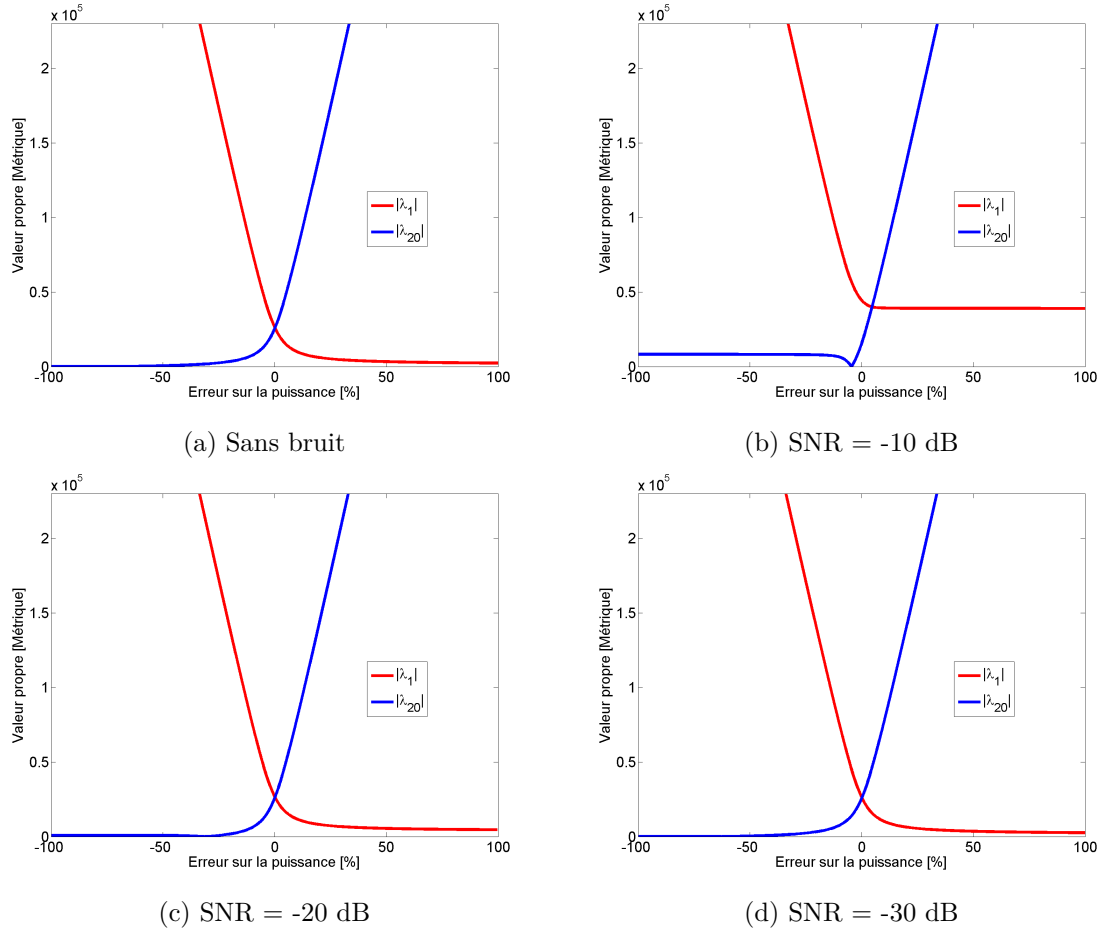


FIGURE 5.3 – Valeurs propres de la matrice de covariance réduite en fonction de l’erreur sur la puissance enlevée pour différentes valeurs de SNR.

effet que l’estimation de la DOA pour la première source est très précise puisque le SNR est calculé à partir de la puissance de la source la plus faible. Ainsi, on ne présentera pas les résultats pour la source 1 dans les graphiques pour mieux visualiser ceux de la source 2. Le réseau simulé est composé de $N = 20$ capteurs, séparés d’une distance $d = 2.5$ cm, ce qui a pour effet que la bande utile du signal généré ne doit pas dépasser 6.8 kHz. Tout comme dans le chapitre 4, le signal modulant de la source est échantillonné à une fréquence de 3000 Hz, et la porteuse est à une fréquence de 4000 Hz. Ainsi, le signal utile est situé entre 2500 Hz et 5500 Hz, ce qui ne dépasse pas la fréquence limite du réseau. Dans le but d’améliorer les résultats en faisant une moyenne, on utilise FBSS avec 8 sous-réseaux pour MUSIC-FP utilisé seul. Les résultats de l’erreur moyenne se retrouvent à la figure 5.4, ceux de l’écart-type à la figure 5.5, et ceux de la probabilité de fausse alarme à la figure 5.6. Tout comme pour les chapitres précédents, on définit une fausse alarme comme un cas où il n’y a pas de source à moins de 10° de la position

théorique.

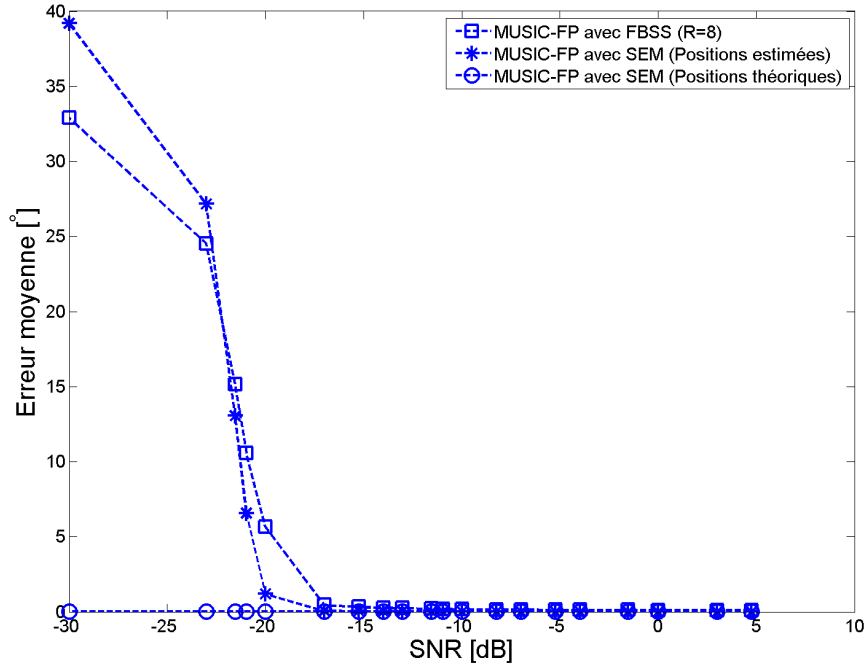
À la figure 5.4, où l'on peut voir l'erreur moyenne pour une grande plage de SNR, les résultats sont légèrement meilleurs avec SEM puisque les résultats sont relativement bons jusqu'à un SNR d'environ -20 dB, tandis que MUSIC-FP avec FBSS directement donne de bons résultats jusqu'à environ -16 dB. Si l'on observe plus attentivement les résultats à fort SNR, présentés à la figure 5.4a, on peut remarquer que MUSIC-FP utilisé seul possède un biais de 0.1° , tandis que SEM élimine ce biais. On remarque aussi que la courbe avec SEM commence à diverger à des valeurs de SNR beaucoup plus faibles que MUSIC-FP seul, ce qui montre l'utilité de la combinaison des méthodes.

En observant maintenant l'erreur sur la position de la première source, on voit que peu importe la méthode, la DOA est pratiquement parfaite pour tous les SNR. Ce résultat est bien ce qui était attendu puisque la différence de puissance entre les sources est tellement élevée qu'un bruit de -30 dB par rapport à la deuxième source reste faible pour la première source.

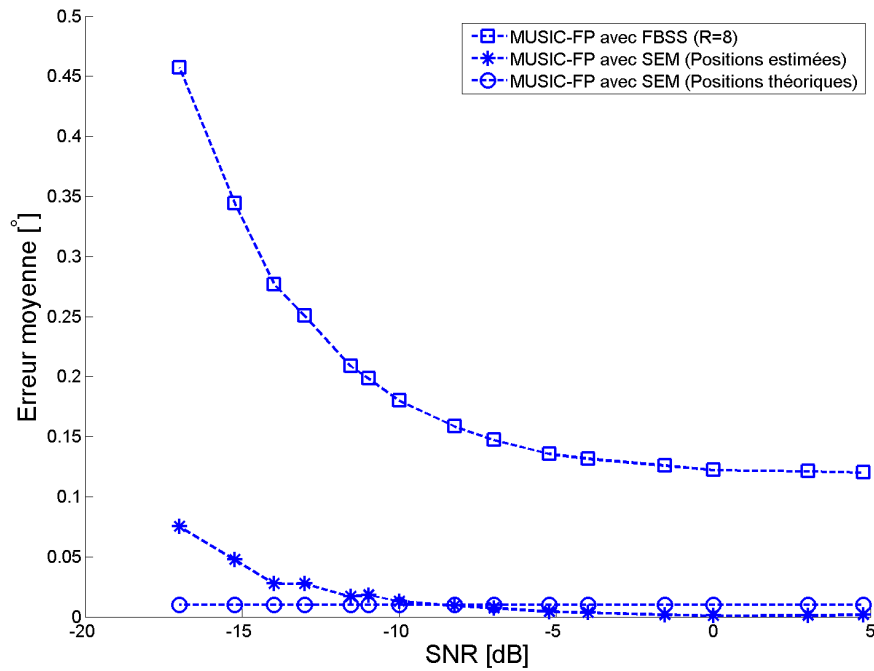
Les résultats pour l'écart-type à la figure 5.5 viennent confirmer ceux obtenus pour la moyenne. En effet, on peut voir qu'à fort SNR, l'utilisation de SEM donne des résultats moins variants que si l'on ne l'utilise pas. Lorsque le SNR commence à diminuer, l'écart-type augmente rapidement si MUSIC-FP est utilisé seul, mais il augmente lentement si on le combine avec SEM. Une observation importante à faire sur cette figure est la diminution de l'écart-type à -30 dB. La raison de ce résultat est que passé un certain SNR, il se produit un pic à une position bien précise, mais qui ne correspond pas à une source. La position de ce pic varie peu, ce qui explique la diminution de l'écart-type à ce point.

À la figure 5.6, on voit un résumé de ce qui est présent pour la moyenne et l'écart-type puisque la probabilité de fausse alarme des sources avec SEM est négligeable jusqu'à ce que le SNR soit plus petit que -20 dB. En deçà de cette valeur, SEM devient de moins en moins précis, comme on a pu le voir à la figure 5.4a.

Tout comme dans le cas à bande étroite, on voit que l'utilisation des valeurs exactes de la position des sources comme première estimation pour SEM, la moyenne et l'écart-type, ainsi que la probabilité de fausse alarme, sont pratiquement nuls, peu importe le SNR. Ce résultat montre à nouveau tout l'intérêt d'avoir la meilleure pré estimation possible.

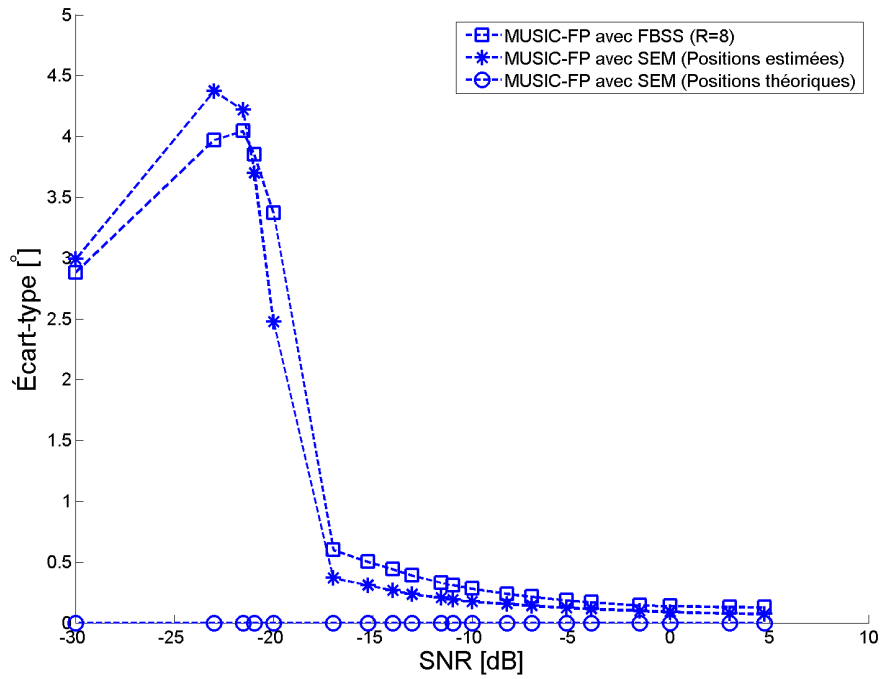


(a) Résultats complets

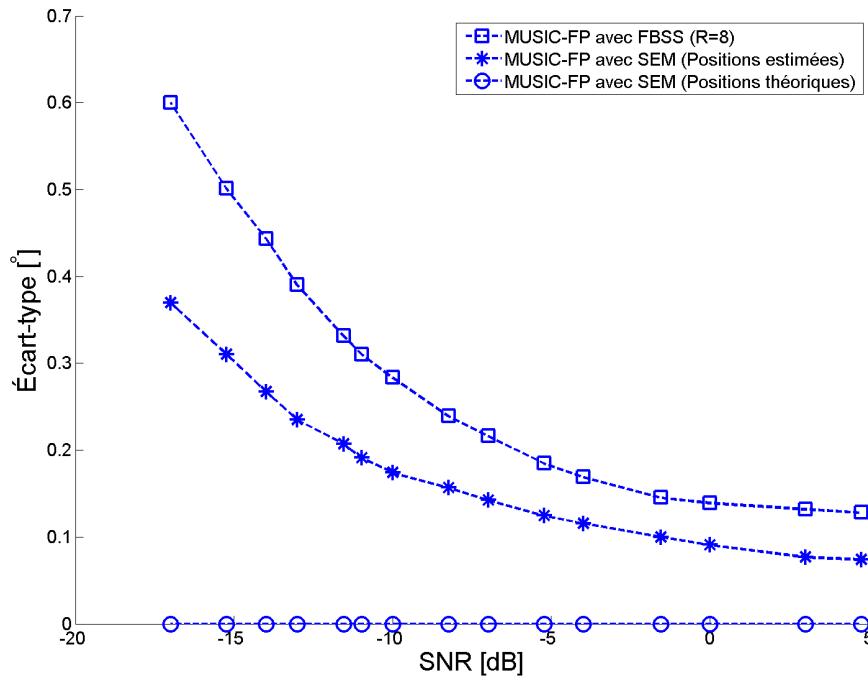


(b) Résultats agrandis pour montrer la différence pour les plus grandes valeurs de SNR.

FIGURE 5.4 – Moyenne de l’erreur de l’estimation de la DOA de la source faible en fonction du SNR de 200 essais de Monte-Carlo dans un cas sans corrélation entre les sources.



(a) Résultats complets



(b) Résultats agrandis pour montrer la différence pour les plus grandes valeurs de SNR.

FIGURE 5.5 – Écart-type de l'estimation de la DOA de la source faible en fonction du SNR de 200 essais de Monte-Carlo dans un cas sans corrélation entre les sources.

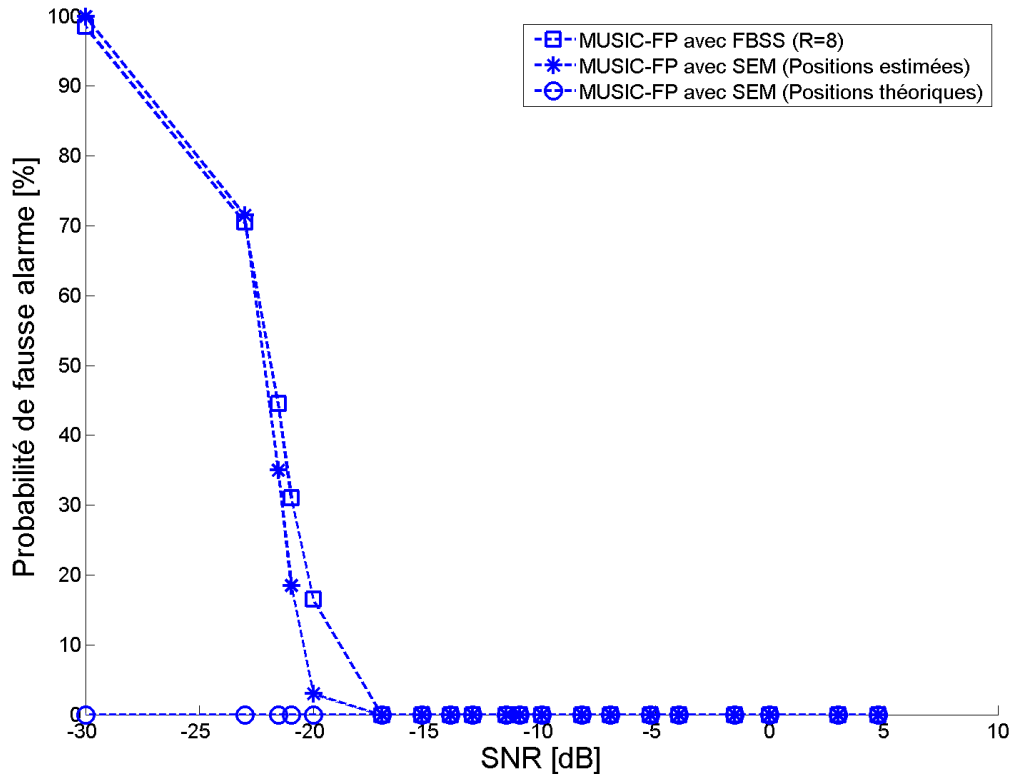
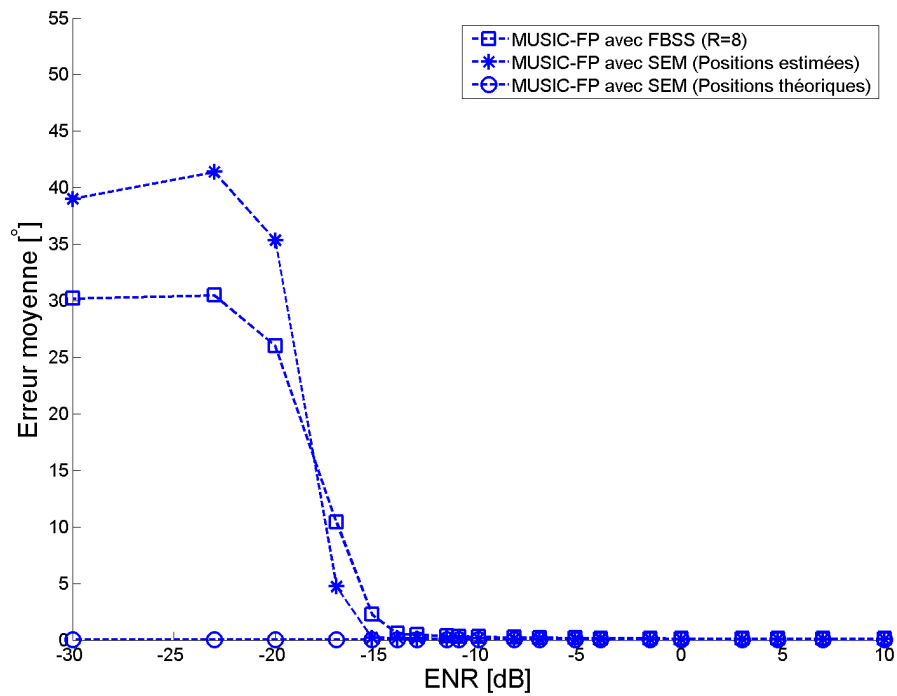


FIGURE 5.6 – Probabilité de fausse alarme de la source faible en fonction du SNR obtenue avec 200 essais de Monte-Carlo pour des sources non corrélées.

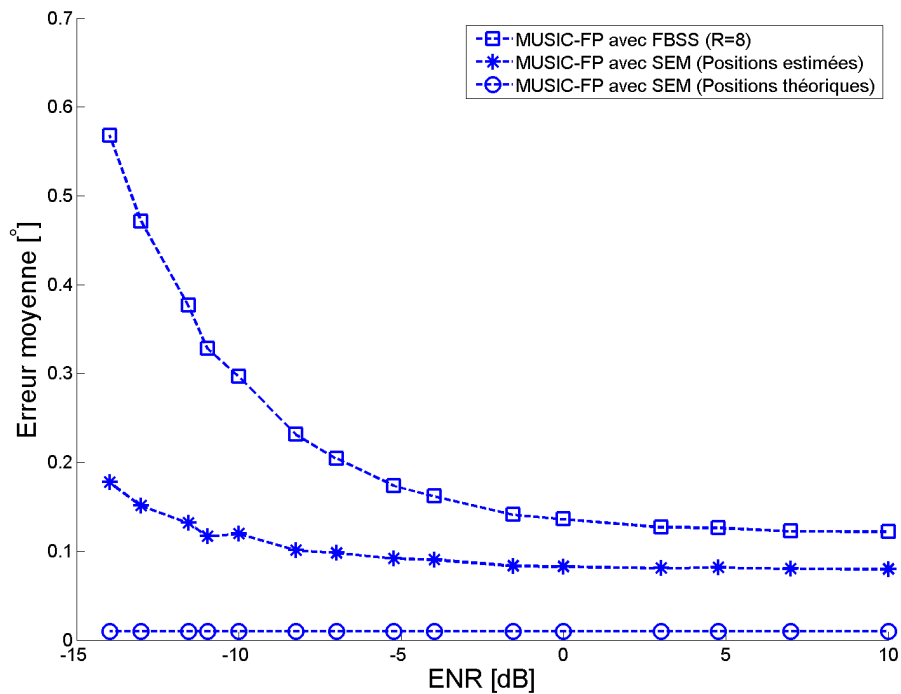
5.3.2 Sources corrélées

Les résultats de la section précédente montrent que SEM améliore les résultats obtenus à l'aide de MUSIC-FP lorsqu'une source est beaucoup plus faible que les autres. Toutefois, l'objectif du projet est de déterminer la DOA d'une source sonore et de ses échos, ce qui implique qu'il est nécessaire d'ajouter de la corrélation entre les sources pour le compléter. On refait donc les simulations avec exactement les mêmes paramètres qu'à la section précédente, sauf que cette fois-ci on utilise des sources corrélées. Pour ce faire, on utilise la méthode présentée à la section 3.2.1. Les résultats obtenus pour l'erreur moyenne de la DOA de la source 2 se retrouvent à la figure 5.7, pour l'écart-type à la figure 5.8, et pour la probabilité de fausse alarme à la figure 5.9.

Sur la figure 5.7a, on voit que l'utilisation de SEM permet de retarder de quelques décibels le ENR à partir duquel l'erreur moyenne commence à devenir importante. En effet, les résultats divergent à partir de -15 dB, et deviennent difficiles à utiliser en bas de -17 dB. Ces résultats sont donc moins précis que ceux obtenus dans le cas



(a) Résultats complets



(b) Résultats agrandis pour montrer la différence pour les plus grandes valeurs de SNR.

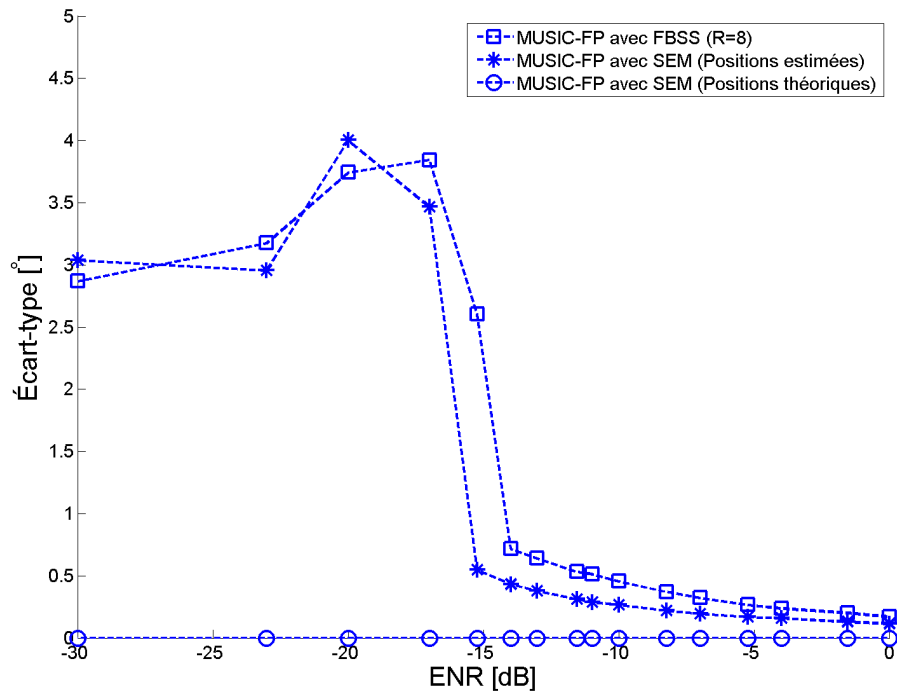
FIGURE 5.7 – Moyenne de l’erreur de l’estimation de la DOA de l’écho en fonction du SNR calculée à partir de 200 essais de Monte-Carlo dans un cas où les sources sont complètement corrélées.

sans corrélation entre les sources, mais pas de beaucoup. Pour MUSIC-FP utilisé avec FBSS, on remarque que les résultats ne sont plus utilisables à partir de -15 dB, ce qui est aussi une régression par rapport au cas sans corrélation. Si l'on regarde les valeurs à hauts ENR présentées à la figure 5.7b, on remarque un léger biais pour SEM qui n'était pas présent dans le cas non corrélé. Ce biais était présent dans le cas à bande étroite présenté au chapitre 3, ce qui montre que la première estimation de la DOA de la seconde source est légèrement décalée à cause de la corrélation entre les sources. Un biais est aussi présent lorsque MUSIC-FP est utilisé avec FBSS, mais son ordre de grandeur est semblable à celui que l'on a vu à la section précédente dans un cas sans corrélation, donc la méthode est robuste à la corrélation lorsque le ENR est élevé.

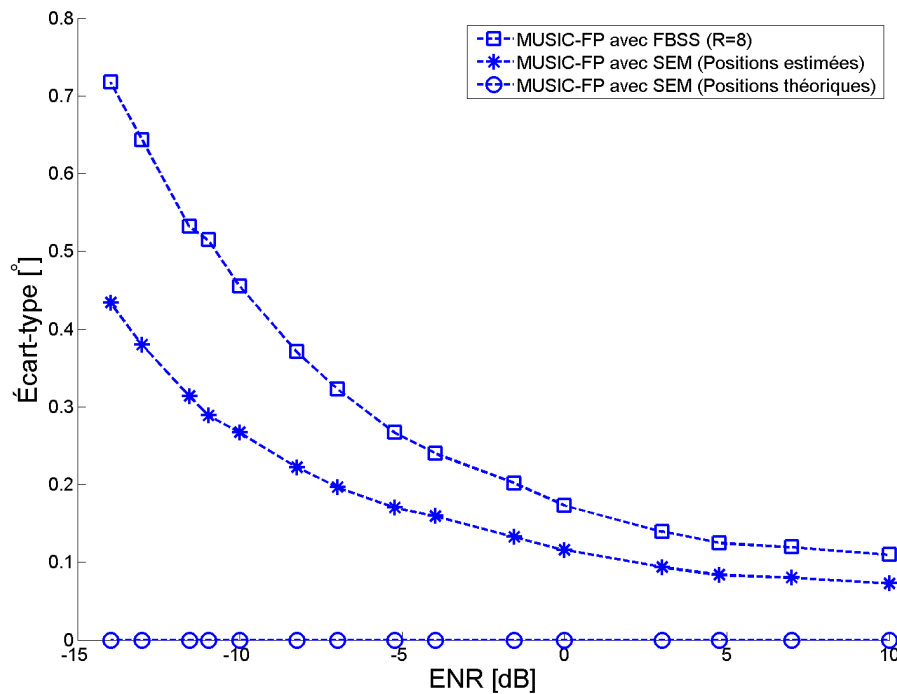
Les résultats présentés à la figure 5.8a expriment essentiellement la même tendance que celle de l'erreur moyenne. On peut y remarquer que l'écart-type est pratiquement nul jusqu'à -10 dB, où elle commence à diverger. Pour ce qui est de la probabilité de fausse alarme, on remarque qu'elle est nulle pour les valeurs plus grandes ou égales à -15 dB, mais augmente très rapidement par la suite. La méthode ne fonctionne donc pas dépassé cette valeur.

L'environnement dans lequel les signaux acoustiques se propagent est souvent composé de plusieurs objets pouvant causer des réflexions. Ainsi, pour se rapprocher des conditions plus réalistes, on ajoute un deuxième écho de même puissance que le premier. Ce second écho permettra d'analyser si le non-respect de la condition sur la puissance des sources peut affecter significativement les résultats pour les échos. Pour faire ces simulations Monte-Carlo, on considère une source principale, située à $\theta_1 = 70^\circ$ et deux échos, positionnés à $\theta_2 = 115^\circ$ et $\theta_3 = 135^\circ$. Le EDR est de -17 dB.

Tout comme au chapitre 3, il est important de comprendre qu'il devient difficile d'associer un pic du pseudospectre obtenu avec une direction d'arrivée, en particulier pour les deux échos à faible ENR. Ainsi, pour les graphiques de l'erreur et de l'écart-type, on ne va conserver que les cas avec une réelle détection. Comme première analyse, on observe donc la probabilité de fausse alarme en fonction du ENR, qui est présent à la figure 5.10. Tout comme dans les chapitres précédents, on définit une fausse alarme si aucun pic n'est présent à moins de 10° de la position théorique de la source. Sur cette figure, on peut voir que peu importe la méthode, les résultats à haut ENR sont fiables. On remarque toutefois que vers -10 dB, il commence à se produire des cas de fausses alarmes si MUSIC-FP est utilisé avec FBSS seulement. Ces cas de fausses alarmes deviennent très significatifs vers -13 dB, où ils se produisent dans plus de 10% des cas



(a) Résultats complets



(b) Résultats agrandis pour montrer la différence pour les plus grandes valeurs de ENR.

FIGURE 5.8 – Écart-type de l'estimation de DOA pour la source faible en fonction du ENR obtenue à l'aide de 200 essais de Monte-Carlo dans un cas où les sources sont complètement corrélées.

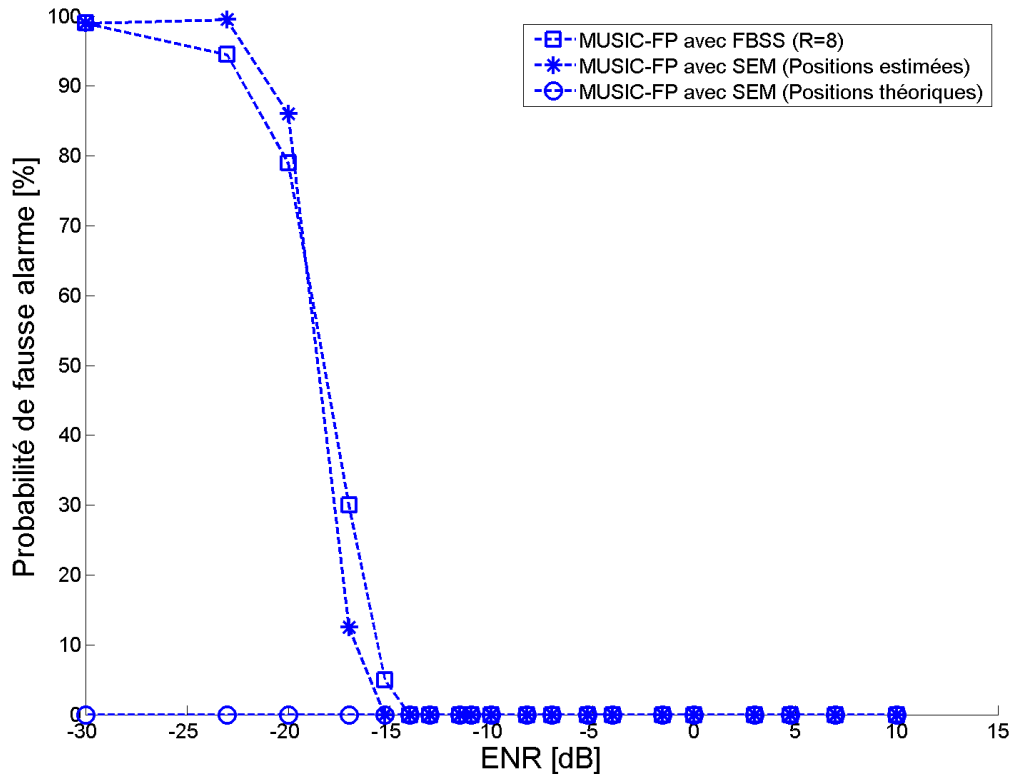


FIGURE 5.9 – Estimation de la probabilité de fausse alarme en fonction du ENR calculée à partir de 200 essais Monte-Carlo pour des sources corrélées.

pour les deux échos. Ces résultats ne sont pas tellement surprenants puisque l’ajout d’un deuxième écho assez près du premier aura tendance à perturber l’espace propre des vecteurs propres 2 et 3, ce qui amène une moins bonne détection. Pour ce qui est de SEM, on peut voir que les résultats sont particulièrement bons puisque même lorsque le ENR est de -23 dB, la probabilité de fausse alarme atteint à peine 20% pour les deux échos. Ces fausses alarmes proviennent de la pré estimation de la DOA puisqu’il n’y a pas de cas de fausse alarme pour le cas de SEM avec positions théoriques.

Pour continuer l’analyse, on observe les résultats à la figure 5.11, qui présente l’erreur moyenne en fonction du ENR. Pour MUSIC avec FBSS, on remarque qu’il y a toujours un biais présent, même à haut ENR. Ce biais est plus important que celui observé lorsqu’il n’y avait qu’un seul écho, ce qui montre qu’il reste de la corrélation entre les sources. Le biais augmente tranquillement par la suite pour finalement atteindre environ 5° à -23 dB. Ce biais peut sembler très faible, mais il faut se rappeler que les cas de fausse alarme ont été enlevés des résultats. Pour ce qui est de l’écart-type, il

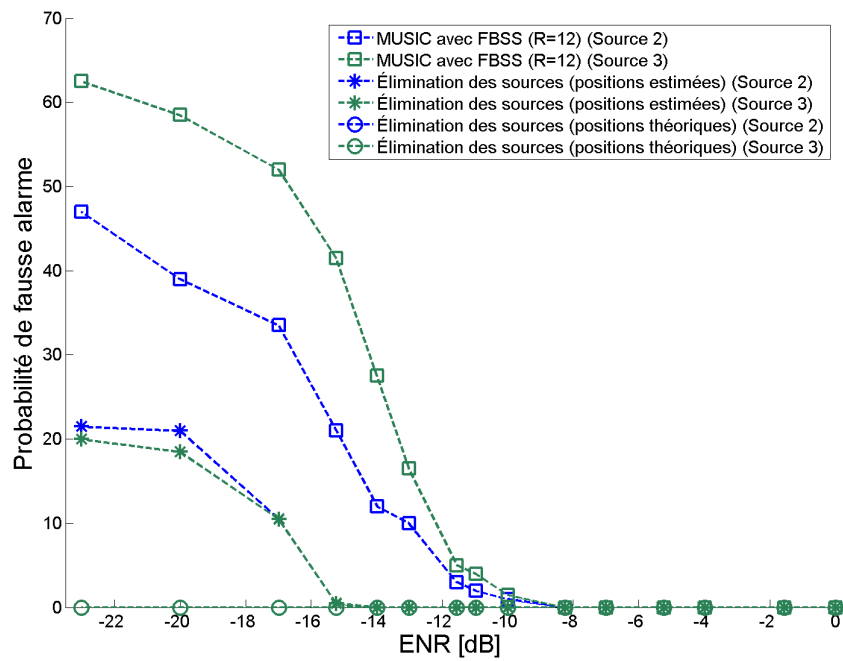


FIGURE 5.10 – Probabilité de fausse alarme en fonction du ENR pour un signal composé d'une source à large bande et de deux échos.

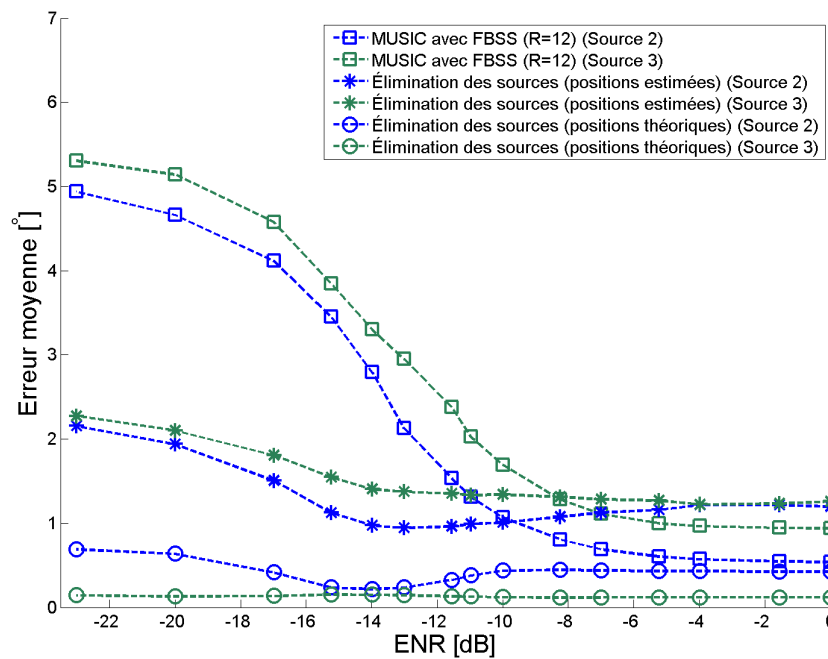


FIGURE 5.11 – Moyenne de l'erreur des bonnes détections de 200 essais de Monte-Carlo en fonction du ENR pour un signal composé d'une source à large bande et de deux échos.

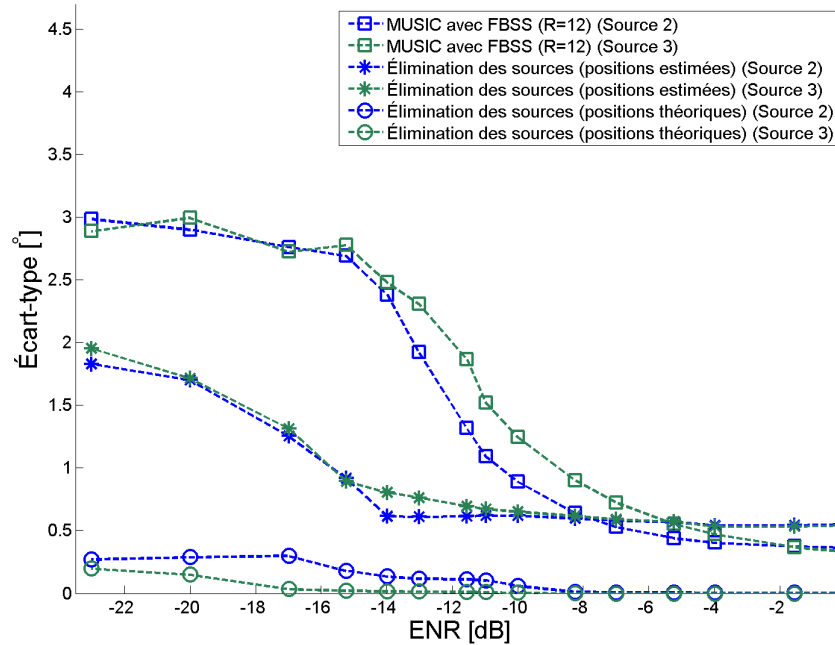


FIGURE 5.12 – Écart-type des bonnes détections de 200 essais de Monte-Carlo en fonction du ENR pour un signal composé d'une source à large bande et de deux échos.

n'est jamais nul, mais assez faible à haut ENR. Lorsque le ENR diminue, il augmente tranquillement jusqu'à atteindre 3° à -23 dB.

L'erreur moyenne de la combinaison MUSIC-FP avec SEM possède elle aussi un biais à haut ENR. La raison principale de ces résultats est que même si on a complètement enlevé la contribution de la première source, les deux échos sont corrélés, donc l'estimation de leur DOA est fortement affectée par le restant de corrélation, ce qui crée un biais. On remarque toutefois que le biais n'augmente pas significativement si le ENR diminue. À nouveau, la principale raison de ce résultat est le fait que l'on enlève les cas de fausses alarmes de la moyenne, ce qui diminue de beaucoup le biais par rapport aux résultats obtenus lorsqu'il n'y avait qu'un écho.

En comparant les résultats des deux échos pour les deux méthodes, on peut voir qu'à haut ENR, MUSIC-FP avec FBSS est meilleur que SEM puisqu'autant son biais que son écart-type sont plus petits. Toutefois, on remarque voir que MUSIC-FP avec FBSS seul diverge plus rapidement que SEM lorsque le ENR diminue. On conclut donc que SEM est plus efficace que MUSIC-FP avec FBSS seul, même dans des conditions de corrélation entre les sources secondaires.

Une dernière série de résultats qu'on n'a pas analysés est celle de SEM en utilisant, comme positions de départ, les positions théoriques. Ce qu'on peut y observer est un biais, plus élevé pour le cas d'un seul écho, et qui est assez constant pour tous les ENR. Le biais pour la source 2 est d'ailleurs plus élevé que celui pour la source 3. Pour expliquer ce résultat, il faut d'abord comprendre que la corrélation entre les sources ne peut pas affecter la DOA de la première itération puisque l'on a utilisé les valeurs théoriques. Toutefois, la présence de corrélation entre les deux échos vient affecter les valeurs propres qui auront tendance à se confondre, ce qui va causer un biais dans l'estimation de la puissance des sources. Ainsi, la matrice de covariance du bruit estimée à la fin de la première itération n'est pas la bonne, ce qui cause une erreur dans la deuxième itération, qui elle n'est pas utilisée avec les valeurs théoriques. On remarque aussi que l'écart-type augmente légèrement lorsque le ENR diminue, mais reste tout de même à un niveau très faible. Contrairement aux autres méthodes, on peut voir que la probabilité de fausse alarme reste nulle pour toutes les valeurs en ENR, ce qui montre que la principale erreur de SEM provient de la pré estimation de la DOA des sources.

Tous les résultats du cas à une source et deux échos montrent que le second écho de même puissance et corrélé vient affecter les résultats de manière significative, peu importe la méthode utilisée. La courbe avec les positions exactes montre les meilleurs résultats possible avec SEM. Pour se rapprocher le plus possible de cette courbe avec une méthode à l'aveuglette, il est possible que l'utilisation de FBSS puisse améliorer les résultats. En particulier, il serait intéressant d'utiliser FBSS avec MUSIC pour déterminer la direction d'arrivée de chaque source, ce qui pourrait améliorer la première estimation de la DOA de chaque source et ainsi améliorer les résultats. Une étude plus approfondie est nécessaire pour vérifier cette hypothèse.

5.4 Conclusion

Dans ce chapitre, on a procédé à la combinaison de SEM à des méthodes d'évaluation de DOA pour des sources à large bande. La méthode CSSM, qui fournit d'excellents résultats lorsque les sources ont une puissance semblable, ne permet pas d'évaluer la DOA de sources dont la puissance diffère de beaucoup. Pour cette raison, on a décidé de laisser de côté cette méthode pour se concentrer sur l'intégration de SEM avec MUSIC-FP. Pour commencer la combinaison de ces deux algorithmes, on a vérifié que l'algorithme développé au chapitre 4 pour déterminer la puissance des sources fonctionne bien avec la matrice de covariance obtenue à l'aide de MUSIC-FP. Par la suite,

on a fait des essais Monte-Carlo pour comparer MUSIC-FP utilisé avec FBSS et avec SEM. Dans un cas sans corrélation des sources, on est arrivé à des résultats intéressants, où la combinaison des deux méthodes élimine un biais à haut SNR et diverge moins rapidement à faible SNR que si l'on utilise FBSS. Dans le cas où les sources sont corrélées, la différence de précision avec FBSS et SEM est plus petite, mais reste quand même présente. Finalement, on a fait une série de 200 essais Monte-Carlo pour un cas où il y a une source puissante et deux échos captés par un réseau de microphones. On a alors observé qu'à haut ENR, MUSIC-FP utilisé avec FBSS est plus précis que SEM puisque cette dernière a un biais causé par la corrélation des sources. Toutefois, SEM conserve de relativement bons résultats pour toutes les valeurs de ENR testés, tandis que l'utilisation de FBSS devient difficilement utilisable à partir de -12 dB. En conclusion, on a pu voir que SEM est, en général, plus efficace que FBSS pour retrouver la direction d'arrivée d'un signal puissant et d'échos beaucoup plus faibles. Au final, comme on a fait lors des chapitres précédents, on écrit l'algorithme 14 dans le but de décrire l'algorithme qui a été créé dans le cadre de ce mémoire à la suite des travaux présentés dans ce chapitre.

Algorithme 14 : MUSIC-FP avec SEM

- 1 Calculer la FFT de \mathbf{X} , la matrice des signaux reçus.
 - 2 Calculer \mathbf{f} , le vecteur des fréquences comprises dans la FFT
 - 3 Couper les fréquences qui sont en dehors de l'intervalle $[f_L, f_H]$
 - 4 Déterminer f_0 , la fréquence de la porteuse
 - 5 Calculer $|\mathbf{X}(f)|$, l'amplitude de tous les points de la FFT qui n'ont pas été coupés
 - 6 Calculer $\phi(f)$, la phase de la FFT pour chaque fréquence pour chaque capteur
 - 7 Calculer $\phi_d(f)$, la phase dépliée de $\phi(f)$.
 - 8 Calculer $\gamma = \frac{f_0}{\mathbf{f}}$
 - 9 Focaliser la phase en faisant la multiplication $\phi_0 = \phi_d(f) \cdot \gamma$
 - 10 Calculer les signaux reçus modifiés avec la formule suivante : $\mathbf{X}_0 = |\mathbf{X}(f)| \exp(\mathbf{j}\phi_0)$
 - 11 Calculer $\mathbf{R}_{xx_0} = \mathbf{X}_0 \mathbf{X}_0^\dagger$
 - 12 **for** $w = 1$ **to** 2 **do**
 - 13 **for** $k=1$ **to** M **do**
 - 14 Faire une décomposition des vecteurs propres pour utiliser MUSIC et trouver
 θ_k
 - 15 Calculer $\mathbf{a}(\theta_k)$
 - 16 **for** $P = 0.5\lambda_k$ **to** $1.5\lambda_k$ **do**
 - 17 Calculer $\hat{\mathbf{R}}_{kP} = P \mathbf{a}(\theta_k) \mathbf{a}(\theta_k)^\dagger$
 - 18 Calculer $\hat{\mathbf{R}}_{redkP} = \hat{\mathbf{R}}_{redk-1} - \hat{\mathbf{R}}_{kP}$
 - 19 Calculer λ_1 et λ_{N-k-1} , les valeurs propres de $\hat{\mathbf{R}}_{redkP}$
 - 20 **end**
 - 21 Déterminer $P = \min(|\lambda_1| - |\lambda_{N-k-1}|)$ Calculer
 $\hat{\mathbf{R}}_{redk} = \hat{\mathbf{R}}_{redk-1} - P \mathbf{a}(\theta_k) \mathbf{a}(\theta_k)^\dagger$
 - 22 **end**
 - 23
 - 24 Calculer $\hat{\mathbf{R}}_{nn} = \mathbf{R}_{xx} - \sum_k \hat{\mathbf{R}}_{redk} = \hat{\mathbf{R}}_{xx} - \hat{\mathbf{R}}_{nn}$
 - 25 **end**
-

Chapitre 6

Mesures expérimentales

Depuis le début de ce mémoire, on s'est contenté de faire des simulations pour vérifier les limites des algorithmes développés. L'utilisation de simulations a comme avantages de permettre un contrôle total et exact de tous les paramètres concernant les sources et le bruit additif, en plus de pouvoir réaliser l'utilisation d'un réseau ayant un espacement inter éléments très petit et composé d'un grand nombre de capteurs. Un autre avantage des simulations est pouvoir réaliser de faire un très grand nombre d'essais Monte-Carlo, ce qui permet d'identifier les limites de l'algorithme avec une bonne précision. Il est toutefois important de prendre des mesures expérimentales pour s'assurer que les algorithmes fonctionnent bien en pratique. Le but de ce chapitre est de vérifier expérimentalement les résultats de tous les algorithmes présentés dans ce mémoire. Pour commencer ce chapitre, on va présenter le montage expérimental qui a été utilisé pour prendre les mesures. Par la suite, on présentera les résultats de mesures faites pour des signaux à bande étroite et pour des signaux musicaux, qui sont à large bande.

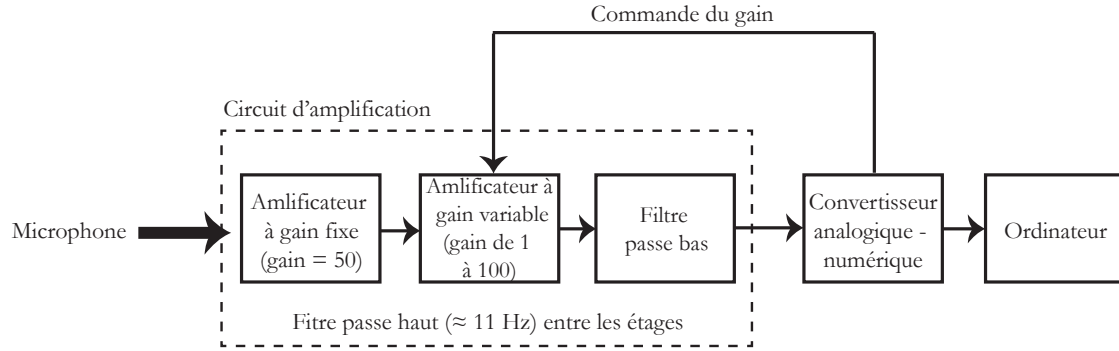


FIGURE 6.1 – Schéma du parcours du signal provenant d'un microphone.

6.1 Montage expérimental

Le montage expérimental utilisé pour les résultats présentés dans ce chapitre a été conçu par Mathieu Gallichand pour son projet de maîtrise [26]. Le modèle de microphones utilisés est CMB-6544PF de la compagnie CUI Inc, qui permet de capter les signaux de fréquence allant de 20 Hz à 20 kHz, ce qui est une bonne plage pour les signaux acoustiques. Les microphones sont omnidirectionnels, donc ils peuvent capter les sons provenant de toutes les directions avec la même précision, ce qui est nécessaire si on veut être capable de mesurer les DOA sur une grande plage d'angles. Les microphones sont montés dans des trous percés sur des tiges de bois, permettant de bien les aligner et de les tenir en place pendant toute la prise de mesures. La distance minimale sur laquelle on peut placer les microphones est de 5 cm, ce qui, selon le tableau 4.1, limite la fréquence maximale du signal à 3.4 kHz. Cette contrainte est assez importante puisqu'elle limite de beaucoup la largeur de bande du signal capté si l'on veut déterminer les DOA.

Une fois que le signal est capté par les microphones, il passe à travers un circuit d'amplification pour ensuite être converti en signal numérique et être enregistré par un ordinateur (figure 6.1). Plus d'informations sur le circuit d'amplification se retrouvent en [26]. Le nombre de canaux de ce circuit d'amplification est 12, ce qui limite à ce nombre la quantité de capteurs composant le réseau.

Pour les mesures de DOA de sources sonores, il est très fréquent d'y avoir de la distorsion du front d'onde du signal à cause de la non-homogénéité de l'air [26]. Dans notre cas, il est important d'avoir un bon contrôle de l'homogénéité de l'air et de la position des échos. La première contrainte oblige de prendre des mesures à l'intérieur pour éviter la distorsion due au vent. La seconde contrainte empêche la prise de mesures dans

une salle de classe puisque les réflexions sur les murs peuvent ajouter des sources non contrôlées dans le signal capté par les microphones. Ainsi, pour faire les mesures, on utilise une salle anéchoïque, qui permet de respecter les deux conditions. Cette salle a été conçue pour faire des mesures de diagramme de rayonnement pour des antennes électromagnétiques, mais les matériaux des parois de la salle sont aussi absorbants aux signaux sonores, en plus d'être composés de cônes pour réduire de beaucoup les réflexions, sans toutefois tous les enlever. Il est donc impossible de connaître le SNR lors de la prise de mesures. Cette salle possède aussi le défaut d'être petite, ce qui limite la distance angulaire entre les sources, la longueur du réseau utilisé et la quantité de sources qui peuvent être placées dans la salle. Il est alors possible de placer deux haut-parleurs dans la chambre anéchoïque et de placer un réseau à l'autre extrémité de la chambre, qui n'est pas fermée sur l'extérieur. Cette extrémité est petite et ne permet pas d'utiliser plus de 6 microphones, ce qui limite de beaucoup la résolution obtenue. De plus, puisque cette extrémité est ouverte, les capteurs ne sont pas isolés des sons qui proviennent de l'extérieur de la salle. La présence d'un mur pas très loin derrière cette ouverture cause des réflexions des ondes sortant de la chambre anéchoïque, qui reviennent directement vers les microphones, ce qui devient une source corrélée à angle droit. Pour limiter ce problème, on place une paroi anéchoïque juste en arrière des capteurs pour diminuer les réflexions. Cette paroi aide aussi à limiter les sons provenant de l'extérieur du laboratoire, mais, par précaution, on fait attention d'éviter de prendre des mesures lorsqu'une personne est près de la porte de la salle. La figure 6.2 montre le montage expérimental.

Pour déterminer la direction d'arrivée théorique, on utilise un niveau laser placé sur un trépied au centre de l'extrémité ouverte de la chambre anéchoïque. On mesure alors la position de l'axe du réseau (l'angle 0°) à l'aide d'un rapporteur d'angle placé sur le trépied, pour ensuite mesurer l'angle de chacune des sources. Le rapporteur d'angle utilisé était gradué à tous les 1° , mais il était souvent difficile de le faire glisser, donc on considère que l'incertitude sur toutes les mesures d'angle est de 1° . À cela s'ajoute l'erreur humaine pour la mesure de l'axe du réseau, qui est de 2° . Ainsi, l'incertitude sur la position exacte des haut-parleurs par rapport au réseau est de 3° . Toutefois, si l'on s'intéresse plutôt à l'incertitude de l'angle entre les deux haut-parleurs, elle est de 2° puisque cette mesure est indépendante de la mesure de l'axe du réseau. Bien que cette incertitude est élevée, elle est suffisamment faible pour qu'il soit possible de séparer les sources.

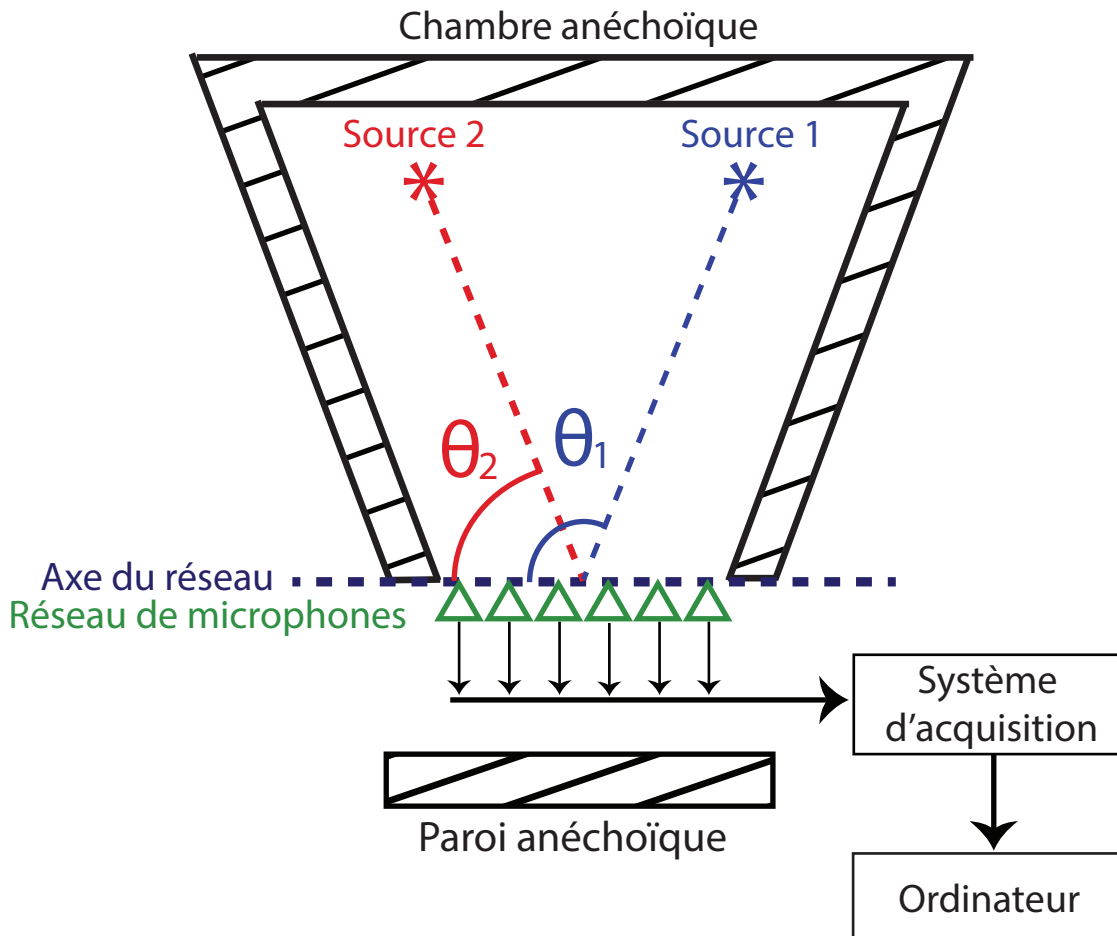


FIGURE 6.2 – Schéma du montage expérimental.

6.2 Mesures de signaux à bande étroite

Une première série de mesures à faire est celle pour des signaux à bande étroite corrélés. Il aurait été intéressant de le faire pour des signaux non corrélés, mais, comme on a vu au chapitre 3, il faut prendre des mesures très longues en durée pour s'assurer leur indépendance, ce qui est difficile à faire en pratique. Avant de faire des mesures avec le réseau, il faut mesurer la position des sources. Pour ce faire, on utilise la méthode vue précédemment, et on trouve que $\theta_1 = 99 \pm 3^\circ$ et $\theta_2 = 85 \pm 3^\circ$. Un autre angle intéressant à mesurer est la distance angulaire entre les deux sources. En prenant la mesure, on trouve que $\Delta\theta = 14 \pm 2^\circ$

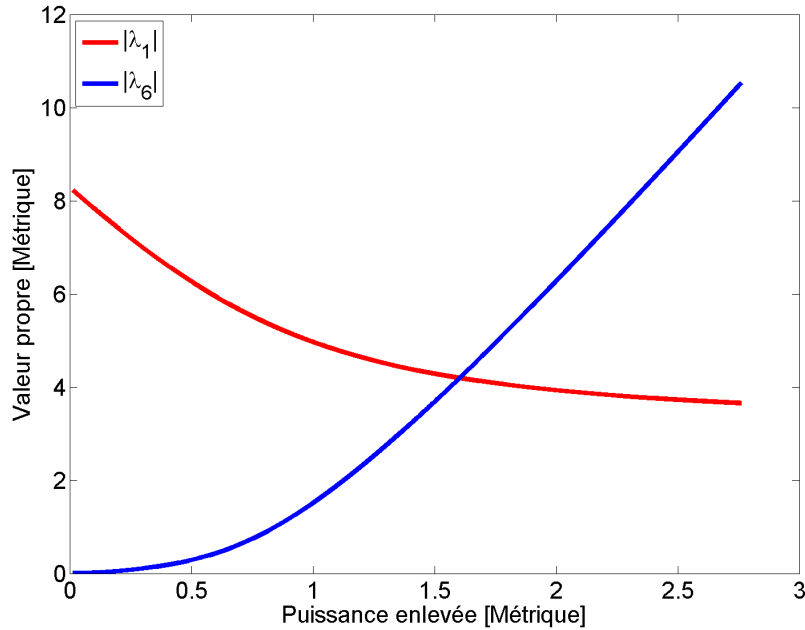


FIGURE 6.3 – Résultat expérimental du test de puissance pour deux sources corrélées dont le rapport de puissance est de -10 dB.

6.2.1 Test de puissance

Comme première série de mesures, il est intéressant de vérifier le test de puissance des sources que l'on a créé au chapitre 3 pour SEM. Dans ce chapitre, les conditions étaient idéales puisque l'on considérait des sources ponctuelles situées très loin d'un réseau composé d'un grand nombre de capteurs. Dans notre cas, le réseau n'est composé que de 6 capteurs, ce qui est suffisant en théorie pour 2 sources, mais qui limite de beaucoup la précision. La figure 6.3 montre le résultat du test de puissance pour deux sources dont le EDR est de -10 dB.

Avant d'analyser la figure, il est important de réaliser qu'il n'est pas possible de connaître exactement la valeur théorique de la puissance, ce qui explique pourquoi l'axe des abscisses n'est pas en fonction de l'erreur comme au chapitre 3, mais en fonction de la puissance enlevée. De plus, il faut comprendre qu'il aurait été arbitraire de mettre des unités dans la figure 6.3, donc c'est pour cette raison que l'on a mis des métriques dans les graphiques.

Sur la figure 6.3, on remarque des courbes de la même forme que celles obtenues avec les simulations au chapitre 3, mais beaucoup moins idéales puisque la dernière valeur propre du système varie fortement dès le début de la réduction de la matrice de covariance

et que la première valeur propre diminue assez lentement en fonction de la puissance enlevée. Puisque le réseau ne comporte que 6 éléments et que le rapport de puissance est de -10 dB, on n'est pas dans les conditions idéales pour l'utilisation du test de puissance. Puisqu'il n'est pas possible de déterminer la puissance théorique des sources, l'idéal est de faire une analyse des pseudospectres obtenue à la fin de l'algorithme.

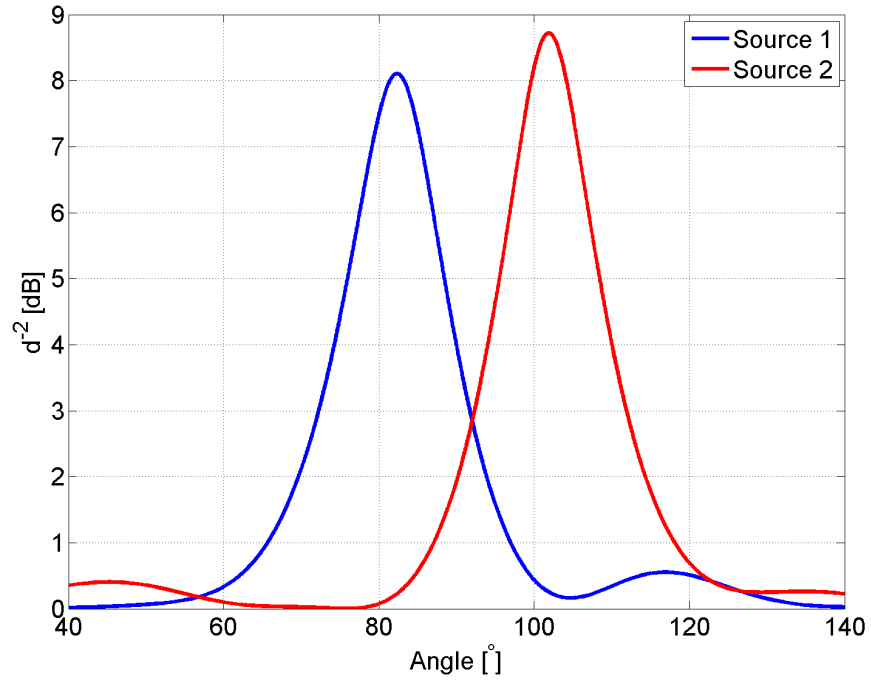
6.2.2 Résultats expérimentaux

Pour commencer les mesures, il est intéressant de s'attarder au cas où l'on utilise qu'une source à la fois. Pour ces mesures, le bruit créé par les réflexions du signal sur le mur arrière et sur les parois imparfaites de la chambre anéchoïque est plus faible. Les résultats dans cette situation sont donc les meilleurs possible que l'on peut obtenir. La figure 6.4a montre les résultats obtenus avec MUSIC seul, et la figure 6.4b montre les résultats si on utilise MUSIC avec SEM.

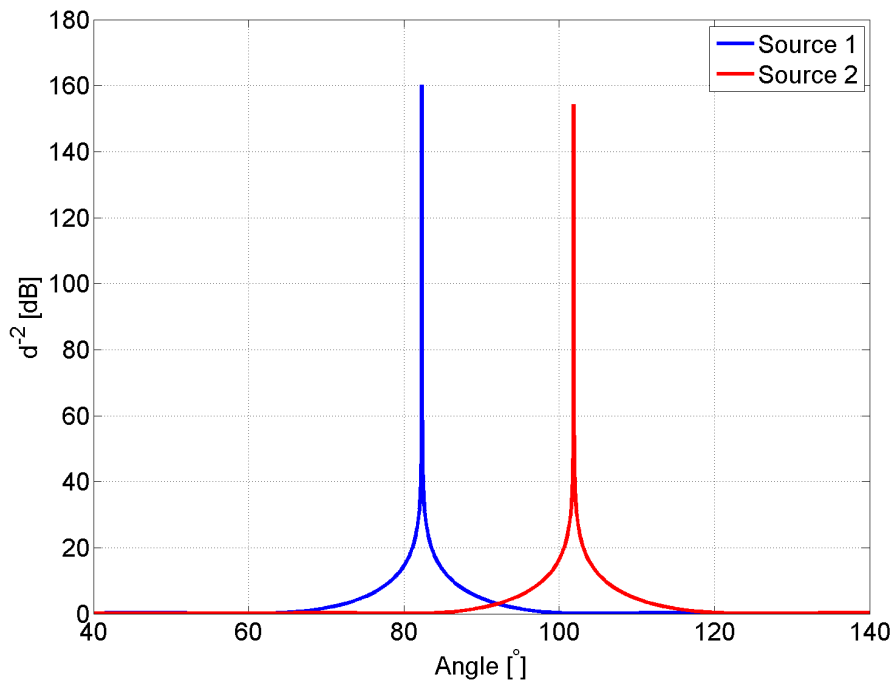
Sur ces figures, on remarque que les pics pour les deux méthodes sont situés aux mêmes positions, soit $\hat{\theta}_1 = 101.9^\circ$ et $\hat{\theta}_2 = 82.3^\circ$. La position de ces pics est aux extrêmes de l'erreur, ce qui semble démontrer l'absence de biais pour les deux méthodes. Toutefois, la distance entre les pics est de $\Delta\hat{\theta} = 19.6^\circ$, ce qui est en dehors de l'intervalle de l'incertitude de quelques degrés. De nombreux éléments peuvent expliquer les résultats, les principaux étant les suivants :

1. Les réflexions sur la paroi anéchoïque ne sont pas totalement bloquées, créant un bruit fortement corrélé. Il faut bien comprendre que la chambre anéchoïque n'a été faite pour des signaux de radiofréquences, donc l'atténuation des réflexions des signaux acoustiques n'est pas parfaite.
2. Les sources ne sont pas assez distantes des capteurs, ce qui a pour effet que le front d'onde n'est pas assez plan pour avoir des résultats parfaits.
3. Le réseau est très petit, amplifiant ainsi l'effet nocif du bruit comparativement à un réseau beaucoup plus long.
4. Le biais que l'on a observé même à haut SNR à la figure 3.9 ajoute certainement une erreur sur la position du pic. Il est possible que le fait que le réseau comporte seulement éléments comparativement à ceux utilisés pour les simulations induise un biais plus grand que les résultats de simulations présentés à la figure 3.9 .

Un autre effet important observé est que l'amplitude des signaux reçus par les différents capteurs varie, même s'il n'y a qu'une seule source. En effet, si on observe la figure 6.5,



(a) MUSIC avec FBSS



(b) MUSIC avec SEM

FIGURE 6.4 – Comparaison des résultats expérimentaux de MUSIC avec FBSS et avec SEM pour deux mesures différentes, chacune avec une seule source.

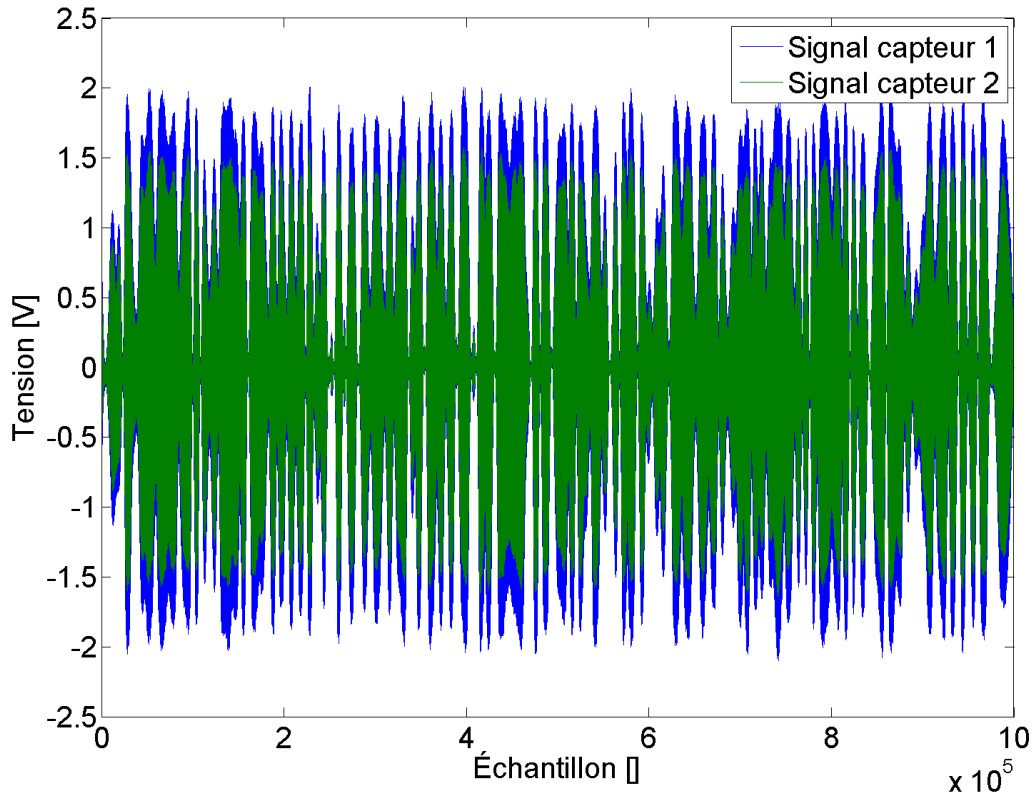
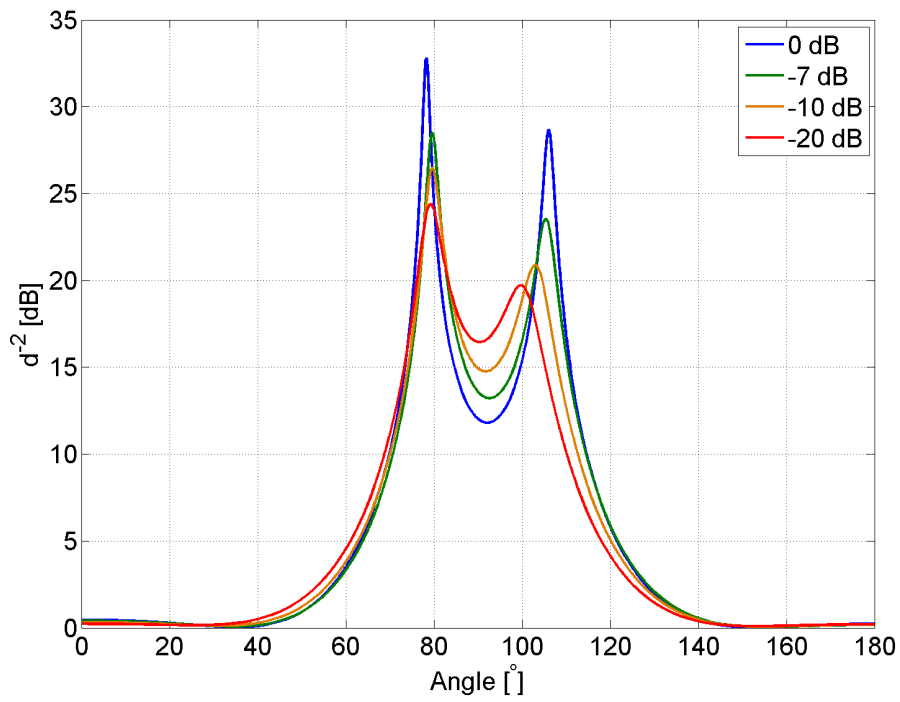


FIGURE 6.5 – Exemple des signaux reçus par les capteurs 1 et 2 en fonction du temps lorsqu’une seule source est présente dans le signal.

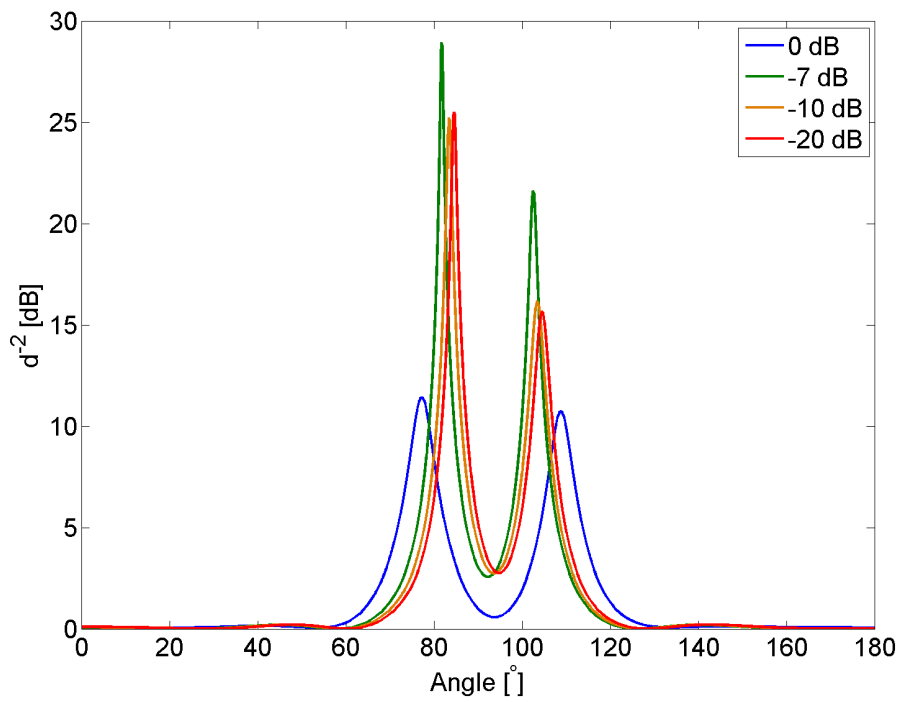
qui présente le signal en tension capté par deux capteurs dans le cas où une seule source est présente dans le signal, on remarque que l’amplitude n’est pas la même pour les deux signaux. Cette variation d’amplitude est probablement causée par le circuit d’amplification de chaque élément du réseau qui n’amplifierait pas les signaux du même facteur. Cette situation peut amener des problèmes pour détecter la DOA de sources sonores et est la principale cause d’erreur dans les mesures.

Pour rendre les mesures plus réalistes, on peut désormais prendre des mesures lorsque les deux sources sont allumées en même temps, et ce, pour différents rapports de puissance entre les sources. Les résultats se retrouvent à la figure 6.6a pour l’utilisation de MUSIC avec FBSS composé de $R = 3$ sous-réseaux, et à la figure 6.6b dans le cas où l’on utilise SEM.

La figure obtenue pour MUSIC avec FBSS montre que les pics sont assez larges et que leurs positions sont plus distantes que celles attendues. Bien que la position pour la seconde source soit assez stable, elle est située à 78° , ce qui est beaucoup trop loin



(a) MUSIC avec FBSS (R=3)



(b) SEM

FIGURE 6.6 – Comparaison des résultats expérimentaux de MUSIC avec FBSS et SEM pour différents rapports de puissance entre deux sources corrélées.

de la position théorique. De plus, on observe que la position de l'autre source varie en fonction du rapport de puissance entre les sources. La principale cause de cette variation de la position de la source 2 est qu'il reste de la corrélation entre les sources après FBSS. Lorsque le rapport de puissance entre les sources est petit, la position de la seconde source est fortement influencée par celle de la première, ce qui fait que sa position se rapproche de celle de la première source. Lorsque le EDR est plus élevé, les deux sources s'influencent moins l'une sur l'autre, ce qui explique la position des sources sur la figure.

Sur la figure 6.6b, on peut voir que lorsque le EDR est de 0 dB, les résultats ne sont pas précis, ce qui est normal puisque la première estimation de la DOA ne sera pas précise due à la présence de corrélation entre les sources. Une grande erreur de cette valeur induit une série d'erreurs et réduit la précision des résultats. Lorsque le rapport de puissance est plus petit, la première valeur propre correspond de plus en plus à la première source, ce qui a pour effet que la première estimation de la DOA est plus près de la réalité, ce qui améliore les résultats par la suite. À partir de ces résultats expérimentaux, il est possible de conclure que l'utilisation de SEM est plus efficace que FBSS si les sources sont corrélées et de puissances différentes. Cette conclusion est encore plus vraie si le réseau est trop petit pour permettre à FBSS de complètement décorréler les sources.

6.3 Signaux à large bande

Dans les chapitres 4 et 5, on a testé MUSIC-FP avec FBSS et avec SEM pour des signaux composés d'un signal modulant et porté à une fréquence bien précise. Ces signaux représentent bien des signaux qui peuvent sortir d'une antenne, puisque la présence d'une porteuse est obligatoire, mais ce n'est pas vrai dans le cas d'une source acoustique, qui peut se propager sans porteuse. Pour avoir des mesures plus près de la réalité, on utilise des pistes musicales pour prendre les mesures, ce qui implique une absence d'onde porteuse dans le signal mesuré. On a vu, dans le chapitre 4, que la fréquence de référence de MUSIC-FP doit être égale à celle de la porteuse pour optimiser les résultats lorsqu'il y a plus d'une source dans le signal, ce qu'il n'est plus possible de faire pour ces mesures. Il est ainsi nécessaire de faire une analyse pour déterminer la meilleure fréquence de référence pour les signaux mesurés. On positionne donc les sources à $\theta_1 = 86 \pm 3^\circ$ et $\theta_2 = 102 \pm 3^\circ$ et on utilise des signaux différents, de mêmes puissances pour obtenir la figure 6.7, qui présente le pseudospectre de MUSIC-

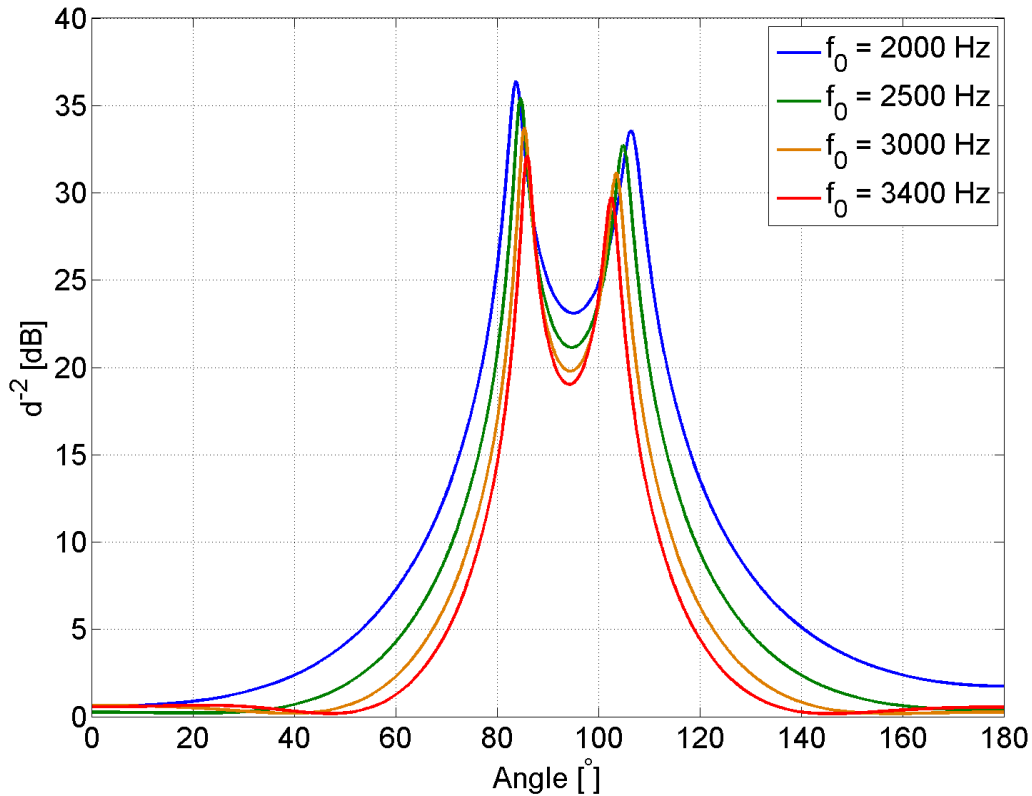


FIGURE 6.7 – Variation du pseudospectre de MUSIC-FP seul pour différentes valeurs de fréquences de référence.

FP avec FBSS ($R=3$) pour différentes valeurs de f_0 . Sur cette figure, on peut voir que la position des pics varie quelque peu en fonction de f_0 , et que la fréquence 3400 Hz, donne des pics très près des positions théoriques. Cette fréquence est d'ailleurs la limite du réseau pour s'assurer qu'il n'y ait pas de recouvrement spatial. La raison pour laquelle cette fréquence donne les meilleurs résultats est le fait qu'à cette fréquence, le déphasage entre les capteurs ajusté ϕ_0 est plus élevé, ce qui fait que le pic pour chacune des positions est plus étroit et limite les erreurs. Pour le reste des mesures, on utilisera cette fréquence comme référence.

Tout comme on a fait dans le cas à bande étroite, il est intéressant de vérifier la DOA de chacune des sources lorsqu'elle est seule puisque ce résultat permet d'avoir une bonne idée des meilleurs résultats possible avec différentes méthodes. La figure 6.8a affiche les résultats de MUSIC-FP seule pour deux mesures différentes, chacune d'elle avec seulement une source ouverte, et la figure 6.8b affiche les résultats dans le cas de MUSIC-FP avec SEM pour les mêmes séries de mesures. On remarque sur ces résultats

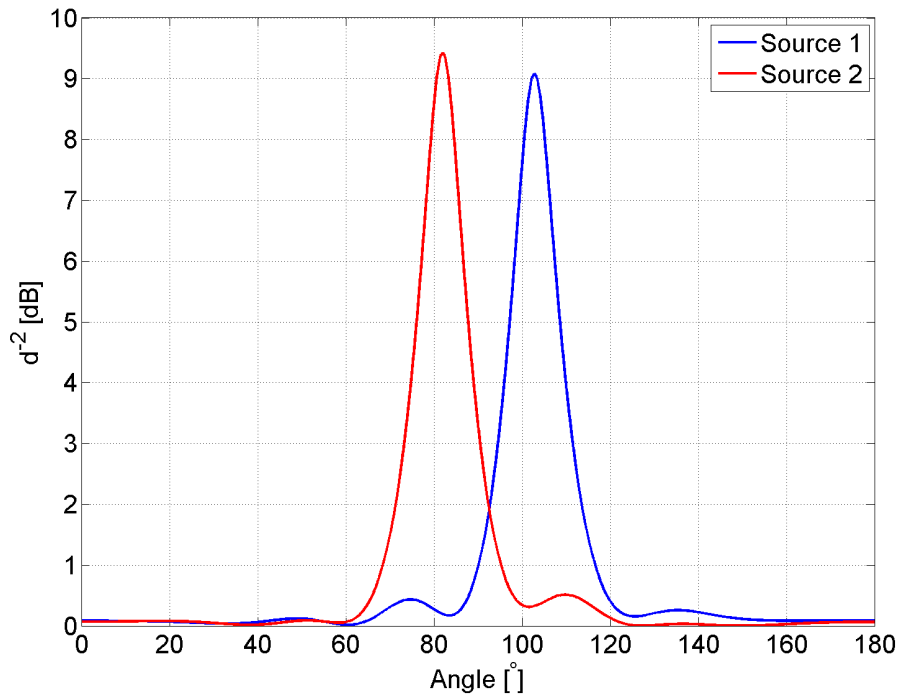
que pour la première source, on obtient des pics à 4° de la valeur théorique, ce qui est légèrement en dehors de la plage d'incertitude. Pour la seconde source, les pics pour les deux méthodes sont exactement à la position théorique. Cette différence est semblable à celle obtenue lors des mesures à bande étroite, ce qui semble confirmer que l'erreur est causée par les conditions expérimentales plutôt qu'un biais dans les mesures de la position exacte des sources, puisque les mesures n'ont pas été prises la même journée et ont donc demandé une nouvelle mesure de la position des sources. Si l'on compare les deux figures, on peut voir jusqu'à quel point SEM est efficace et permet d'améliorer les résultats puisque ses pics sont très hauts et minces par rapport à MUSIC-FP.

Maintenant que l'on a fait des mesures avec les deux sources séparément, il est intéressant de faire quelques séries avec les sources allumées en même temps, et ce, pour différents rapports de puissance entre les sources. Pour ces séries, on utilise de la musique de styles différents pour chacun des haut-parleurs dans le but d'éviter le plus possible la corrélation entre les sources. La figure 6.9a montre les résultats de MUSIC-FP avec FBSS fait avec 3 sous-réseaux, et la figure 6.9b ceux de MUSIC-FP avec SEM.

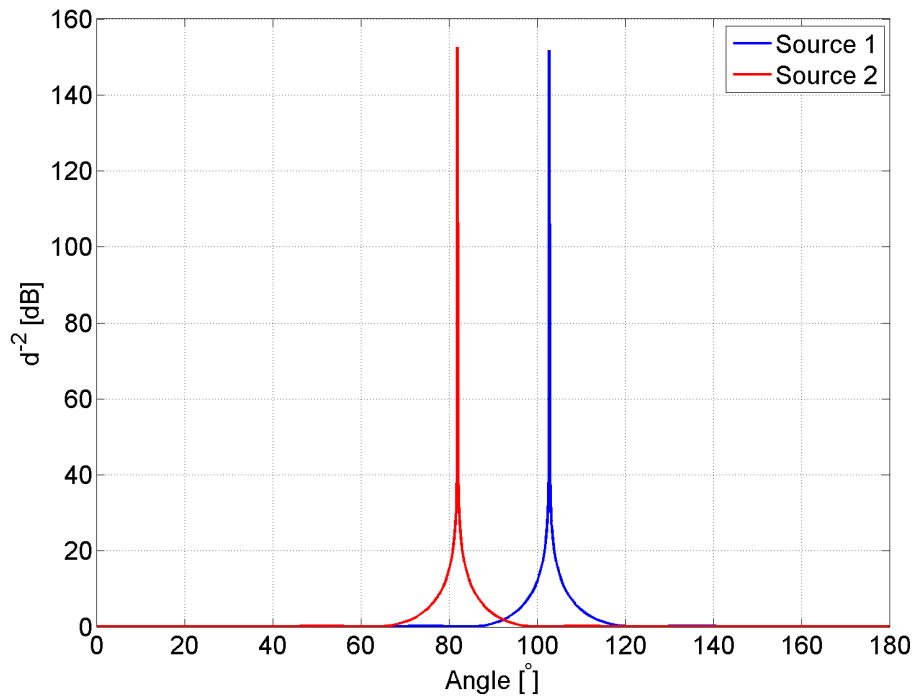
Sur la figure 6.9a, on voit que les pics du pseudospectre obtenu sont assez larges et se recourent assez rapidement. Cela s'explique par le fait qu'on utilise un réseau composé de peu de capteurs. De plus, on voit que la position des pics varie légèrement lorsque la puissance entre les sources passe de 0 dB à -7 dB. Cette différence provient du bruit présent dans le local et du bruit électronique du système d'acquisition, qui viennent ajouter une variance dans les données. Il est aussi intéressant de noter que la distance entre les pics reste la même, ce qui montre que la méthode fonctionne dans tous les cas, mais que les conditions expérimentales ont causé un biais. Lorsque le rapport de puissance entre les sources tombe à -20 dB, on remarque qu'un seul pic est présent, ce qui montre qu'on a dépassé la limite de résolution de l'algorithme. Il est très probable que cela résulte du fait que la seconde source est noyée dans le bruit et que MUSIC-FP n'est plus capable de la distinguer.

Les résultats présents à la figure 6.9b présentent le même biais lorsque les puissances des sources sont égales que pour la figure 6.8b, ce qui montre qu'une erreur présente lors de la prise de mesure a biaisé les résultats. Les résultats montrent aussi que les pics sont plus hauts, plus étroits et que les positions sont plus constantes que ceux obtenus avec la combinaison de MUSIC-FP et de FBSS. De plus, on peut voir que même à -20 dB, les pics sont à la bonne position. SEM est donc plus efficace que FBSS.

Les pistes diffusées par chacun des haut-parleurs lors des mesures présentes à la figure

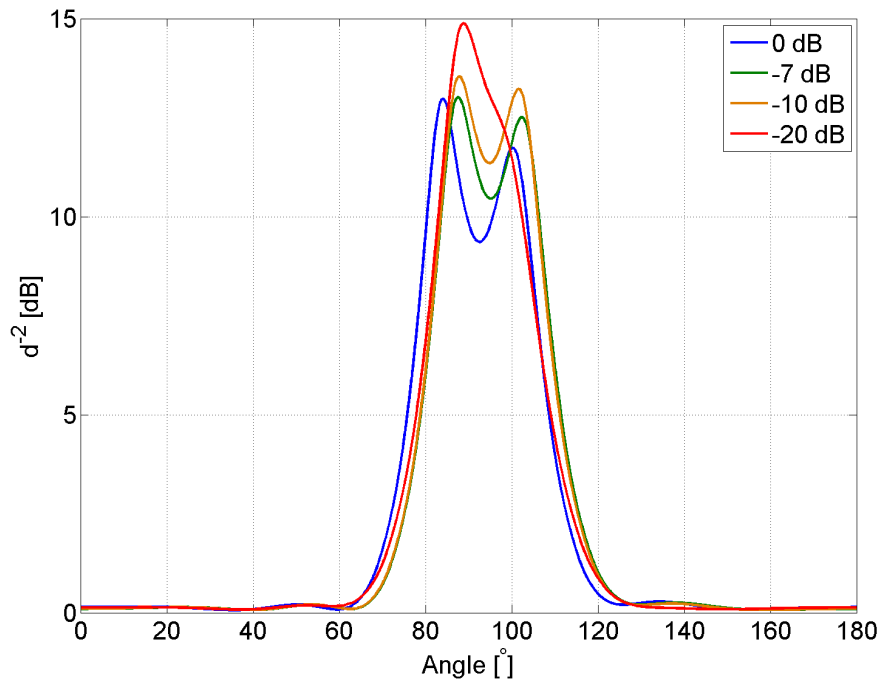


(a) MUSIC-FP seul

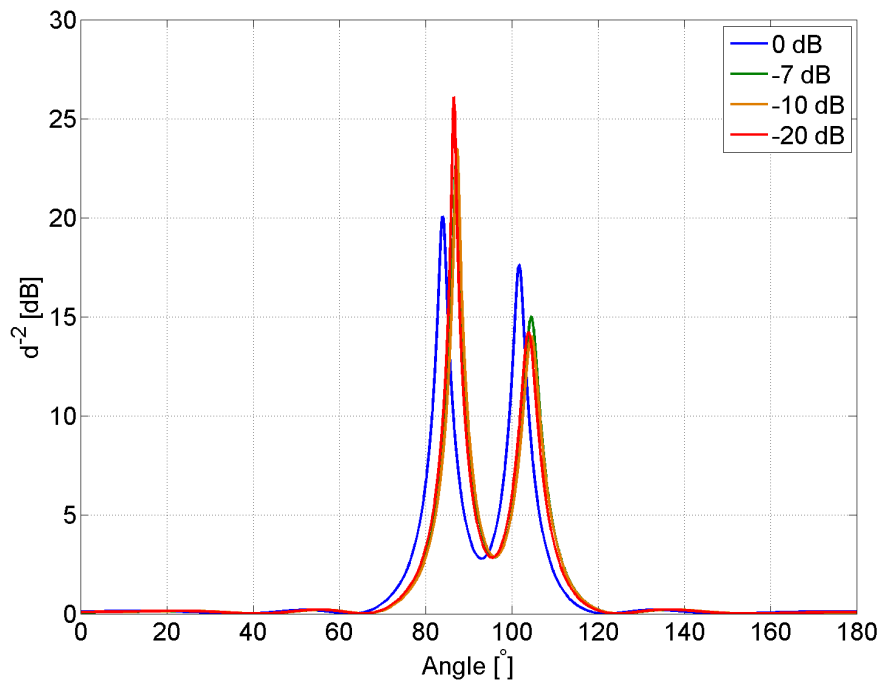


(b) MUSIC-FP avec SEM

FIGURE 6.8 – Résultats expérimentaux de MUSIC avec et sans SEM pour deux mesures différentes, chacune avec une seule source à large bande.



(a) MUSIC-FP avec FBSS (R=3)



(b) MUSIC-FP avec SEM

FIGURE 6.9 – Résultats expérimentaux de MUSIC-FP avec FBSS et avec SEM pour différents rapports de puissance entre les sources, dans le cas où les signaux à large bande sortant des haut-parleurs sont différents

6.9 sont différentes pour limiter la corrélation entre les sources. Pour se rapprocher de l'objectif de ce mémoire, on prend une nouvelle série de mesures, mais cette fois-ci en utilisant la même piste pour chacun des haut-parleurs, ce qui fait que les sources sont complètement corrélées. Les résultats apparaissent à la figure 6.10.

Sur la figure 6.10b, on remarque que l'utilisation de MUSIC-FP avec SEM est très précise. En effet, on y voit que les pics du pseudospectre sont très stables sur toutes les courbes, sauf celles où la puissance des sources est la même. Dans ce cas, on sait que la première estimation de la direction d'arrivée est faussée par la corrélation entre les sources, ce qui cause une suite d'erreurs qui conduisent au résultat apparaissant à la figure 6.10b.

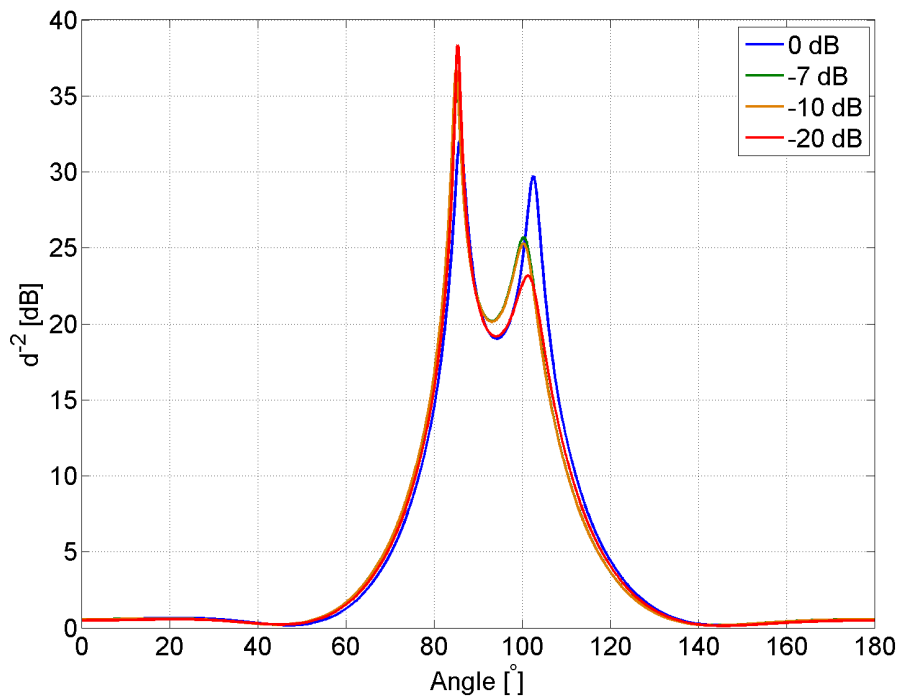
Les résultats obtenus pour l'utilisation de MUSIC-FP avec SEM sont aussi très précis, sauf dans le cas où les sources ont la même puissance. Ce résultat n'est pas très surprenant puisque ce cas ne respecte pas l'approximation faite au chapitre 3 par rapport à la puissance relative des sources, ce qui a pour effet que la méthode pour trouver les DOA individuelles n'est pas précise. Une idée pour améliorer l'algorithme est d'utiliser FBSS pour faire la première estimation de la DOA.

6.4 Conclusion

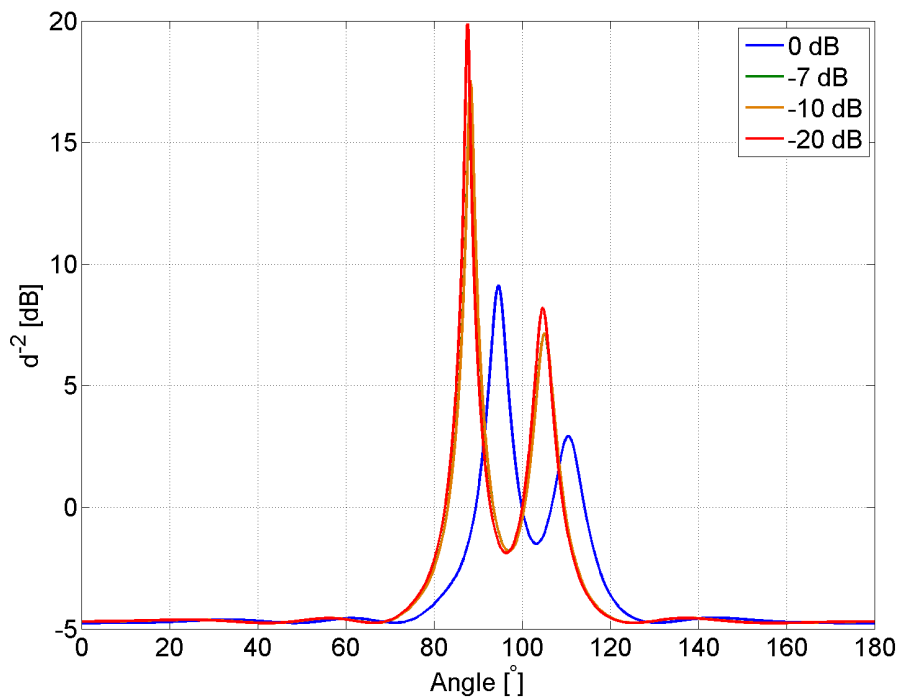
Dans ce chapitre, on a présenté le montage expérimental utilisé pour prendre des mesures de signaux acoustiques. Bien que ce montage est sensible à la présence de bruit provenant de l'extérieur de la salle et qu'il impose l'utilisation d'un petit nombre de capteurs, il permet de limiter la présence d'échos non contrôlés et de limiter le bruit si l'on prend les mesures au bon moment. Le montage expérimental est donc capable de se rapprocher des conditions des simulations faites dans les autres chapitres.

Une première série de mesures a été faite dans l'objectif de vérifier expérimentalement le comportement de SEM dans le cas de signaux à bande étroite corrélés. Les résultats obtenus montrent que si le rapport de puissance des sources n'est pas unitaire, SEM est plus efficace que MUSIC utilisé avec FBSS. Il faut toutefois comprendre que FBSS est difficile à utiliser lorsque le réseau est composé d'un petit nombre de capteurs, ce qui était le cas lors de la prise de mesures. On peut donc conclure que lorsque le réseau est petit, il vaut mieux utiliser SEM que FBSS. Des mesures dans une chambre anéchoïque plus grande seraient nécessaires pour conclure pour des réseaux plus grands.

Une deuxième série de mesures a été prise, cette fois-ci avec des sources musicales, qui



(a) MUSIC-FP avec FBSS (R=3)



(b) MUSIC-FP avec SEM

FIGURE 6.10 – Résultats expérimentaux de MUSIC-FP avec FBSS et avec SEM pour différents rapports de puissance entre deux sources à large bande.

ont la particularité d'être à large bande. Toutefois, ces sources ne sont pas modulées par une porteuse, donc il fut nécessaire de déterminer quelle était la meilleure fréquence de référence. On a alors trouvé que f_{0max} fourni les meilleurs résultats. Par la suite, on a pris des mesures lorsque les pistes étaient différentes l'une de l'autre et que l'on faisait varier la puissance de la seconde source. Les résultats obtenus ont montré que SEM permet de diminuer l'erreur de l'estimation de la DOA. En utilisant des sources complètement corrélées, on a obtenu de meilleurs résultats que si les pistes étaient différentes, ce résultat surprenant provient des conditions expérimentales, qui ne sont pas idéales.

Au final, les résultats présents dans ce chapitre viennent confirmer ceux des simulations. En plus de cela, les résultats de ce chapitre ont montré que SEM est particulièrement meilleure que FBSS lorsque le réseau utilisé est trop court pour créer une grande quantité de sous-réseaux.

Chapitre 7

Conclusion

Ce mémoire a couvert l'estimation de directions d'arrivée (DOA) de sources sonores et de ses échos en traitant les trois problèmes principaux rencontrés avec ce type de signaux, soit la corrélation entre les sources, la largeur de bande du signal ainsi que la différence de puissance entre la source principale et les échos.

Dans le chapitre 2, on a présenté les bases théoriques de l'estimation de la DOA ainsi que de nombreux algorithmes présents dans la littérature. En particulier, on a développé un modèle du signal à bande étroite et à large bande. Par la suite, on a présenté MVDR et MUSIC, deux algorithmes qui permettent de déterminer la DOA de sources à bande étroite. Deux méthodes pour généraliser les algorithmes à bande étroite ont été présentées, soit la *Incoherent Signal Subspace Method*, qui utilise des méthodes à bande étroite sur chaque bande de fréquence pour ensuite combiner les résultats, et la *Coherent Signal Subspace Method*, qui transforment les matrices de covariance de chaque fréquence pour les ramener à une fréquence de référence et ensuite les combiner pour utiliser une méthode d'estimation de la direction d'arrivée sur la matrice de covariance combinée. Finalement, on a présenté le lissage spatial, *Covariance Differencing and Iterative Spatial Smoothing*, qui permet de limiter les effets néfastes de la corrélation entre les sources, et le critère AICe, qui permet d'estimer le nombre de sources dans le signal.

Au chapitre 3, on a développé l'algorithme s'appelant méthode par élimination des sources, qui sert à améliorer les performances de MUSIC lorsqu'une source est très puissante par rapport à ses échos. Cette méthode, comme son nom l'indique, élimine la contribution de chacune des sources à la matrice de covariance des signaux reçus jusqu'à ce qu'il ne reste qu'une estimation de la contribution du bruit. Une rétroaction est alors

ajoutée dans l'algorithme dans le but de recommencer le processus d'élimination des sources après avoir enlevé l'estimation de la contribution du bruit dans la matrice de covariance des signaux reçus, ce qui augmente la résolution à faible SNR. Par la suite, on a développé un test pour estimer la DOA de chaque source individuellement et un test pour estimer la puissance de chaque source. Le critère pour estimer la puissance est le point d'intersection de la première et de la dernière valeur propre lors de la réduction de la matrice de covariance. Par la suite, on a effectué des essais Monte-Carlo dans des cas non corrélés et corrélés. Ces tests ont montré que si les sources ne sont pas corrélées, il est mieux d'utiliser FBSS que d'utiliser SEM. Toutefois, s'il y a présence de corrélation entre les sources, il est plus utile d'utiliser SEM, qui fournit de meilleurs résultats pour toutes les valeurs de ENR testés. Toutefois, on a remarqué que peu importe la méthode, il y a une probabilité de fausse alarme importante avec un ENR plus faible que -15 dB. L'objet de ce chapitre est d'ailleurs le sujet d'un article soumis au Journal Canadien en Génie Électrique. Une copie de l'article est présente à l'annexe A.

Pour contourner le problème de la largeur de bande du signal, on a développé au chapitre 4 un nouvel algorithme, que l'on appelle MUSIC à focalisation de phase, qui effectue une transformation sur la phase de chaque fréquence pour la ramener à une fréquence de référence. Après un test, on a déterminé que la meilleure fréquence de référence est exactement la même que celle de la porteuse. Pour tester la robustesse de l'algorithme, on a fait des essais Monte-Carlo en utilisant MUSIC-FP et CSSM. Les résultats obtenus montrent la supériorité de CSSM, en particulier lorsque le SNR est élevé.

L'intégration de SEM pour des algorithmes pour des signaux à large bande a été faite au chapitre 5. Pour commencer, on a vérifié le test de puissance des sources pour CSSM et MUSIC-FP, ce qui a montré que ce test n'était pas possible à faire pour CSSM. Pour la suite du chapitre, on a décidé de laisser tomber l'intégration de SEM avec CSSM pour se concentrer à faire l'intégration avec MUSIC-FP. Pour vérifier les performances de l'utilisation de SEM, on a effectué des essais Monte-Carlo. Autant dans le cas non corrélé que corrélé, l'ajout de SEM améliore les résultats obtenus pour toutes les valeurs de SNR, mais surtout lorsque le SNR est faible, ce qui montre l'utilité de la méthode.

Finalement, au chapitre 6, on a vérifié expérimentalement les résultats obtenus par simulations des chapitres précédents. On a commencé par présenter le montage expérimental qui, à défaut d'être installé dans une salle anéchoïque acoustique, a été installé dans une salle anéchoïque pour des ondes électromagnétiques. Les matériaux des parois de la salle diminuent bien les échos sonores et le bruit ambiant, mais ne les annulent

pas totalement. De plus, l'espace étant très limité, on doit se contenter d'utiliser un réseau de 6 capteurs et d'avoir seulement deux sources. Pour commencer, on a vérifié le test de puissance pour le cas à bande étroite, et l'on s'est rendu compte que dans un cas pratique, les courbes obtenues ont la bonne forme, mais il est difficile de statuer sur le test puisqu'on ne connaît pas la puissance théorique du signal. Ensuite, on a utilisé MUSIC-FBSS et SEM sur les signaux reçus par le réseau de capteurs. Les résultats obtenus ont montré la supériorité de l'ajout de SEM lorsque le réseau est petit par rapport à FBSS pour le cas à bande étroite. Pour les mesures à large bande, on a utilisé des pièces musicales. Ces signaux ne comportent pas de porteuse, alors il a fallu faire à nouveau des tests pour la fréquence de référence. On a déterminé que d'utiliser la fréquence limite du réseau fournit les meilleurs résultats. En appliquant MUSIC-FP avec FBSS et avec SEM, on a obtenu des résultats plus précis lorsque l'on utilise SEM, ce qui confirme les résultats de simulations.

7.1 Travaux futurs

Dans ce travail, l'intégration de la SEM n'a pas été faite avec CSSM. L'une des principales raisons est la difficulté de déterminer une puissance théorique de chaque source. Il serait alors très pertinent d'effectuer plusieurs tests dans le but d'établir un critère pour la puissance de la source. Tel que l'on a vu au chapitre 4, CSSM fournit d'excellents résultats, même à faible SNR, alors il est fort possible que l'ajout d'élimination des sources puisse être très efficace.

Les résultats obtenus à l'aide de SEM lorsque l'on utilise les positions théoriques montrent que la première estimation de la direction d'arrivée est extrêmement importante. On a vu, autant dans le cas à bande étroite que pour le cas à large bande, que la première estimation est beaucoup plus difficile à effectuer due à la corrélation entre les sources. Il serait possible que l'utilisation de FBSS au moment de faire l'estimation d'arrivée d'une des sources permette une meilleure estimation de la DOA de chaque source, ce qui par la suite améliorera les résultats finaux.

Tel que l'on a vu à la section 4.2, le choix de la fréquence de référence est très critique à l'obtention d'une estimation de DOA précise. On a aussi fait l'hypothèse que la principale raison du choix optimal dans notre cas provient de la symétrie du spectre, sans toutefois l'analyser avec plus de précision. En particulier, dans le cadre de ce mémoire, on s'est limité à des sources modulées, mais il existe bien d'autres types de sources qui ne sont pas modulées par amplitude, donc pour lesquelles on n'a pas

déterminé de fréquence de référence. C'est ainsi qu'une analyse plus poussée de la phase et de sa transformation durant MUSIC-FP pour différents types de sources pourrait aider à déterminer une fréquence de référence universelle.

Finalement, les mesures faites au chapitre 6 ont confirmé que les méthodes fonctionnent bien dans la très grande majorité des cas malgré les conditions expérimentales qui n'étaient pas les meilleures. Si les mesures avaient été faites dans une salle anéchoïque faite pour les ondes acoustiques, assez grande pour accueillir un réseau d'au moins 12 capteurs et de nombreuses sources, il aurait été possible d'obtenir des résultats utiles pour faire une comparaison plus complète entre SEM et FBSS.

Bibliographie

- [1] A.-A. SAUCAN. « Student research highlight enhanced sonar bathymetry : An echo DOA tracking approach ». In : *Aerospace and Electronic Systems Magazine, IEEE* 30.6 (2015), p. 34–36.
- [2] Chen PENG et al. « New wideband DOA estimation algorithm for planar arrays ». In : *Intelligent Signal Processing and Communication Systems, 2007. ISPACS 2007. International Symposium on*. Nov. 2007, p. 76–79. DOI : 10.1109/ISPACS.2007.4445827.
- [3] D. KHAYKIN et B. RAFAELY. « Coherent signals direction-of-arrival estimation using a spherical microphone array : Frequency smoothing approach ». In : *Applications of Signal Processing to Audio and Acoustics, 2009. WASPAA '09. IEEE Workshop on*. Sept. 2009, p. 221–224. DOI : 10.1109/ASPAA.2009.5346492.
- [4] J. CAPON. « High-resolution frequency-wavenumber spectrum analysis ». In : *Proceedings of the IEEE* 57.8 (août 1969), p. 1408–1418. ISSN : 0018-9219. DOI : 10.1109/PROC.1969.7278.
- [5] G BIENVENU et L KOPP. « Principe de la goniométrie passive adaptative ». In : *7° Colloque sur le traitement du signal et des images, FRA, 1979*. GRETSI, Groupe d'Etudes du Traitement du Signal et des Images. 1979.
- [6] Ralph O SCHMIDT. « Multiple emitter location and signal parameter estimation ». In : *Antennas and Propagation, IEEE Transactions on* 34.3 (1986), p. 276–280.
- [7] F. COLONE, R. CARDINALI et P. LOMBARDO. « Cancellation of clutter and multipath in passive radar using a sequential approach ». In : *Radar, 2006 IEEE Conference on*. 2006, pp.698–722.
- [8] L.-P. GILL, D. GRENIER et J.-Y. CHOUINARD. « Deletion of the signal in direct way by clean analysis in a system of coherent passive detection ». In : *Electrical*

and Computer Engineering, 2009. CCECE '09. Canadian Conference on. Mai 2009, p. 1099–1104.

- [9] Guaning SU et M. MORF. « The signal subspace approach for multiple wide-band emitter location ». In : *Acoustics, Speech and Signal Processing, IEEE Transactions on* 31.6 (déc. 1983), p. 1502–1522. ISSN : 0096-3518. DOI : 10.1109/TASSP.1983.1164233.
- [10] P. HELIN et al. « Likelihood based combining of subband estimates for wideband DOA ». In : *Image and Signal Processing and Analysis (ISPA), 2013 8th International Symposium on*. Sept. 2013, p. 320–325. DOI : 10.1109/ISPA.2013.6703760.
- [11] Petri HELIN. *Wideband Direction of Arrival Estimation and Sparse Modeling for Underwater Surveillance*. Mémoire de maîtrise, Tampere University of Technology, Finlande, 2013.
- [12] M.R. AZIMI-SADJADI, A. PEZESHKI et N. ROSEVEARE. « Wideband DOA estimation algorithms for multiple moving sources using unattended acoustic sensors ». In : *Aerospace and Electronic Systems, IEEE Transactions on* 44.4 (oct. 2008), p. 1585–1599. ISSN : 0018-9251. DOI : 10.1109/TAES.2008.4667733.
- [13] Jian LI, Petre STOICA et Zhisong WANG. « On robust Capon beamforming and diagonal loading ». In : *Acoustics, Speech, and Signal Processing, 2003. Proceedings. (ICASSP '03). 2003 IEEE International Conference on*. T. 5. Avr. 2003, V–337–40 vol.5. DOI : 10.1109/ICASSP.2003.1199947.
- [14] H. WANG et M. KAVEH. « Coherent signal-subspace processing for the detection and estimation of angles of arrival of multiple wide-band sources ». In : *Acoustics, Speech and Signal Processing, IEEE Transactions on* 33.4 (août 1985), p. 823–831. ISSN : 0096-3518. DOI : 10.1109/TASSP.1985.1164667.
- [15] H. HUNG et Mostafa KAVEH. « Focussing matrices for coherent signal-subspace processing ». In : *Acoustics, Speech and Signal Processing, IEEE Transactions on* 36.8 (août 1988), p. 1272–1281. ISSN : 0096-3518. DOI : 10.1109/29.1655.
- [16] JE EVANS, JR JOHNSON et DF SUN. « High resolution angular spectrum estimation techniques for terrain scattering analysis and angle of arrival estimation ». In : *Proc. 1st ASSP Workshop Spectral Estimation*. 1981, p. 134–139.
- [17] S.U. PILLAI et B.H. KWON. « Forward/backward spatial smoothing techniques for coherent signal identification ». In : *Acoustics, Speech and Signal Processing, IEEE Transactions on* 37.1 (jan. 1989), p. 8–15.

- [18] E.M. AL-ARDI, R.M. SHUBAIR et M.E. AL-MUALLA. « Computationally efficient DOA estimation in a multipath environment using covariance differencing and iterative spatial smoothing ». In : *Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on*. 2005, 3805–3808 Vol. 4.
- [19] K.P. RAY, R.K. KULKARNI et B. KASYAP RAMKRISHNAN. « DOA estimation in a multipath environment using covariance differencing and iterative Forward and Backward spatial smoothing ». In : *Recent Advances in Microwave Theory and Applications, 2008. MICROWAVE 2008. International Conference on*. 2008, p. 794–796. DOI : 10.1109/AMTA.2008.4763086.
- [20] H. AKAIKE. « A new look at the statistical model identification ». In : *Automatic Control, IEEE Transactions on* 19.6 (déc. 1974), p. 716–723. ISSN : 0018-9286. DOI : 10.1109/TAC.1974.1100705.
- [21] Paul E HOWLAND, D MAKSIMIUK et G REITSMA. « FM radio based bistatic radar ». In : *IEE Proceedings-Radar, Sonar and Navigation* 152.3 (2005), p. 107–115.
- [22] Rajesh SAINI, M CHERNIAKOV et V LENIVE. « Direct path interference suppression in bistatic system : DTV based radar ». In : *Radar Conference, 2003. Proceedings of the International*. IEEE. 2003, p. 309–314.
- [23] KS KULPA et Z CZEKAŁA. « Masking effect and its removal in PCL radar ». In : *IEE Proceedings-Radar, Sonar and Navigation* 152.3 (2005), p. 174–178.
- [24] Mireille KUBICA et al. « Optimum target detection using illuminators of opportunity ». In : *Radar, 2006 IEEE Conference on*. IEEE. 2006, 8–pp.
- [25] Alberto LEON-GARCIA. *Probability, statistics, and random processes for electrical engineering*. Upper Saddle River, NJ : Pearson/Prentice Hall, 2008. ISBN : 9780131471221 0131471228.
- [26] Mathieu GALLICHAND. *Réalisation d'un réseau linéaire de capteurs acoustiques pour la localisation de sources sonores distordues*. Mémoire de maîtrise, Quebec, Canada, 2014.

Annexe A

Article soumis

Résumé

Cet article présente une nouvelle méthode d'estimation de direction d'arrivée un signal composé d'une source acoustique unique et de ses échos, qui sont beaucoup plus faibles. Un développement des relations mathématiques entre la matrice de covariance des signaux reçus, les vecteurs directeurs et la puissance entre les sources est présenté dans le but d'exposer la contribution de toutes les sources. Deux algorithmes sont alors développés, un premier pour estimer la direction d'arrivée de chaque source individuellement et un deuxième pour déterminer la puissance de chaque source. L'idée derrière la méthode par élimination des sources (SEM) est d'éliminer, une par une, la contribution de chaque source à la matrice de covariance des signaux reçus pour améliorer l'estimation de la direction d'arrivée. Des résultats de simulations Monte Carlo sont présentés, montrant que SEM est meilleur que la méthode *Multiple Signal Classification* (MUSIC) avec un lissage spatial de type *forward-backward* pour positionner les échos dont le ratio écho-à-bruit est situé entre -13 dB et -17 dB. Des résultats expérimentaux montrent que dans le cas d'un petit réseau de capteurs et des sources de puissances différentes, SEM donne de meilleurs résultats que MUSIC avec un lissage spatial de type *forward-backward*.

Direction of arrival estimation of acoustic echoes using Source Elimination Method

Marc-André Guérard, Dominic Grenier

Abstract—A new method for direction of arrival (DOA) estimation for a single narrowband acoustic source with multiple low power echoes is proposed. The mathematical relations between the signal covariance matrix, the sources steering vectors, and the power of the sources are developed in order to expose the contribution of each source. Algorithms to find each of the sources DOA and their respective power are presented. The Source Elimination Method (SEM), based on the elimination of the contribution of each source to improve the DOA estimation, is developed. Monte Carlo simulations are presented, showing that SEM yields more accurate results than Multiple Signal Classification (MUSIC) with forward-backward spatial smoothing to find the echoes DOA with an echo-to-noise ratio between -13 dB to -17 dB. Experimental results show that for a small array and two sources with different power, SEM outperforms MUSIC with forward-backward spatial smoothing.

Index Terms—Acoustic arrays, Acoustic signal processing, Direction of arrival estimation, Acoustic position measurement

I. INTRODUCTION

The direction of arrival (DOA) estimation of an acoustic source and its echoes is a subject of great interest in bathymetry [1] and in the acoustic analysis of theaters, arenas or other live music rooms. However, the DOA estimation of echoes by high resolution algorithms are difficult to extract for two reasons:

- 1) The power ratio between the echoes and the direct signal is small because a big part of the power is lost in the reflections.
- 2) Since the echoes are a copy of the direct signal, the echoes are highly correlated with the direct signal

One of the most used algorithms to find the DOA of a narrowband source is Multiple Signal Classification (MUSIC) [2] [3]. This algorithm uses the noise subspace of the covariance matrix calculated from the signals received by a microphone array. It yields accurate results if the sources are the same power and are not correlated. In the presence of acoustic echoes, these criteria do not hold.

This paper is divided into four sections. In section 2 we develop a signal model to find a relationship between the signals covariance matrix, the sources steering vectors and the sources power. In section 3, we develop the source

elimination method (SEM) and then we find two unknown parameters for the algorithm, the DOA of the first source and the sources power. In section 4, we present Monte Carlo simulation results and, in section 5, we show and analyze experimental results.

II. SIGNAL MODEL

Consider a uniform linear array (ULA) composed of N microphones with inter-elements distance d measuring M sources coming from the DOA $[\theta_1, \theta_2, \dots, \theta_M]$ (fig 1). The array output for the time t_k can be described as:

$$\mathbf{x}(t_k) = \sum_{m=1}^M s_m(t_k) \mathbf{a}(\theta_m) + \mathbf{n}(t_k) \quad (1)$$

where s_m is the signal delivered by the source m , $\mathbf{a}(\theta_m)$ is

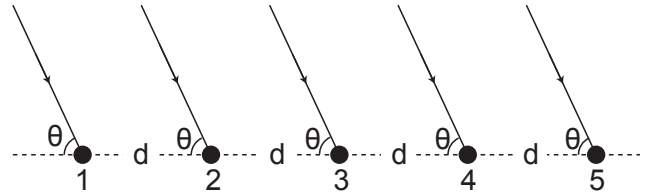


Fig. 1. Linear array of microphones.

the steering vector at the direction θ_m , $\mathbf{n}(t_k)$ is the additive Gaussian white noise array, and $\mathbf{x}(t_k)$ is the received signal array for the snapshot k . With multiple snapshots, equation (1) can be rewritten as:

$$\mathbf{X} = \mathbf{A} \mathbf{S} + \mathbf{N} \quad (2)$$

where

$$\mathbf{A} = [\mathbf{a}_1 \ \mathbf{a}_2 \ \dots \ \mathbf{a}_M] \quad (3)$$

$(N \times M)$

$$\mathbf{S} = [s_1 \ s_2 \ \dots \ s_K] \quad (4)$$

$(M \times K)$

$$\mathbf{N} = [\mathbf{n}_1 \ \mathbf{n}_2 \ \dots \ \mathbf{n}_K] \quad (5)$$

$(N \times K)$

The covariance matrix of the received signals can be described as:

$$\mathbf{R}_{xx} = E \left\{ \mathbf{x}_k \mathbf{x}_k^\dagger \right\} \quad (6)$$

$$= \mathbf{A} \underbrace{E \left\{ \mathbf{s}_k \mathbf{s}_k^\dagger \right\}}_{\mathbf{R}_{ss}} \mathbf{A}^\dagger + E \left\{ \mathbf{n}_k \mathbf{n}_k^\dagger \right\} \quad (7)$$

\mathbf{R}_{nn}

$$= \underbrace{\mathbf{A} \mathbf{R}_{ss} \mathbf{A}^\dagger}_{\mathbf{R}_{xx-}} + \mathbf{R}_{nn} \quad (8)$$

Manuscript submitted on May 31, 2016. This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) under Discovery Grant.

Marc-André Guérard is with the Electrical and Computer Engineering of Université Laval, 2325 Rue de l'Université, Québec, QC G1V 0A6, Canada (email: marc-andre.guerard.1@ulaval.ca)

Dominic Grenier is with the Electrical and Computer Engineering of Université Laval, 2325 Rue de l'Université, Québec, QC G1V 0A6, Canada (email: dgrenier@gel.ulaval.ca)

where \dagger is the conjugate-transpose operator, \mathbf{R}_{ss} is the sources covariance matrix, \mathbf{R}_{nn} is the noise covariance matrix, \mathbf{R}_{xx} is the covariance matrix from the received signals and \mathbf{R}_{xx-} is the covariance matrix of the received signals without the noise. If the sources are uncorrelated, then the diagonal elements of \mathbf{R}_{ss} are the sources power and the extra-diagonal elements are null. The same goes for \mathbf{R}_{nn} , so if the noise is independent between every microphones, then \mathbf{R}_{nn} is diagonal.

If we diagonalise \mathbf{R}_{xx-} , we obtain the following equation:

$$\mathbf{R}_{xx-} = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^\dagger = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^\dagger \quad (9)$$

where \mathbf{V} is a matrix composed of the M eigenvectors corresponding to the M greatest eigenvalues, given by the diagonal matrix $\mathbf{\Lambda}$. In this paper, we arrange the eigenvalues in ascending order, so we will refer the biggest one as the first eigenvalue, and the smallest one as the last eigenvalue. It is important to note that a negative eigenvalue is considered as smaller than any positive eigenvalues. The column vectors of \mathbf{A} and \mathbf{V} defines the M dimensional sources subspace. The complementary subspace of the signal subspace is called the noise subspace. The steering matrix \mathbf{A} is an unknown, but \mathbf{R}_{xx} can be estimated from the K snapshots $\mathbf{x}(t_k)$, and with its estimation we can calculate the eigenvector matrix \mathbf{V} and the eigenvalues. It is then possible to do an orthogonal projection with the approximated noise subspace and the steering vector for every angle possible to obtain a pseudospectrum with the zeros being the DOA.

If the sources are correlated, then \mathbf{R}_{ss} is not diagonal, and equation (9) does not hold. This is a problem when using MUSIC because the signal subspace defined by the steering matrix and the subspace defined by the M firsts eigenvectors does not define the same subspace. To work around this problem, it is possible to perform MUSIC with correlated sources by performing the Forward Spatial Smoothing Technique (FSS) [4] beforehand. This algorithm separates the N microphones in R sub-arrays, then compute the covariance matrix of each sub-array and then calculate a new covariance matrix by taking the mean of all the sub-arrays covariance matrices. A more efficient algorithm, called the Forward-Backward Spatial Smoothing (FBSS) [5], which uses the same steps as the spatial smoothing technique, but adds another step by inverting all the microphones and perform another sub-array separation to calculate their covariance matrices to take the mean of these covariance matrices. This technique was proven to be more efficient than FSS [6]. More recent methods such as FBSS with augmented arrays (FBSSOA) [7] has been shown to be more efficient than FBSS in the case of small arrays. Another method, called differencing and iterative Forward and Backward spatial smoothing, developed by Raye et al. [8], is used in three steps. First, it finds the direct path sources DOA first to then eliminate their contributions to the covariance matrix and finally uses FSS and MUSIC on the remaining covariance matrix. Ray et al [9] improved this method by using FBSS instead of FSS. Another recent method, called

iterative MUSIC, used for Magnetoencephalography (MEG) and Electroencephalography (EEG) [10], uses a suppression matrix to eliminate the common components of the correlated signals and then using an iterative process to find two sources. However, this method cannot be used in the case of small echoes, since eliminating the common component would mean eliminating the sources completely.

These methods are designed to work when the sources power are of the same magnitude, which is not necessarily true in the case of echoes detection. Colone [11] and Gill [12] created a new method, using CLEAN analysis to eliminate the contribution of the direct path source in the time signals received by an antenna by projecting the time signals in the complementary subspace of the direct path sources subspace. We can express this projection by using the following equation:

$$\mathbf{x}_c(t_k) = \mathbf{E}_c^H \mathbf{x}(t_k) \quad (10)$$

where

$$\mathbf{E}_c = [\mathbf{v}_{M+1} \ \mathbf{v}_{M+2} \ \dots \ \mathbf{v}_N] \quad (11)$$

This method is used to eliminate the contribution of the direct path source for passive-coherent radar detection. In this case, it is possible to make the assumption that the sources are uncorrelated because the Doppler effect is important and because the time delay between the direct path and the echoes is important. This way, the first eigenvector contains the direct path signal completely. This assumption cannot be made in the case of acoustic sources because the Doppler effect is not present for stationary sources and because the time delay is not necessarily long compared to the decorrelation time.

Other recent methods use a learning-based approach by extracting information from the generalized cross correlation vectors [13]. These methods, however, are used only to find the direct path source DOA and to cancel the reverberations.

III. SOURCE ELIMINATION METHOD

The idea of eliminating the direct path signals present in [11], [12], and in [8] is interesting because it can help to recover the information lost in the direct path signals. However, the CLEAN analysis method, as discussed in [12], cannot be used for low power acoustic echoes, and [8] does not explain how to calculate the covariance matrix for the direct path sources.

The source elimination method (SEM) is based on the idea to remove, one by one, the contribution of each source from \mathbf{R}_{xx} . When all the contributions are eliminated, we can approximate the noise covariance matrix by subtracting the contribution of every source from \mathbf{R}_{xx} . Once \mathbf{R}_{nn} is approximated, we can restart the process for another iteration with the noise contribution eliminated.

Before we carry on, it is necessary to define a few elements:

- 1) **First order reduced covariance matrix** (\mathbf{R}_{red_1}): The covariance matrix obtained when the exact contribution (power) of the first source is removed from the original covariance matrix. Higher order reduced covariance matrices are obtained when the contribution of more than one source is removed from the original covariance matrix.
- 2) **Partially reduced covariance matrix** : Matrix obtained when the contribution of a source is not completely removed from the original covariance matrix.
- 3) **Residual eigenvector** : Eigenvector of the partially reduced covariance matrix which is parallel to the eigenvector of the first source.
- 4) **Residual eigenvalue**: Eigenvalue corresponding to the residual eigenvector.

The first order reduced covariance matrix can be written as:

$$\mathbf{R}_{red_1} = \mathbf{R}_{xx} - \mathbf{R}_1 = \mathbf{R}_{xx} - \mathbf{a}_1 \sigma_1^2 \mathbf{a}_1^\dagger \quad (12)$$

where σ_1^2 and \mathbf{a}_1 are, respectively, the power and the steering vector of source 1 and \mathbf{R}_1 is the source 1 contribution to the covariance matrix.

The goal of SEM is to retrieve the best approximation of \mathbf{R}_{red_1} from the estimation $\hat{\mathbf{R}}_1$. Equation (12) shows that to calculate $\hat{\mathbf{R}}_{red_1}$, we need to estimate \mathbf{R}_{xx} , \mathbf{a}_1 and σ_1^2 . A good approximation of \mathbf{R}_{xx} , noted $\hat{\mathbf{R}}_{xx}$, can be calculated from the snapshots, as indicated previously. Two unknown remains in the equation 12, \mathbf{a}_1 and σ_1^2 are more difficult to find. Approximation methods for both elements are the subjects of the next two sections.

A. Finding the steering vector

The steering vector for the source m is given by:

$$\mathbf{a}(\theta_m) = \begin{bmatrix} 1 & e^{-j\phi_m} & e^{-j2\phi_m} & \dots & e^{-j(N-1)\phi_m} \end{bmatrix}^T \quad (13)$$

(Nx1)

with

$$\phi_m = \frac{2\pi d \cos(\theta_m)}{f} \quad (14)$$

The only unknown in these equations is θ_m , which is the parameter we want to estimate in this section. MUSIC considers the signal subspace and the steering vectors subspace to define the same subspace, but it does not consider that individually, the eigenvectors to be parallel to the steering vectors. This implies that we have to use MUSIC assuming the right number of sources, we cannot find only one source, especially when the sources are correlated.

If we expand equation 9, we get:

$$\mathbf{R}_{xx-} = \mathbf{A}\mathbf{R}_{ss}\mathbf{A}^\dagger = \mathbf{a}_1 \sigma_1^2 \mathbf{a}_1^\dagger + \dots + \mathbf{a}_M \sigma_M^2 \mathbf{a}_M^\dagger \quad (15)$$

$$= \sum_{m=1}^M \mathbf{a}(\theta_m) \sigma_m^2 \mathbf{a}^\dagger(\theta_m) \quad (16)$$

In this equation, \mathbf{a}_m is the steering corresponding to the m^{th} source, and σ_m is the power of the m^{th} source. Adding the noise contribution, we obtain:

$$\mathbf{R}_{xx} = \sum_{m=1}^M \mathbf{a}(\theta_m) \sigma_m^2 \mathbf{a}^\dagger(\theta_m) + \mathbf{R}_{nn} \quad (17)$$

Equation (17) considers non correlated sources, which is not true in our case. However, it is possible to see \mathbf{R}_{ss} as a sum of a diagonal matrix, \mathbf{R}_{ssdiag} , which diagonal elements are the power of each source, and a null-diagonal matrix, $\mathbf{R}_{ssdiag0}$, which extra-diagonal elements are $\mu_{ij}\sigma_i\sigma_j$, μ_{ij} being the correlation coefficient between the sources i and j . We obtain a general case of equation (17):

$$\mathbf{R}_{xx} = \underbrace{\mathbf{A}\mathbf{R}_{ssdiag}\mathbf{A}^\dagger}_{\mathbf{R}_{xx0-}} + \mathbf{A}\mathbf{R}_{ssdiag0}\mathbf{A}^\dagger + \mathbf{R}_{nn} \quad (18)$$

In this equation, \mathbf{R}_{xx0-} is the covariance matrix received by the array without noise and correlation between sources. This matrix contains the important information that is needed to use MUSIC. If the power ratio between the direct path source and the echoes is large, then the matrix $\mathbf{R}_{ssdiag0}$ is small compared to \mathbf{R}_{xx0-} , so that it does not significantly affects \mathbf{R}_{xx-} . In this case, it is then possible to use MUSIC to approximate the DOA of the first source only, and then the steering vector \mathbf{a}_1 .

B. Power testing

It is possible to find σ_1^2 by observing the behavior of the first and last eigenvalues of the partially reduced covariance matrix for a range of different powers removed from the original covariance matrix (fig. 2). Since the signals are computer generated, they have no physical meaning, so we decided to use metric, a dummy unit, to express the signal power that could either be W or kW . The results in fig. 2 show that the intersection between the absolute value of the first and the last eigenvalues is close to the real power of the source, within 1% error.

The main interpretation of these results is that with increasing power, the residual eigenvalue gets smaller and smaller, down to a point where its value is of the same magnitude of the echoes eigenvalues. In this case, the matrix $\mathbf{R}_{ssdiag0}$ cannot be neglected. Consequently, the direct path source and the echoes are combined in a single eigenvalue and a single eigenvector, which is not parallel to any steering vector. By continuing subtracting the contribution of the first source, the last eigenvalue gets negative because it becomes the residual eigenvalue, so we are subtracting a vector now included in the noise subspace. The intersection point shows that nearly all of the first source contribution has been removed and that the last eigenvalue is now the residual eigenvalue, which is an acceptable approximation the power in our case. It is important to note that when the first source power is removed, then the first eigenvalue now corresponds to the second source.

C. Algorithm definition

Now that we can approximate every parameters in equation (12), it is possible to formulate an algorithm that can eliminate the contribution of the direct path source and its echoes.

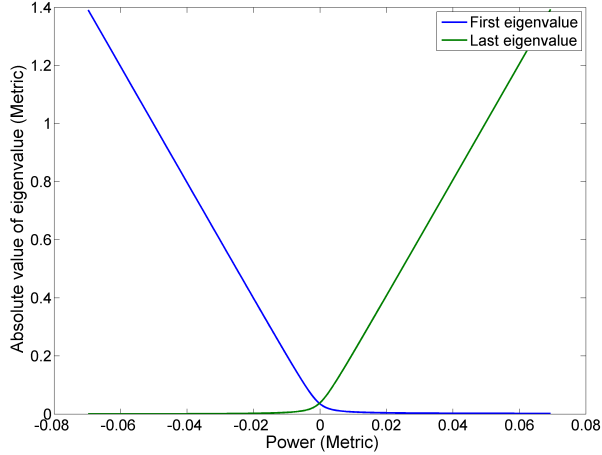


Fig. 2. Absolute value of the first and last eigenvalue of the partially reduced covariance matrix of multiple estimated power for a signal composed of a source and two echoes without noise.

- 1) Calculate, from the K snapshots, an approximation of \mathbf{R}_{xx} .
- 2) Using MUSIC assuming one source, calculate θ_1 , the DOA of the direct path source.
- 3) Estimate the power of the sources source by determining the intersection point between the curves of the absolute value of the biggest and smallest eigenvalue of partially reduced covariance matrix. Since the curves are differentiable, it is possible to use Newton's method to find their intersection instead of scanning for all powers.
- 4) Calculate an approximation of \mathbf{R}_{red_1} using equation (12).
- 5) Restart steps 2-5 for each source of the signal.
- 6) Estimate the noise covariance matrix \mathbf{R}_{nn} by subtracting $\sum_{m=1}^M \tilde{\mathbf{R}}_{red_m}$ from \mathbf{R}_{xx} .
- 7) Make a second iteration of steps 2-5 by using $\tilde{\mathbf{R}}_{xx} - \tilde{\mathbf{R}}_{nn}$ instead of $\tilde{\mathbf{R}}_{xx}$.

The last step in the algorithm is important because it helps to partially eliminate the noise and the correlation from the original covariance matrix before doing a new DOA estimation for every source. It is then possible to use multiple iterations for steps 2-7 to get each time a new estimation of the noise covariance matrix. In this paper, we will only use two iterations since preliminary simulation results have shown that more iterations do not necessarily yield to better results.

IV. SIMULATIONS

We assume a ULA $N = 20$ microphones, distant by 5 cm. We also assume a single narrowband source at 2000 Hz, situated at $\theta_1 = 70^\circ$ and two echoes situated at $\theta_2 = 115^\circ$ and $\theta_3 = 135^\circ$. The Echo-to-Direct path signal Ratio (EDR) is -17 dB.

Fig. 3 shows the probability of false alarm (PFA) for two echoes (sources 2 and 3) of 200 Monte Carlo trials in function of the ENR for three different cases:

- 1) MUSIC used with forward-backward spatial smoothing with the use of $R = 8$ sub arrays, which is optimal for this system.
- 2) SEM algorithm using MUSIC to find the initial DOA estimation for each source (step 2 in the algorithm definition).
- 3) SEM algorithm using the theoretical DOA in the first iteration. This curve is only made to give the theoretical limitations of SEM algorithm. Also, this curve shows the bias due to the power estimation algorithm by eliminating the bias due to the preliminary DOA estimation.

We define the PFA as the probability that no source is found within a 10° range around the theoretical source position. The reason this choice of range is the ambiguity between the two echoes that can occurs if the range is greater than 10° . In all the simulation results shown in this article, we do not show the results for the direct path source because the goal of this article is to detect the echoes and since the direct-to-noise ratio is high, the bias for this DOA is always less than 0.1° .

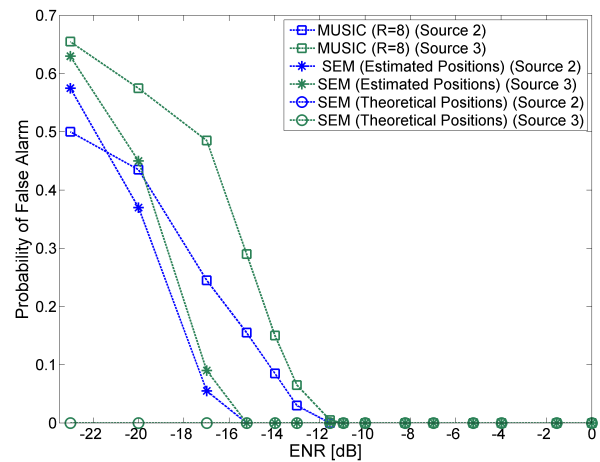


Fig. 3. Probability of false alarm for different ENR of 200 Monte Carlo trials for multiple DOA estimation methods.

The results presented in fig. 3 show that both methods can be used to find the source at high ENR 100% of the time. The PFA for MUSIC with spatial smoothing starts to increase at an ENR of -13 dB (then the direct path signal to noise ratio is equal to 4 dB) and continues to grow as the ENR gets lower, whereas for the SEM method the PFA starts to increase at -17 dB. The theoretical positions curves show a null PFA for all ENR, showing the importance of the first DOA estimation for each source.

Fig. 4 shows the mean error and fig. 5 shows the standard deviation when a true detection occurs for the same 200 Monte Carlo trials used in fig. 3. We did not include the results in the case of a false detection because when it happens, the choice of the peak representing each echoes is arbitrary. Both curves for MUSIC show the same trends, an exponential growth, going from 0 dB to -16 dB after which

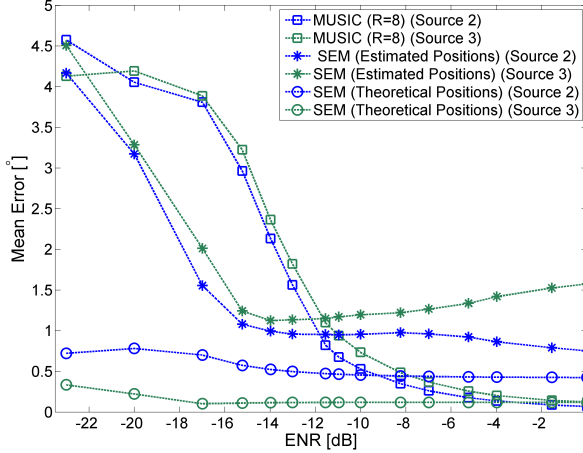


Fig. 4. Mean error for different ENR of 200 Monte Carlo trials for multiple DOA estimation methods in the case of a real detection.

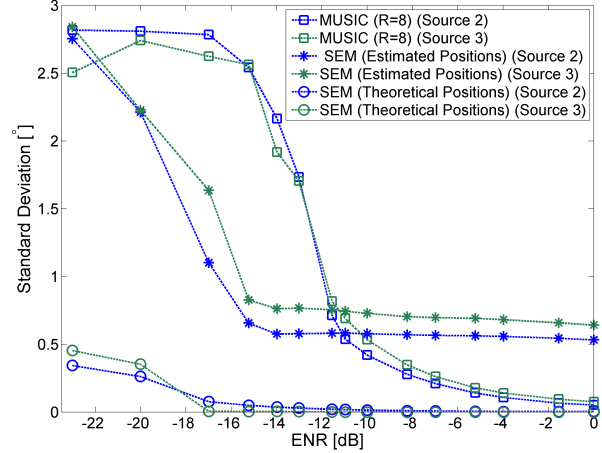


Fig. 5. Standard deviation for different ENR of 200 Monte Carlo trials for multiple DOA estimation methods in the case of a real detection.

the growth stabilizes. It is important to realize that in low ENR, only a small portion of the 200 trials is included in the results, so both curves just represent the trends. The curves for SEM with estimated are constant from 0 dB to -16 dB and then increasing linearly. Finally, the curves for SEM using theoretical positions for the first DOA estimation show a small constant bias for every ENR with a small increase at -16 dB. For its standard deviation, the curve is near 0° up to -20 dB, where we can see a slight increase. These results show that SEM algorithm, as defined in this paper, is better than MUSIC with forward-backward spatial smoothing in the mid ENR range, but is worse in the high ENR range. The explanation for the results in low ENR is that the noise level gets too high to find the echoes DOAs precisely enough to get an appropriate approximation of the noise covariance matrix. These errors trigger a domino effect in the second iteration, which degenerate deeply the results in low ENR. The curves using the theoretical positions for the first DOA estimation show that SEM depends strongly on the DOA estimation in the first iteration. This result means that most of the false detections, the bias, and the standard deviation comes from the algorithm for the preliminaries DOAs and not from the algorithm for the power estimation.

V. EXPERIMENTAL RESULTS

To further test SEM, measurements were taken with the experimental setup showed in fig. 6. Two speakers are placed in an anechoic chamber and an ULA composed of $N = 6$ microphones distant by $d = 5$ cm is placed outside of the chamber. The speakers positions are $\theta_1 = 99 \pm 3^\circ$ and $\theta_2 = 85 \pm 3^\circ$. The uncertainty comes from a 2° uncertainty from the array axis measurement (the 0° reference) and an additional 1° from the measurement of the speakers angular position. The distance between the speakers and the number of microphones were limited by the size of the anechoic chamber. An anechoic wall is placed behind the array to prevent sound reflections on the wall behind the array to negatively affect the results. The use an anechoic chamber

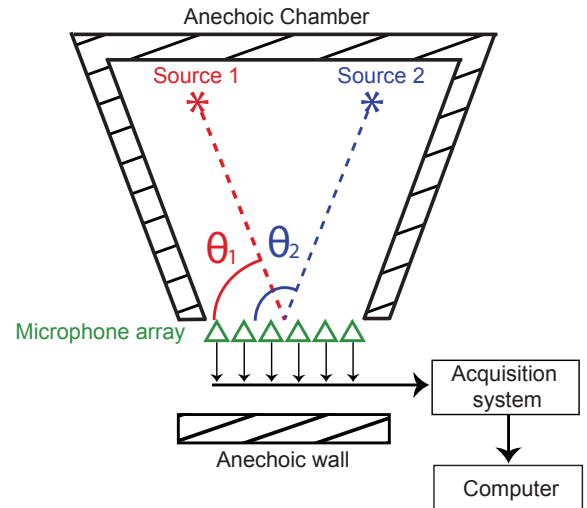


Fig. 6. Experimental setup

reduce the noise level so that we can test the method in the best case scenario.

Once detected by the microphones, the resulting signals go to an acquisition system to be digitized and sent to a computer. The sound coming out of the speakers is narrowband at $f = 2$ kHz and completely correlated between both speakers. To test the small echoes approximation, we vary the power of the second source to change the EDR. Pseudospectrums obtained with MUSIC algorithm with spatial smoothing ($R = 3$) for different EDR between the sources are shown in fig. 7 and the pseudospectrums obtained using SEM for the same measurements are found in fig. 8.

The results presented in fig. 7 shows that MUSIC with forward-backward spatial smoothing can efficiently find the DOA for the first source, but has more problems finding the

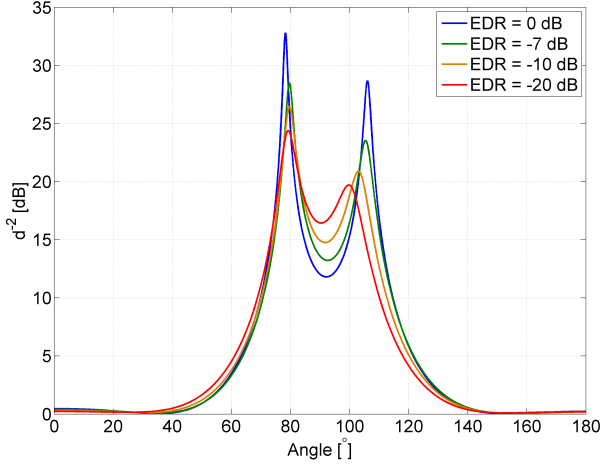


Fig. 7. Experimental MUSIC with forward-backward spatial smoothing ($R = 3$) pseudospectrums for different power ratio between the sources for measurements taken in an anechoic chamber.

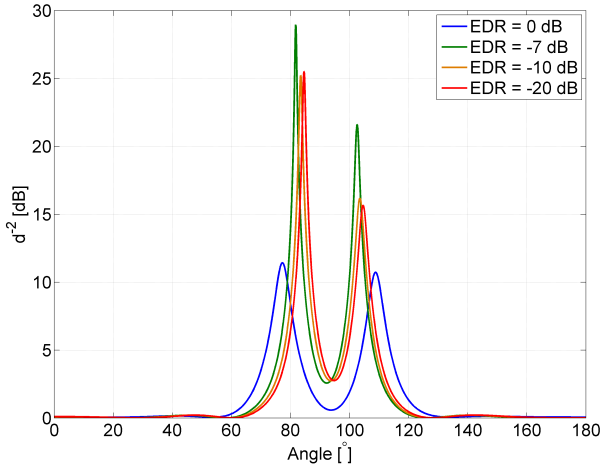


Fig. 8. Experimental SEM pseudospectrums for different power ratio between the sources for measurements taken in an anechoic chamber.

DOA of the second source. Whilst the peak of the second source in the high EDR is high, its position is outside the certainty range. The opposite is true for the low EDR. The main reason for these results is that the correlation between the sources is not completely removed with the forward-backward spatial smoothing with a small array.

In fig. 8, we can see that if the power of the sources are the same, SEM pseudospectrum has peaks outside the certainty range. The main reason for these results comes from the correlation between the sources, which creates an important bias in the initial DOA estimation. This bias creates a series of errors to finally make a pseudospectrum with peaks at the wrong position. If the second source is less powerful than the first one, we can see that the peaks positions are all in the uncertainty range and that their separation is constant.

From these results, we can conclude that if the EDR is

lower than 0 dB, and if the array is composed of a small number of microphones, SEM outperforms MUSIC with forward-backward spatial smoothing because its results are more accurate with strong source correlation. However, due to limited space, it was impossible to test the algorithm with a bigger array.

VI. CONCLUSION

In this paper, we presented a new algorithm to find the direction of arrival of acoustic echoes by eliminating the contribution in the covariance matrix of each source. We have first shown that it is possible to find the DOA of the direct path source with MUSIC if the echoes are low power compared to the direct path source. We then developed a test to find, with an error of 1%, the power of each source. Once these two parameters are estimated, we can eliminate the contribution of each source in the covariance matrix. The last step is to approximate the noise covariance matrix to make another iteration, without most of the noise contribution. Using Monte Carlo simulations, we have found that this new algorithm yields better results than MUSIC with forward-backward spatial smoothing in mid ENR range, but worse results in high and low ENR range. The small bias in the DOA of echoes for low ENR range comes from a small inaccuracy in DOA estimation during the first iteration, whereas the large bias in the low ENR range comes from a large bias in the DOA estimation during the first iteration. These results show the importance of the preliminary DOA estimation in the algorithm. Experimental results have shown that in the case of high ENR, a small EDR, and small array, SEM yields more stable results than MUSIC.

REFERENCES

- [1] A.-A. Saucan. “Student research highlight enhanced sonar bathymetry: An echo DOA tracking approach”. In: *Aerospace and Electronic Systems Magazine, IEEE* 30.6 (2015), pp. 34–36.
- [2] Ralph O Schmidt. “Multiple emitter location and signal parameter estimation”. In: *Antennas and Propagation, IEEE Transactions on* 34.3 (1986), pp. 276–280.
- [3] G Bienvenu and L Kopp. “Principe de la goniométrie passive adaptative”. In: *7 Colloque sur le traitement du signal et des images, FRA, 1979*. GRETSI, Groupe d’Etudes du Traitement du Signal et des Images. 1979.
- [4] JE Evans, JR Johnson, and DF Sun. “High resolution angular spectrum estimation techniques for terrain scattering analysis and angle of arrival estimation”. In: *Proc. 1st ASSP Workshop Spectral Estimation*. 1981, pp. 134–139.
- [5] S.U. Pillai and B.H. Kwon. “Forward/backward spatial smoothing techniques for coherent signal identification”. In: *Acoustics, Speech and Signal Processing, IEEE Transactions on* 37.1 (1989), pp. 8–15.
- [6] J.S. Thompson, Peter M. Grant, and B. Mulgrew. “Performance of spatial smoothing algorithms for correlated sources”. In: *Signal Processing, IEEE Transactions on* 44.4 (1996), pp. 1040–1046.
- [7] T. Iwai et al. “DOA estimation by MUSIC algorithm using forward-backward spatial smoothing with overlapped and augmented arrays”. In: *Antennas and Propagation (ISAP), 2014 International Symposium on*. 2014, pp. 375–376.
- [8] E.M. Al-Ardi, R.M. Shubair, and M.E. Al-Mualla. “Computationally efficient DOA estimation in a multipath environment using covariance differencing and iterative spatial smoothing”. In: *Circuits and Systems, 2005. ISCAS 2005. IEEE International Symposium on*. 2005, 3805–3808 Vol. 4.
- [9] K.P. Ray, R.K. Kulkarni, and B. Kasyap Ramkrishnan. “DOA estimation in a multipath environment using covariance differencing and iterative Forward and Backward spatial smoothing”. In: *Recent Advances in Microwave Theory and Applications, 2008. MICROWAVE 2008. International Conference on*. 2008, pp. 794–796. DOI: 10.1109/AMTA.2008.4763086.
- [10] Zhongping Tao and Junpeng Zhang. “Iterative MUSIC for Highly Correlated MEG Source Localization: A Simulation Study”. In: *Software Engineering (WCSE), 2013 Fourth World Congress on*. 2013, pp. 217–220.
- [11] F. Colone, R. Cardinali, and P. Lombardo. “Cancellation of clutter and multipath in passive radar using a sequential approach”. In: *Radar, 2006 IEEE Conference on*. 2006, 7 pp.–.
- [12] L.-P. Gill, D. Grenier, and J.-Y. Chouinard. “Deletion of the signal in direct way by clean analysis in a system of coherent passive detection”. In: *Electrical and Computer Engineering, 2009. CCECE '09. Canadian Conference on*. 2009, pp. 1099–1104.
- [13] Xiong Xiao et al. “A learning-based approach to direction of arrival estimation in noisy and reverberant environments”. In: *Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on*. IEEE. 2015, pp. 2814–2818.