



HAL
open science

**Contributions à la mise au point de méthodes
adaptatives de reproduction de champs sonores
multi-zone pour les auditeurs en mouvement : Sound
zones pour auditeurs en mouvement**

Georges Roussel

► **To cite this version:**

Georges Roussel. Contributions à la mise au point de méthodes adaptatives de reproduction de champs sonores multi-zone pour les auditeurs en mouvement : Sound zones pour auditeurs en mouvement. Acoustique [physics.class-ph]. Le Mans Université, 2019. Français. NNT : 2019LEMA1018 . tel-02383240

HAL Id: tel-02383240

<https://theses.hal.science/tel-02383240>

Submitted on 27 Nov 2019

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

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

THÈSE DE DOCTORAT DE

LE MANS UNIVERSITE
COMUE UNIVERSITÉ BRETAGNE LOIRE

ECOLE DOCTORALE N° 602
Sciences pour l'Ingénieur
Spécialité : « *Acoustique* »

Par

Georges ROUSSEL

Contributions à la mise au point de méthodes adaptatives de reproduction de champs sonores multi-zone pour les auditeurs en mouvement

« Sound zones pour auditeurs en mouvement »

Thèse présentée et soutenue à « Le Mans », le « 3 Juillet 2019 »
Unité de recherche : LAUM, UMR CNRS 6613
Thèse N° : 2019LEMA1018

Rapporteurs avant soutenance :

Philippe-Aubert Gauthier
Hervé Lissek

Professeur, Université de Sherbrooke, Sherbrooke, Canada
Professeur, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Suisse

Composition du Jury :

Président : Philippe Herzog
Examineurs : Etienne Corteel
Philippe Herzog

Directeur de Recherche, Arteac Labs, Marseille, France
Ingénieur de Recherche, L-Acoustics, Marcoussis, France
Directeur de Recherche, Marseille, France

Dir. de thèse : Laurent Simon
Co-dir. de thèse : Rozenn Nicol
Co-dir. de thèse : Guy Plantier

Professeur, LAUM, Le Mans, France
Ingénieure de Recherche, Orange Labs, Lannion, France
Professeur, LAUM, ESEO, Angers, France

Remerciements

Avant toutes choses, je souhaite remercier mes encadrants, Rozenn, Guy et Laurent qui m'ont fait confiance et m'ont soutenu avec une admirable bienveillance tout au long de ce parcours fascinant et instructif qu'est la thèse, tant au niveau scientifique que personnel. Ce fut un réel plaisir de travailler avec eux, même dans les moments difficiles, et pour ça je les remercie chaleureusement.

Je remercie aussi, bien entendu, mes rapporteurs, Hervé Lissek et Philippe-Aubert Gauthier, qui ont accepté d'évaluer mon travail de thèse. Je remercie également les membres du jury, Etienne Corteel et Philippe Herzog, de prendre le temps de participer à la soutenance de cette thèse. Je remercie également les membres de mon CSI, Finn Agerkvist et Philippe Herzog, de m'avoir conseillé lors de ces réunions annuelles.

J'en profite également pour remercier ceux qui m'ont entouré à Orange, l'équipe d'origine, mais aussi la seconde équipe qui m'a fait un accueil très chaleureux. Bien entendu, les thésards et amis (La liste est bien trop longue pour ne pas en oublier!), les postdocs et les stagiaires, particulièrement les deux que j'ai eu la chance d'encadrer : Nicolas et Christie, tous les deux des stagiaires formidables, chacun à leur façon.

Je tiens aussi particulièrement à remercier Rozenn, pour sa joie de vivre et sa bienveillance qu'elle partage avec tous, mais aussi pour ses coups de fouet lorsque c'était nécessaire. Aux mercredis gastronomiques et ensoleillés avec elle et Julian. Aux histoires sur les improbables colis de l'illustre inconnu(e). Aux plantes vertes du désespoir. Aux fou-rires...

Julian, mon co-bureau pendant près de deux ans, Loïc, mon co-bureau de la seconde équipe, et aussi Tinaboo, autre co-bureau, à distance cette fois-ci, pour leur joie de vivre et aussi leur capacité à me distraire et à me faire rêver avec leurs aventures ! Leny, mon compagnon de galère outre-manche, avec qui j'ai pu échanger sur notre lot de petits malheurs du quotidien et le fil à retordre que nous a donné le formidable sujet des *Sound Zones*.

Je remercie également Lionel pour ses conseils et pour sa poésie scientifique qu'il partage avec passion et que j'ai reçue avec bonheur ; Régis, l'homme aux 1001 brevets, qui ne manque jamais une occasion pour blaguer, même quand il s'agit d'aller jouer les éboueurs ; Serge, grand amoureux de la mer, qui m'a fait partager sa passion, mais aussi homme de mémoire, avec ses anecdotes les plus drôles et ses histoires les plus folles.

Ali, qui m'a accompagné, hébergé, nourri, motivé et supporté sur la fin de rédaction, dans le superbe cadre du bord de mer Perrosien.

Passons maintenant aux à-côtés de la thèse, qui donnent à ce parcours quelque chose de complet. Commençons par la Lilann, troupe d'impro, qui m'a fait découvrir cette pratique formidable et vivre des moments inoubliables dans cette grande famille. Continuons par les bars et leurs petites mains, qui n'ont jamais failli à reconforter les amis de tous bords. Je pense à Elise et William du Truc café, Jean-Marc des Valseuses ou encore Denis, du café Théodore.

Il y a bien entendu tous les Koupaings que je remercie... Les Koupaings D'Orange et de Lannion, bien sûr. Les Koupaings de promo, et particulièrement ceux qui se sont lancés dans cette même aventure, tous dans le même bateau, et toujours partant pour se retrouver dès que l'occasion se présente. Il y a aussi les Koupaings de longue date : Clo, Ambrichou, Jules et Val, même si on est pas toujours d'accord, on s'aime quand même ! Mais aussi la Team des Affrontaux, pour leur débilité notoire dans de nombreuses situations, qui ne manque pas de créer tout un tas d'aventures mémorables, plus improbables les unes que les autres ! Dis moi, ça ferait pas une bonne histoire à raconter ça ?...

Il y a bien entendu aussi Flo et Léo, les Koupaings de la montagne, qui m'ont fait et me feront encore voyager, pour m'évader toujours plus haut, toujours plus loin, toujours plus fort !

Je remercie aussi ma famille, particulièrement ma mère et ma soeur pour m'avoir soutenu. Je remercie tout particulièrement Julie. Ma Zulie, qui, depuis un petit bout de temps déjà, m'accompagne dans ma vie, me soutient, me supporte, me comprend (parfois difficilement... Mais bon, on comprend pas toujours ce que je raconte...).

Passons maintenant à ceux, qui, de près ou de loin, ont participé à ma thèse, à leur manière. Je pense en particulier à Edward A. Murphy dont l'esprit n'a pas hésité à m'accompagner tout au long de la thèse, particulièrement dans les moments difficiles, et ce, jusqu'à la fin. Je tiens en particulier à le remercier pour ses cadeaux lors de la rédaction : Une jambe... Une épaule... Comme d'habitude, je n'ai vraiment pas hâte de le croiser à nouveau, mais il risque de m'accompagner encore longtemps... Je ferai avec...

Avec plus d'affection cette fois-ci, tous ceux sans qui tout ça n'aurait pas été possible, Jean Baptiste Joseph Fourier, ainsi que la Dream team de la physique, tels que H. Helmholtz, Lord Rayleigh, I. Newton et j'en passe des meilleurs ! Ainsi que tous les chercheurs à travers les âges, qui ont rendu possible la petite goutte d'eau que représente mon travail de thèse dans cet océan de savoirs. On navigue ici humblement sur les épaules de ces géants.

Je ne manquerai pas non plus de remercier ceux qui m'ont donné le goût des sciences : tous mes profs de physique, qui, chacun à leur façon, m'ont permis d'arriver ici, Fred et Jamy pour m'avoir inculqué cette curiosité scientifique dès le plus jeune âge, et bien sûr mes profs de l'université du Mans, pour m'avoir fait découvrir et aimer ce merveilleux domaine qu'est l'acoustique.

"Même placée entre de bonnes mains, l'évidence est parfois ce qu'il y a de plus difficile à percevoir."
Frank Thilliez

Table des matières

1	Contexte et problématiques	3
1.1	L'audio spatialisé	3
1.1.1	Une brève histoire de l'audio spatialisé	3
1.1.2	Éléments de perception sonore	4
1.1.3	Méthodes basées sur la perception	6
1.1.4	Méthodes de reproduction de champs sonores	7
1.2	Les <i>Sound Zones</i>	11
1.2.1	Méthodes de focalisation	14
1.2.2	Méthodes d'annulation d'énergie	16
1.2.3	Méthodes de synthèse de champ sonore	18
1.2.4	Améliorations des <i>Sound Zones</i>	20
1.2.5	Audio spatialisé pour utilisateurs mobiles	22
2	Application des méthodes adaptatives à la reproduction sonore dans une zone en mouvement	25
2.1	Problèmes traités	26
2.1.1	<i>Bright zone</i> seule en contexte mobile	26
2.1.2	Cas de plusieurs auditeurs	27
2.2	Méthodes adaptatives	28
2.2.1	Filtre LMS	29
2.2.2	Filtre NLMS	30
2.2.3	Filtre APA (<i>Affine Projection Adaptive</i>)	31
2.2.4	Filtre VLLMS (<i>Variable Leaky LMS</i>)	33
2.3	Évaluation du champ sonore	35
2.3.1	Niveau sonore	35
2.3.2	Contraste	36
2.3.3	L'effort de contrôle	36
2.3.4	Homogénéité de la zone	37
2.4	Spécification du champ sonore en <i>Dark zone</i>	37
2.4.1	<i>Weighted Least Squares</i> avec résolution LMS	38
2.4.2	<i>Dark zone</i> choisie itérativement : méthodes DZIG et DZII	38
2.5	Simulation numérique dans le domaine fréquentiel	40
2.5.1	Principe et méthode	40
3	Reproduction sonore pour une seule zone en mouvement	45
3.1	Évaluation des algorithmes en configuration statique	46
3.1.1	Ajustement du paramètre de régularisation	46
3.1.2	Conditions de comparaison	47
3.1.3	Temps de convergence	48

3.1.4	Effort	49
3.1.5	Champ sonore reproduit	50
3.1.6	Niveau sonore	52
3.1.7	Effets de la régularisation	53
3.2	Déplacement d'un auditeur unique	60
3.2.1	Déplacement selon l'axe Y	61
3.2.2	Déplacement selon l'axe X	66
3.3	L'effet mémoire	68
3.3.1	Comparaison des performances des méthodes LMS et VLLMS	69
3.4	Conclusion	74
4	Vers des Sound Zones en mouvement	77
4.1	Généralisation du problème pour inclure une <i>Dark zone</i>	78
4.2	Influence de la stratégie de contrôle de la <i>Dark zone</i>	78
4.2.1	Paramètres expérimentaux	80
4.2.2	Résultats	80
4.2.3	Influence du contraste pour les méthodes DZIG et DZII	82
4.3	Déplacement de l'utilisateur	87
4.3.1	Déplacement de la Bright zone selon l'axe Y	88
4.3.2	Problème d'occlusion asymétrique	91
4.3.3	Problème d'occlusion symétrique	94
5	Discussions et Perspectives	99
5.1	Bilan des différentes solutions explorées dans les travaux de thèse	99
5.2	Amélioration de la méthode PM/VLLMS	100
5.2.1	Convergence de l'algorithme VLLMS	100
5.2.2	Sélection des paramètres VLLMS	101
5.3	Vers d'autres solutions	102
5.3.1	Hybridation avec d'autres méthodes de <i>Sound Zones</i>	102
5.3.2	Autres méthodes itératives	102
5.4	Interaction des méthodes dynamiques avec l'effet de salle	103
5.5	Dispositif microphonique pour le contrôle du champ sonore	104
5.6	Quid de la perception ?	105
6	Conclusion	107
7	Annexes	109
A	Algorithmes pour les <i>Sound Zones</i>	109
A.1	Algorithme <i>Brightness Control</i>	109
A.2	Algorithme ACC	111
A.3	Algorithmes PM	112
A.4	Algorithme LMS	113
A.5	Algorithme NLMS	113
A.6	Algorithme APA	114
A.7	Algorithme VLLMS	115
B	Théorie des algorithmes adaptatifs	116
B.1	Problème des moindres carrés	116
C	Notes sur la complexité algorithmique	125
D	Influence des paramètres de la méthode VLLMS	126

D.1	Paramètres de pas de transition : ld et lu	126
D.2	Paramètre de vitesse de transition : α	129
E	Champs sonores NLMS et APA	130
F	Effet mémoire des méthodes NLMS et APA	132

Notations

Lettres Grecques

α	Exposant vitesse d'oubli VLLMS
α_0	Angle par rapport à θ_{PW}
β	Paramètre de régularisation
γ	Paramètre d'oubli
γ_{max}	Paramètre d'oubli maximal
∇^2	Laplacien
ζ	Paramètre de réglage AEDM
η	Paramètre de réglage Weighted Least Square
θ_{PW}	Angle d'une onde plane incidente
ι	Coefficient d'amplitude pour la formulation intégrale liée à la position de la source
κ_{max}	Valeur propre maximale des matrices de fonction de transfert
λ	Multiplicateur de Lagrange
μ_L	Pas pour les algorithmes LMS, Leaky LMS et VLLMS
μ_{Lmax}	Pas maximal pour les algorithmes LMS, Leaky LMS et VLLMS
$\tilde{\mu}$	Constante d'adaptation NLMS et APA
ν	Ordre des fonctions de Bessel et Hankel pour la méthode HOA
ξ	Coefficient d'amplitude VLLMS pour dark zone
π	Nombre $\pi \approx 3.1415926535897932384626433832795028841971 \dots$
ρ	Masse volumique de l'air $\rho = 1.2 \text{ kg.m}^{-3}$
σ_{SPL}	Ecart-type des niveaux sonores dans une zone
τ_l	Décalage temporel du haut-parleur l pour le beamforming
χ	Maximum du ratio d'oubli
ω	Pulsation $\omega = 2\pi f$
Ω	Frontière du domaine de reproduction acoustique
Ω_e	Domaine de reproduction acoustique extérieur
Ω_i	Domaine de reproduction acoustique intérieur

Lettres Romanes

a	Constante de proportionnalité (pour les poids $\hat{\mathbf{q}}$)
a_W	Fenêtre d'activation des sources
c	Célérité du son $c = 343 \text{ m.s}^{-1}$
Δd	Différence de marche entre les deux oreilles
d_1 / d_2	Différence de marche de l'onde plane
f	Fréquence [Hz]
j	Opérateur complexe $j = \sqrt{-1}$
k	Nombre d'onde $k = \frac{\omega}{c}$
ℓ_1 / ℓ_2	Normalisation en norme 1 ou 2
ld	Pas de réduction méthode VLLMS

lu	Pas d'augmentation méthode VLLMS
n	Variable d'itération
o	Origine
p	Pression acoustique
p_0	Pression acoustique de référence $2 \times 10^{-5} \text{Pa}$
p_{ref}	Pression de référence (calcul de l'effort)
p_T	Pression cible
p_{Tm}	Pression cible au $m^{\text{ième}}$ microphone
q_l	Poids filtre pour le $l^{\text{ième}}$ haut parleur
q_r	Source de référence en débit (calcul de l'effort)
r_l	Position de focalisation pour le Beamforming
r_{max}	Distance maximale de focalisation pour le Beamforming
t	Variable temps continu
x_l	Position du haut parleur l
x_m	Position du microphone m
\mathbb{A}	Zone de reproduction \mathbb{A}
\mathbb{B}	Zone de reproduction \mathbb{B}
$C_{B \rightarrow D}$	Contraste entre la <i>Bright zone</i> et la <i>Dark zone</i>
C_ξ	Contraste résiduel
D_{HOA}	Fonction de contrôle HOA
D_{WFS}	Fonction de contrôle WFS
E	Effort de réseau
E_C	Effort cible à reproduire
e_{errC}	Erreur cible à reproduire
G_{ml}	Fonction de transfert entre le $l^{\text{ième}}$ haut parleur et le $m^{\text{ième}}$ microphone
$G(\vec{x} \vec{x}_0)$	Fonction de Green pour les vecteurs d'espace \vec{x} et \vec{x}_0
$H_\nu^{(2)}$	Fonction de Hankel de type 2 et d'ordre ν
J_ν	Fonction de Bessel d'ordre ν
J_{ACC}	Fonction de coût de la méthode <i>Acoustic Contrast Control</i>
J_{BC}	Fonction de coût de la méthode <i>Brightness Control</i>
J_{LMS}	Fonction de coût de la méthode LMS
J_{PM}	Fonction de coût de la méthode <i>Pressure Matching</i>
J_{WLS}	Fonction de coût de la méthode <i>Weighted Least Squares</i>
K	Ordre du filtre APA
L_{HP}	Nombre total de haut-parleurs
L	Niveau sonore
L_C	Niveau de contraste cible
L_{TB}	Niveau sonore cible dans la <i>Bright zone</i>
LL / RR	Trajet acoustique ipsilatéral
LR / RL	Trajet acoustique controlateral
M	Nombre de microphones dans une zone
Q	Contrainte de débit
R_H	Rayon de la tête
R_{ref}	Distance à la source de référence (calcul de l'effort)
R_{WFS}	Rayon du réseau circulaire de haut-parleurs pour la méthode WFS
S	Source rayonnante
X	Axe X
Y	Axe Y

Matrices et Vecteurs

\vec{n}_0	Normale sortante à une surface
κ	Vecteur des valeurs propres des matrices de fonction de transfert
ξ	Matrice de coefficients d'amplitude VLLMS pour la <i>Dark zone</i>
\mathbf{e}	Erreur de reproduction
\mathbf{p}	Vecteur des pressions reproduites
\mathbf{p}_B	Vecteur des pressions reproduites en <i>Bright zone</i>
\mathbf{p}_D	Vecteur des pressions reproduites en <i>Dark zone</i>
\mathbf{p}_T	Vecteur des pressions cibles
\mathbf{p}_{TF}	Vecteur des pressions cibles pour la méthode APA
\mathbf{q}	Vecteur des poids des sources de débit
$\hat{\mathbf{q}}$	Vecteur des poids prototypes
\vec{x}	Vecteur d'espace
\vec{x}_0	Vecteur d'espace
\mathbf{D}	Vecteur des pressions désirées pour plusieurs itérations
\mathbf{E}_{rrF}	Erreur de reproduction pour la méthode APA
\mathbf{G}	Matrice des fonctions de transfert entre l'ensemble des haut parleurs et microphones
\mathbf{G}_B	Matrice des fonctions de transfert pour la <i>Bright zone</i>
\mathbf{G}_D	Matrice des fonctions de transfert pour la <i>Dark zone</i>
\mathbf{G}_F	Matrice des fonctions de transfert pour la méthode APA
\mathbf{I}	Matrice identité

Indices et exposants

\bullet_A	Indicateur de zone \mathbb{A}
$\bullet_{\mathbb{B}}$	Indicateur de zone \mathbb{B}
\bullet_ϕ	Indicateur pseudo paramètre
\bullet_{aud}	Position du centre du réseau de microphones
\bullet_i	Indice du microphone
\bullet_l	Indice du haut-parleur
\bullet_m	Indicateur microphone
\bullet_B	Indicateur <i>Bright zone</i>
\bullet_D	Indicateur <i>Dark zone</i>
\mathbf{x}^T	Matrice transposée de \mathbf{x}
\mathbf{x}^H	Matrice transposée conjuguée de \mathbf{x}

Autres symboles

∇^2	Laplacien
$ \bullet $	Module / Valeur absolue
\Re	Opérateur de partie réelle
$\mathbf{0}$	Vecteur nul
$\delta\mathbf{q}$	Variation des poids entre deux itérations
\S	Partie / Section

Abréviations

iter	Itération
ACC	<i>Acoustic Contrast Control</i>
AEDM	<i>Acoustic Energy Difference Maximization</i>
APA	<i>Affine Projection Adaptive</i>
BC	<i>Brightness Control</i>
DZII	<i>Dark zone Itérative Individuelle</i>

DZIG	<i>Dark zone Itérative Globale</i>
Eq.	<i>Equation</i>
HOA	<i>High Order Ambisonics</i>
HRTF	<i>Head Related Transfer Function</i>
ILD	<i>Interaural Level Difference</i>
ITD	<i>Interaural Time Difference</i>
LASSO	<i>Least Absolute Shrinkage and Selection Operator</i>
LMS	<i>Least Mean Square</i>
NLMS	<i>Normalized Least Mean Square</i>
PC	<i>Planarity Control</i>
PM	<i>Pressure Matching</i>
RI	<i>Réponse impulsionnelle</i>
RSB	<i>Rapport Signal à Bruit</i>
RLS	<i>Recursive Least Squares</i>
VBAP	<i>Vector Based Amplitude Panning</i>
VLLMS	<i>Variable Leaky Least Mean Square</i>
SPL	<i>Sound Pressure Level</i>
WDPF	<i>Wave Domain Point Focusing</i>
WFS	<i>Wave Field Synthesis</i>
WLS	<i>Weighted Least Squares</i>

Introduction

La reproduction de contenus audio spatialisés est aujourd'hui en expansion dans l'audiovisuel grand public, que ce soit au cinéma, à la télévision, sur internet ou par le biais de la réalité virtuelle. La diffusion de tels contenus nécessite l'utilisation de systèmes de reproduction spécifiques pour que ceux-ci soient idéalement perçus. Cependant, lorsque plusieurs contenus sont en concurrence dans la même pièce, ils ne peuvent pas être reproduits directement pour chaque auditeur sans utilisation de casque audio, isolant ainsi les auditeurs de leur environnement. Ce type de situation est d'autant plus fréquent que le nombre d'appareils permettant la reproduction sonore est en augmentation permanente. Il devient alors essentiel de contrôler le son de chacun de ces appareils pour que tous les utilisateurs disposent d'une qualité de contenu optimale en minimisant les interférences entre ceux-ci.

Cette capacité à reproduire un contenu audio de façon indépendante à différents endroits de l'espace est connue sous le nom de *Sound Zones* (zones sonores) [Betlehem *et al.*, 2015]. C'est grâce aux méthodes de reproduction de champs sonores basées sur le contrôle physique d'un ensemble de sources audio, qu'il est actuellement possible de reproduire des zones sonores spatialement séparées à partir d'un réseau de haut-parleurs [Choi *et al.*, 2008, Choi et Kim, 2002, Elliott *et al.*, 2010, Kirkeby *et al.*, 1996, Poletti, 2008]. Le but est de reproduire un champ sonore défini dans une première zone de l'espace, appelée *Bright zone* et de minimiser la pression dans la seconde zone, appelée *Dark zone*. Ces méthodes ne nécessitent pas l'utilisation d'un casque audio ce qui permet d'être conscient de son environnement extérieur. Elles peuvent donc être utilisées même dans des conditions où la sécurité l'exige comme par exemple dans les véhicules [Cheer *et al.*, 2013, Elliott et Jones, 2006].

Cependant, jusqu'à présent, la plupart de ces méthodes sont destinées à une utilisation pour des auditeurs statiques. Leur coût de calcul ainsi qu'un réglage réalisé pour une configuration donnée ne permet pas leur utilisation dans le contexte d'un problème dynamique assujetti aux mouvements d'auditeurs et pour lequel on souhaite une solution temps réel [Betlehem *et al.*, 2015, Trevino *et al.*, 2016]. La capacité à s'adapter à différentes positions d'auditeurs en temps réel permettrait l'utilisation de ces systèmes dans un plus grand nombre de contextes. Pour ce faire, les méthodes actuelles peuvent être reformulées à l'aide de méthodes adaptatives de sorte à réduire le coût de calcul ainsi qu'à offrir l'automatisation des réglages associés à la production de zones sonores.

Cette thèse s'intéresse au développement d'une nouvelle méthode de calcul, plus rapide que les méthodes actuelles, pour le déplacement en temps réel de *Sound Zones*, en s'appuyant sur des méthodes adaptatives. Il est montré qu'il est possible de traiter le problème de *Sound Zones* (i.e. production d'une *Bright zone* et d'une *Dark zone*), de la même façon que le problème de synthèse de champ sonore, i.e. une *Bright zone* unique sans *Dark zone*. Plusieurs algorithmes sont proposés puis évalués pour différentes configurations. Dans un premier chapitre, un aperçu des méthodes existantes

pour la reproduction sonore spatialisée est donné. Un intérêt particulier est apporté aux méthodes multi-auditeurs ainsi qu'à quelques réalisations existantes pour les auditeurs mobiles. Les éléments théoriques nécessaires à la conception de zones sonores individuelles par des méthodes adaptatives sont introduits. Dans un second chapitre, les concepts fondamentaux des méthodes développées pour répondre au problème de *Sound Zones* s'adaptant aux déplacements des auditeurs sont présentées pour un ou plusieurs auditeurs. Les paramètres d'évaluation ainsi que la méthode de simulation utilisée sont décrits. Le troisième chapitre décrit l'évaluation de la méthode mono-auditeur, i.e. *Bright zone* unique sans *Dark zone*, en s'intéressant à l'influence du déplacement de celui-ci au sein de l'espace. Un effet mémoire de l'algorithme est mis en évidence et corrigé pour permettre son utilisation dans des conditions imposées par les auditeurs mobiles. Le chapitre 4 est consacré à l'évaluation de la méthode dans le cas de deux utilisateurs mobiles, c'est à dire en incluant une *Bright zone* et une *Dark zone*, en s'intéressant à plusieurs aspects, d'une part l'influence du champ sonore dans la *Dark zone* ainsi que le lien avec les méthodes de l'état de l'art, d'autre part l'influence des positions relatives des utilisateurs sur les performances des systèmes de *Sound Zones*. Par ailleurs, une étude de convergence des algorithmes utilisés est proposée afin de comparer les paramètres les plus adaptés à une utilisation dans le cas d'auditeurs mobiles. Les principales limitations des solutions développées ainsi que les perspectives sur les différents aspects de ces méthodes sont discutées dans le dernier chapitre.

Chapitre 1

Contexte et problématiques

1.1 L'audio spatialisé

*"To understand a science, it's necessary to know its history."
Auguste Comte*

1.1.1 Une brève histoire de l'audio spatialisé

La conception actuelle de l'audio spatialisé remonte au XVIII^{ième} avec les premières utilisations des équations différentielles et la formulation en harmoniques sphériques, qui a permis de poser les bases mathématiques de la formulation physique et plus particulièrement de l'acoustique. Le siècle suivant a vu l'avènement des séries de Fourier et du théorème de Green permettant la formulation des équations intégrales de Kirchhoff-Helmholtz. La première application au contrôle d'un objet sonore est réalisée sur des diapasons couplés électriquement en 1877 par Lord Rayleigh pour montrer un exemple de vibrations forcées [Rayleigh, 1896]. La même année, la première reproduction audio est réalisée par Thomas Edison à l'aide d'un cylindre de cire et d'amplification mécanique. Pour ce qui est des haut-parleurs tels que nous les connaissons aujourd'hui, il convient de citer travail de Kellogg et de Rice en 1925 [Rice, 1929]. La première reproduction d'audio spatialisée est réalisée par Clément Ader à l'aide d'un ensemble de transmissions appelé thêatrophone lors de l'exposition universelle de 1881 à Paris. Cette idée a alors été reprise par la suite pour le développement de la stéréophonie, grâce à l'utilisation de deux sources audio comportant des signaux avec des différences de phase et d'amplitude. Cette idée est ensuite approfondie par Steinberg [Steinberg et Snow, 1934] qui propose le concept du rideau acoustique qui permet de reproduire avec un ensemble de haut-parleurs une source audio dans une autre pièce en captant le son avec la même disposition de microphones. Cette idée pose les bases de ce que sera la Wave Field Synthesis (WFS) [Berkhout, 1988]. Depuis cette époque, un grand nombre de méthodes a été développé pour la reproduction de contenus audio spatialisés, tel que le Vector Based Amplitude Panning (VBAP) [Pulkki, 1997], une généralisation de la stéréophonie à une sphère de haut-parleurs, la Wave Field Synthesis (WFS) [Berkhout, 1988] basée sur le principe de Huygens, les méthodes Ambisoniques [Gerzon, 1973] qui projettent le champ sonore sur une base d'harmoniques sphériques ou encore les méthodes de décompositions spectrales [Ahrens et Spors, 2010] qui réalisent un filtrage spatial.

Ces méthodes utilisent toutes des réseaux de haut-parleurs pour reproduire des champs sonores sous un ensemble de conditions. D'autres méthodes basées sur la re-

production des indices de localisation perçus sont également utilisées pour reproduire un contenu audio spatialisé. C'est le cas par exemple des méthodes binaurales [Nicol, 2010] et transaurales [Schroeder *et al.*, 1962]. La spatialisation est alors rendue possible grâce à l'utilisation de HRTF (Head Related Transfer Function), en reproduisant la propagation acoustique entre une source et les deux oreilles de l'utilisateur. Ces fonctions sont individuelles pour la morphologie de chaque utilisateur.

Durant les dernières décennies, de nombreux travaux ont cherché à reproduire les contenus spatialisés pour plusieurs utilisateurs dans un environnement commun. Ces méthodes, regroupées sous le nom de *Sound Zones*, réalisent le contrôle de champ sonore dans plusieurs zones de l'espace. Ces méthodes s'appuient à la fois sur les travaux d'audio spatialisé, et sur des méthodes de traitement du signal développées au cours de xx^{ième} siècle. Les travaux de P.Lueg en 1936 [Lueg, 1936] décrivent le principe physique du contrôle actif à l'aide d'une source secondaire. Ces avancées ont mené aux algorithmes de contrôle en boucle fermée de May et Olson en 1953 permettant par la suite de réaliser le premier système de casque audio avec réduction active du bruit dans les années 80 [Bose et Carter, 1984].

Tout ceci conduit à l'apparition des méthodes dédiées à la réalisation de *Sound Zones*. Des méthodes empruntées à l'électromagnétisme tel que le beamforming [Van Veen et Buckley, 1988] ont tout d'abord été utilisées pour focaliser des sources dans une zone de l'espace, puis des méthodes d'optimisation de champ sonore ont été introduites ([Choi et Kim, 2002],[Choi *et al.*, 2008]).

Les méthodes de *Sound Zones* résultent de la combinaison de ces idées. Elles réalisent le contrôle physique d'un champ sonore en utilisant un ensemble de haut-parleurs pour reproduire un contenu audio spatialisé tout en permettant le contrôle indépendant de ces champs sonores pour plusieurs utilisateurs¹.

1.1.2 Éléments de perception sonore

La perception de notre environnement sonore détermine notre relation au monde avec l'ensemble des stimuli extérieurs. Un certain nombre d'animaux se servent principalement de leur audition pour se repérer dans l'espace grâce à l'écholocation. C'est le cas des chauve-souris ou des dauphins par exemple. Pour l'humain cette relation à l'audio est moins présente, mais permet tout de même d'obtenir un grand nombre d'informations sur son environnement. Ces informations peuvent être séparées en plusieurs catégories. Les informations temporelles, les informations fréquentielles et les informations spatiales. L'ensemble de ces informations, obtenues par le biais des oreilles, est ensuite traité par le cerveau. Les informations temporelles et fréquentielles permettent de reconnaître les sons de l'environnement en fonction de leur signature sonore [Blauert, 1983]. Pour ce qui est des informations spatiales, elles s'appuient aussi sur les indices temporels et fréquentiels qui, en les combinant, reconstruisent la direction d'arrivée de l'onde sonore. Cela correspond aux différences interaurales de temps (ITD), aux différences interaurales de niveau (ILD) ainsi que des variations spectrales.

La différence interaurale de temps correspond à la différence de temps obtenue entre les deux oreilles lorsqu'une source sonore arrive dans une direction donnée. Cet effet est principalement exploité par le système auditif en basses fréquences.

1. Ce voyage historique au coeur de l'audio spatialisé est fortement inspiré d'une présentation réalisée par Philip Nelson lors de la Sound Field Control Conference en 2016 à Guilford [Conference, 2016].

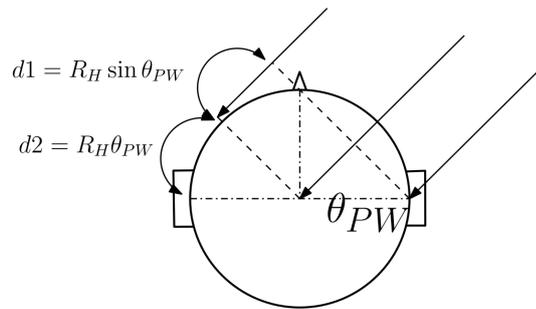


Figure 1.1 – Différence de temps et de niveau entre les deux oreilles. La différence de marche d'une onde plane entre les deux oreilles correspond à la distance $\Delta d = d1 + d2 = R_H(\sin\theta_{PW} + \theta_{PW})$. On en déduit donc la différence de temps entre les deux oreilles comme étant $\Delta t = R_H(\sin\theta_{PW} + \theta_{PW})/c$. [Blauert, 1983]

Lorsqu'une onde sonore arrive dans une direction, il existe une différence de temps d'arrivée entre l'oreille droite et l'oreille gauche directement liée à l'angle d'arrivée de l'onde (**Figure 1.1**). Cette différence de temps est alors interprétée par le cerveau comme une direction d'arrivée. De plus, la tête réalise une "ombre" et ne transmet donc pas toute l'énergie acoustique jusqu'à l'autre oreille, ce qui produit une différence de niveau sonore. Ces deux paramètres sont les deux principaux facteurs permettant la localisation sonore dans le plan horizontal. Cependant, cette information n'est pas unique pour toutes les directions d'arrivée. D'autres spécificités entrent en jeu comme les indices spectraux.

Lorsqu'une onde acoustique arrive au niveau de la tête, elle est diffractée (**Figure 1.2**). Le buste ainsi que le pavillon de l'oreille contribuent aussi à cette diffraction.

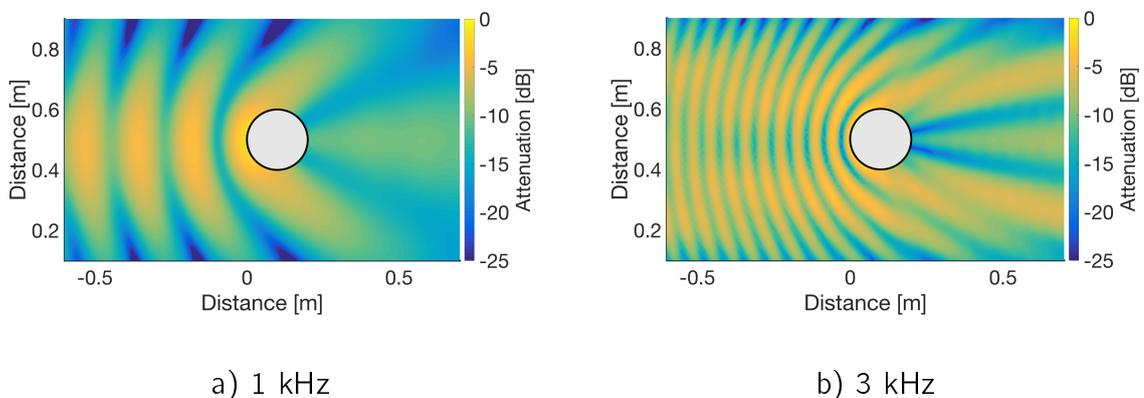


Figure 1.2 – Figure de diffraction pour un cylindre simulant une tête pour différentes fréquences. On observe que le niveau sonore est différent autour du cylindre. a) Atténuation pour la diffraction par un cylindre de 20 cm d'une onde plane à 1kHz venant par la gauche. b) Atténuation pour la diffraction par un cylindre de 20 cm d'une onde plane à 3kHz venant par la gauche.

Le phénomène de diffraction est différent pour chaque fréquence, ce qui implique une variation du niveau sonore et de phase qui est différente pour chaque fréquence. Ce filtrage, induit par la morphologie de l'individu, ajoute une information supplémentaire dans les indices cités précédemment. L'influence de la diffraction dépend bien entendu de la taille de la sphère. Dans le cas d'un auditeur humain, elle se produit à partir de 300Hz [Blauert, 1983].

Un autre paramètre influe sur ces indices spectraux : la forme de l'oreille. Lorsque l'onde acoustique arrive au niveau du pavillon, elle est en effet déformée par les différentes cavités [Blauert, 1983]. Ce phénomène permet principalement de reconnaître les directions d'arrivée en élévation. L'ensemble de ces phénomènes permet donc d'avoir une connaissance assez précise de l'environnement sonore (précision minimale d'environ $\pm 3^\circ$ en azimut et $\pm 10^\circ$ en élévation pour les sources en face de l'auditeur [Blauert, 1983]).

Le but de la reproduction audio spatialisée est de restituer cette écoute naturelle comme elle aurait été perçue mais en l'absence de sources primaires à l'origine du champ sonore.

1.1.3 Méthodes basées sur la perception

Parmi les méthodes de reproduction audio spatialisée, les méthodes perceptives s'intéressent à reproduire une identité perceptive de l'environnement sonore par la reproduction des indices de localisation et non par la reproduction du champ sonore. Ces méthodes s'appuient sur les indices perceptifs précédemment présentés. L'application la plus directe de ces indices perceptifs est l'utilisation de la stéréo. En jouant sur l'amplitude et le délai des signaux envoyés à deux haut-parleurs, elle induit une différence de niveaux et de temps entre les deux haut-parleurs et peut donc être perçue comme une différence sur l'ILD et l'ITD. Cette méthode souffre cependant de nombreux inconvénients, en particulier la spatialisation est assez peu précise et limitée dans l'espace entre les deux haut-parleurs et peu robuste dans le cas d'une stéréophonie utilisant uniquement l'ITD. De plus le signal de chaque haut-parleur n'arrive pas directement à l'oreille à laquelle il est destiné. Il existe des trajets croisés, du haut-parleur droit vers l'oreille gauche et du haut-parleur gauche vers l'oreille droite (**Figure 1.3**). Lorsque l'on souhaite reproduire un environnement sonore, il est pourtant nécessaire que les signaux des oreilles droite et gauche soient indépendants.

Par ailleurs, les indices perceptifs sont différents pour chaque utilisateur [Blauert, 1983]. Pour obtenir une connaissance des paramètres pour un utilisateur, il est possible de mesurer les fonctions de transfert liée à la tête : les HRTF (Head Related Transfer Functions). Ces fonctions de transfert représentent, pour un utilisateur donné, la façon dont l'ensemble tête, buste et oreilles modifie l'onde sonore. La mesure des HRTF pouvant être longue à réaliser, plusieurs approches existent pour pallier le problème de leur individualisation. L'une d'entre elles consiste à approximer les HRTF de l'auditeur à l'aide d'une présélection de HRTF existantes dans une base de données en laissant l'auditeur faire sa propre sélection, plutôt que de faire une nouvelle mesure [Nicol *et al.*, 2018]. D'autres approches s'intéressent plutôt à la génération d'un jeu de HRTF sur la base de mesures anthropométriques, plus facile à mettre en œuvre ([Gupta *et al.*, 2004]).

La méthode la plus facile pour la mise en place des méthodes perceptives reste l'utilisation d'un casque. Cette méthode est appelée reproduction binaurale. L'usage d'un système de haut-parleurs pour reproduire directement le signal binaural aux oreilles présente quant à lui l'inconvénient principal que chaque oreille perçoit les signaux des deux haut-parleurs. Chaque signal doit alors être traité pour compenser la perception à l'oreille contro-latérale.

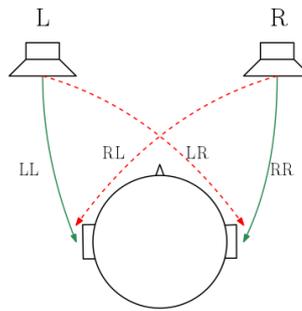


Figure 1.3 – Principe de l'écoute transaurale. Le signal provenant du haut-parleur gauche L est restitué uniquement à l'oreille gauche et le signal provenant du haut-parleur droit R est restitué uniquement à l'oreille droite. Les trajets croisés RL et LR doivent être annulés.

Pour que le signal soit uniquement transmis à une seule oreille, le son du haut parleur droit destiné à l'oreille droite et perçu par l'oreille gauche doit être annulé pour l'oreille gauche et le son venant du haut-parleur gauche destiné à l'oreille gauche doit être annulé pour l'oreille droite (**Figure 1.3**). La restitution binaurale à partir de haut-parleurs est appelée méthode transaurale [Atal et Schroeder, 1966, Møller, 1992] et des variantes ont été proposées ; telles que le stéréo-dipôle [Tokuno *et al.*, 1996, Kirkeby *et al.*, 1997]. Pour obtenir une bonne reproduction transaurale, il est également possible d'utiliser des méthodes de contrôle de champ sonore pour focaliser le son directement sur une oreille. Les méthodes utilisées dans ce cas sont assez proches de celles utilisées dans le cas des *Sound Zones* (§1.2).

1.1.4 Méthodes de reproduction de champs sonores

A l'inverse des approches perceptives, ayant pour but de reproduire l'environnement sonore tel qu'il a été perçu, les approches physiques consistent à reproduire le champ sonore tel qu'il a été produit par l'ensemble des sources acoustiques primaires présentes dans la scène sonore. En reproduisant le contenu audio de cette façon, les effets de localisation sont théoriquement perçus par l'utilisateur comme si la source réelle était présente.

Le principe de reproduction de champs sonores s'appuie généralement sur les équations de propagation des ondes acoustiques. Ces méthodes ont l'avantage de ne pas dépendre de la morphologie de l'utilisateur pour reproduire le contenu audio. Les signaux sont généralement reproduits par un ensemble de haut-parleurs dont on cherche à estimer les signaux d'entrée nécessaires à la reproduction du champ sonore cible.

Formulation acoustique

Les méthodes de reproduction de champ sonore s'appuient sur les équations de propagation des ondes pour contrôler un ensemble de sources permettant la reconstruction d'un champ sonore donné. Cette équation de propagation est définie, dans le cas d'un problème sans sources, par

$$\left(\nabla^2 - \frac{1}{c^2} \frac{\partial^2}{\partial t^2} \right) p(\vec{x}, t) = 0, \quad (1.1)$$

avec $p(\vec{x}, t)$, la pression pour le vecteur position \vec{x} et au temps t , $c = 343 \text{m.s}^{-1}$ la célérité du son dans l'air, et ∇^2 le Laplacien. Pour résoudre cette équation, en considérant un

problème stationnaire harmonique, il est possible de traduire le problème dans le domaine de Fourier, l'équation (1.1) peut alors s'écrire sous la forme de l'équation de Helmholtz :

$$(\nabla^2 + k^2) p(\vec{x}, \omega) = 0, \quad (1.2)$$

avec $k = \frac{\omega}{c}$, le nombre d'onde et $\omega = 2\pi f$ la pulsation à la fréquence f . En considérant la reproduction d'un champ sonore dans un domaine borné sans sources contenues dans le domaine de reproduction, comme illustré en **Figure 1.4**, on définit un problème intérieur avec \vec{n}_0 la normale sortante du domaine Ω_i , correspondant au volume intérieur, (c'est à dire au volume de reproduction du champ sonore), Ω_e le volume à l'extérieur du volume de reproduction et Ω la surface de frontière entre le domaine intérieur et extérieur. La résolution de l'équation (1.2) dans le cas de conditions aux frontières inhomogènes le long du domaine Ω , peut s'écrire sous la forme de l'intégrale de Kirchhoff-Helmholtz [Williams, 1999, p.252-257]

$$i(\vec{x}) p(\vec{x}, \omega) = - \oint_{\Omega} \left(G(\vec{x}|\vec{x}_0, \omega) \frac{\partial p(\vec{x}_0, \omega)}{\partial \vec{n}_0} - p(\vec{x}_0, \omega) \frac{\partial G(\vec{x}|\vec{x}_0, \omega)}{\partial \vec{n}_0} \right) d\Omega. \quad (1.3)$$

Le coefficient $i(\vec{x})$ est déterminé par la position de la source rayonnante [Williams, 1999, p.252-257],

$$i(\vec{x}) = \begin{cases} 1 & \text{si } \vec{x} \in \Omega_i \\ \frac{1}{2} & \text{si } \vec{x} \in \Omega \\ 0 & \text{sinon.} \end{cases} \quad (1.4)$$

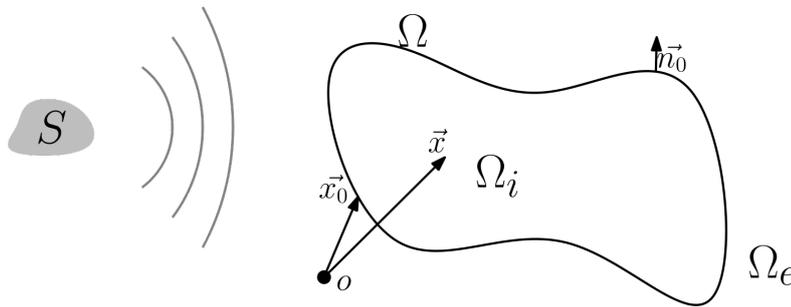


Figure 1.4 – Source rayonnante S dans le domaine Ω_e extérieur et reproduite dans le domaine intérieur Ω_i . Les domaines intérieur et extérieur sont délimités par une frontière Ω .

Cette formulation permet de définir complètement et de façon unique un champ sonore dans le domaine Ω_i à l'aide d'une continuité de sources infinitésimales sur la frontière Ω . Le champ sonore d'une source primaire peut donc être reproduit par un ensemble de sources secondaires placées à la frontière (**Figure 1.4**). Cette solution (**Eq. 1.3**) comporte deux parties, l'une correspondant aux sources monopolaires (termes en $G(\vec{x}|\vec{x}_0, \omega) \frac{\partial p(\vec{x}_0, \omega)}{\partial \vec{n}_0}$), l'autre aux sources dipolaires (termes en $p(\vec{x}_0, \omega) \frac{\partial G(\vec{x}|\vec{x}_0, \omega)}{\partial \vec{n}_0}$). Ces sources dipolaires sont fortement redondantes avec les sources monopolaires, et il est possible de ne conserver qu'un seul type de sources [Nicol, 1999]. De plus, pour une implémentation pratique de ces méthodes, les haut-parleurs sont une assez bonne approximation des monopôles, surtout en basse fréquence. Le but des méthodes HOA et WFS consiste alors à déterminer les poids de ces sources monopolaires réparties sur la frontière afin de reproduire un champ sonore donné, sans considérer de sources dipolaires dans la solution.

A noter que la principale différence entre le WFS et le HOA est l'approche considérée pour l'élimination et la compensation de ces sources dipolaires, menant à une résolution différente du problème².

Wave Field Synthesis (WFS)

La méthode WFS s'appuie sur la solution de l'équation intégrale de Kirchhoff-Helmoltz (**Eq.** (1.3)), permettant de définir un champ sonore à l'intérieur d'un volume en connaissant la pression ainsi que la vitesse particulière à sa frontière. Comme définie dans la partie précédente, la formulation de la WFS nécessite l'élimination des sources dipolaires de l'équation (1.3). Pour ce faire, on considère que $\frac{\partial G(\vec{x}|\vec{x}_0, \omega)}{\partial \vec{n}_0} = 0$. La fonction de Green est donc modifiée de sorte que les sources dipolaires soient nulles [Spors et Ahrens, 2008]. Une des conséquences de cette formulation est que le champ sonore n'est pas reproduit de façon exacte dans le domaine de reproduction Ω_i dans toutes les géométries. La disposition des sources à la frontière doit en outre être convexe, de sorte que les sources ne soient pas disposées dans le sens inverse de propagation du champ sonore. Plutôt que de se limiter à certaines géométries, une solution généralement retenue pour éviter ce type de problèmes pour des géométries arbitraires est l'utilisation d'une fenêtre $a_W(\vec{x})$ ayant pour rôle la désactivation des sources gênantes ou inutiles à la reproduction sonore en l'absence de dipôles [Spors et Ahrens, 2008].

En pratique, il est nécessaire de placer un réseau de haut-parleurs discrets à la frontière de reproduction. Cette discrétisation a pour effet de limiter la fréquence de reproduction maximale du réseau de haut-parleurs. Cet effet correspond à l'effet d'aliasing spatial [Berkhout, 1988]. Enfin les poids nécessaires à la reproduction d'un champ sonore par la méthode WFS dans le cas d'un réseau circulaire de haut-parleurs, correspondent à [Spors et Ahrens, 2008] :

$$D_{WFS}(\alpha_0, R_{WFS}, \omega) = 2j a_W(\alpha_0) \cos(\alpha_0 - \theta_{PW}) e^{-j \frac{\omega}{c} R_{WFS} \cos(\alpha_0 - \theta_{PW})}, \quad (1.5)$$

avec θ_{PW} la direction de l'onde plane incidente à reproduire, R_{WFS} le rayon du réseau de haut-parleurs et α_0 l'angle de la frontière par rapport à θ_{PW} . Cela signifie qu'il suffit de connaître la direction d'une source à reproduire ainsi que la répartition des sources secondaires (définie ici par R_{WFS}), pour connaître les poids associés à chaque haut-parleur.

High Order Ambisonics (HOA)

De la même façon que pour la méthode WFS, la méthode HOA réalise une élimination des sources dipolaires de l'équation (1.3), mais s'appuie sur un principe différent. L'élimination de ces termes est réalisée en considérant une formulation uniquement monopolaire. Cette hypothèse a pour conséquence une modification des fonctions de Green, de sorte à compenser l'absence des sources dipolaires. Pour obtenir les poids à l'aide de la méthode HOA, il convient de résoudre l'équation (1.3) de façon analytique. Cette résolution est généralement réalisée sur la base d'une décomposition du champ sonore en harmoniques sphériques dans le cas d'une reproduction sur une sphère de haut-parleurs. Cette base est orthonormale et il devient alors possible de recréer n'importe quel champ sonore à l'aide d'un nombre suffisamment important de haut-parleurs,

2. L'approche proposée pour les méthodes HOA et WFS est celle considérée dans [Spors et Ahrens, 2008]. D'autres formulations existent, mais celles proposées ici permettent une comparaison aisée des deux méthodes.

en considérant ce champ comme une combinaison linéaire de ces fonctions. La résolution est généralement conditionnée par des géométries simples. Les poids peuvent alors s'écrire sous la forme [Spors et Ahrens, 2008] :

$$D_{HOA}(\nu, R, \omega) = -\frac{2}{j\pi R} \frac{j^{-\nu} J_{\nu}(kr)}{J_{\nu}(kr) H_{\nu}^{(2)}(kr)} e^{-j\nu\theta_{PW}}, \quad (1.6)$$

avec J_{ν} la fonction de Bessel de première espèce d'ordre ν et $H_{\nu}^{(2)}$ la fonction de Hankel de second type, d'ordre ν .

Un des problèmes liés à cette formulation est qu'elle n'est pas définie de façon unique pour toutes les fréquences. En effet, lorsque les fréquences à reproduire correspondent aux fréquences propres du volume de reproduction, les fonctions de Bessel s'annulent et les poids de l'équation divergent [Moreau, 2006]. Il devient alors nécessaire d'ajouter une régularisation à l'équation (1.6).

Pressure Matching

Les méthodes de reproduction de champs sonores peuvent aussi être vues comme un problème d'optimisation, l'objectif étant d'estimer les sources monopolaires optimales pour reproduire un champ sonore dans le domaine de reproduction. Les méthodes d'optimisation cherchent alors à minimiser une fonction de coût qui définit un système donné. La résolution du problème d'optimisation peut être réalisée par une grande variété de méthodes qui sont à la fois dépendantes de l'application, de l'approche considérée et des contraintes de ressource liées à cette résolution. Le but est ici de minimiser l'erreur $\mathbf{e}_{rr} = \mathbf{p} - \mathbf{p}_T$ entre la pression cible \mathbf{p}_T et la pression reproduite \mathbf{p} sur un ensemble de positions, appelées points de contrôles. Cette minimisation fournit la valeur optimale de chaque source monopolaire reproduisant le champ sonore cible. Pour poser le problème, on discrétise l'espace de reproduction et on répartit les sources sur le domaine Ω . L'objectif est d'estimer les poids \mathbf{q} des sources permettant de minimiser l'erreur de reproduction. Le problème est basé sur la minimisation de l'erreur au sens des moindres carrés [Kirkeby *et al.*, 1996], ce qui correspond à la fonction de coût :

$$J_{PM} = [\mathbf{p} - \mathbf{p}_T]^H [\mathbf{p} - \mathbf{p}_T] + \beta \mathbf{q}^H \mathbf{q}, \quad (1.7)$$

avec H la transposée hermitienne. Le premier terme correspond à l'erreur quadratique de reproduction $\|\mathbf{e}_{rr}\|^2$, et le second terme permet de régulariser le problème. Le paramètre de régularisation β est introduit pour limiter la puissance délivrée par le réseau de haut-parleurs. Il évite que le problème ne diverge en donnant trop d'énergie au système. Cette régularisation correspond à une régularisation de Tikhonov [Elliott *et al.*, 2012, Coleman *et al.*, 2013] (régularisation en norme ℓ_2). D'autres régularisations sont possibles en fonction du problème considéré. Ainsi les méthodes nécessitant la réduction du nombre de sources s'appuient plutôt sur une normalisation en norme ℓ_1 [Lilis *et al.*, 2010]. Des méthodes hybrides de type *elastic-net*, combinant une régularisation en norme ℓ_1 et ℓ_2 tirent parti des avantages de chacune des régularisations [Radmanesh et Burnett, 2013, Gauthier et Berry, 2015, Gauthier *et al.*, 2017].

La valeur du paramètre de régularisation β est choisie de sorte à obtenir les meilleures performances possibles selon un critère donné. Il existe une grande diversité de stratégies de sélection de ce paramètre de régularisation selon les auteurs [Coleman *et al.*, 2014b]. Il peut être obtenu par exemple par optimisation d'un critère physique du système ou en utilisant des règles empiriques liées au conditionnement d'une matrice donnée

[Kirkeby *et al.*, 1996]. L'optimisation du paramètre β , est généralement réalisée de façon indépendante pour toutes les fréquences jusqu'à obtenir une valeur satisfaisant toutes les contraintes [Coleman *et al.*, 2014b]. Cette partie du calcul des poids peut être relativement coûteuse car la valeur de régularisation peut être très variable.

Pour obtenir la solution optimale associée à la fonction de coût (**Eq.** (1.7)), et donc les poids des sources correspondantes, la méthode de résolution la plus utilisée est la méthode Directe, correspondant alors à un filtre de Wiener (**Eq.** (1.30)). Un ensemble de contraintes peuvent également être ajoutées dans le cas des méthodes d'optimisation pour répondre à un problème donné. Ces contraintes supplémentaires définissent de nouveaux problèmes permettant de modifier le comportement de la méthode. La résolution est alors différente en fonction de l'approche et des contraintes considérées. Il est par exemple possible de poser le problème de *Sound Zones* à partir de cette méthode. La plupart des méthodes de *Sound Zones* s'appuient sur la résolution du problème par la méthode des multiplicateurs de Lagrange. Cette méthode résout un problème d'optimisation sous contraintes d'égalité en le reformulant de sorte à ce qu'il corresponde à un problème non contraint [Haykin, 2002]. Plus récemment, [Gauthier *et al.*, 2016] a proposé une méthode de décomposition en valeurs singulières généralisée pour la réalisation de *Sound Zones*, ouvrant la porte à une meilleure compréhension du fonctionnement des algorithmes de *Sound Zones*.

1.2 Les Sound Zones

Les méthodes d'audio spatialisé peuvent être mises à contribution pour la reproduction sonore de champs sonores cibles dans des zones précises de l'espace, les *Sound Zones*. Ces méthodes se répartissent suivant deux catégories : les méthodes analytiques, et les méthodes optimales. Les méthodes analytiques permettent de générer les signaux des haut-parleurs en s'appuyant directement sur les solutions de l'équation de propagation, comme c'est le cas des méthodes WFS ou HOA. Ces formulations ont l'avantage de décrire précisément le champ sonore et offrent donc potentiellement une reproduction de très bonne qualité - même si la séparation entre les contenus audio des différentes zones de l'espace n'est pas très élevée - tandis que les méthodes optimales ont quant à elles été développées principalement dans le but de réaliser des *Sound Zones* reproduisant un champ sonore cible sous un certain nombre de contraintes. Ces contraintes permettent par exemple de focaliser le son dans une direction, de limiter le niveau sonore dans une zone ou bien encore de limiter l'énergie délivrée par l'ensemble des haut-parleurs.

Cette partie présente les différentes méthodes de *Sound Zones* utilisées pour les auditeurs fixes, puis des méthodes d'audio spatialisé proposées dans la littérature pour les auditeurs mobiles. Leur utilisation est discutée et mène à la proposition de méthodes adaptatives permettant de répondre aux contraintes imposées par la création de *Sound Zones* mobiles.

Le principe fondamental des *Sound Zones* est de partager le même espace sans partager le même espace sonore. De la même façon qu'un ensemble de spots lumineux permet d'éclairer de petites zones de l'espace en laissant le reste de l'espace dans l'ombre, le but des *Sound Zones* est de proposer un contenu audio indépendant dans différents endroits de l'espace et donc une séparation des contenus audio (**Figure** 1.5).

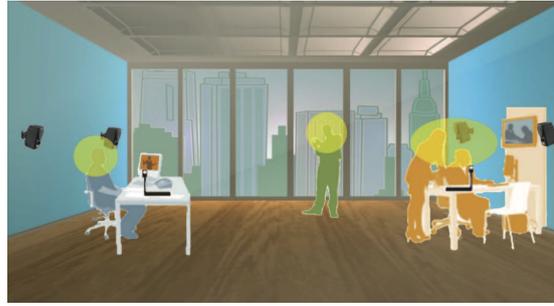


Figure 1.5 – Principe des *Sound Zones*. Chaque auditeur peut disposer d'un contenu audio différent dans sa bulle sonore sans perturber l'écoute des autres utilisateurs présents dans le même espace sonore avec la possibilité de bouger librement dans la pièce [Betlehem *et al.*, 2015].

La solution la plus intuitive à ce problème est l'utilisation d'un casque audio. Cependant, cette méthode réalise une séparation physique entre l'auditeur et son environnement, ce qui peut poser des problèmes de sécurité, par exemple dans le cas de l'acoustique intérieur des véhicules. De plus, cela limite les interactions entre les différents utilisateurs. On s'intéresse donc ici à l'utilisation de haut-parleurs directifs ou à la reproduction d'un champ sonore correspondant aux caractéristiques désirées.

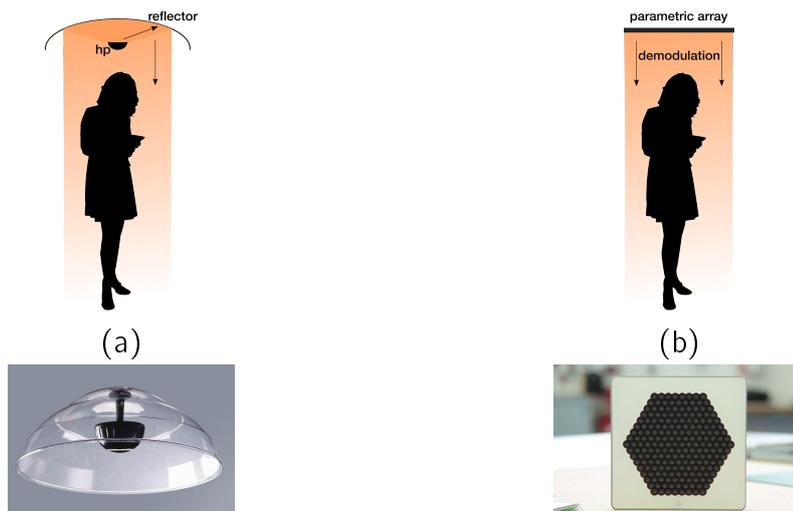


Figure 1.6 – Exemple de haut-parleurs directifs : (a) Utilisation d'un réflecteur parabolique pour focaliser le son en un faisceau. (b) Utilisation d'une antenne paramétrique, fonctionnant sur le principe de la démodulation non-linéaire d'ultrasons [Robin et Kaddouch, 2018].

La mise en place de haut-parleurs directifs peut être réalisée de différentes façons. Il est possible d'utiliser des réflecteurs paraboliques permettant de diriger un faisceau directement vers l'utilisateur (**Figure 1.6 (a)**). Une autre solution consiste à utiliser des antennes paramétriques (**Figure 1.6 (b)**) composées de transducteurs ultrasonores permettant d'obtenir un faisceau très directif. Les utilisateurs disposent d'un contenu audio grâce à une méthode de démodulation non-linéaire des ondes ultrasonores [Westervelt, 1963, Robin et Kaddouch, 2018]. Cette méthode nécessite de forts niveaux acoustiques, qui peuvent avoir des effets sur le corps humain, particulièrement dans la gamme des ultrasons [Smagowska et Pawlaczyk-Łuszczynska, 2013]. Le principal inconvénient de

ces méthodes est qu'elles nécessitent de disposer d'un haut-parleur pour chaque utilisateur et qu'il n'est pas possible de suivre un auditeur sans déplacer l'ensemble du système. Récemment, une méthode a permis de reproduire un champ sonore directement à proximité des oreilles d'un auditeur en utilisant un laser [Sullenberger *et al.*, 2019]. Cette méthode réalise un échauffement local de l'air induisant sa dilatation et par conséquent une onde acoustique. Cette méthode est encore pour l'instant à ses débuts. Une dernière solution consiste à utiliser un réseau de haut-parleurs (**Figure 1.7**) et des méthodes d'audio spatialisé comme celles présentées précédemment (§1.1.4) pour permettre la reproduction d'un champ sonore cible. Seule cette dernière solution sera étudiée dans ce document, les autres méthodes ne seront pas développées.

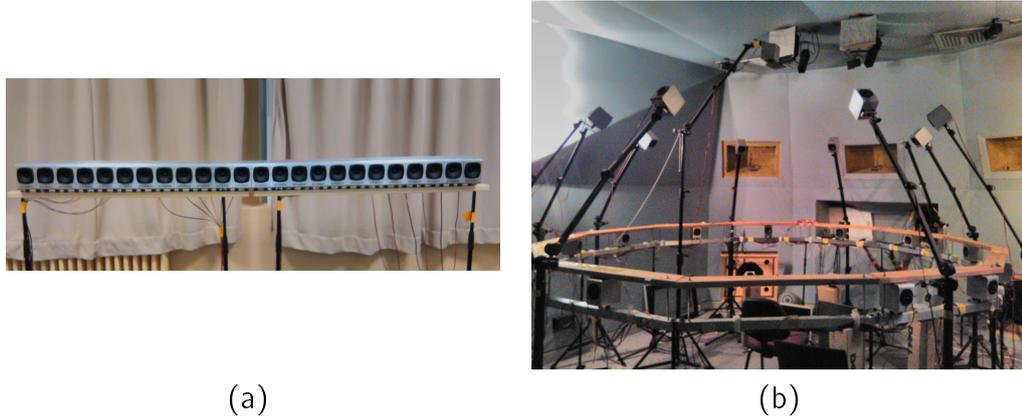


Figure 1.7 – Exemple de réseaux de haut-parleurs mis en place à Orange Labs : (a) Réseau linéaire de 24 haut-parleurs (WFS), (b) Réseau sphérique de 22 haut-parleurs (HOA/22.2)

La littérature sur les *Sound Zones* définit 2 zones de l'espace : une première zone dans laquelle le contenu audio doit être reproduit, appelée *Bright zone*, et une seconde zone, appelée *Dark zone*, dans laquelle la pression acoustique doit être minimale (**Figure 1.8**). Imposer un champ acoustique cible $p_{\mathbb{A}}$ dans une zone \mathbb{A} de l'espace et un champ acoustique cible $p_{\mathbb{B}}$ dans une zone \mathbb{B} de l'espace, revient à décomposer le problème en 2 étapes :

1. $\mathbb{A} \equiv \textit{Bright zone}$ cible et $\mathbb{B} \equiv \textit{Dark zone}$ pour le premier auditeur,
2. $\mathbb{A} \equiv \textit{Dark zone}$ et $\mathbb{B} \equiv \textit{Bright zone}$ pour le second auditeur.

Ces 2 étapes étant équivalentes, seule l'une d'entre elles sera analysée par la suite. On note par la suite \mathbb{B} la *Bright zone* et \mathbb{D} la *Dark zone*. Le champ de pression mesuré par des microphones³ dans la *Bright zone* est noté $\mathbf{p}_{\mathbb{B}}$ et celui de dans la *Dark zone* est noté $\mathbf{p}_{\mathbb{D}}$. Les trajets acoustiques entre les haut-parleurs et les microphones de la *Bright zone* et de la *Dark zone* sont déterminés respectivement par les matrices de fonction de transfert $\mathbf{G}_{\mathbb{B}}$ et $\mathbf{G}_{\mathbb{D}}$. Enfin le vecteur des poids des sources pour un système à L sources est noté $\mathbf{q} = [q_1, \dots, q_L]^T$, avec T la transposée. Les pressions sont alors définies par

$$\mathbf{p}_{\mathbb{B}} = \mathbf{G}_{\mathbb{B}}\mathbf{q} \quad (1.8)$$

3. Les microphones correspondent à des points de mesure du champ sonore. Ils peuvent être réels ou simulés. La simulation de ces derniers permet alors d'obtenir un contrôle en "boucle ouverte".

et

$$\mathbf{p}_D = \mathbf{G}_D \mathbf{q}. \quad (1.9)$$

Les différentes approches proposées dans la littérature pour la réalisation de *Sound Zones* sont développées par la suite en se basant sur ce formalisme.

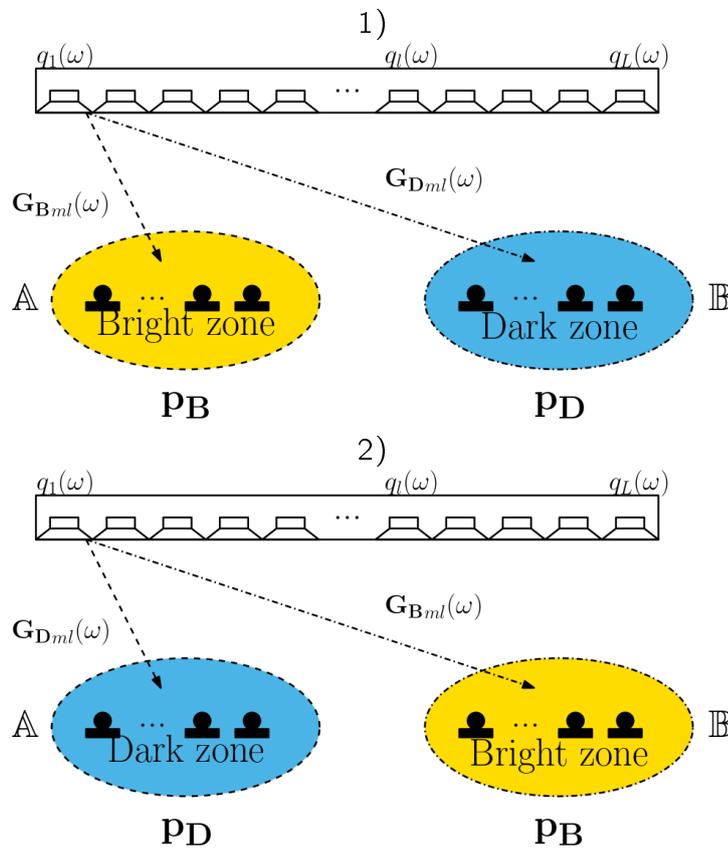


Figure 1.8 – Définition des deux sous-problèmes acoustique pour la génération de *Sound Zones* . 1) Le premier sous-problème définit une *Bright zone* en zone \mathbb{A} et une *Dark zone* en zone \mathbb{B} . 2) Le second-sous problème définit une *Dark zone* en zone \mathbb{A} et une *Bright zone* en zone \mathbb{B} . En combinant les deux sous-problèmes, on obtient des zones de reproduction individuelles pour chaque auditeur. La matrice de fonctions de transfert $\mathbf{G}_B(\omega, n)$ définit les trajets acoustiques entre les M microphones et les L_{HP} haut-parleurs de la *Bright zone*, de la même façon, la matrice de fonctions de transfert $\mathbf{G}_D(\omega, n)$ définit les trajets acoustiques pour la *Dark zone*. Chaque microphone m délivre une pression $p_m(\omega, n)$, avec $p_{Bm}(\omega, n)$ pour les pressions en *Bright zone* et $p_{Dm}(\omega, n)$ pour les pressions en *Dark zone*. Les haut-parleurs sont pilotés par des poids $q_l(\omega, n)$.

1.2.1 Méthodes de focalisation

Beamforming

Une première approche consiste à réaliser une focalisation de l'énergie acoustique dans la *Bright zone* B sans se préoccuper de la *Dark zone* D , par exemple à l'aide des méthodes de formation de voies par délai et somme (*Beamforming*, [Van Veen et Buckley, 1988]). On obtient alors une focalisation de l'intensité sonore dans une zone donnée

de l'espace. Les premiers travaux réalisés dans ce domaine sont ceux de [Van Veen et Buckley, 1988]. Les poids des sources peuvent s'écrire $\forall l \in [1, L_{HP}]$ sous la forme :

$$\mathbf{q}(\omega) = [e^{j\omega\tau_1}, e^{j\omega\tau_2}, \dots, e^{j\omega\tau_{L_{HP}}}]^T \quad (1.10)$$

avec $\tau_l = \frac{r_{max} - r_l}{c}$, r_l correspondant à la distance entre le haut-parleur et la position de focalisation et $r_{max} = \max\{r_l\}$.

L'énergie acoustique est alors focalisée sur un seul point de contrôle et n'offre pas la possibilité d'annuler le champ dans une autre zone de l'espace. Elle représente la solution la plus simple et se calcule de façon analytique. D'autres méthodes de sources focalisées sont développées et se présentent sous forme de problèmes optimaux [Elliott *et al.*, 2010]. Enfin la méthode de *Beamforming* peut être proposée comme une ancre basse dans les systèmes d'évaluation des *Sound Zones* [Wen *et al.*, 2005].

Brightness control

Pour réaliser une focalisation acoustique sur une zone plutôt qu'en un point, le problème peut être posé sous forme d'un problème d'optimisation correspondant à la maximisation de la pression \mathbf{p}_B dans la *Bright zone*.

Dans le cas de la méthode de *Brightness Control* (BC), le but est de maximiser la pression quadratique $\mathbf{p}_B^H \mathbf{p}_B$ de la *Bright zone* tout en limitant l'énergie $\mathbf{q}^H \mathbf{q}$ rayonnée par le réseau de haut-parleurs. Une fonction de coût proposée par [Elliott *et al.*, 2010] permet de poser le problème d'optimisation sous la forme :

$$J_{BC} = \mathbf{p}_B^H \mathbf{p}_B - \beta(\mathbf{q}^H \mathbf{q} - Q), \quad (1.11)$$

avec une contrainte $Q = |q_r|^2 \times 10^{\frac{E}{10}}$, le poids d'une source de débit. Elle est définie à l'aide d'une source de débit de référence q_r arbitrairement choisie et un effort de réseau E en dB (Eq. (2.37)). Cette contrainte permet de limiter le débit acoustique des sources dans le but de contrôler l'effort de réseau E .

Le problème est généralement résolu à l'aide de la méthode des multiplicateurs de Lagrange. La maximisation de la fonction de coût J_{BC} s'obtient en cherchant les points stationnaires de cette fonction en annulant les dérivées par rapport à \mathbf{q} . Le terme de régularisation β correspond à la contrainte dans le problème d'optimisation de Lagrange et devient dans ce cas un multiplicateur de Lagrange noté β . On cherche donc :

$$\frac{\partial J_{BC}}{\partial \beta} = 2[\mathbf{G}_B^H \mathbf{G}_B - \lambda] \mathbf{q} = 0, \quad (1.12)$$

et

$$\frac{\partial J_{BC}}{\partial \mathbf{q}} = \mathbf{q}^H \mathbf{q} - Q = 0. \quad (1.13)$$

Par ailleurs en partant de l'équation (1.11) il est alors possible d'écrire le multiplicateur de Lagrange sous la forme :

$$\beta = \frac{\mathbf{p}_B^H \mathbf{p}_B}{\mathbf{q}^H \mathbf{q}} = \frac{\mathbf{q}^H \mathbf{G}_B^H \mathbf{G}_B \mathbf{q}}{\mathbf{q}^H \mathbf{q}}. \quad (1.14)$$

La maximisation de β correspond à la pression maximale possible pour une puissance $\mathbf{q}^H \mathbf{q}$ donnée. D'autre part l'équation 1.12 peut être posée comme un problème aux valeurs propres, sous la forme

$$\mathbf{G}_B^H \mathbf{G}_B \mathbf{q} = \beta \mathbf{q}. \quad (1.15)$$

Le vecteur optimal \mathbf{q} est alors proportionnel au vecteur prototype $\hat{\mathbf{q}}$, à une constante multiplicative a près, $\mathbf{q} = a \hat{\mathbf{q}}$. Ce vecteur de poids prototype $\hat{\mathbf{q}}$ correspond au vecteur propre associé à la plus grande valeur propre de $\mathbf{G}_B^H \mathbf{G}_B$.

Pour obtenir un niveau sonore cible L_{TB} dans la *Bright zone*, sachant que l'on veut une pression cible p_T , i.e. $\mathbf{p}_B^H \mathbf{p}_B = p_T^2$, on peut alors écrire

$$\mathbf{p}_B^H \mathbf{p}_B = a_{BC}^2 \hat{\mathbf{q}}^H \mathbf{G}_B^H \mathbf{G}_B \hat{\mathbf{q}}. \quad (1.16)$$

On en déduit la valeur de la constante multiplicative

$$a = \sqrt{\frac{M_B |p_0|^2 \times 10^{\frac{L_{TB}}{10}}}{\hat{\mathbf{q}}^H \mathbf{G}_B^H \mathbf{G}_B \hat{\mathbf{q}}}}, \quad (1.17)$$

$p_T = M_B |p_0| \times 10^{\frac{L_{TB}}{10}}$ la pression cible prenant en compte le nombre de microphones M_B de la *Bright zone*, $p_0 = 2 \times 10^{-5}$ Pa la pression de référence.

On note que l'ajustement de la valeur de a a une influence aussi bien sur la valeur des poids \mathbf{q} que sur la valeur de la pression quadratique dans la *Bright zone*. La valeur de la pression dans la *Bright zone* est donc dépendante de l'effort de réseau. Cependant, le poids de la source de référence q_r correspondant à une source unique monopolaire générant un niveau sonore L_B dans la *Bright zone*, on est assuré que les poids obtenus respectent la contrainte d'effort. Un algorithme pour l'obtention des poids avec la méthode BC est proposé en **Annexe A.1**

Les méthodes telles que le *Brightness Control* ou le *Beamforming* correspondent à des méthodes de focalisation et ne permettent pas d'obtenir une bonne séparation des contenus audio et s'avèrent donc assez peu adaptées au cas multi-utilisateur [Coleman, 2014].

1.2.2 Méthodes d'annulation d'énergie

Acoustic Contrast Control

Les méthodes de Contrôle de Contraste Acoustique (ACC) proposées par [Elliott *et al.*, 2012] cherchent à maximiser le contraste entre la *Bright zone* et la *Dark zone*, donc à maximiser la séparation de contenu entre les deux zones. Pour cela, la pression quadratique $\mathbf{p}_B^H \mathbf{p}_B$ dans la *Bright zone* doit tendre vers une valeur de pression désirée p_T^2 tout en minimisant la pression quadratique $\mathbf{p}_D^H \mathbf{p}_D$ dans la *Dark zone* avec une contrainte limitant l'énergie $\mathbf{q}^H \mathbf{q}$ rayonnée par le réseau, ce qui peut s'écrire sous forme d'un problème d'optimisation avec la fonction de coût :

$$J_{ACC} = \mathbf{p}_D^H \mathbf{p}_D + \lambda (\mathbf{p}_B^H \mathbf{p}_B - p_T) - \beta (\mathbf{q}^H \mathbf{q} - Q). \quad (1.18)$$

En résolvant le problème avec la méthode des multiplicateurs de Lagrange, cette fonction de coût est minimale pour ses points stationnaires correspondant aux dérivées partielles nulles de J_{ACC} respectivement par rapport à \mathbf{q} , λ et β . Dans ce cas, λ est un

multiplicateur de Lagrange mais ne correspond pas à un paramètre de régularisation. La régularisation doit s'effectuer de façon indépendante. On a donc

$$\frac{\partial J_{ACC}}{\partial \mathbf{q}} = 2[\mathbf{G}_D^H \mathbf{G}_D \mathbf{q} + \lambda \mathbf{G}_B^H \mathbf{G}_B \mathbf{q} + \beta \mathbf{q}] = 0, \quad (1.19)$$

$$\frac{\partial J_{ACC}}{\partial \lambda} = \mathbf{p}_B^H \mathbf{p}_B - p_T^2 = 0, \quad (1.20)$$

$$\frac{\partial J_{ACC}}{\partial \beta} = \mathbf{q}^H \mathbf{q} - Q = 0. \quad (1.21)$$

Pour minimiser la fonction de coût, on récupère dans un premier temps des poids prototypes $\hat{\mathbf{q}}$ correspondant à la résolution du problème aux valeurs propres de l'équation (1.19) qui s'écrit :

$$[\mathbf{G}_B^H \mathbf{G}_B]^{-1} [\mathbf{G}_D^H \mathbf{G}_D + \beta \mathbf{I}] \mathbf{q} = \lambda \mathbf{q}. \quad (1.22)$$

Les poids prototypes $\hat{\mathbf{q}}$ sont proportionnels au vecteur propre correspondant à la plus grande valeur propre de $[\mathbf{G}_B^H \mathbf{G}_B]^{-1} [\mathbf{G}_D^H \mathbf{G}_D + \beta \mathbf{I}] \mathbf{q}$.

L'inversion de la matrice $\mathbf{G}_B^H \mathbf{G}_B$ est problématique car celle-ci peut être mal conditionnée. Il est donc nécessaire de régulariser l'inversion pour obtenir des résultats acceptables. Pour réaliser la régularisation, il est préférable de résoudre le problème équivalent

$$[\mathbf{G}_D^H \mathbf{G}_D + \beta \mathbf{I}]^{-1} [\mathbf{G}_B^H \mathbf{G}_B] \mathbf{q} = \frac{1}{\lambda} \mathbf{q}, \quad (1.23)$$

en cherchant les poids prototypes $\hat{\mathbf{q}}$ proportionnels aux vecteurs propres correspondant aux valeurs propres minimales de $[\mathbf{G}_D^H \mathbf{G}_D + \beta \mathbf{I}]^{-1} [\mathbf{G}_B^H \mathbf{G}_B]$.

En plus de régulariser l'inversion de la matrice, β agit comme un facteur permettant de donner plus d'importance, soit à la maximisation de la pression dans la *Bright zone* lorsqu'il augmente, soit à l'annulation d'énergie dans la *Dark zone* lorsqu'il diminue.

En pratique, pour réaliser l'ensemble de ces contraintes, on calcule un premier prototype de poids avec $\beta_0 = 0$ (sans régularisation), puis a est ajusté comme défini par l'équation (1.17) pour obtenir la valeur de pression désirée dans la *Bright zone*. L'estimation de β est ensuite réalisée de façon itérative par descente de gradient, de sorte à satisfaire à la fois la contrainte $\mathbf{q}^H \mathbf{q} = Q$ et la contrainte sur le niveau sonore cible dans la *Bright zone* à l'aide du coefficient a . Un algorithme pour l'obtention des poids avec la méthode ACC est proposé en **Annexe A.2**.

Un des problèmes majeurs de cette méthode réside dans l'inversion des matrices de fonction de transfert, lorsque le problème est mal conditionné. Une solution pour pallier ce problème est proposée par la méthode AEDM.

Acoustic Energy Difference Maximization (AEDM)

Pour éviter l'inversion non régularisée de l'ACC, une formulation alternative a été proposée par [Elliott *et al.*, 2012]. La minimisation dans la *Dark zone* dépend d'un paramètre $\zeta \in \mathbb{R}$, au lieu d'être liée au multiplicateur de Lagrange. ζ est donc un paramètre qui règle le ratio entre le niveau sonore de la *Bright zone* et la *Dark zone*. Ce paramètre doit être ajusté pour chaque fréquence et constitue en soi un autre problème d'optimisation. Lorsque $\zeta = 0$, le problème se ramène au cas de *Brightness Control*. Lorsque $\zeta = \beta$

(Eq. (1.23)) on se retrouve dans le cas de l'ACC. ζ nécessite donc d'être optimisé de sorte à respecter les contraintes de contraste qui peuvent être imposées. Cette méthode permet par ailleurs de relâcher des contraintes sur l'inversion. Les meilleures performances obtenues en termes de contraste correspondent aux performances de l'ACC [Coleman, 2014]. La fonction de coût s'écrit sous la forme :

$$J_{AEDM} = \mathbf{p}_B^H \mathbf{p}_B - \zeta \mathbf{p}_D^H \mathbf{p}_D - \beta (\mathbf{q}^H \mathbf{q} - Q) = 0. \quad (1.24)$$

Pour minimiser la fonction de coût J_{AEDM} , de la même façon que précédemment, le problème est résolu par la méthode des multiplicateurs de Lagrange. Le paramètre de régularisation définit alors la contrainte et s'écrit β . On calcule alors la dérivée selon \mathbf{q} pour l'annuler, ce qui donne :

$$\frac{\partial J}{\partial \mathbf{q}} = 2 (\mathbf{G}_B^H \mathbf{G}_B \mathbf{q} - \zeta \mathbf{G}_D^H \mathbf{G}_D \mathbf{q} - \beta \mathbf{q}). \quad (1.25)$$

Il est alors possible de calculer β de sorte à ce que la dérivée soit nulle, pour obtenir les poids optimaux, sans avoir à réaliser une inversion. On a donc

$$\lambda \mathbf{q} = [\mathbf{G}_B^H \mathbf{G}_B - \zeta \mathbf{G}_D^H \mathbf{G}_D] \mathbf{q}. \quad (1.26)$$

La solution du problème aux valeurs propres est alors liée au vecteur propre correspondant à la plus grande valeur propre de $[\mathbf{G}_B^H \mathbf{G}_B - \zeta \mathbf{G}_D^H \mathbf{G}_D]$. De la même façon que précédemment, on obtient les poids prototypes. Ces derniers sont alors proportionnels à la solution du problème. Il reste donc à choisir un facteur de proportionnalité de sorte à respecter la contrainte sur les poids \mathbf{q} .

1.2.3 Méthodes de synthèse de champ sonore

Sound Zones avec WFS et HOA

Les méthodes de synthèse de champ sonore analytique permettent aussi la réalisation de *Sound Zones*, mais nécessitent quelques adaptations pour pouvoir être utilisées sous cette forme. Les méthodes HOA et WFS permettent de définir un champ sonore de façon globale, sur l'ensemble de la zone de reproduction. L'intérêt pour les *Sound Zones* est de pouvoir définir de façon indépendante les champs sonores dans les différentes zones. Pour cela, une méthode proposée par [Wu et Abhayapala, 2009, Wu et Abhayapala, 2011] retranscrit les champs sonores locaux (dans chaque zone), dans le champ sonore global, par l'utilisation d'une matrice appelée matrice de translation. Cette matrice contient les coefficients de chaque champ sonore retranscrit en fonction des coefficients du champ sonore global. Pour le cas de HOA, le nombre limite de modes⁴ à reproduire dépend du rayon de chaque zone. Le nombre de modes est donc limité d'une part par la dimension de chaque zone, mais aussi par la dimension de la zone englobante [Wu et Abhayapala, 2011, Kennedy *et al.*, 2007]. En appliquant ces méthodes aux *Sound Zones*, les méthodes de synthèse de champ donnent autant de poids à la minimisation de la pression dans la *Dark zone* que dans la *Bright zone*. Elle n'offrent donc pas un aussi bon contraste que les méthodes de types ACC mais permettent à l'inverse de reproduire un champ sonore de bonne qualité.

4. Le nombre de modes correspond à l'ordre de la fonction HOA.

Pressure Matching

Une autre solution utilisant une méthode de synthèse de champ sonore s'appuie sur la méthode de *Pressure Matching*. Le but est donc de minimiser l'erreur $\mathbf{e} = \mathbf{p} - \mathbf{p}_T$ entre la pression cible \mathbf{p}_T et la pression reproduite \mathbf{p} en chaque point de l'espace x défini par une position microphonique. Le problème est basé sur la minimisation de l'erreur au sens des moindres carrés, la formulation la plus simple de la fonction de coût s'écrit [Kirkeby *et al.*, 1996] :

$$J_{LS} = \mathbf{e}^H \mathbf{e} = (\mathbf{p} - \mathbf{p}_T)^H (\mathbf{p} - \mathbf{p}_T). \quad (1.27)$$

En ajoutant une contrainte sur l'énergie rayonnée par le réseau de haut-parleurs, on obtient l'équation équivalente à (1.7)

$$J_{PM} = [\mathbf{p} - \mathbf{p}_T]^H [\mathbf{p} - \mathbf{p}_T] + \beta \mathbf{q}^H \mathbf{q}. \quad (1.28)$$

La solution du problème peut être trouvée en calculant la dérivée de la fonction de coût

$$\frac{\partial J_{PM}}{\partial \mathbf{q}} = 2\mathbf{G}^H [\mathbf{G}\mathbf{q} - \mathbf{p}_T] + 2\beta\mathbf{I}\mathbf{q}, \quad (1.29)$$

avec \mathbf{I} , la matrice identité. En annulant la dérivée, il est possible de calculer directement les poids des sources

$$\mathbf{q} = [\mathbf{G}^H \mathbf{G} + \beta \mathbf{I}]^{-1} \mathbf{G}^H \mathbf{p}_T. \quad (1.30)$$

Cette solution est appelée méthode Directe dans la suite du document et correspond à la solution optimale de la fonction de coût. Cette solution est équivalente à un filtrage de Wiener [Haykin, 2002].

Dans le cas de la réalisation de *Sound Zones*, il est possible de définir deux champs cibles indépendants \mathbf{p}_{T_B} et \mathbf{p}_{T_D} respectivement pour la *Bright zone* et la *Dark zone*. Par ailleurs, on définit aussi indépendamment les pressions reproduites \mathbf{p}_B et \mathbf{p}_D . Pour se ramener au problème de reproduction de champ sonore, il est possible d'écrire

$$\mathbf{p}_T = \begin{bmatrix} \mathbf{p}_{T_B} \\ \mathbf{p}_{T_D} \end{bmatrix} \quad \text{et} \quad \mathbf{p} = \begin{bmatrix} \mathbf{p}_B \\ \mathbf{p}_D \end{bmatrix}. \quad (1.31)$$

En définissant une pression nulle dans la *Dark zone*, i.e. $\mathbf{p}_{T_D} = \mathbf{0}$, on définit alors le problème de *Sound Zones* en utilisant la méthode PM. Cependant, dans ce cas, un effort d'optimisation équivalent est donné dans chacune des zones pour reproduire le champ sonore dans la *Bright zone* et pour minimiser la pression en *Dark zone*. Dans le cas de *Sound Zones*, il peut être préférable de donner plus d'importance à la minimisation du champ sonore en *Dark zone* qu'à la précision de reproduction du champ sonore en *Bright zone*. Pour ce faire, une approche équivalente combinant la méthode PM dans la *Bright zone* et la méthode ACC dans la *Dark zone* peut être utilisée. On reproduit alors le champ uniquement en *Bright zone* tout en minimisant l'énergie acoustique en *Dark zone* sans préciser explicitement de champ cible. Une méthode proposée par [Poletti, 2008] permet de poser le problème sous la forme de la fonction de coût :

$$J_{PM} = \mathbf{p}_D^H \mathbf{p}_D + (\mathbf{p}_B - \mathbf{p}_{T_B})^H (\mathbf{p}_B - \mathbf{p}_{T_B}) + \beta (\mathbf{q}^H \mathbf{q} - Q), \quad (1.32)$$

les pressions cibles dans la *Bright zone* aux positions microphoniques x_m étant définies comme

$$\mathbf{p}_{T_B} = p_T e^{-jkx_m \cos(\theta_{PW})}. \quad (1.33)$$

Le minimum de cette fonction est obtenu en dérivant J_{PM} par rapport à \mathbf{q} et β . On obtient alors les poids sous la forme

$$\mathbf{q} = (\mathbf{G}_B^H \mathbf{G}_B + \mathbf{G}_D^H \mathbf{G}_D + \beta \mathbf{I})^{-1} \mathbf{G}_B^H \mathbf{p}_{T_B}, \quad (1.34)$$

avec une contrainte sur les poids

$$\mathbf{q}^H \mathbf{q} = Q. \quad (1.35)$$

La méthode PM permet donc d'obtenir directement les poids pour l'ensemble du réseau. De la même façon que précédemment, on adapte β de sorte à répondre à la contrainte de puissance des poids.

Une dernière déclinaison de la méthode PM permettant de régler plus finement le contraste entre les deux zones est la méthode *Weighted Least Square* (WLS). Proposée par [Chang et Jacobsen, 2012], l'équation (1.36) ajoute un coefficient η à la formulation de l'équation (1.32), ce qui permet de choisir le poids attribué à chaque partie de l'optimisation. Lorsque $\eta > 0.5$, plus d'importance est donnée à la minimisation de la *Dark zone*; si $\eta < 0.5$, plus d'importance est donnée à la minimisation de l'erreur dans la *Bright zone*. Lorsque $\eta = 0.5$, la formulation est alors équivalente à l'équation (1.32).

$$J_{WLS} = \eta \mathbf{p}_D^H \mathbf{p}_D + (1 - \eta) (\mathbf{p}_B - \mathbf{p}_{T_B})^H (\mathbf{p}_B - \mathbf{p}_{T_B}) + \beta (\mathbf{q}^H \mathbf{q} - Q). \quad (1.36)$$

A noter que la résolution de l'approche *Weighted Least Square* est équivalente à la résolution classique de PM.

1.2.4 Améliorations des *Sound Zones*

Les méthodes pour la reproduction de *Sound Zones* présentées dans la partie précédente donnent un aperçu de la diversité des solutions pour résoudre ce problème. Ces méthodes, combinées entre elles, donnent à leur tour naissance à de nouvelles approches qui permettent d'obtenir de meilleures performances ou de répondre à un problème plus spécifique. Parmi ces différentes méthodes, celles qui présentent la meilleure qualité audio sont les méthodes de focalisation sonore qui définissent des filtres simples, mais avec très peu de contraste [Coleman *et al.*, 2014b]. Les méthodes de synthèse de champ disposent elles aussi d'une bonne qualité audio due à la reproduction assez fidèle du champ avec un contraste acceptable, mais les interférences entre les programmes audio restent présents [Coleman *et al.*, 2014b, Baykaner *et al.*, 2015]. Les méthodes ACC quant à elles ont un très bon contraste et sont assez robustes dans les environnements réverbérants, mais ont une très mauvaise distribution spatiale du son à l'intérieur de la zone, il en résulte de nombreux artefacts audibles et des pré-échos dans le contenu audio [Coleman *et al.*, 2014b, Baykaner *et al.*, 2015].

Les principaux travaux réalisés dans les dernières années proposent des fonctions de coût qui combinent donc différentes méthodes pour améliorer certains aspects en fonction de différents critères. Quelques-unes de ces méthodes sont succinctement présentées ici.

Amélioration de la qualité audio

Un des problèmes rencontrés lors de la conception des *Sound Zones* est que le contraste entre les deux zones se fait au détriment de la qualité de reproduction du champ sonore. Soit parce que cela n'est pas prévu par la méthode, comme c'est le

cas pour l'ACC, soit parce que l'on préfère donner plus de poids à l'augmentation de contraste qu'à l'erreur de reproduction, comme dans la méthode *Weighted Least Square*. Au lieu de se baser sur une solution classique de reproduction du champ sonore, en approchant la pression à chaque microphone, la méthode de *Planarity Control* (PC) [Coleman *et al.*, 2014a] propose de représenter le champ sonore comme une onde plane dans la *Bright zone*. Pour cela, les auteurs utilisent dans la *Bright zone* la méthode *Wave Domain Point Focusing* (WDPF) qui a la particularité de représenter une onde plane sous forme d'un point dans le domaine des nombres d'onde et la méthode ACC dans la *Dark zone*, qui confère un bon contraste. Cette méthode a donc pour vocation de reproduire une onde plane de façon optimale dans la *Bright zone*, et donc d'obtenir une meilleure reproduction du champ sonore⁵ ainsi qu'un bon contraste entre les zones. De plus, l'onde plane n'est pas contrainte à être reproduite dans une seule direction, mais dans une gamme de directions, permettant au système de choisir la direction d'onde plane la plus adaptée à un problème donné, limitant ainsi les problèmes d'occlusion [Coleman, 2014].

Une autre approche proposée par [Møller *et al.*, 2012] permet de combiner la méthode PM dans la *Bright zone* avec la méthode AEDM dans la *Dark zone*. Cela permet alors de combiner la qualité audio de la méthode PM avec la possibilité de régler le contraste entre les deux zones en fonction d'un critère donné. Des études sur la prédiction de contrastes obtenus par les différentes méthodes [Møller et Olsen, 2016] ainsi que des études perceptives sur la gêne induite entre deux zones et leur perception [Francombe *et al.*, 2012, Francombe *et al.*, 2013b, Francombe *et al.*, 2013a] servent de base au réglage du contraste désiré en fonction du contexte d'écoute. D'autres méthodes s'intéressent plus particulièrement à l'optimisation de la reproduction vocale dans les *Sound Zones* en utilisant les effets de masquage perceptif pour limiter la gêne entre les deux zones [Donley *et al.*, 2016, Donley et Ritz, 2015].

Méthodes de résolution et d'optimisation

D'autres améliorations possibles dans le contexte des *Sound Zones* portent sur la méthode de résolution et d'optimisation. Parmi les problèmes de résolutions des méthodes de *Sound Zones*, le paramètre de régularisation revêt une importance particulière aussi bien dans les performances que dans la robustesse du système. Il en découle un grand nombre de méthodes pour sélectionner ce paramètre [Elliott *et al.*, 2012, Coleman *et al.*, 2013, Coleman *et al.*, 2014b]. D'autres améliorations portent par exemple sur la sélection des haut-parleurs afin de réduire le nombre nécessaire à la reproduction. C'est par exemple le cas des méthodes LASSO (Norme ℓ_1) et de leurs dérivées [Lilis *et al.*, 2010, Radmanesh et Burnett, 2013, Gauthier *et al.*, 2017]. D'autres méthodes sont aussi proposées afin d'améliorer la résolution ainsi que la compréhension du fonctionnement des *Sound Zones*, comme par exemple la méthode GSVD (*Generalized Singular Value Decomposition*) proposée par [Gauthier *et al.*, 2016].

Prise en compte de l'effet de salle

L'ensemble des méthodes proposées précédemment s'appuie sur une hypothèse de champ libre ainsi que sur l'utilisation de sources monopolaires. Ces hypothèses permettent de faciliter l'écriture des problèmes et donc leur résolution. Cependant, les

5. Dans ce cas, la qualité de champ sonore correspond à une bonne planéité dans la *Bright zone*, plutôt que l'erreur de reproduction. La planéité définit la ressemblance du champ sonore à une onde plane.

méthodes de restitution de champs sonores ont vocation à être utilisées en présence d'un effet de salle. Ce dernier doit alors être pris en compte pour permettre une reproduction du champ sonore désiré avec les meilleures performances possibles. L'effet de salle peut être compensé de différentes manières en fonction des méthodes de reproduction. L'approche la plus évidente consiste à réduire directement les réflexions des murs en ajoutant des matériaux absorbants. Cependant, cela n'a qu'un effet limité et devient rapidement coûteux ou impossible à mettre en œuvre dans certains contextes, tel que l'environnement domestique. Une autre approche est l'utilisation de haut-parleurs directs permettant de limiter l'énergie acoustique en direction des murs et donc d'obtenir une plus faible influence de la salle [Poletti *et al.*, 2010, Poletti et Abhayapala, 2011] ou d'utiliser la réflexion de ces murs pour aider la reproduction du champ sonore [Betlehem *et al.*, 2010]. Une autre solution consiste à prendre en compte l'effet de la salle directement dans le calcul des *Sound Zones* en utilisant les réponses impulsionnelles aux positions des microphones de contrôle. D'autres approches s'appuient sur des méthodes de compensation de salle telles que le *Modal-Space Processing* [Betlehem et Abhayapala, 2005, Talagala *et al.*, 2014]. Enfin, des méthodes telles que le *Wave Domain Adaptive Filtering* (WDAF) [Gauthier et Berry, 2006, Spors *et al.*, 2007] utilisent un filtre adaptatif pour compenser l'effet de salle en séparant les composantes du champ sonore lié à la salle et celles du champ sonore à reproduire.

1.2.5 Audio spatialisé pour utilisateurs mobiles

L'ensemble des méthodes de *Sound Zones* proposées dans les parties précédentes est généralement utilisé dans le cas d'utilisateurs fixes à des positions connues. Certaines de ces méthodes ont déjà été adaptées pour le cas des utilisateurs mobiles multiples, mais, à notre connaissance, il existe encore peu de littérature à ce sujet.

La méthode la plus utilisée dans le cas d'un unique utilisateur mobile est basée sur des méthodes transaurales dynamiques, qui se rapprochent dans le principe des méthodes de *Sound Zones* mais pour un nombre de points de reproduction limité. On cherche à reproduire un signal à une oreille tout en cherchant à minimiser ce même signal à l'autre oreille (**Figure** 1.3). Pour le transaural, le nombre de haut-parleurs ainsi que le nombre de points de contrôle est restreint. Il est donc possible d'appliquer la méthode Directe pour des utilisateurs mobiles [Song *et al.*, 2010] à condition de prendre en compte les effets d'artefacts causés par les changements dynamiques de filtres. Ces méthodes ont même été utilisées dans le cas de plusieurs utilisateurs avec un réseau de haut-parleurs [Simón Gálvez et Fazi, 2016, Simon Galvez et Fazi, 2017] et sont très proches des méthodes PM, ce qui se rapproche sensiblement d'un problème de *Sound Zones*. Cependant, dans le cas de deux utilisateurs [Simon Galvez et Fazi, 2017], l'auteur ne contrôle pas les 4 points correspondant aux oreilles de l'ensemble des utilisateurs, mais chaque utilisateur est considéré indépendamment sans réaliser de *Dark zone* pour le second utilisateur. De plus, la capacité à augmenter le nombre de points de contrôle et de haut-parleurs reste limitée principalement à cause du coût de calcul engendré par la résolution de tels systèmes en temps réel. Pour ce qui est du contrôle de champ sonore, une approche proposée par [Trevino *et al.*, 2016] permet l'utilisation de la méthode HOA dans le cas d'utilisateurs mobiles. Elle est basée sur une modification de la matrice de décodage de sorte à ce qu'elle soit déplacée à la position attendue de l'utilisateur, grâce à une prédiction de son déplacement, ce dernier restant toujours à l'intérieur de la zone. De plus, l'auteur propose un nouvel ensemble d'opérateurs sur la matrice de translation pour simplifier le problème. Cette méthode n'a pour l'instant été testée qu'en simulation

et l'auteur suggère qu'il n'est pour l'instant pas envisageable de l'implémenter en temps réel.

Conclusion

L'audio spatialisé est un vaste domaine étudié depuis de nombreuses années. Il en découle un ensemble de solutions variées et adaptées à des problèmes spécifiques. Ainsi, les méthodes perceptives cherchent à reproduire une identité perceptive d'une source sonore primaire alors que les méthodes de champ sonore reproduisent une identité physique du champ sonore induit par cette source. Les méthodes de champ sonore ouvrent la possibilité d'une mise en œuvre des *Sound Zones* en ajoutant des contraintes sur le champ à reproduire. Les *Sound Zones* cherchent à reproduire un contenu audio pour différents auditeurs partageant un environnement commun. Les différentes méthodes de *Sound Zones* révèlent également une grande diversité en termes de performances et de méthodes de résolution. Les méthodes de focalisation acoustique, malgré leur simplicité et leur bonne qualité audio ne sont pas adaptées à l'utilisation de *Sound Zones* à cause de leur contraste trop faible. Les méthodes d'annulation d'énergie, à l'inverse, permettent une très bonne séparation des contenus entre les différents utilisateurs, mais disposent d'une moins bonne qualité de reproduction dans la *Bright zone*. Enfin, les méthodes de synthèse de champ sonore permettent d'obtenir une très bonne qualité de reproduction du champ sonore, mais avec un contraste plus faible que les méthodes d'annulation. Des hybridations de ces méthodes existent, principalement en combinant les différentes méthodes de reproduction de champs sonores en *Bright zone* et les méthodes d'annulation d'énergie en *Dark zone*.

Parmi les méthodes de champs sonore, la méthode *Pressure Matching* dispose d'une grande flexibilité dans son écriture et permet d'être facilement utilisée dans de nombreux cas. C'est cette méthode qui est retenue dans ce document car elle offre à la fois une bonne qualité de reproduction dans la *Bright zone* et un contraste acceptable. De plus, différentes approches peuvent être considérées pour le choix de la *Dark zone*. Cependant, la plupart des méthodes s'intéressent au cas d'auditeurs statiques et ne sont pas adaptées à une utilisation dans un contexte d'utilisateurs mobiles à cause de leur coût de calcul élevé, en particulier lié à la sélection du paramètre de régularisation. Seules quelques applications ont été développées pour une utilisation pour des utilisateurs mobiles, mais avec de fortes contraintes sur le problème traité.

L'approche développée au cours de la thèse propose une méthode de résolution compatible avec une implémentation en temps réel en s'appuyant sur les méthodes de *Sound Zones* existantes. La méthode de *Pressure Matching* est une méthode de filtrage optimal et permet donc d'être reformulée à l'aide d'une résolution adaptative plutôt que d'utiliser une approche Directe. Ces méthodes adaptatives sont compatibles avec l'utilisation dans le cas de systèmes variant dans le temps et constituent l'approche choisie pour répondre au problème de *Sound Zones* pour des utilisateurs mobiles.

Chapitre 2

Application des méthodes adaptatives à la reproduction sonore dans une zone en mouvement

L'utilisation des méthodes de *Sound Zones* dans un contexte d'auditeurs mobiles requiert une résolution en temps réel du problème. Les méthodes actuelles de *Sound Zones* ont un coût de calcul élevé, qu'il s'agisse des méthodes d'optimisation ou des méthodes analytiques de synthèse de champs sonores présentées dans le chapitre précédent. Ce coût de calcul est principalement lié aux dimensions du problème, qui nécessite un grand nombre de haut-parleurs pour reproduire correctement le champ sonore ainsi qu'à la méthode de résolution. Dans notre cas d'étude, la méthode choisie étant la méthode de *Pressure Matching* (PM), elle nécessite également un nombre élevé de points de contrôle du champ sonore. Même si, dans certains cas, comme pour le transaural, le nombre de points de contrôle peut être diminué [Simón Gálvez et Fazi, 2016], le coût de calcul pour un champ sonore complet reste important, ce particulièrement à cause de l'inversion régularisée nécessaire à la résolution Directe du problème. Pour calculer cette régularisation, il est nécessaire de poser un problème d'optimisation indépendant dans le but d'obtenir des résultats exploitables. A la différence de [Simón Gálvez et Fazi, 2016], nous nous focalisons sur la reproduction d'un champ sonore sur un ensemble de points de contrôle plutôt qu'à deux points de l'espace. De ce fait, il n'est pas possible d'utiliser la méthode Directe comme proposée dans la littérature. Pour réaliser des *Sound Zones* adaptées à des auditeurs mobiles, il devient alors nécessaire de réduire ce coût de calcul. Le choix de la méthode PM est motivé par sa simplicité ainsi que par sa capacité d'adaptation à de nombreux problèmes de *Sound Zones* sans modification majeure. De plus, elle peut être combinée aux méthodes ACC par exemple, pour obtenir un meilleur contraste que la méthode PM seule. L'approche proposée pour réduire le coût de calcul consiste en l'utilisation de méthodes de résolution adaptatives. Dans ce chapitre, nous décrivons pour commencer les méthodes adaptatives classiques de type LMS (*Least Mean Square*) [Widrow et Hoff, 1960], NLMS (*Normalized LMS*) [Nagumo et Noda, 1967, Albert et Gardner, 1967] ainsi que APA (*Affine Projection Adaptive*) [Ozeki et Umeda, 1984] qui peuvent être adaptées au problème de PM. Ce chapitre définit l'ensemble des notions ainsi que les méthodes utilisées pour permettre l'évaluation de champs sonores et de *Sound Zones* mobiles. Les configurations utilisées sont présentées dans un premier temps, puis le problème PM (§1.2.3) est reformulé de sorte à s'intégrer dans le formalisme des filtres LMS, NLMS et APA. Les critères d'évaluation pour les zones mobiles sont ensuite définis avant de décrire le protocole de

simulation utilisé.

2.1 Problèmes traités

Deux problèmes sont traités dans ce document. Dans un premier temps, on s'intéresse au cas de la reproduction d'un champ sonore seul pour un auditeur mobile, ce qui correspond au problème de reproduction d'un champ sonore dans une *Bright zone* uniquement sans s'intéresser au problème de la *Dark zone*. Dans un second temps, le cas des *Sound Zones* pour plusieurs auditeurs est traité, c'est à dire la reproduction d'un champ sonore en *Bright zone* et la réduction du niveau sonore en *Dark zone*. Les solutions proposées se basent sur la méthode PM dans le domaine fréquentiel avec une résolution itérative dans un contexte quasi-statique (peu de variation entre deux itérations). Il existe alors une dépendance temporelle du système liée au déplacement de l'auditeur. Cette dépendance est définie pour l'ensemble du document comme étant liée à l'itération n (temps discret) d'un algorithme. Le choix de cette dépendance en temps discret et non en temps continu permet de s'affranchir du temps de calcul de chaque algorithme et permet de ne s'intéresser qu'aux performances de la méthode en termes de reproduction de champ sonore. Par ailleurs, on ne cherche pas ici à améliorer la qualité de reproduction des méthodes existantes, mais il s'agit plutôt de conserver leurs performances tout en permettant une résolution plus rapide du problème. D'autre part, nous considérons le problème de localisation de l'auditeur dans l'espace résolu puisque de nombreux dispositifs commerciaux existent (Kinect (Microsoft), Xtion (Asus), Carmine (PrimeSense), etc.) et peuvent être utilisés dans ce but .

2.1.1 Bright zone seule en contexte mobile

Pour évaluer le comportement des méthodes adaptatives dans le contexte de *Sound Zones* pour des auditeurs mobiles, nous nous intéressons dans un premier temps à l'utilisation des ces méthodes dans le cas d'un seul auditeur. La configuration pour un unique auditeur mobile est présentée en **Figure 2.1**.

La reproduction de champ sonore a pour but de reproduire le champ cible

$$\mathbf{p}_T(\omega, n) = [p_{T_1}, \dots, p_{T_m}, \dots, p_{T_M}]^T$$

pour chaque instant n et pour une pulsation ω , où T correspond à l'opérateur de transposition. Un groupe de M microphones virtuels¹ est disposé aux positions $x_m(n)$, $m = [1, \dots, M]$ et capture un ensemble de pressions $\mathbf{p}(\omega, n)$. Le champ sonore est reproduit par un ensemble de L_{HP} haut-parleurs fixes aux positions x_l , $l = [1, \dots, L_{HP}]$ pilotés par un ensemble de poids

$$\mathbf{q}(\omega, n) = [q_1(\omega, n), \dots, q_l(\omega, n), \dots, q_L(\omega, n)]^T.$$

Les trajets de propagation entre chaque haut-parleur l et chaque microphone m sont définis par un ensemble de fonctions de transfert $G_{ml}(\omega, n)$ assemblées dans une

1. Le terme microphone virtuel désigne ici le fait que même lors d'une implémentation pratique, il n'est pas prévu de positionner des micros à l'emplacement de l'auditeur, pour des raisons pratiques évidentes. Cet aspect sera discuté dans le Chapitre 5.

matrice définie par :

$$\mathbf{G}(\omega, n) = \begin{bmatrix} G_{11}(\omega, n) & \cdots & G_{1L}(\omega, n) \\ \vdots & \ddots & \vdots \\ G_{M1}(\omega, n) & \cdots & G_{ML}(\omega, n) \end{bmatrix}. \quad (2.1)$$

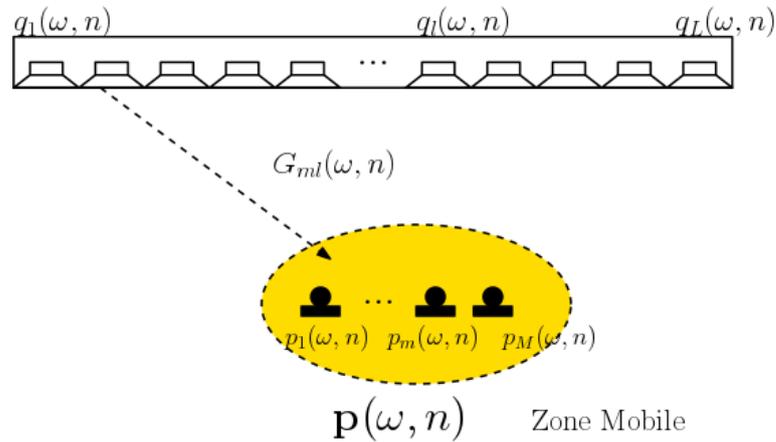


Figure 2.1 – Définition du problème acoustique pour la génération d'un champ sonore mono-auditeur dans un contexte d'un unique auditeur mobile. La matrice de fonctions de transfert $\mathbf{G}(\omega, n)$ définit les trajets acoustiques entre les M microphones et les L_{HP} haut-parleurs. Chaque micro m délivre une pression $p_m(\omega, n)$. Les haut-parleurs sont pilotés par des poids $q_l(\omega, n)$.

L'ensemble des pressions microphoniques est défini comme

$$\mathbf{p}(\omega, n) = \mathbf{G}(\omega, n)\mathbf{q}(\omega, n). \quad (2.2)$$

Le champ sonore qui doit être reproduit dépend lui aussi du temps ; toutes les notations dépendent alors du temps discret n . Comme défini en §1.2.3, la résolution Directe de la méthode PM permet d'obtenir les poids suivants pour l'ensemble des haut-parleurs :

$$\mathbf{q}(n) = [\mathbf{G}^H(n)\mathbf{G}(n) + \beta(n)\mathbf{I}]^{-1} \mathbf{G}^H(n)\mathbf{p}_T(n), \quad (2.3)$$

en ajoutant la variable temporelle n signifiant la variation temporelle du système. Cette formulation sert dans ce qui suit de référence à laquelle seront comparées les méthodes adaptatives de reproduction de champ sonore.

2.1.2 Cas de plusieurs auditeurs

Après avoir défini le problème de reproduction pour un auditeur mobile unique, nous nous intéressons maintenant à la formulation de ce même problème pour des utilisateurs multiples. C'est à dire dans le cas des *Sound Zones* combinant une *Bright zone* et une *Dark zone*. De la même façon que dans la partie 1.2.3, le but est de reproduire un champ sonore dans la *Bright zone* et de minimiser la pression acoustique dans la *Dark zone* (**Figure 2.2**). Les pressions dans la *Bright zone* et dans la *Dark zone* sont définies indépendamment au cours des itérations n et sont réunies dans une même matrice pour la pression cible $\mathbf{p}_T(n)$ et la pression observée $\mathbf{p}(n)$:

$$\mathbf{p}_T(n) = \begin{bmatrix} \mathbf{p}_{TB}(n) \\ \mathbf{p}_{TD}(n) \end{bmatrix} \quad \text{et} \quad \mathbf{p}(n) = \begin{bmatrix} \mathbf{p}_B(n) \\ \mathbf{p}_D(n) \end{bmatrix}. \quad (2.4)$$

Avec cette notation, la résolution se ramène à la résolution pour le problème mono-auditeur en équation 2.2. C'est pourquoi dans tout ce chapitre, les méthodes adaptatives sont présentées uniquement pour le problème mono-auditeur. A l'instar des méthodes hybrides PM, il est possible de définir le champ sonore cible $\mathbf{p}_{\text{TD}}(n)$ en Dark zone de différentes façons. Cet aspect sera développé dans la suite du document plus spécifiquement en relation avec les méthodes adaptatives (§2.4).

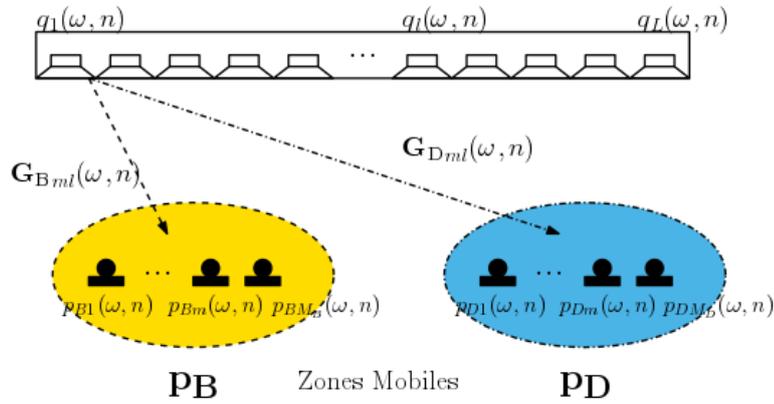


Figure 2.2 – Définition du problème acoustique pour la génération de *Sound Zones* dans un contexte de plusieurs auditeurs mobiles combinant une *Bright zone* et une *Dark zone*. La matrice des fonctions de transfert $\mathbf{G}_{\mathbf{B}}(\omega, n)$ définit les trajets acoustiques entre les M microphones et les L_{HP} haut-parleurs de la *Bright zone*. De la même façon, la matrice des fonctions de transfert $\mathbf{G}_{\mathbf{D}}(\omega, n)$ définit les trajets acoustiques pour la *Dark zone*. Chaque microphone m délivre une pression $p_m(\omega, n)$, avec $p_{Bm}(\omega, n)$ pour les pressions en *Bright zone* et $p_{Dm}(\omega, n)$ pour les pressions en *Dark zone*. Les haut-parleurs sont pilotés par des poids $q_l(\omega, n)$.

2.2 Méthodes adaptatives

La solution proposée pour réduire le coût de calcul considère une approche adaptative. Pour obtenir la solution du problème optimal à chaque itération, au lieu de calculer intégralement la solution au nouveau problème, l'algorithme se base sur les résultats des calculs aux itérations précédentes. Il faut cependant avoir à l'esprit que le problème peut changer entre deux itérations, si l'auditeur a bougé. Ceci permet, sous certaines conditions, de réduire le coût de calcul. Pour cela, il est nécessaire que le problème change peu entre deux itérations successives. Dans ce contexte, il est important de noter que la définition du problème est quasi-statique, et qu'en conséquence, il peut être exprimé dans le domaine fréquentiel à chaque itération temporelle. On considère donc un problème stationnaire variant lentement dans le temps. Cela implique que l'on considère une propagation instantanée de l'onde acoustique. Dans le contexte d'auditeur mobile, on considère que la vitesse de déplacement des auditeurs n'excède pas 1% de la vitesse de propagation de l'onde, la vitesse moyenne de marche étant de environ 1.4 ms^{-1} [Bohannon, 1997]. Il est donc raisonnable d'utiliser cette définition. Cette hypothèse est implicite dans les méthodes de transaural mobile telles que [Simón Gálvez et Fazi, 2016, Simon Galvez et Fazi, 2017, Song *et al.*, 2010]. Il est considéré que les filtres changent suffisamment peu entre deux itérations pour que cela ne soit pas audible par l'auditeur. En pratique, il existe des artefacts lors du changement de filtre. Cependant, dans le cas de [Simón Gálvez et Fazi, 2016], lors du filtrage en temps

réel, un fondu enchaîné est réalisé entre les filtres pour limiter ces artefacts [Wefers et Vorländer, 2014].

Les trois méthodes de résolution adaptatives qu'on se propose d'appliquer au problèmes d'auditeurs mobiles, à savoir, la méthode LMS (Least Mean Square) [Widrow et Hoff, 1960], la méthode NLMS (Normalized LMS) [Bitmead et Anderson, 1980] et la méthode APA (Affine Projection Adaptive) [Ozeki et Umeda, 1984] sont présentées par la suite. La théorie détaillée de ces algorithmes est proposée en **Annexe B**. Une quatrième méthode, dérivée du LMS et intégrant un paramètre d'oubli est également présentée : la méthode VLLMS (Variable Leaky LMS). L'intérêt de cette méthode sera mis en avant lors de l'évaluation des méthodes LMS, NLMS et APA dans le chapitre 3.

Seul le cas mono-auditeur est développé pour les méthodes itératives, la formulation étant équivalente pour le problème mono et multi-auditeur.

2.2.1 Filtre LMS

Le filtre LMS est le filtre adaptatif le plus simple. Il s'appuie sur le calcul de la dérivée de la fonction de coût par la méthode du gradient stochastique. La fonction de coût utilisée est la même que celle utilisée dans la méthode PM, sans le terme de régularisation, ce qui s'écrit :

$$J_{LMS} = \mathbf{e}^H \mathbf{e} = (\mathbf{p} - \mathbf{p}_T)^H (\mathbf{p} - \mathbf{p}_T), \quad (2.5)$$

A chaque itération, le poids est calculé comme suit : [Haykin, 2002]

$$\mathbf{q}(n+1) = \mathbf{q}(n) - \frac{1}{2} \mu_L \frac{\partial J_{LMS}}{\partial \mathbf{q}(n)}, \quad (2.6)$$

avec $\mathbf{q}(n+1)$ le poids mis à jour et $\mathbf{q}(n)$ le poids à l'itération courante. μ_L définit le pas de l'algorithme correspondant à la quantité de changement apportée à chaque itération.

La dérivée de la fonction de coût (**Eq.** (2.5)) s'écrit pour chaque itération :

$$\frac{\partial J_{LMS}}{\partial \mathbf{q}(n)} = \mathbf{G}^H(n) \mathbf{G}(n) \mathbf{q}(n) - 2 \mathbf{G}^H(n) \mathbf{p}_T. \quad (2.7)$$

En remplaçant la dérivée de J_{LMS} par son expression (**Eq.** (2.7)) dans l'équation (2.6) le nouveau poids s'obtient alors comme

$$\mathbf{q}(n+1) = \mathbf{q}(n) - \mu_L \mathbf{G}^H(n) [\mathbf{G}(n) \mathbf{q}(n) - \mathbf{p}_T(n)]. \quad (2.8)$$

La régularisation β n'est pas considérée pour le filtre LMS dans un premier temps car elle n'est plus utile puisque le conditionnement de l'inversion de matrice ne pose plus de problème contrairement au cas de la résolution Directe. En première approximation, le terme β n'est pas nécessaire au fonctionnement du filtre et n'apparaît donc pas ici. Cependant, l'absence de β n'impose aucune limite sur l'effort de reproduction et ne permet donc pas de le contrôler. Cet aspect est discuté dans la partie 2.2.4 et 3.1.1.

La sortie du système peut être identifiée par $\mathbf{p}(n) = \mathbf{G}(n) \mathbf{q}(n)$ correspondant aux pressions microphoniques observées. L'erreur de reproduction est donnée par $\mathbf{e}_{rr}(n) = \mathbf{G}(n) \mathbf{q}(n) - \mathbf{p}_T(n)$. En l'introduisant dans l'équation (2.8), il vient :

$$\mathbf{q}(n+1) = \mathbf{q}(n) - \mu_L \mathbf{G}^H(n) \mathbf{e}_{rr}(n). \quad (2.9)$$

Pour assurer la stabilité du filtre, μ_L doit être choisi de sorte à ce que [Haykin, 2002]

$$\mu_{Lmax} < \frac{2}{\kappa_{max}}, \quad (2.10)$$

où κ_{max} correspond à la valeur propre maximale de la matrice de corrélation spatiale $\mathbf{G}^H(n)\mathbf{G}(n)$. La valeur de μ_L a une influence sur la stabilité du filtre, la vitesse de convergence de l'algorithme et son bruit de gradient. Le bruit de gradient correspond à l'erreur résiduelle autour de la valeur optimale et susceptible d'être amplifiée par μ_L [Haykin, 2002]. Le paramètre μ_L conditionne le pas d'adaptation. Plus il est grand, plus l'algorithme converge rapidement, mais plus le bruit de gradient est important. Plus le pas est faible, plus le bruit de gradient est faible, et la convergence lente. En effet, vitesse de convergence et bruit de gradient sont liés. Lorsque l'algorithme est proche de la solution optimale et que le pas μ_L est grand, une grande partie de l'erreur résiduelle est prise en compte à chaque itération et se retrouve amplifiée. De plus, la matrice des fonctions de transfert $\mathbf{G}^H(n)$ prend en compte l'amplitude des pressions microphoniques : l'amplitude de cette matrice varie en fonction de la distance entre le réseau de haut-parleurs et le réseau de microphones. Plus le réseau de microphones (et donc l'auditeur) est éloigné du réseau de haut-parleurs, plus l'algorithme converge lentement, et plus il est proche, plus il converge vite. Le choix de μ_L est donc essentiel dans l'utilisation du filtre LMS car il définit à la fois son efficacité et sa stabilité. Dans le but de faciliter le choix de μ_L , il est possible de l'automatiser, comme c'est le cas des méthodes NLMS et APA, mais au prix d'un coût de calcul plus élevé. Ces méthodes sont développées dans les sections suivantes.

2.2.2 Filtre NLMS

Les méthodes APA et NLMS permettent d'ajuster la valeur du paramètre de pas μ_L entre chaque itération de sorte à maximiser la vitesse de convergence et minimiser le bruit de gradient. Le principe du filtre NLMS est de minimiser la variation des coefficients du filtre entre chaque itération, ce qui s'écrit [Haykin, 2002] :

$$\delta\mathbf{q}(n+1) = \mathbf{q}(n+1) - \mathbf{q}(n). \quad (2.11)$$

Dans le cas du filtre NLMS, la contrainte est portée sur l'égalité entre les poids $\mathbf{q}(n+1)$ de sortie du filtre et la pression cible $\mathbf{p}_T(n)$

$$\mathbf{G}(n)\mathbf{q}(n+1) = \mathbf{p}_T(n). \quad (2.12)$$

Le problème NLMS peut être posé comme un problème d'optimisation de Lagrange [Haykin, 2002], ce qui conduit à la fonction de coût suivante :

$$J_{NLMS} = \|\delta\mathbf{q}(n+1)\|^2 + \Re(\lambda^H(n) [\mathbf{G}(n)\mathbf{q}(n+1) - \mathbf{p}_T(n)]), \quad (2.13)$$

avec $\lambda(n) = [\lambda_0, \dots, \lambda_0]^T$ les multiplicateurs de Lagrange pour chaque pression microphonique, \Re l'opérateur partie réelle et

$$\|\delta\mathbf{q}(n+1)\|^2 = [\mathbf{q}(n+1) - \mathbf{q}(n)]^H \times [\mathbf{q}(n+1) - \mathbf{q}(n)]. \quad (2.14)$$

Dans le but de minimiser la fonction de coût, la dérivée est calculée comme

$$\frac{\partial J_{NLMS}}{\partial \mathbf{q}(n+1)} = 2 [\mathbf{q}^H(n+1) - \mathbf{q}^H(n)] + \lambda^H(n)\mathbf{G}(n). \quad (2.15)$$

La valeur optimale du poids mis à jour est obtenue en annulant cette dérivée :

$$\mathbf{q}(n+1) = \mathbf{q}(n) - \frac{1}{2} \mathbf{G}^H(n) \lambda(n). \quad (2.16)$$

La valeur de $\mathbf{q}(n+1)$ est alors remplacée dans la contrainte définie par l'équation (2.12) pour en déduire la valeur du multiplicateur de Lagrange

$$\lambda(n) = 2 [\mathbf{G}^H(n) \mathbf{G}(n) + \beta \mathbf{I}]^{-1} [\mathbf{e}_{\mathbf{r}\mathbf{r}}(n)], \quad (2.17)$$

avec $\beta > 0$, le paramètre de régularisation, réintroduit ici pour stabiliser l'inversion de matrice. De la même façon que pour le filtre LMS, il est à noter que le paramètre de régularisation n'a pas d'influence sur l'effort de reproduction et ne permet donc pas de limiter l'effort, mais régularise uniquement l'inversion de matrice. En combinant les équations (2.16) et (2.17), le filtre mis à jour s'écrit

$$\mathbf{q}(n+1) = \mathbf{q}(n) - [\mathbf{G}^H(n) \mathbf{G}(n) + \beta \mathbf{I}]^{-1} [\tilde{\mu} \mathbf{G}^H(n) \mathbf{e}_{\mathbf{r}\mathbf{r}}(n)], \quad (2.18)$$

avec $\tilde{\mu}$ une constante d'adaptation [Haykin, 2002]. Une condition sur $\tilde{\mu}$ est aussi nécessaire pour assurer la stabilité du filtre et a également une influence sur la vitesse de convergence. Dans notre cas, comme son nom l'indique, le filtre NLMS est une version normalisée du filtre LMS, le pas est réglé en fonction de l'énergie des données d'entrée du problème, ici les fonctions de transfert $\mathbf{G}(n)$. Cet aspect permet au filtre d'avoir une vitesse de convergence optimale et indépendante de l'amplitude des données d'entrée puisque le critère de stabilité est calculé à chaque pas par l'inversion de matrice, ce qui n'est pas le cas du filtre LMS. Si le système est non bruité, il n'est pas nécessaire de compenser le pas puisque le filtre est stable. Dans ce cas, la valeur optimale est $\tilde{\mu} = 1$ [Haykin, 2002]. Une valeur de $\tilde{\mu} < 1$ permet de réduire la vitesse de convergence pour assurer sa stabilité dans le cas d'un système bruité. Dans notre cas, le bruit d'entrée du système peut exister à cause de l'erreur de positionnement de l'auditeur, produite par les capteurs de position, et retranscrite dans les fonctions de transfert. On notera par ailleurs que le coût de calcul induit par cette méthode est équivalent à celui de la résolution directe du système puisqu'il nécessite l'inversion régularisée de la matrice de corrélation spatiale, ainsi que le calcul du paramètre de régularisation β équivalent à la solution Directe. On le vérifie en remarquant que $\lambda(n)$ est équivalent à la solution Directe du problème.

2.2.3 Filtre APA (Affine Projection Adaptive)

Le filtre APA peut être vu comme la généralisation du filtre NLMS. Comme le filtre NLMS, le filtre APA minimise à chaque itération la variation des coefficients du filtre avec une contrainte sur la pression reproduite. À la différence du filtre NLMS, qui ne prend en compte que la contrainte à l'itération courante, cette contrainte est réalisée, pour l'algorithme APA, sur les K sorties précédentes, définissant l'ordre du filtre [Haykin, 2002]. La contrainte équivalente de l'équation (2.12) pour le filtre APA s'écrit :

$$\sum_{i=0}^K [\mathbf{G}(n-i) \mathbf{q}(n+1) - \mathbf{p}_{\mathbf{r}}(n-i)]. \quad (2.19)$$

En écrivant l'équation (2.13) avec la contrainte de l'équation (2.19), la fonction de coût pour le filtre APA s'exprime sous la forme :

$$J_{APA} = \|\delta \mathbf{q}(n+1)\|^2 + \sum_{i=0}^K \Re(\lambda_i^H [\mathbf{G}(n-i) \mathbf{q}(n+1) - \mathbf{p}_{\mathbf{r}}(n-i)]). \quad (2.20)$$

Pour simplifier les écritures, le problème peut être écrit sous forme matricielle avec

$$\mathbf{G}_F(n) = \begin{bmatrix} \mathbf{G}(n) \\ \mathbf{G}(n-1) \\ \mathbf{G}(n-2) \\ \vdots \\ \mathbf{G}(n-K) \end{bmatrix}, \quad (2.21)$$

$$\mathbf{p}_{T_F}(n) = \begin{bmatrix} \mathbf{p}_T(n) \\ \mathbf{p}_T(n-1) \\ \mathbf{p}_T(n-2) \\ \vdots \\ \mathbf{p}_T(n-K) \end{bmatrix}, \quad (2.22)$$

et

$$\boldsymbol{\lambda} = \underbrace{\left[\underbrace{\lambda_0, \dots, \lambda_0}_M, \underbrace{\lambda_1, \dots, \lambda_1}_M, \dots, \underbrace{\lambda_K, \dots, \lambda_K}_M \right]}_{K \times M}^T. \quad (2.23)$$

Il en suit

$$J_{APA} = \|\delta \tilde{\mathbf{q}}(n+1)\|^2 + \Re(\boldsymbol{\lambda}^H [\mathbf{G}_F(n) \mathbf{q}(n+1) - \mathbf{p}_{T_F}(n)]), \quad (2.24)$$

et l'erreur sur les K précédentes itérations vaut

$$\mathbf{E}_{rr}(n) = \mathbf{G}_F(n) \mathbf{q}(n+1) - \mathbf{p}_{T_F}(n). \quad (2.25)$$

De la même façon que pour le filtre NLMS, en calculant la dérivée de l'équation (2.24), on obtient

$$\frac{\partial J_{APA}}{\partial \mathbf{q}(n+1)} = 2[\mathbf{q}^H(n+1) - \mathbf{q}^H(n)] + \boldsymbol{\lambda}^H \mathbf{G}_F(n). \quad (2.26)$$

Les valeurs des multiplicateurs de Lagrange sont obtenues en annulant la dérivée de la fonction de coût (**Eq.** (2.24)). En prenant en compte la contrainte de l'équation (2.19), il vient :

$$\boldsymbol{\lambda} = 2[\mathbf{G}_F^H(n) \mathbf{G}_F(n) + \beta \mathbf{I}]^{-1} \times [\mathbf{G}_F^H(n) \mathbf{E}_{rr}(n)]. \quad (2.27)$$

En reportant les multiplicateurs de Lagrange ainsi obtenus dans l'expression de la dérivée de la fonction de coût (**Eq.** (2.24)), les poids optimaux sont donnés par l'annulation de cette dernière :

$$\mathbf{q}(n+1) = \mathbf{q}(n) - [\mathbf{G}_F^H(n) \mathbf{G}_F(n) + \beta \mathbf{I}]^{-1} [\tilde{\boldsymbol{\mu}} \mathbf{G}_F^H(n) \mathbf{E}_{rr}(n)]. \quad (2.28)$$

Un filtre APA d'ordre K est généralement appelé filtre APA K . Il est à noter que la solution du filtre NLMS correspond à un APA 1, c'est à dire un filtre APA à l'ordre $K = 1$. La remarque apportée sur la régularisation β du filtre NLMS reste valable pour le filtre APA. C'est à dire que le paramètre de régularisation n'a pas d'influence sur l'effort de reproduction et ne permet donc pas de limiter l'effort, mais régularise uniquement l'inversion de matrice. Le coût de calcul induit par ce type de filtre est supérieur à celui du filtre NLMS et donc aussi à celui de la méthode Directe. Son intérêt dans le cas d'auditeurs mobiles est limité. Cependant, il reste intéressant d'évaluer ce filtre pour en comparer les performances de reproduction avec les deux autres filtres présentés précédemment.

2.2.4 Filtre VLLMS (Variable Leaky LMS)

À la différence des filtres précédemment proposés (LMS, NLMS et APA), les filtres Leaky LMS permettent de ne garder qu'une solution partielle des poids calculés à l'itération précédente². Les poids de l'itération précédente sont atténués d'un coefficient appelé facteur d'oubli, noté γ par la suite. L'intérêt de cette méthode dans le cas des *Sound Zones* est de réduire l'effet mémoire [Roussel et Nicol, 2018a] qui sera mis en évidence en partie 3.3, et se caractérise par une "zone de trainée" (**Figure 3.24**) aux positions précédemment occupées par l'auditeur. Dans ce type de filtres, les poids plus éloignés dans le temps ne sont plus pris en compte. Cependant, l'estimation des poids du filtre est biaisée et ne permet pas de converger à la valeur optimale. Une méthode [Kamenetsky et Widrow, 2004] propose d'utiliser un facteur d'oubli $\gamma(n)$ variable dans le temps. Cette méthode est appelée ici Variable Leaky LMS (VLLMS). Elle permet de réduire le biais induit par le facteur d'oubli de la méthode Leaky LMS en ne l'appliquant que lorsqu'il est nécessaire. Ce facteur d'oubli $\gamma(n)$ est introduit au moyen d'un paramètre de régularisation en injectant une contrainte sur la puissance des poids du filtre LMS. L'effet et l'intérêt de cette régularisation doivent être distingués de la régularisation appliquée dans le cas des filtres NLMS et APA. En effet, la régularisation β des filtres NLMS et APA permet d'assurer le bon conditionnement des matrices à inverser, tandis que la régularisation $\gamma(n)$ du filtre VLLMS limite l'effort de réseau et réduit la mémoire de l'algorithme. À la différence des méthodes LMS, NLMS et APA présentées précédemment, il est possible de limiter l'effort de réseau, ce qui revêt un avantage particulier dans le cas des *Sound Zones*.

Dans le cas d'un unique champ sonore, sans *Dark zone*, le problème peut s'écrire en minimisant l'erreur quadratique de reproduction de champ sonore à l'aide de la fonction de coût :

$$J = (\mathbf{G}(n)\mathbf{q}(n) - \mathbf{p}_T(n))^H (\mathbf{G}(n)\mathbf{q}(n) - \mathbf{p}_T(n)) + \gamma(n)\mathbf{q}^H(n)\mathbf{q}(n). \quad (2.29)$$

De la même façon que pour le filtre LMS, la solution au problème est recherchée de façon itérative par la méthode du gradient stochastique. La dérivée de l'équation (2.29) s'écrit :

$$\frac{\partial J}{\partial \mathbf{q}(n)} = \mathbf{G}^H(n)\mathbf{G}(n)\mathbf{q}(n) - 2\mathbf{G}^H(n)\mathbf{p}_T + \gamma(n)\mathbf{q}(n). \quad (2.30)$$

En remplaçant l'expression de l'équation (2.30) dans la méthode de descente de gradient (**Eq. (2.6)**), on obtient les poids à chaque itération de la méthode VLLMS

$$\mathbf{q}(n+1) = [1 - \mu\gamma(n)]\mathbf{q}(n) - \mu\mathbf{G}^H(n) [\mathbf{G}(n)\mathbf{q}(n) - \mathbf{p}_T], \quad (2.31)$$

avec $\gamma(n)$ correspondant au facteur d'oubli. Ce facteur d'oubli est variant au cours des itérations n . Lorsque $\gamma(n) = 0$ on se retrouve dans le cas d'un LMS classique. Lorsque $\gamma(n) > 0$ les poids de l'itération précédente sont alors en partie oubliés, on se retrouve alors dans le cas Leaky LMS [Kamenetsky et Widrow, 2004]. Cependant, cet oubli rend l'estimation des coefficients du filtre biaisée, il ne donne pas le résultat optimal. La variation du facteur d'oubli doit être liée à un paramètre reflétant le niveau requis d'oubli de sorte à doser le processus VLLMS. Dans le cas de [Kamenetsky et Widrow, 2004], il est proposé de le relier à l'erreur de reproduction. Pour choisir la valeur du paramètre $\gamma(n)$, les auteurs comparent la valeur de l'erreur a posteriori du filtre LMS et du filtre VLLMS pour chaque itération. Lorsque l'erreur du filtre VLLMS est supérieure

2. La solution est dite partielle dans le sens où une partie de l'amplitude des poids est oubliée, et non en tant que solution incomplète.

à celle de l'erreur du filtre LMS, alors le paramètre $\gamma(n)$ est augmenté ce qui permet à l'algorithme d'oublier plus vite les valeurs précédentes. Si la valeur de l'erreur du filtre VLLMS est inférieure ou égale à celle de l'erreur du filtre LMS, alors $\gamma(n)$ est diminué pour permettre au filtre de se stabiliser. Dans le cas de [Kamenetsky et Widrow, 2004], l'objectif est d'améliorer la vitesse de convergence pour les problèmes avec des entrées ayant une grande dispersion de valeurs propres. Cette méthode comporte deux principaux inconvénients, d'une part, elle nécessite le calcul des deux filtres (LMS et VLLMS), ce qui accroît le temps de calcul, d'autre part, dans notre cas, le but est de limiter l'effet mémoire de l'algorithme lié au déplacement de l'auditeur (§3.3) et de limiter l'énergie délivrée par le réseau. Il n'est donc pas pertinent d'utiliser ce même critère, ce pourquoi un autre paramètre est proposé. L'erreur de reproduction n'est quantifiée que dans la zone de reproduction (*Bright zone*) et non pas à l'extérieur de celle-ci. Si un champ sonore est présent dans la zone non contrôlée, il n'est pas pris en compte dans l'erreur de reproduction et n'a donc pas d'influence directe sur les poids calculés à l'itération suivante. Le champ sonore présent dans la trainée quantifie l'effet mémoire de l'algorithme. La prise en compte de ce champ sonore n'est pourtant pas envisageable car il nécessiterait d'évaluer la pression sur l'intégralité de l'espace, ce qui est coûteux en temps de calcul.

Afin d'identifier une stratégie plus appropriée à notre problème, nous évaluons l'intérêt d'autres critères pour contrôler le facteur d'oubli. L'utilisation de l'effort de réseau semble a priori un critère pertinent puisque l'effet mémoire influe sur ce paramètre (§3.3). L'inconvénient principal est qu'il n'existe pas de critère objectif pour déterminer quelle doit être la valeur de l'effort optimal pour une configuration donnée [Coleman et al., 2014b]. L'effort pouvant être variable au sein de l'espace pour des performances équivalentes d'un algorithme (§3.2.2), il n'est donc pas non plus intéressant de se baser sur sa variation.

La nouvelle stratégie que nous proposons [Roussel et Nicol, 2018a] consiste à se baser directement sur la cause de la variation de l'effet mémoire : le déplacement de l'auditeur (§3.3). Dans le cas où l'auditeur est fixe, l'algorithme n'a pas besoin d'oublier les filtres précédents, on impose donc $\gamma = 0$ pour que le filtre soit non biaisé et se comporte comme un filtre LMS. Dans le cas où l'auditeur bouge, il risque d'y avoir une "trace" de l'ancien emplacement de l'auditeur et il est donc nécessaire d'augmenter le facteur d'oubli en augmentant la valeur de γ .

Pour pouvoir réguler le facteur d'oubli, il est choisi dans [Kamenetsky et Widrow, 2004] de définir

$$\gamma(n) = \gamma_{max} \times \left(\frac{m}{\chi} \right)^\alpha, \quad (2.32)$$

avec $m = 0, \dots, \chi$ une variable indépendante ayant pour maximum χ . Cette formulation est choisie de sorte à permettre un contrôle indépendant pour la vitesse de montée ou de descente du paramètre d'oubli tout en bornant les valeurs de $\gamma(n)$. Pour simplifier le contrôle, il est intéressant de choisir $\chi = \mu$, ce qui permet de rendre le paramètre d'oubli indépendant du pas d'adaptation μ . De plus, il est utile d'imposer $0 < m < 1$, ce qui permet de choisir un paramètre d'oubli borné par $0 < \gamma(n) < \gamma_{max}$. Les bornes $\gamma(n) = 0$ et $\gamma(n) = \gamma_{max}$ correspondent respectivement au filtre LMS classique non biaisé et au facteur d'oubli le plus fort désiré du filtre Leaky LMS. Avec cette définition, γ_{max} s'identifie à un pourcentage maximal de poids à oublier entre chaque itération, ce qui rend son interprétation aisée.

Le paramètre α permet quant à lui de choisir la vitesse d'augmentation ou de diminution de la valeur de $\gamma(n)$ (Voir **Annexe D**).

Le choix de m varie au cours du temps. Dans notre cas, il est choisi tel que, s'il existe un déplacement de l'auditeur en position $x_{aud}(n)$, alors le facteur d'oubli augmente et s'il n'y a pas de déplacement, il diminue [Roussel et Nicol, 2018a].

$$\begin{aligned} \text{si } x_{aud}(n) \neq x_{aud}(n+1) \text{ alors } m &= \min(m + l_u, 1) \\ \text{si } x_{aud}(n) = x_{aud}(n+1) \text{ alors } m &= \max(m - l_d, 0), \end{aligned}$$

avec $0 < l_u \ll 1$ et $0 < l_d \ll 1$ des pas d'incrément respectivement dans le sens de l'augmentation de l'effet d'oubli et dans le sens de la diminution de l'effet d'oubli. La définition de deux paramètres différents l_u et l_d permet de choisir les vitesses de réaction du système en fonction du début ou de la fin du déplacement de l'auditeur, ce qui est un autre aspect original de notre solution. Leur influence est évaluée en **Annexe D**.

2.3 Evaluation du champ sonore

Pour permettre une comparaison des différents algorithmes, il est nécessaire de disposer de critères d'évaluation objectifs. Les paramètres les plus largement utilisés dans la littérature sont présentés ici [Coleman, 2014]. Un nouveau paramètre est introduit afin de rendre compte de l'homogénéité du niveau sonore dans la zone. Tous les paramètres sont définis en prenant en compte le caractère variant dans le temps des méthodes utilisées.

2.3.1 Niveau sonore

Le niveau sonore permet de rendre compte de la pression acoustique présente dans la zone considérée. Pour un ensemble de microphones, le vecteur des pressions microphoniques est défini comme

$$\mathbf{p}(\omega, n) = \mathbf{G}(\omega, n)\mathbf{q}(\omega, n). \quad (2.33)$$

Le niveau sonore pour un seul microphone est défini par :

$$L(n) = 20 \log_{10} \left(\frac{|p(n)|}{p_0} \right), \quad (2.34)$$

avec $p_0 = 2 \times 10^{-5}$ Pa, la pression acoustique de référence correspondant au seuil d'audibilité.

Dans le cas de la pression quadratique moyenne sur une zone (ensemble des pressions contenues dans le vecteur)

$$L(n) = 10 \log_{10} \left(\frac{\mathbf{p}^H(n)\mathbf{p}(n)/M}{p_0^2} \right), \quad (2.35)$$

où la transposée hermitienne notée \bullet^H correspond à la transposée conjuguée du vecteur des pressions microphoniques et où M correspond au nombre de microphones dans la zone considérée.

2.3.2 Contraste

Le contraste est défini comme le rapport des pressions quadratiques moyennes en décibel (dB) entre la *Bright zone* et la *Dark zone*. Il évalue la différence de niveau sonore sur l'ensemble de la *Bright zone* par rapport à la *Dark zone*. Il est équivalent à un RSB (Rapport Signal à Bruit) entre les deux zones, et constitue un paramètre objectif de la gêne occasionnée par la *Bright zone* dans la *Dark zone*. Il est défini à partir des pressions microphoniques suivantes :

$$C_{B \rightarrow D}(n) = 10 \log_{10} \left(\frac{M_D \mathbf{P}_B^H(n) \mathbf{P}_B(n)}{M_B \mathbf{P}_D^H(n) \mathbf{P}_D(n)} \right). \quad (2.36)$$

2.3.3 L'effort de contrôle

L'effort de contrôle E correspond à la quantité d'énergie rayonnée par l'ensemble du réseau de haut-parleurs par rapport au débit d'une source de référence q_r . Il s'exprime également en décibel et s'écrit :

$$E(n) = 10 \log_{10} \left(\frac{\mathbf{q}^H(n) \mathbf{q}(n)}{|q_r(n)|^2} \right). \quad (2.37)$$

La valeur de q_r est arbitraire, mais elle est généralement choisie comme étant le poids d'une source monopolaire de débit délivrant un niveau de pression équivalent à celui de la *Bright zone* lorsque cette source de référence est située dans le réseau de façon équidistante aux deux zones [Coleman, 2014]. Cette définition permet une interprétation physique simple de ce critère. Cependant, celui-ci est donné dans le cas de zones fixes définies pour un réseau circulaire de haut-parleurs et doit être adapté. Dans le cas d'auditeurs mobiles et d'un réseau linéaire de haut-parleurs, la position de la source doit être adaptée en fonction de la position des auditeurs. Nous proposons de définir le poids de cette source de référence comme étant celui d'une source monopolaire rayonnant le même niveau sonore que le niveau cible dans la *Bright zone*. La position de cette source est définie par la ligne de référence matérialisant le réseau linéaire de haut-parleurs à une distance R_{ref} perpendiculaire au centre de la *Bright zone* (**Figure 2.3**). En d'autres termes, on projette le centre de la *Bright zone* sur le réseau de haut-parleurs.

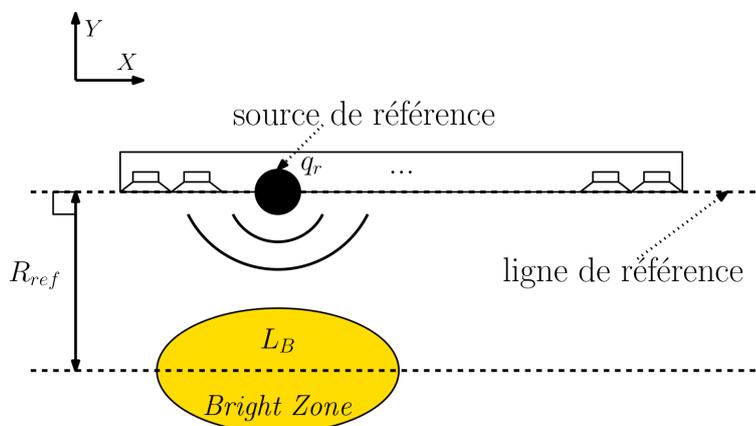


Figure 2.3 – Configuration de la source de débit de référence q_r pour le calcul de l'effort de réseau pour un niveau sonore L_B dans la *Bright zone* dans le cas d'un réseau linéaire de haut-parleurs adapté au problème d'auditeurs en mouvement.

L'interprétation physique reste simple dans ce cas et est utilisée pour l'ensemble des évaluations. Un effort de 0 dB exprime donc le fait que l'ensemble du réseau délivre la même énergie qu'une source monopolaire seule positionnée au même endroit. Il est préférable de garder cet effort en dessous de cette limite de 0 dB pour ne pas rayonner inutilement de l'énergie acoustique sur l'ensemble de l'espace de reproduction.

A partir de la définition de la pression rayonnée par une source de débit, le poids de cette source de référence est donné par

$$q_r(n) = \frac{p_{ref}(n)}{\frac{j\rho ck}{4\pi R_{ref}(n)} e^{-jkR_{ref}(n)}}, \quad (2.38)$$

avec $p_{ref}(n) = p_0^2 10^{\frac{L_{TB}(n)}{20}}$ la pression sonore correspondant au niveau sonore de référence, c'est à dire le niveau sonore cible L_{TB} de la *Bright zone*.

2.3.4 Homogénéité de la zone

Le niveau sonore au sein d'une zone n'est généralement caractérisé que par son niveau sonore moyen dans la zone. Cette mesure ne prend pas en compte sa répartition spatiale au sein de la zone. Cette information de répartition spatiale peut être utile soit lorsque les méthodes ne sont définies que par le niveau sonore moyen à reproduire dans la zone (par exemple ACC, chapitre 1), soit dans la *Dark zone*, qui ne dispose que très rarement d'un niveau sonore homogène. Même si l'homogénéité de la *Dark zone*, n'est pas directement relié à un critère de qualité dans la littérature, il peut être pertinent d'intégrer ce critère [Roussel et Nicol, 2018b]. Nous proposons de définir cette homogénéité par l'écart-type de niveau sonore pour l'ensemble de microphones définissant une zone, ce qui s'écrit :

$$\sigma_{SPL}(n) = \sqrt{\sum_{m=1}^M \frac{(L_m - L)^2}{M}}, \quad (2.39)$$

avec L_m le niveau sonore du $m^{\text{ième}}$ microphone.

2.4 Spécification du champ sonore en Dark zone

La pression cible en *Dark zone* peut être définie de plusieurs façons pour améliorer les performances de celle-ci. La solution la plus simple est de considérer le champ sonore cible comme nul sur l'ensemble des microphones de la zone ($\mathbf{p}_{\mathbf{T}_D} = \mathbf{0}$). Cette approche simple à mettre en place correspond à la formulation originelle de la méthode PM. Cependant, le contraste de cette méthode peut être amélioré par la méthode *Weighted Least Squares* (WLS), de sorte à donner plus d'importance soit à la minimisation de la *Dark zone* soit à la minimisation de l'erreur dans la *Bright zone*. Une formulation de cette méthode sous forme d'une résolution LMS est développée ici. Par ailleurs, nous proposons de développer une autre solution pour l'amélioration du contraste entre les zones, en tirant parti de la formulation itérative des algorithmes en utilisant le champ sonore déjà présent en *Dark zone* à l'itération précédente.

2.4.1 Weighted Least Squares avec résolution LMS

La méthode WLS (*Weighted Least Squares*) correspond à la fonction de coût de l'équation (1.36). Cette formulation permet de pondérer chaque terme de l'équation, c'est à dire soit la minimisation de la pression en *Dark zone*, soit la qualité de reproduction du champ sonore en *Bright zone*. Nous proposons d'adapter cette méthode de la littérature [Chang et Jacobsen, 2012] à la résolution LMS présentée précédemment (§2.2.1). En reprenant la fonction de coût de la méthode WLS (**Eq.** (1.36)), reportée ici sans le terme de régularisation :

$$J_{WLS} = \eta \mathbf{p}_D^H \mathbf{p}_D + (1 - \eta) (\mathbf{p}_B - \mathbf{p}_{T_B})^H (\mathbf{p}_B - \mathbf{p}_{T_B}). \quad (2.40)$$

Pour obtenir la formulation WLS avec une résolution itérative de type LMS, il suffit de calculer la dérivée de la fonction de coût (**Eq.** (2.40)), ce qui donne :

$$\frac{\partial J_{WLS}}{\partial \mathbf{q}(n)} = \eta \mathbf{G}_D^H(n) \mathbf{G}_D(n) \mathbf{q}(n) + (1 - \eta) (\mathbf{G}_B^H(n) [\mathbf{G}_B \mathbf{q}(n) - \mathbf{p}_T]). \quad (2.41)$$

En remplaçant l'expression de la dérivée dans l'équation de la descente de gradient (**Eq.** (2.6)), il vient l'expression des poids pour la méthode WLS avec une résolution de type LMS :

$$\mathbf{q}(n+1) = \mathbf{q}(n) + \mu_L [\eta \mathbf{G}_D^H(n) \mathbf{G}_D(n) \mathbf{q}(n) + (1 - \eta) (\mathbf{G}_B^H(n) [\mathbf{G}_B \mathbf{q}(n) - \mathbf{p}_T])]. \quad (2.42)$$

Ce résultat offre une alternative dans le choix de la fonction de coût afin d'améliorer le contraste entre les zones.

2.4.2 Dark zone choisie itérativement : méthodes DZIG et DZII

L'utilisation de méthodes itératives permet également de tirer avantage de la connaissance des itérations précédentes pour permettre d'adapter le champ sonore en *Dark zone*. Plutôt que de chercher à reproduire un champ cible nul dans la *Dark zone*, nous proposons d'utiliser le champ sonore présent à l'itération précédente dans la *Dark zone* et de l'atténuer à un niveau de contraste cible L_C [Roussel et Nicol, 2018b]. Deux solutions sont possibles. La première que nous appellerons DZIG (*Dark zone Itérative Globale*) consiste à réaliser une atténuation du niveau sonore global dans la *Dark zone*. La seconde appelée DZII (*Dark zone Itérative Individuelle*) consiste à contrôler le contraste de façon indépendante pour chaque microphone, de sorte à améliorer l'homogénéité du champ sonore en *Dark zone* [Roussel et Nicol, 2018b]. En contrôlant le contraste de façon indépendante pour chaque microphone, il est possible de diminuer le niveau sonore dans une partie spécifique de la zone. À l'inverse, lorsque le niveau sonore est trop faible dans la zone par rapport au contraste cible, il est intéressant de l'augmenter de sorte à ne pas fournir trop d'énergie pour la minimisation dans une partie de la zone. Le contraste cible L_C est défini de façon arbitraire de sorte à correspondre à un critère de contraste donné. Ce contraste cible peut éventuellement être choisi sur une base de critères perceptifs, ou en fonction du contenu audio [Francombe *et al.*, 2013b].

Le contraste résiduel C_ξ , correspond à la différence entre le contraste L_C cible et le contraste effectif observé $C_{B \rightarrow D}$ à l'instant n .

Il est défini comme :

$$C_\xi = C_{B \rightarrow D} - L_C = L_B - L_D - L_C. \quad (2.43)$$

A partir de ce contraste résiduel, on calcule pour chacun des microphones le coefficient d'atténuation correspondant

$$\xi = 10^{\frac{C_\xi}{20}}. \quad (2.44)$$

Ce coefficient correspond à l'atténuation à appliquer aux microphones. La méthode DZII correspond au calcul des coefficients ξ de la *Dark zone* avec le contraste résiduel C_ξ calculé pour chaque microphone, autorisant un niveau sonore homogène. La méthode DZIG quant à elle correspond au calcul de ξ avec le contraste résiduel C_ξ calculé pour l'ensemble de la zone.

Lorsque le contraste observé atteint la valeur cible L_C pour un microphone, alors $C_\xi \approx 0$ donc $\xi \approx 1$. Cela signifie que la pression à ce microphone correspond à la valeur de pression cible dans la *Dark zone*.

Lorsque le contraste résiduel est négatif $C_\xi < 0$, le contraste cible n'est pas encore atteint, et on doit donc obtenir une amplitude de pression plus faible à ce microphone qu'elle ne l'est à l'itération en cours.

Lorsque le contraste résiduel est positif $C_\xi > 0$, la pression à ce point est trop faible. Cette dernière doit donc être augmentée pour correspondre au niveau sonore souhaité dans la *Dark zone*, de sorte à améliorer l'homogénéité de la *Dark zone*.

L'objectif est d'utiliser le champ de pression présent en *Dark zone*, induit par la *Bright zone* pour définir le champ cible de la *Dark zone*, et d'atténuer (ou d'amplifier) les valeurs individuelles de pression à chaque microphone de la *Dark zone* pour les faire correspondre au niveau sonore cible dans la *Dark zone* sur l'ensemble des microphones, améliorant ainsi l'homogénéité de la *Dark zone* [Roussel et Nicol, 2018b].

Pour tous les microphones, on définit le vecteur à composante réelle

$$\boldsymbol{\xi} = [\xi_1, \dots, \xi_m, \dots, \xi_M]^T. \quad (2.45)$$

Ce coefficient ayant une partie imaginaire nulle, il ne contraint pas la phase dans la *Dark zone*. L'information de phase est donnée par le rayonnement induit par la *Bright zone* et peut donc varier librement au cours des itérations dans la *Dark zone*. De plus, ce coefficient est calculé pour chaque itération et peut donc s'écrire $\boldsymbol{\xi}(n)$ en affichant explicitement sa dépendance temporelle.

Pour les méthodes itératives, le champ sonore en *Dark zone* $\mathbf{p}_{\text{TD}}(n)$ est défini comme suit [Roussel et Nicol, 2018b] :

- Pour la première itération, $\mathbf{p}_{\text{TD}}(0) = \mathbf{0}$ pour initialiser l'algorithme.
- Pour les itérations suivantes, on calcule le champ de pression $\mathbf{p}_{\text{D}}(n)$ reproduit aux positions des microphones virtuels de la *Dark zone*, correspondant au champ de pression dans la *Dark zone* induit par le rayonnement des haut-parleurs permettant de générer la *Bright zone* à l'itération n :

$$\mathbf{p}_{\text{D}}(n) = \mathbf{G}_{\text{D}}(n)\mathbf{q}(n). \quad (2.46)$$

On définit alors le nouveau champ de pression cible de la *Dark zone* pour l'itération $n+1$ comme étant ce champ atténué

$$\mathbf{p}_{\text{TD}}(n+1) = \boldsymbol{\xi}(n) \times \mathbf{p}_{\text{D}}(n), \quad (2.47)$$

où \times désigne une multiplication terme à terme.

2.5 Simulation numérique dans le domaine fréquentiel

Malgré le contexte variant dans le temps du système à étudier, c'est à dire les *Sound Zones* pour auditeurs mobiles, le choix est fait de réaliser les simulations dans le domaine fréquentiel, correspondant à un problème quasi-statique comme expliqué en §2.2.

2.5.1 Principe et méthode

La simulation est réalisée pour des sources monopolaires de débit en champ libre dans le domaine fréquentiel. La fonction de transfert entre le haut-parleur l et un point de l'espace \mathbf{x} , est donnée par :

$$G_{xl}(\omega, n) = \frac{j\rho\omega}{4\pi|\mathbf{x} - \mathbf{x}_l|} e^{-jk|\mathbf{x} - \mathbf{x}_l|}, \quad (2.48)$$

avec $\rho = 1.2 \text{ kg.m}^{-3}$ la masse volumique de l'air. Par extension, les fonctions de transfert entre un haut-parleur l et un microphone m sont définies par

$$G_{ml}(\omega, n) = \frac{j\rho\omega}{4\pi R(n)} e^{-jkR(n)}, \quad (2.49)$$

où $R(n) = \|x_m(n) - x_l\|$ correspond à la distance entre le haut-parleur l et le microphone m à l'itération n . La simulation est réalisée en 3D et observée en 2D. Les axes X et Y sont définis en **Figure 2.4**.

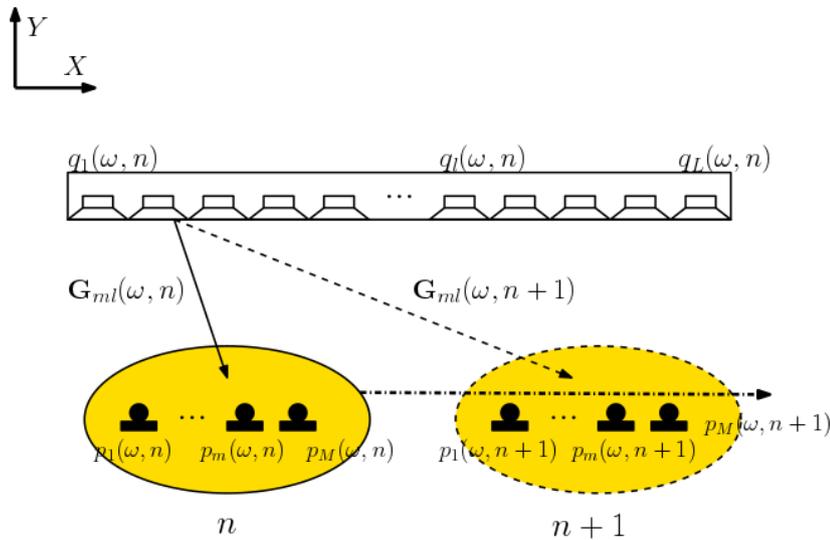


Figure 2.4 – Configuration de la simulation 2D en champ libre pour un ensemble de L_{HP} haut-parleurs fixes et un groupe de M microphones mobiles. La matrice de fonction de transfert est calculée à chaque itération n .

La configuration proposée pour l'évaluation des algorithmes est un réseau linéaire de $L_{HP} = 24$ haut-parleurs espacés de 0.09 m. Cette configuration est choisie pour deux raisons. Premièrement, le but final de l'implémentation est de permettre une réalisation dans un cadre domestique, c'est à dire avec une contrainte d'espace et un nombre de haut-parleurs restreint. Même si l'espace nécessaire pour ce réseau reste important, il ne nécessite pas d'avoir un ensemble de haut-parleurs autour de l'auditeur comme c'est généralement le cas pour les réseaux circulaires utilisés dans la littérature. De plus, ce type de réseau existe à Orange Labs, permettant à terme, la mise en œuvre expérimentale.

Les pressions cibles $\mathbf{p}_T(n)$ doivent être reproduites pour un groupe de $M = 400$ microphones arrangés en un réseau carré³ de 20×20 microphones espacés de 0.02 m. La position du centre du réseau de microphones est définie par la position de l'auditeur $\mathbf{x}_{aud}(n) = [X_{aud}, Y_{aud}]$ avec X_{aud} et Y_{aud} ses coordonnées respectivement selon l'axe X et Y . La pression cible $\mathbf{p}_{TB}(n)$ est une onde plane se propageant selon l'axe $-Y$ avec un niveau sonore de cible $L_{TB} = 76\text{dB SPL}$. Cette configuration est utilisée pour l'ensemble des simulations de ce document⁴. Un exemple de cette simulation pour la reproduction d'une onde plane à 1kHz en *Bright zone* est donné en **Figure 2.5**. Chaque haut-parleur est défini par un triangle noir. La zone de reproduction, définie par le réseau de microphone, est représenté par le carré pointillé blanc.

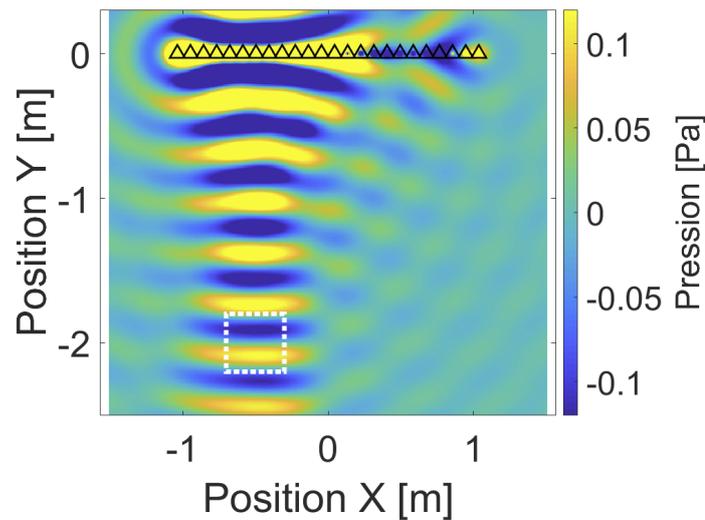


Figure 2.5 – Exemple de champ sonore pour la reproduction d'une onde plane à 1kHz en *Bright zone*. Les haut-parleurs sont représentés par les triangles noirs et le réseau de microphones par le carré pointillé blanc.

Notes sur les limitations de la simulation

Les simulations proposées approximent un réseau de haut-parleurs par un ensemble de sources monopolaires. En effet, le comportement d'un haut-parleur n'est monopolaire qu'en basse fréquence en raison de la surface rayonnante de sa membrane. Ce dernier devient directif au-delà d'une certaine fréquence. De plus, les haut-parleurs étant proches les uns des autres, il peut éventuellement exister un couplage au sein du réseau. Ces aspects ne sont pas pris en compte afin de simplifier la simulation mais permettraient d'obtenir un modèle plus réaliste d'un dispositif expérimental. Par ailleurs, il est à noter que les simulations sont réalisées en conditions anéchoïques et ne prennent donc pas en compte les effets liés à la salle. Cet aspect est discuté en partie 5.4. De plus le choix d'un réseau linéaire de haut-parleurs dans le cas d'une écoute en salle ne permet pas un bon contrôle des réflexions sur l'ensemble de la salle, ces dernières arrivant de toutes les directions.

3. Ce qui correspond à une fréquence d'aliasing d'environ 8.5 kHz pour une onde plane dans un des axes du réseau de microphones et 6 kHz dans le pire des cas, c'est à dire avec un angle de 45° .

4. Le niveau sonore choisi ici est considéré comme un niveau sonore confortable d'écoute [Francombe *et al.*, 2012, Francombe *et al.*, 2013a]

Conclusion

Les méthodes de résolution classiques pour les *Sound Zones* ne permettent pas une implémentation directe dans le cas d'auditeurs mobiles et nécessitent donc d'être modifiées. Le principal frein à leur utilisation est le coût de calcul engendré par la résolution et la régularisation du problème en temps réel. En s'appuyant sur la formulation *Pressure Matching* pour la création de champs sonores et de *Sound Zones*, nous avons développé un ensemble d'algorithmes adaptatifs pour la résolution du problème de *Sound Zones*. Ces algorithmes sont basés sur une résolution itérative du problème, en s'appuyant sur les résultats obtenus à l'itération précédente pour calculer les nouveaux poids des haut-parleurs. Quatre méthodes ont été mises en œuvre pour résoudre le problème : les filtres LMS, NLMS, APA et VLLMS. Le filtre LMS est le filtre le plus simple à obtenir et s'appuie sur la méthode de descente de gradient stochastique. Il dispose d'une faible complexité, adaptée à une utilisation en temps réel. L'inconvénient de ce filtre est la nécessité de réglage du pas μ_L qui détermine à la fois sa vitesse de convergence, son bruit de gradient ainsi que sa stabilité. Les filtres NLMS et APA permettent de pallier le problème du réglage du pas en utilisant les données d'entrée de l'itération précédente pour définir le pas de l'algorithme. Le filtre APA, à la différence du filtre NLMS utilise les K itérations précédentes et peut être vu comme une généralisation du filtre NLMS. Cependant, ces méthodes sont plus coûteuses que le filtre LMS. Le filtre NLMS est équivalent à la méthode Directe en termes de coût de calcul. La complexité algorithmique de la méthode APA est par contre plus élevée. Ces méthodes sont évaluées dans le Chapitre 3 où nous comparons également leurs performances avec celles de l'algorithme LMS qui sert donc de référence. Malgré leur caractère adaptatif, ces filtres ne permettent pas de répondre complètement au problème de *Sound Zones* en mouvement à cause de l'effet mémoire intrinsèque à ces filtres. Cet effet mémoire, analysé dans le Chapitre 3, engendre une traînée sur le passage de l'auditeur. Afin de réduire l'effet mémoire de l'algorithme LMS, nous proposons d'utiliser le filtre VLLMS qui réduit l'influence des itérations précédentes à l'aide d'un facteur d'oubli variable.

Par ailleurs, dans le cas du problème global de *Sound Zones* la spécification du champs dans la *Dark zone* peut se faire de différentes façons. Plusieurs solutions existent pour le champ sonore de la *Dark zone* et sont directement adaptables pour la formulation itérative de type LMS. De plus, l'utilisation de cette formulation itérative rend possible de nouvelles approches pour la définition du champ sonore en *Dark zone*. A la différence des méthodes de résolution Directe de la littérature, nous proposons de tirer parti de l'aspect itératif en utilisant le champ sonore induit par la *Bright zone* et déjà présent en *Dark zone* pour servir de champ cible en *Dark zone* à l'itération suivante. Ce champ cible en *Dark zone* est alors atténué pour correspondre à une valeur désirée tout en rendant possible un niveau sonore homogène dans la *Dark zone*.

Les Chapitres 3 et 4 s'intéressent à l'évaluation des nouvelles solutions que nous venons de décrire. Pour comparer les différents algorithmes, plusieurs critères d'évaluation ont été adaptés de la littérature pour correspondre au problème d'auditeurs mobiles. Un nouveau paramètre a également été proposé pour rendre compte de l'homogénéité du niveau sonore dans la zone, qui n'est pas prise en compte dans les algorithmes actuels. Enfin, la méthode de simulation utilisée dans la suite du document a été présentée. Elle s'appuie sur une formulation quasi-statique dans le domaine fréquentiel. Ce choix est motivé par l'hypothèse de faible variation du système entre deux itérations. C'est sur l'ensemble de ces formulations que s'appuie la suite du travail de thèse présenté. Le Chapitre 3 porte sur l'évaluation des méthodes LMS, NLMS, APA et VLLMS dans le

cas d'un unique auditeur mobile, mais nous avons montré que cette simplification ne restreint pas la généralité. Le Chapitre 4 quant à lui est consacré à l'évaluation de la méthode VLLMS dans le contexte de *Sound Zones* mobiles.

Principales avancées

- **Adaptation de la méthode PM** aux algorithmes **LMS, NLMS et APA**.
- **Proposition de la méthode VLLMS** pour **réduire l'effet mémoire**.
- **Adaptation** des méthodes PM et WLS à l'algorithme **VLLMS**.
- **Proposition** de deux méthodes de **Dark zone adaptatives**, les méthodes DZII et DZIG.
- **Modification** du calcul de l'effort au contexte d'auditeurs en mouvement.
- **Proposition** d'un paramètre d'évaluation **d'homogénéité de la zone**.

Chapitre 3

Reproduction sonore pour une seule zone en mouvement

Le problème de *Sound Zones* peut se décomposer en deux sous-problèmes : d'une part optimiser la reproduction d'un champ sonore dans une zone limitée de l'espace pour l'auditeur cible (*Bright zone*), et d'autre part minimiser ce même champ sonore dans une seconde zone (*Dark zone*), afin de réduire le champ parasite induit pour d'autres auditeurs. C'est au premier sous-problème que ce chapitre s'intéresse : la reproduction de champ sonore dans une zone limitée de l'espace dont la localisation peut changer en fonction du temps. La méthode de reproduction de champ sonore retenue dans ce chapitre est la méthode *Pressure Matching* (PM). Le chapitre précédent a présenté différentes stratégies de résolution du problème : d'une part, la méthode Directe issue de l'état de l'art, et d'autre part les méthodes de résolution adaptative appliquées à une reproduction sonore par la méthode PM qui constituent l'une des contributions de la présente thèse. La méthode LMS et ses améliorations Leaky LMS et VLLMS ainsi que les méthodes NLMS et APA sont ici comparées à la méthode Directe et ce pour plusieurs configurations. L'objectif principal est de choisir la méthode la plus appropriée à une utilisation pour des *Sound Zones* avec des auditeurs en mouvement (Voir Chapitre 4).

Dans un premier temps, les algorithmes sont évalués dans une configuration statique pour la reproduction d'une onde plane afin de déterminer leurs performances sans modification de l'objectif de l'algorithme. Les algorithmes NLMS, APA et la méthode Directe nécessitent le choix d'un paramètre de régularisation β , qui influence leurs performances et leur stabilité ainsi que leur comportement à long terme. Son effet est évalué dans ce chapitre. Même si les méthodes LMS, Leaky LMS et VLLMS ne disposent pas d'un paramètre de régularisation au même titre que les méthodes Directe, NLMS et APA, le comportement de ces filtres est également influencé par le paramètre de pas μ_L et le paramètre d'oubli $\gamma(n)$ dont l'effet est également examiné dans ce chapitre. L'ensemble des algorithmes est ensuite évalué dans le cas de déplacements simples selon deux directions perpendiculaires. L'effet mémoire de ces algorithmes est alors mis en évidence. Il se caractérise par un effet de trainage qui altère la qualité de reproduction hors de la zone de reproduction lorsque l'auditeur se déplace. Il importe donc de corriger cet effet pour améliorer la qualité de l'ensemble du champ sonore reproduit, ainsi que la réactivité de l'algorithme aux déplacements. Pour pallier ce problème, la méthode VLLMS proposée dans le chapitre précédent a pour but de diminuer cet effet mémoire en supprimant une partie des poids de l'itération précédente.

3.1 Évaluation des algorithmes en configuration statique

Dans le contexte de la reproduction d'un champ sonore, la méthode Pressure Matching (PM) permet d'obtenir les poids nécessaires à la reproduction de ce champ de manière optimale. Ces poids peuvent être obtenus par différentes méthodes de calcul, dont les méthodes adaptatives qui ont été retenues ici et qui consistent à approcher le résultat par itérations successives. Dans cette partie, la qualité de reproduction du champ sonore ainsi que l'effort de reproduction de chaque méthode sont comparés à la méthode Directe. Un intérêt particulier est porté au temps de convergence correspondant (temps que met l'algorithme à converger). Dans un premier temps (§3.1), les algorithmes sont comparés dans un contexte statique, d'une part afin d'analyser le comportement de chaque méthode en contexte stationnaire, d'autre part pour estimer le temps de convergence de chaque méthode. Dans une seconde partie (§3.2), on s'intéresse au comportement de ces méthodes dans le cas d'un déplacement de la zone.

3.1.1 Ajustement du paramètre de régularisation

Les méthodes Directe, NLMS et APA, nécessitent le calcul d'un paramètre de régularisation β pour assurer l'inversion des différentes matrices qui apparaissent dans les calculs (**Eq.** (2.3), (2.17) et (2.27)). La méthode LMS ne nécessitant pas d'inversion de matrice, elle n'est pas concernée par cette nécessité de régulariser le problème. Les effets du paramètre de régularisation β ainsi que le choix de ce dernier ont déjà été étudiés dans la littérature des *Sound Zones* pour la méthode Directe [Coleman *et al.*, 2014b]. S'il n'y a pas de consensus sur le choix de la méthode de sélection de ce paramètre qui influe grandement sur les performances (en particulier sur le contraste et l'effort de réseau). Ce paramètre est généralement choisi en fonction d'un critère arbitraire (minimisation de l'effort de réseau, maximisation du contraste, respect d'un ratio donné entre plus grande et plus petite valeur propre de la matrice $\mathbf{G}^H\mathbf{G}$, etc.) [Coleman *et al.*, 2014b]. L'ajustement de β nécessite généralement une optimisation complémentaire et ajoute donc du temps de calcul au problème. La méthode de sélection du paramètre β retenue ici pour la méthode Directe est celle proposée par [Coleman *et al.*, 2013]. Si cette méthode est appliquée à l'origine sur l'effort, elle est ici appliquée à l'erreur de reproduction. Le paramètre de régularisation β est déterminé selon la procédure suivante. β est d'abord augmenté jusqu'à ce que le conditionnement de la matrice à inverser $[\mathbf{G}^H(n)\mathbf{G}(n) + \beta(n)\mathbf{I}]$ soit inférieur à 10^{10} (ce qui permet de garantir l'inversion), puis si l'erreur de reproduction dépasse une certaine valeur d'erreur cible e_{rrc} , le paramètre est de nouveau augmenté afin de réduire l'erreur à une valeur inférieure ou égale à e_{rrc} .

Note sur la régularisation des méthodes NLMS et APA

Les méthodes NLMS et APA, présentées dans le chapitre 2, nécessitent la sélection du paramètre de régularisation β qui assure une inversion correcte de matrice respectivement dans les équations (2.17) et (2.27).

Pour la méthode NLMS (**Eq.** (2.18)), l'inversion est identique à celle réalisée par la méthode Directe (**Eq.** 2.3). La principale différence entre les deux méthodes est que pour la méthode Directe, la pression à obtenir est directement basée sur la pression cible \mathbf{p}_T , tandis que pour la méthode NLMS, la pression à obtenir est basée sur la différence

entre les pressions entre deux itérations successives (correspondant à l'erreur \mathbf{e}_{rr} en **Eq.** (2.18)). Le paramètre de régularisation calculé pour la méthode Directe est utilisé de façon identique pour la méthode NLMS. Les algorithmes des méthodes Directe et NLMS sont présentés en **Annexe A.3** et **Annexe A.5**.

Pour ce qui est de la méthode APA, la matrice à inverser est différente de celle des méthodes Directe et NLMS. Afin de disposer d'une méthode de régularisation équivalente pour les méthodes Directe, NLMS et APA (en vue de pouvoir les comparer), nous proposons de définir pour la méthode APA un problème équivalent à celui de la méthode Directe.

Pour obtenir ce problème APA équivalent, nous proposons de remplacer dans l'équation (2.27) le terme \mathbf{E}_{rr} par la pression cible \mathbf{p}_{T_F} ce qui donne :

$$\mathbf{q}_\psi(n) = 2 [\mathbf{G}_F^H(n)\mathbf{G}_F(n) + \beta\mathbf{I}]^{-1} \times [\mathbf{G}_F^H(n)\mathbf{p}_{T_F}(n)]. \quad (3.1)$$

De ce fait on définit un problème permettant d'obtenir des pseudo-poids \mathbf{q}_ψ . Ces paramètres sont qualifiés de pseudo-paramètres (pseudo-effort, pseudo-erreur, pseudo-niveau sonore, etc.) car ils correspondent aux résultats du pseudo-problème APA de l'équation (3.1). Ces pseudo-paramètres ne sont utilisés que lors de la régularisation de la méthode et sont notés avec l'indice ψ .

Lorsque le filtre APA d'ordre K est statique (l'auditeur ne bouge pas pendant au moins K itérations), le problème à résoudre est quasiment identique à celui du filtre NLMS ou de la méthode Directe, à la seule différence que l'on prend en compte K fois le même problème aux mêmes positions. Lorsque le filtre APA est utilisé dans le cas dynamique, il permet de reproduire le champ sonore sur l'ensemble des positions occupées par le réseau de microphones aux K itérations précédentes.

Pour obtenir le paramètre de régularisation β , la méthode est donc équivalente aux méthodes Directes et NLMS, mais avec la matrice $[\mathbf{G}_F^H(n)\mathbf{G}_F(n) + \beta\mathbf{I}]$. On augmente β afin que le conditionnement de la matrice soit inférieur à 10^{10} , puis on ajuste le paramètre de régularisation en vue de rapprocher la valeur du pseudo-paramètre d'une valeur cible donnée. Les pseudo-paramètres sont ensuite calculés à l'aide des pseudo-poids de l'équation (3.1). Une fois la valeur de β obtenue, on l'utilise pour régulariser l'inversion de la matrice de la méthode APA.

3.1.2 Conditions de comparaison

Dans ce qui suit, la comparaison des méthodes Directe, LMS, NLMS et APA est réalisée en imposant à chaque méthode d'atteindre une erreur de reproduction équivalente d'un même champ sonore produit avec la méthode PM. Cette comparaison permet d'illustrer les différences entre ces méthodes pour une précision de reproduction du champ sonore équivalente, ceci sans contraindre plusieurs paramètres. Par ailleurs, l'erreur de reproduction est ici le seul paramètre vraiment pertinent à fixer car il n'aurait pas été possible de réaliser cette comparaison à effort équivalent, les algorithmes LMS, NLMS et APA ne permettant pas d'être contrôlés en effort (Voir §2.2).

L'erreur cible de reproduction est fixée arbitrairement à $E_C = -65dB$ ¹. Pour ce qui est des algorithmes itératifs LMS, NLMS et APA, ils sont lancés sur 200 itérations. Lorsque le critère d'erreur cible E_C est égal à la valeur du paramètre (ou du pseudo-paramètre) de la méthode, on calcule le niveau sonore, l'erreur de reproduction, l'effort

1. Le choix de cette valeur correspond à une valeur que tous les algorithmes sont capables d'atteindre dans les conditions de l'étude. Cette valeur est déterminée empiriquement.

ainsi que le temps de convergence de chaque algorithme sur la gamme de fréquences [1 Hz-8 kHz].

Par ailleurs, la méthode LMS nécessite le réglage du pas μ_L . Afin de garantir la stabilité du filtre, le pas μ_L est choisi de sorte à respecter le critère de stabilité de l'équation (2.10). Deux valeurs de pas sont proposées et notées μ_{L1} et μ_{L2} . Pour ce faire, μ_{Lmax} est arbitrairement divisé par 4 et par 10. On a donc :

$$\mu_{L1} = \frac{\mu_{Lmax}}{4} \quad \text{et} \quad \mu_{L2} = \frac{\mu_{Lmax}}{10}. \quad (3.2)$$

La valeur de μ_{max} définissant la stabilité du filtre est calculée grâce à l'équation (2.10) où κ_{max} correspond à la valeur propre maximale de la matrice de corrélation spatiale $\mathbf{G}^H(n)\mathbf{G}(n)$ pour la position initiale.

La *Bright zone* est positionnée en $X_{aud} = -0.5$ et $Y_{aud} = -2$ (**Figure 2.4**). Le choix de cette position répond à deux critères. La position selon l'axe X correspond à la position de départ de l'ensemble des évaluations dans ce chapitre. La position selon l'axe Y correspond quant à elle à la position la plus éloignée de toutes les positions utilisées dans la suite du chapitre. L'intérêt de choisir la zone la plus éloignée du réseau de haut parleur est qu'elle correspond à la zone dont la convergence sera a priori la plus lente dans le cas d'un paramètre μ_L fixe², en particulier pour le filtre LMS comme précisé au §2.2.2. Le choix du paramètre μ_L en fonction de la position d'un auditeur peut faire l'objet de nombreuses approches et est discuté en Chapitre 5.

3.1.3 Temps de convergence

La **Figure 3.1 a)** reproduit le temps de convergence d'une part, de la méthode LMS pour deux pas différents μ_{L1} et μ_{L2} , et d'autre part des méthodes NLMS et APA, pour l'ensemble des fréquences allant de 1Hz à 8kHz. Le résultat de la méthode Directe n'est pas représenté car il ne nécessite aucune itération.

Les méthodes NLMS et APA convergent beaucoup plus rapidement que la méthode LMS. Cela est lié au choix du pas des algorithmes. Les méthodes NLMS et APA étant basées sur un pas optimal pour chaque itération ont donc un avantage sur la méthode LMS qui impose un pas fixe n'évoluant pas au cours du temps. Cependant, le temps de convergence de la méthode LMS peut être amélioré en augmentant la valeur de son pas μ_L , ceci sans affecter significativement ses performances en termes d'effort (**Figure 3.2**) et de niveau sonore (**Figure 3.5**). En effet, le filtre LMS utilisant μ_{L1} offre un temps de convergence moins élevé tout en gardant un comportement similaire sur l'ensemble de la gamme de fréquences.

2. Dans le cas statique, le calcul de μ_L est réalisé à la position d'intérêt, il n'est pas exact de considérer que cette position est la plus lente à converger.

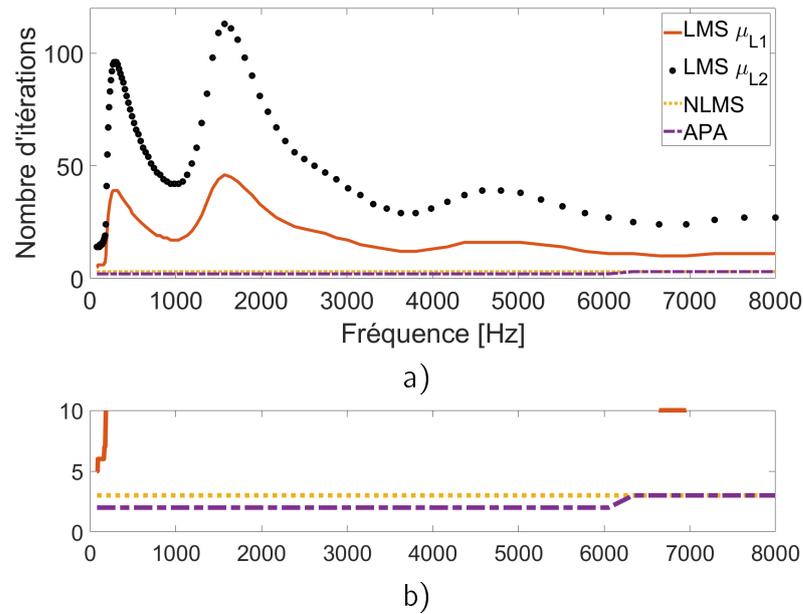


Figure 3.1 – Temps de convergence des algorithmes LMS, NLMS, et APA pour une erreur cible $E_C = -65\text{dB}$. La figure b) représente un zoom pour observer le comportement des méthode NLMS et APA.

Par ailleurs, la convergence du filtre LMS est variable en fonction de la fréquence. On observe deux zones de convergence plus lente. La seconde zone se situe autour de la fréquence d'aliasing spatial du réseau de haut-parleurs qui vaut 1715Hz et la première zone se situe à la moitié de celle-ci. Cependant, la cause de cette zone de convergence lente à ces fréquences particulières n'apparaît pas clairement sur l'observation des champs sonores (**Figure 3.3**). De plus, cette variation n'est pas visible sur le zoom en **Figure 3.1 b)** dans le cas des méthodes NLMS et APA. Cette observation ne nous permet pas de déduire un lien direct entre ces fréquences particulières et la convergence des algorithmes adaptatifs.

Enfin, la convergence du filtre APA est obtenue en une seule itération sur une grande partie de la gamme de fréquences, ce qui la rend plus rapide que la méthode NLMS qui converge en deux itérations. Cette différence n'est que peu significative dans ce cas, et ne lui confère qu'un avantage négligeable, en comparaison de son coût de calcul supérieur.

3.1.4 Effort

La **Figure 3.2** présente l'effort de réseau en fonction de la fréquence pour les différentes méthodes. L'ensemble des méthodes montrent une augmentation de l'effort avec la fréquence, à l'exception de la méthode APA. La méthode APA présente un effort plus constant, mais plus élevé sur toute la gamme de fréquences. La **Figure 3.3** illustre, après convergence, l'allure du champ sonore dans l'espace considéré. On observe que les méthodes LMS et Directe sont caractérisées par un nombre élevé de lobes secondaires présents sur les bords du réseau de haut-parleurs, ce qui explique l'augmentation de l'effort. L'énergie acoustique totale délivrée est donc plus importante avec l'augmentation des lobes secondaires pour une même erreur de reproduction. Les champs sonores reproduits par les méthodes NLMS et APA sont visibles en **Annexe E**.

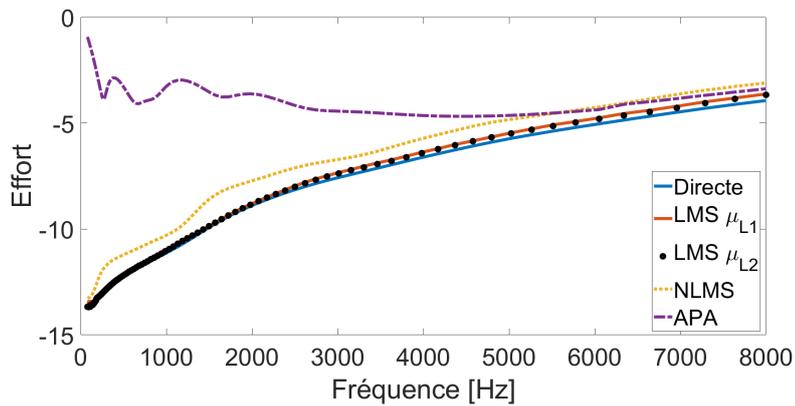


Figure 3.2 – Effort de réseau après convergence des algorithmes LMS, NLMS, et APA en comparaison avec la méthode Directe pour une erreur cible $E_C = -65dB$.

3.1.5 Champ sonore reproduit

L'impact de l'effort sur le champ sonore reproduit est représenté en **Figure 3.4** pour l'ensemble des méthodes. La solution qui reproduit le champ sonore le plus proche de la méthode de référence (Directe) est l'algorithme LMS. Pour la méthode NLMS, l'énergie acoustique totale délivrée est plus importante que pour la méthode Directe, ce qui est cohérent avec un effort de réseau plus élevé. La méthode APA délivre quant à elle beaucoup d'énergie acoustique en dehors de la zone de reproduction, énergie qui ne sert pas directement à la reproduction du champ sonore cible correspondant également à l'effort de réseau élevé décrit en **Figure 3.2**. Ce comportement de la méthode PM trahit en général une régularisation insuffisante du problème [Coleman, 2014]. La question de la régularisation est développée en §3.1.7.

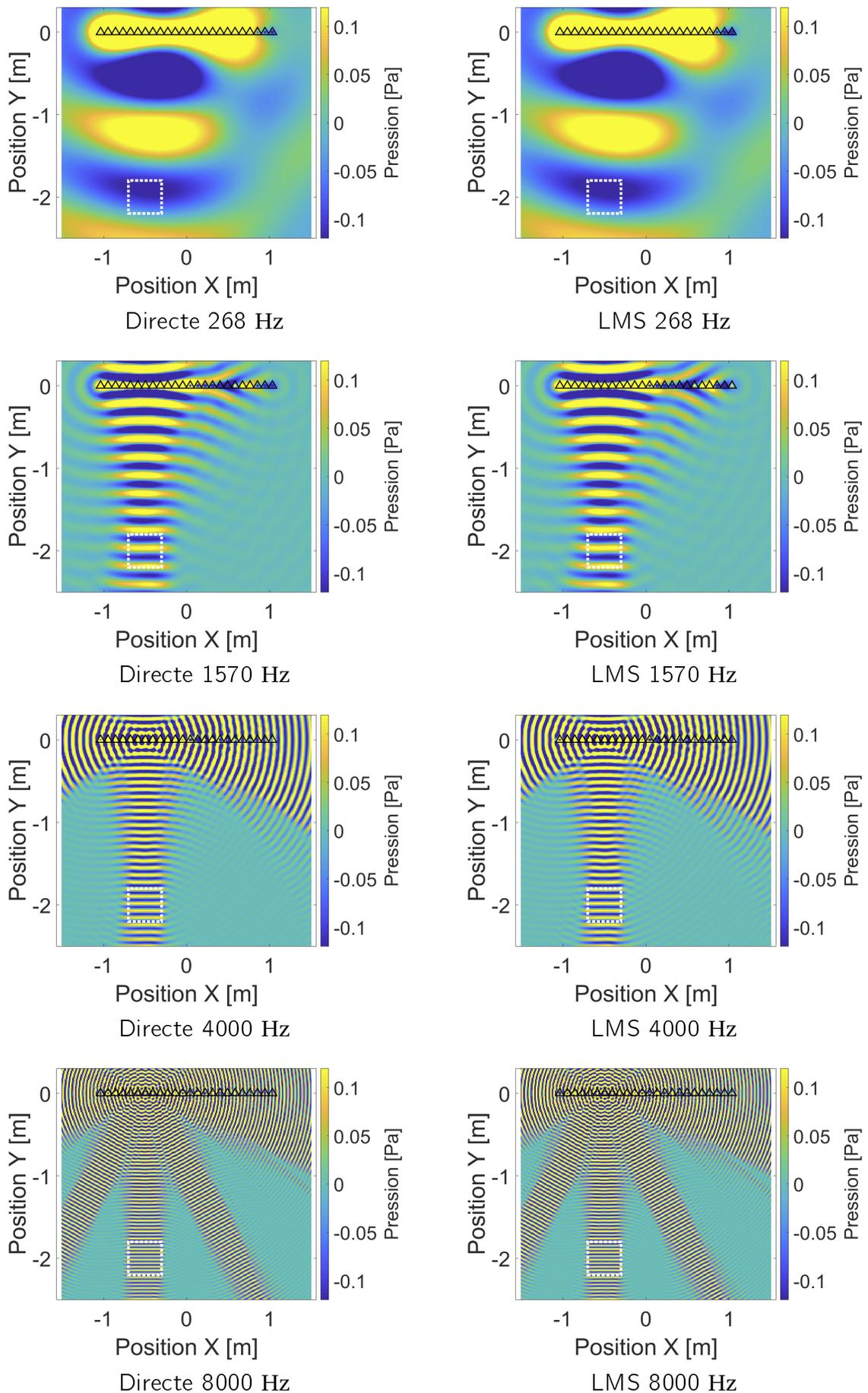


Figure 3.3 – Champ sonore après la convergence pour différentes fréquences pour les méthodes Directe (gauche) et LMS (droite). La position de la *Bright zone* est définie par le carré pointillé blanc.

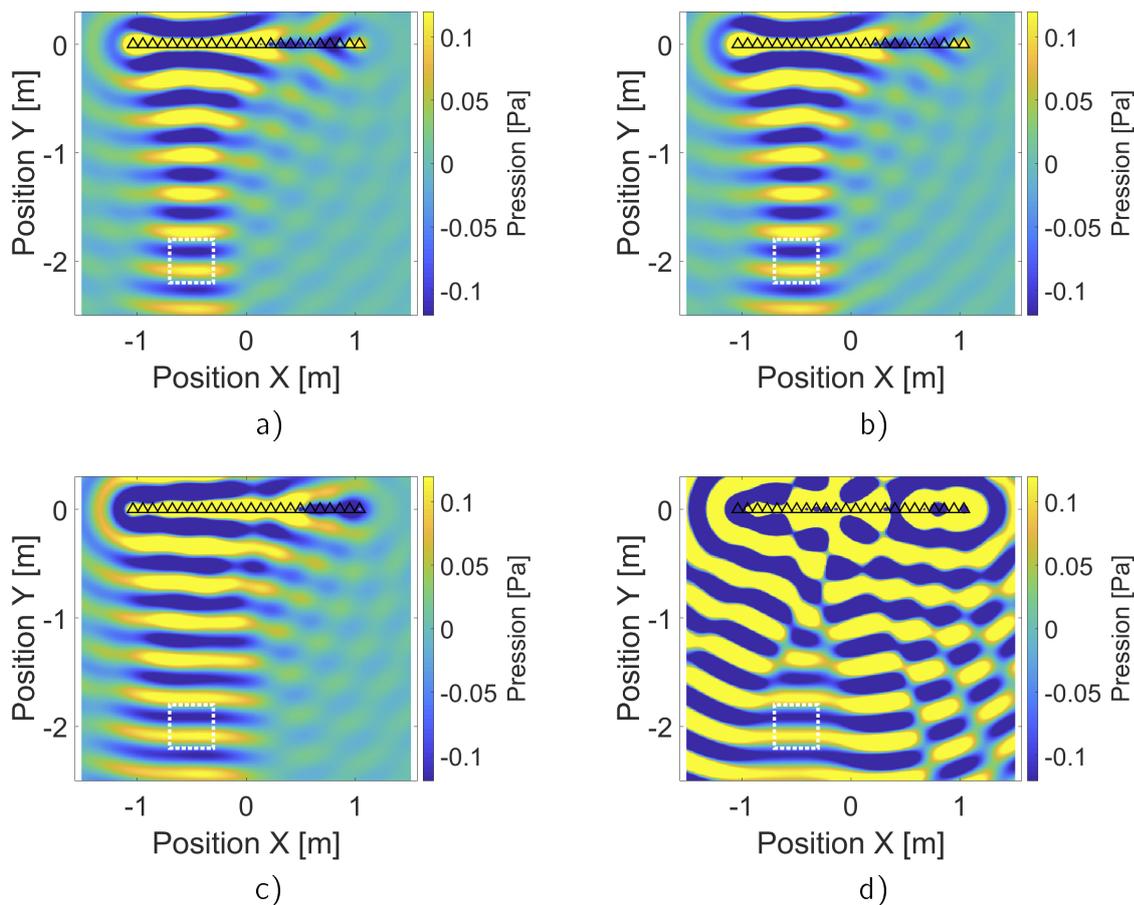


Figure 3.4 – Champ sonore à 1 kHz après la convergence pour les méthodes Directe a), LMS b), NLMS c) et APA d). La position de la *Bright zone* est définie par le carré pointillé blanc.

3.1.6 Niveau sonore

La **Figure 3.5** représente le niveau sonore reproduit par chaque algorithme en fonction de la fréquence. Le niveau sonore cible étant de 76dB SPL, et la dynamique en ordonnées étant très réduite, la différence entre les méthodes peut être considérée comme négligeable. Toutes les méthodes parviennent donc à accéder au niveau sonore cible.

Ainsi, toutes les méthodes adaptatives proposées reproduisent donc le champ sonore cible, dans une zone donnée de l'espace, avec un nombre raisonnable d'itérations, tout en respectant une contrainte d'erreur. La méthode LMS nécessite un nombre plus élevé d'itérations afin d'atteindre le même niveau d'erreur que les autres algorithmes, mais le champ sonore reproduit est plus proche de celui obtenu par la méthode Directe. Par ailleurs, les méthodes NLMS et APA présentent un effort de réseau plus élevé qui se caractérise sur le champ sonore par une énergie acoustique plus importante en dehors de la zone de reproduction. Cette énergie est inutile à la reproduction du champ sonore dans la *Bright zone* et ne se retrouve pas autant dans le cas des méthodes Directe et LMS.

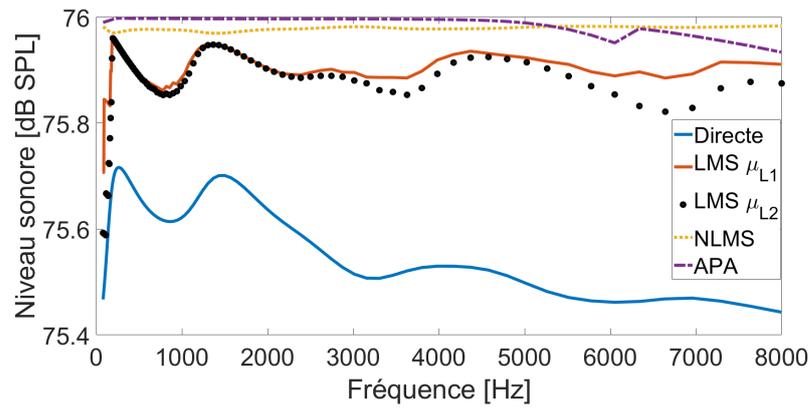


Figure 3.5 – Niveau sonore après la convergence des algorithmes LMS, NLMS, et APA en comparaison à la méthode Directe pour une erreur cible $E_C = -65dB$.

3.1.7 Effets de la régularisation

Le rôle de la régularisation est déterminant dans le fonctionnement des méthodes nécessitant l'inversion d'une matrice car il permet d'obtenir une solution numérique stable. De manière générale, le rôle de ce paramètre est double. D'une part, il améliore le conditionnement des matrices lors de leur inversion dans les méthodes Directe, NLMS et APA. D'autre part il évite la divergence du filtre en limitant l'énergie à reproduire par le réseau de haut-parleurs (effort de réseau). Comme précisé dans la partie 2.2, les méthodes APA et NLMS ne sont pas limitées en énergie par le paramètre de régularisation β . Celui-ci n'a une influence que sur le temps de convergence ainsi que la stabilité du filtre. Pour les filtres NLMS et APA, la régularisation ne permet donc que de s'approcher d'un temps optimal de convergence. On s'intéresse ici aux conséquences de la régularisation sur les différentes méthodes, et sur le comportement "à long terme" des algorithmes, puisque ceux-ci ne sont pas limités en effort. Comme précisé dans la partie 2.2, puisque le filtre LMS ne nécessite pas de paramètre de régularisation, il n'est pas inclus dans cette étude.

L'évaluation de la régularisation des méthodes NLMS et APA est réalisée à la position fixe $X_{aud} = -0.5\mathbf{m}$ et $Y_{aud} = -1\mathbf{m}$, pour une durée de 1000 itérations, à la fréquence de 1kHz. Les résultats pour plusieurs paramètres de régularisation sont comparés : absence de régularisation ($\beta = 0$), peu régularisé ($\beta = 10^3$), régularisation quasi optimale ($\beta = 10^6$) et problème fortement régularisé ($\beta = 10^9$). Les valeurs de régularisation ont été préalablement classées comme sous-régularisée, quasi-optimale et fortement régularisée selon les critères sur le champ sonore reproduit, définis par [Coleman, 2014] pour la méthode PM en résolution Directe. Les **Figures** 3.6, 3.7 et 3.8 représentent respectivement le niveau sonore, l'erreur de reproduction et l'effort de réseau.

En l'absence de régularisation, on observe une augmentation rapide du niveau sonore ainsi que de l'erreur de reproduction, aussi bien pour le filtre NLMS que pour le filtre APA. Cette divergence correspond à une instabilité des filtres NLMS et APA. Dans ce cas, le pas entre deux itérations est trop grand car le résultat de l'inversion non régularisée de matrice génère une très grande amplitude. Ce problème transparaît également sur la valeur très importante de l'effort obtenue par la méthode Directe (**Figure** 3.8 a)) ce qui rend impossible l'utilisation de cette solution en raison des niveaux sonores élevés demandés à chaque haut-parleur.

Dans le cas d'une faible régularisation, $\beta = 10^3$, le niveau sonore (**Figure** 3.6 b))

ainsi que l'erreur (**Figure 3.7 b**)) sont stables et convergent rapidement vers le niveau sonore cible tout en assurant une erreur de reproduction faible. Cependant, l'effort de réseau reste élevé et rend difficile l'utilisation dans un cas réel, car l'énergie acoustique rayonnée dans l'environnement est trop élevée.

La régularisation quasi optimale $\beta = 10^6$ donne le niveau sonore souhaité sur l'ensemble des itérations (**Figure 3.6 c**)). Dans le même temps, l'erreur de reproduction reste faible (**Figure 3.7 c**)). En comparaison des résultats pour le cas faiblement régularisé ($\beta = 10^3$) les résultats sont similaires pour ces deux critères d'évaluation. Pour ce qui est de l'effort de réseau (**Figure 3.8**), la valeur d'effort à la première itération de l'algorithme reste relativement faible. Cependant, l'effort des algorithmes NLMS et APA n'étant pas limité, il continue à augmenter, ce qui peut poser problème lors d'une implémentation.

Le cas fortement régularisé $\beta = 10^9$, présente une convergence lente du niveau sonore (**Figure 3.6 d**)) ainsi que de l'erreur de reproduction (**Figure 3.7 d**)) mais les algorithmes finissent par atteindre la valeur cible et se stabiliser. De la même façon que dans le cas de la régularisation quasi-optimale, les valeurs d'effort de réseau sont raisonnables mais l'effort augmente au fur et à mesure des itérations, et peut atteindre des valeurs incompatibles avec une utilisation en conditions réelles, menant par exemple à la destruction des haut-parleurs.

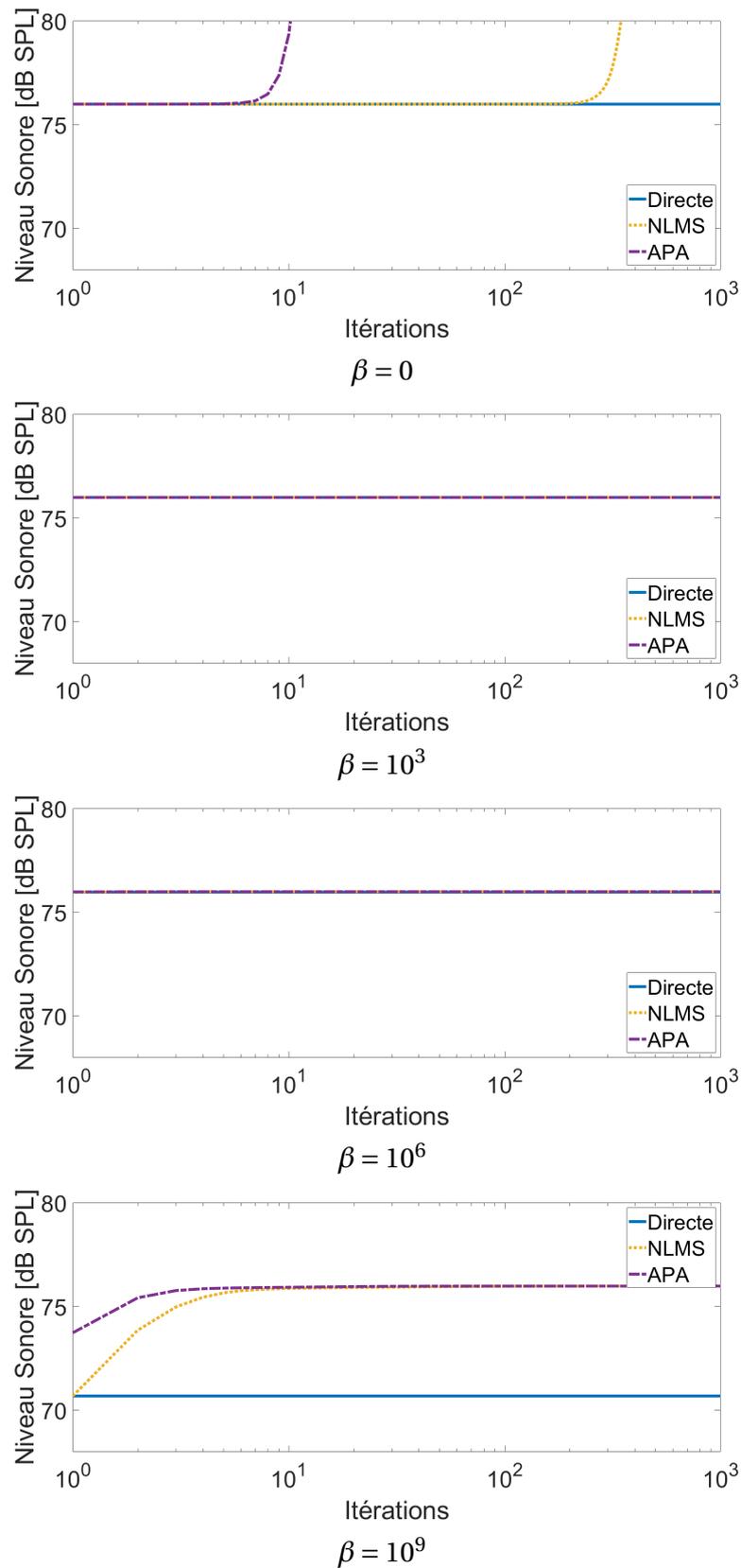


Figure 3.6 – Niveau sonore dans la *Bright zone* pour différents paramètres de régularisation β pour un auditeur fixe pour les méthodes Directe, NLMS et APA sur 1000 itérations. (représentation est logarithmique en abscisse)

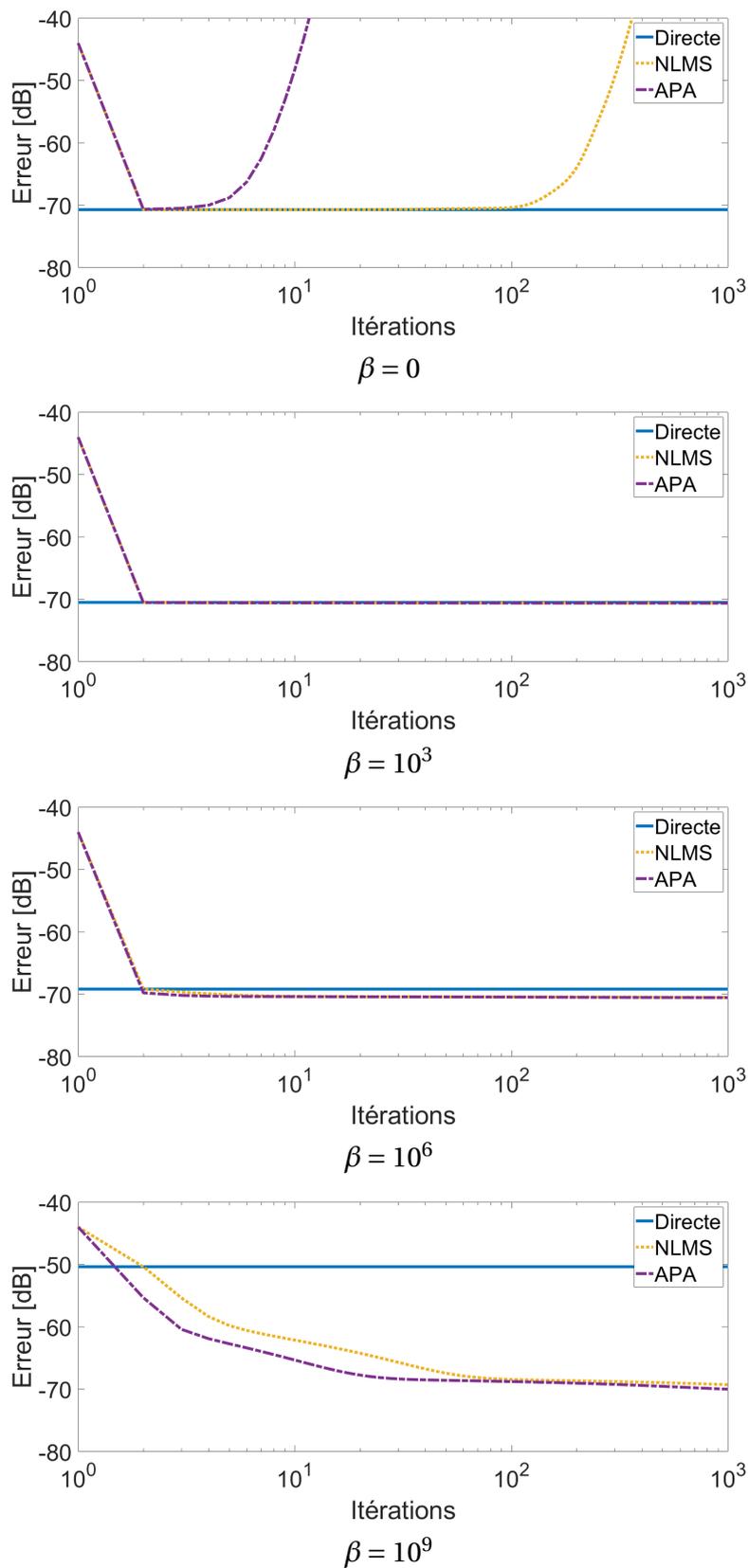


Figure 3.7 – Erreur de reproduction dans la *Bright zone* pour différents paramètres de régularisation β pour un auditeur fixe pour les méthodes Directe, NLMS et APA sur 1000 itérations. (représentation est logarithmique en abscisse)

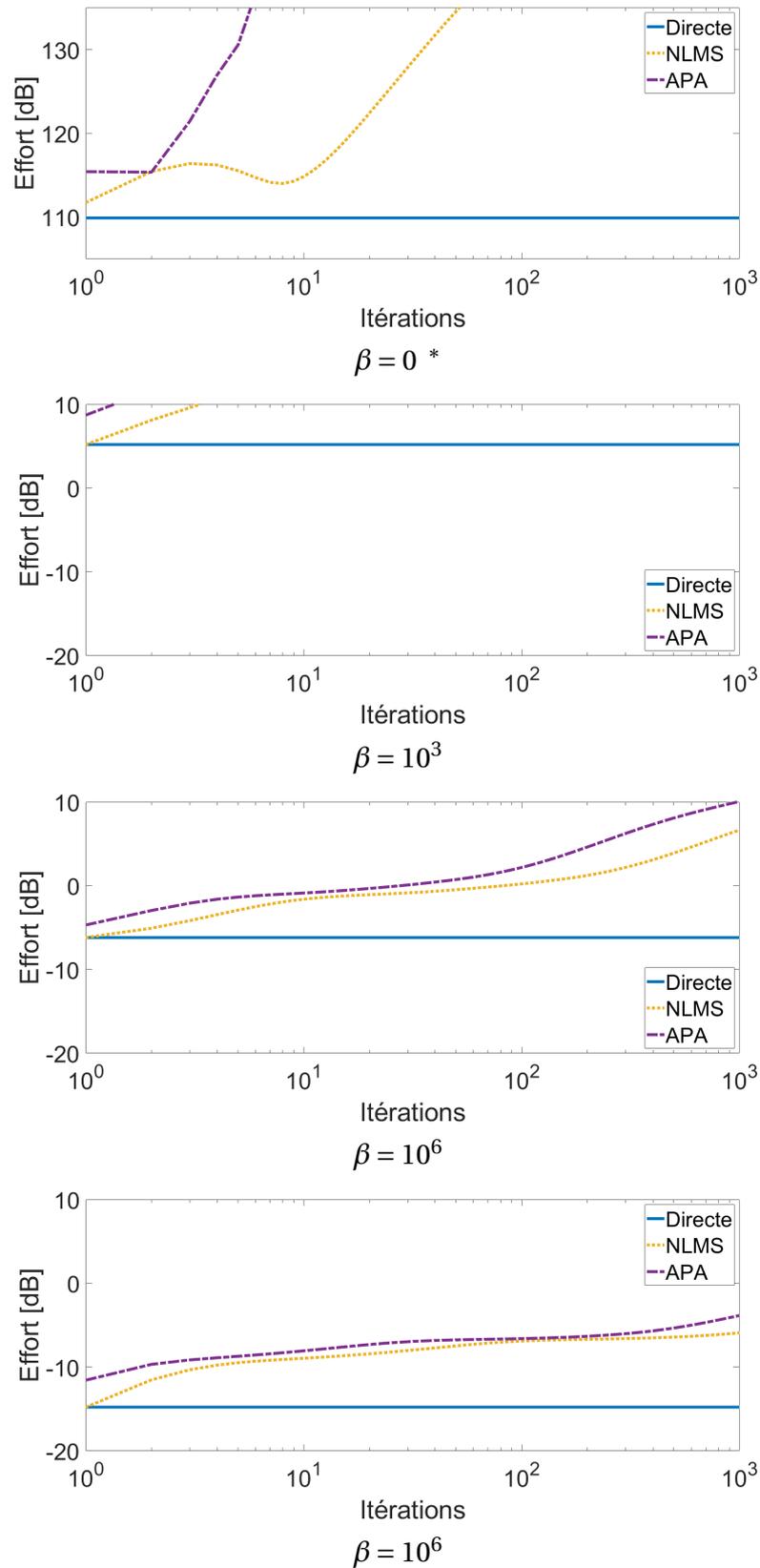


Figure 3.8 – Effort de Réseau la *Bright zone* pour différents paramètres de régularisation β pour un auditeur fixe pour les méthodes Directe, NLMS et APA sur 1000 itérations. (représentation est logarithmique en abscisse). * Pour $\beta = 0$, l'axe des ordonnées est translaté, dans un souci de meilleur lecture de l'ensemble des figures.

La régularisation des méthodes NLMS et APA est indispensable pour garantir leur stabilité ainsi que la possibilité de les exploiter avec un faible nombre d'itérations. Pour un paramètre de régularisation β élevé, on observe une réduction du temps de convergence des algorithmes. Il est par ailleurs à noter que l'effort de la méthode NLMS reste inférieur à celui de la méthode APA à niveau de régularisation identique. Même les valeurs quasi-optimales et fortement régularisées du paramètre β ne permettent pas de limiter l'effort de réseau, ce qui confirme l'hypothèse selon laquelle les algorithmes ne sont pas limités en effort par la régularisation. Dans la suite du document, les méthodes NLMS et APA sont régularisées par la procédure décrite en §3.1.1 à l'aide de leurs pseudo-paramètres. Cependant, l'effort de réseau ne pouvant pas être directement maîtrisé, on ne cherchera pas une régularisation optimale des méthodes NLMS et APA, ce d'autant que leurs performances n'en sont pas affectées de façon significative. De plus, il apparaît qu'il est particulièrement difficile d'obtenir une régularisation optimale à l'aide de la méthode des pseudo-paramètres pour la méthode APA.

Comportement à long terme des méthodes LMS, Leaky LMS et VLLMS

Dans cette partie, le filtre LMS ainsi que la variante Leaky LMS sont évalués dans leur comportement à long terme dans les mêmes conditions que les algorithmes NLMS et APA. L'algorithme VLLMS n'est pas présenté dans cette partie puisqu'il est la combinaison du filtre LMS et du filtre Leaky LMS. Par construction, il alterne entre le fonctionnement de l'un ou de l'autre des algorithmes en fonction des besoins. De plus, l'évaluation n'étant réalisée que dans le cas statique, le filtre VLLMS se comporte comme le filtre LMS, puisque l'augmentation du facteur d'oubli, et donc sa capacité à se comporter comme un filtre Leaky LMS est directement reliée au déplacement de l'auditeur dans notre définition (Voir §2.2.4).

Pour mémoire, les méthodes LMS et Leaky LMS ne nécessitent pas de terme de régularisation (pas d'inversion de matrice). Cependant, la méthode LMS n'étant pas limitée en effort, il est donc nécessaire de vérifier si son comportement tend à faire augmenter l'effort, comme observé pour les filtres NLMS et APA (§3.1.7). La méthode Leaky LMS, quant à elle, a pour but de limiter l'effort de réseau en oubliant une partie des poids calculés à l'itération précédente.

Les conditions d'évaluation sont identiques à l'évaluation des méthodes NLMS et APA : 1000 itérations pour un auditeur positionné en $X_{aud} = -0.5\mathbf{m}$ et $Y_{aud} = -1\mathbf{m}$, pour une fréquence de 1kHz. Le facteur d'oubli est arbitrairement fixé à la valeur $\gamma_{max} = 0.05$, ce qui correspond à un oubli de 5% des poids de l'itération précédente.

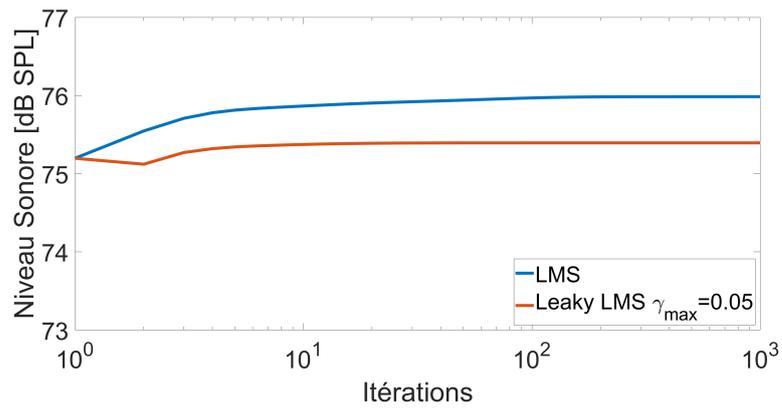


Figure 3.9 – Niveau sonore dans la *Bright zone* pour les méthodes LMS et Leaky LMS avec $\gamma_{max} = 0.05$ pour 1000 itérations.

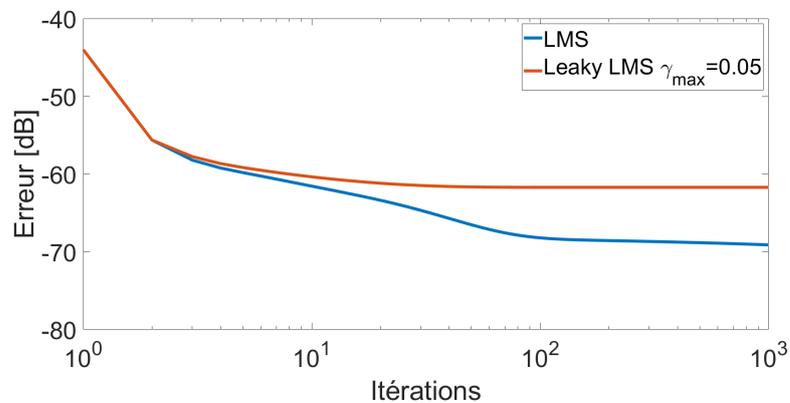


Figure 3.10 – Erreur de reproduction dans la *Bright zone* pour les méthodes LMS et Leaky LMS avec $\gamma_{max} = 0.05$ pour 1000 itérations.

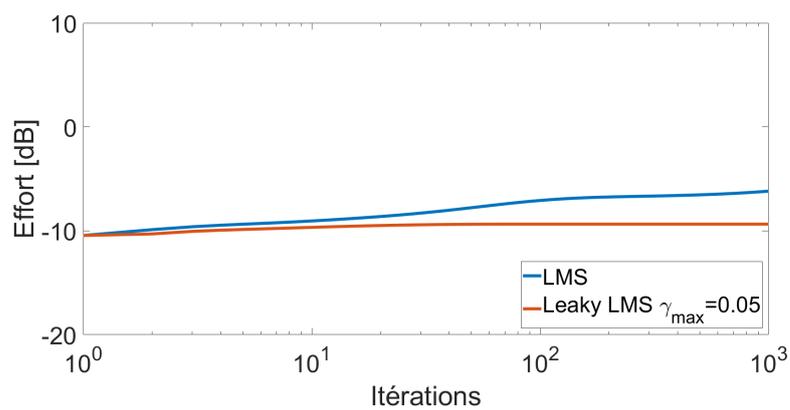


Figure 3.11 – Effort de réseau dans la *Bright zone* pour les méthodes LMS et Leaky LMS avec $\gamma_{max} = 0.05$ pour 1000 itérations.

La figure 3.9 présente le niveau sonore pour la méthode LMS et la méthode Leaky LMS. On observe que la méthode LMS atteint rapidement le niveau sonore cible et que l'erreur de reproduction peut-être qualifiée de satisfaisante (**Figure 3.10**). Pour ce qui est du filtre Leaky LMS, l'erreur de reproduction est plus élevée ce qui s'explique en partie par le fait que le niveau sonore cible n'est pas atteint. La méthode Leaky LMS est donc biaisée. Cependant, malgré un niveau sonore ne correspondant pas à la cible et une erreur de reproduction plus élevée, cette solution permet de réguler l'effort de réseau, contrairement à la méthode LMS dont l'effort continue d'augmenter. Le facteur d'oubli permet donc de limiter l'effort de réseau, mais au prix d'un biais sur le champ sonore reproduit. C'est justement l'intérêt de l'utilisation de la méthode VLLMS qui permet de passer du mode LMS non biaisé à la méthode Leaky LMS dans le cas où une limitation de l'effort devient nécessaire. Cette méthode permet également de limiter l'effet mémoire qui est discuté en partie 3.3.

3.2 Déplacement d'un auditeur unique

Après avoir évalué dans la section précédente les méthodes itératives dans le cas statique, nous nous intéressons au cas d'un auditeur unique en mouvement. Lorsque l'auditeur se déplace, les fonctions de transfert entre les haut-parleurs et les microphones varient au cours des itérations. Le problème est donc différent entre deux itérations et les algorithmes doivent donc s'adapter à ce changement. Les trajectoires proposées pour l'ensemble des évaluations, sont définies sur $n = 500$ itérations comme suit :

- les poids $\mathbf{q}(0)$ sont initialisés à $\mathbf{0}$,
- les 200 premières itérations servent à stabiliser le filtre,
- les 100 suivantes réalisent un déplacement de 1 cm entre chaque itération,
- les 200 dernières permettent d'observer la réaction du filtre lors de l'arrêt du déplacement.

Le pas de déplacement entre deux itérations est choisi de sorte à considérer un déplacement lent. Cette vitesse est exprimée par la suite en m/itération, et non en m/s^{-1} pour s'affranchir du temps de calcul qui peut varier entre les différents algorithmes, chaque algorithme ayant un coût de calcul différent.

Une comparaison des algorithmes est proposée pour un déplacement selon l'axe Y (**Figure 3.12**) dans un premier temps, puis selon l'axe X (**Figure 3.20**) dans un second temps, en étudiant l'effet mémoire avec un intérêt particulier porté au filtre LMS. Dans une seconde étape, les variantes Leaky LMS et VLLMS sont testées en vue de réduire l'effet mémoire.

Sur l'ensemble de l'étude, pour garantir la stabilité du filtre et bénéficier d'une convergence optimale, le pas μ_{Lmax} est calculé à chaque itération comme défini en équation (2.10). Cela ne contribue pas à réduire le coût de calcul de la méthode LMS et d'autres solutions seraient plus pertinentes de ce point de vue. Cet aspect est discuté dans le Chapitre 5. Le pas choisi pour le filtre LMS est $\mu_{L2} = \mu_{max}/10$, actualisé à chaque itération. Les filtres NLMS et APA sont régularisés par la méthode décrite en début de chapitre [Coleman *et al.*, 2013], avec pour objectif d'effort de réseau des pseudo-paramètres $E_{\psi} < E_C$ avec $E_C = -5dB$, ce qui permet de réduire l'effort de réseau final des méthodes. Toutes les évaluations sont réalisées à la fréquence 1kHz.

3.2.1 Déplacement selon l'axe Y

La première partie de l'étude porte sur la reproduction d'un champ sonore pour un auditeur mobile se déplaçant selon les Y négatifs, ce qui signifie que l'auditeur s'éloigne du réseau de haut-parleurs. Les algorithmes LMS, NLMS et APA sont comparés pour la trajectoire définie en **Figure 3.12**.

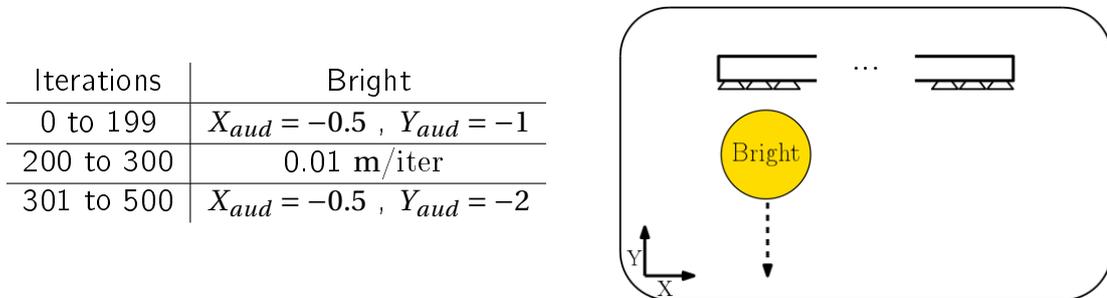


Figure 3.12 – Déplacement de l'auditeur selon l'axe $-Y$. Le déplacement s'effectue entre la 200^{ième} et la 300^{ième} itération.

La **Figure 3.13** présente le niveau sonore pour les différentes méthodes LMS, NLMS et APA en comparaison de la méthode Directe (qui sert de référence). Les lignes pointillées verticales rouges délimitent les itérations durant lesquelles l'auditeur se déplace. Dans le cas du filtre LMS, on remarque une importante diminution du niveau sonore pendant le déplacement. Cette variation est aussi présente dans le cas des méthodes NLMS et APA mais dans une moindre mesure. Cette diminution du niveau sonore s'accompagne d'une importante augmentation de l'erreur de reproduction lors du déplacement (**Figure 3.14**). Cette augmentation brutale de l'erreur de reproduction s'explique par le fait que la zone de reproduction définie par le réseau de microphones a une partie commune entre deux itérations (**Figure 3.15**). En effet, la zone de reproduction est de dimension $0.4\text{m} \times 0.4\text{m}$ alors que le déplacement entre deux itérations est de 0.01m . Comme illustré en **Figure 3.15**, il existe trois zones distinctes lors du déplacement de l'auditeur entre l'itération n représentée en trait plein et l'itération $n+1$ représentée en pointillés. La zone A est une zone abandonnée à l'itération $n+1$, elle ne fait plus partie du problème et est laissée sans contrainte comme le reste du champ sonore. La zone B est la zone commune à l'itération n et l'itération $n+1$, elle constitue une partie de l'erreur liée au mouvement. En effet, il s'agit d'une zone de transition entre les cibles de l'itération n et $n+1$. La reproduction y est dégradée par la nouvelle cible. De plus cette variation du champ sonore cible \mathbf{p}_T entre deux itérations dépend également de la direction de propagation de l'onde. L'onde se propage selon l'axe $-Y$. Lorsque l'on se déplace selon le même axe, aucune des pressions microphoniques à reproduire \mathbf{p}_T ne sera identique au champ déjà présent dans la zone. La différence entre deux itérations est maximisée par les variations propres à la direction de propagation de l'onde. La zone C représente la nouvelle zone qui n'est contrôlée qu'à l'itération $n+1$. Cette partie n'étant pas contrainte à l'itération précédente, le champ sonore présent dans cette zone est donc soit nul, soit un champ sonore quelconque induit par les poids des itérations précédentes. L'impact du déplacement sur le niveau sonore et sur l'erreur de reproduction est moindre dans le cas des filtres NLMS et APA car ceux-ci disposent d'un temps de convergence plus élevé que l'algorithme LMS.

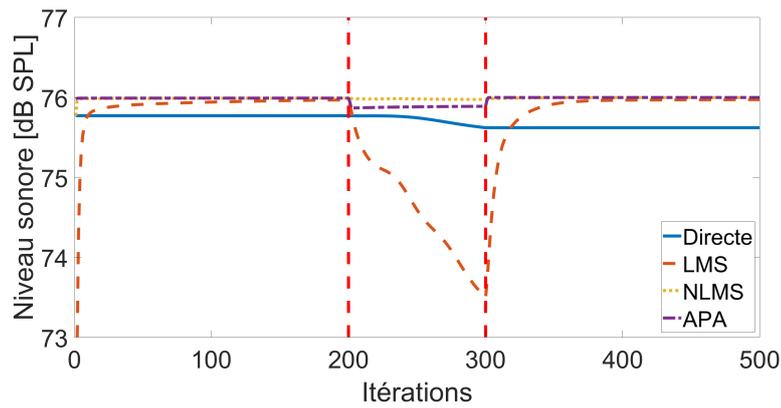


Figure 3.13 – Niveau sonore dans la *Bright zone* pour les différents algorithmes pour un auditeur se déplaçant selon l'axe -Y entre les itérations 200 et 300. Les lignes verticales pointillées rouges repèrent le début et la fin du déplacement.

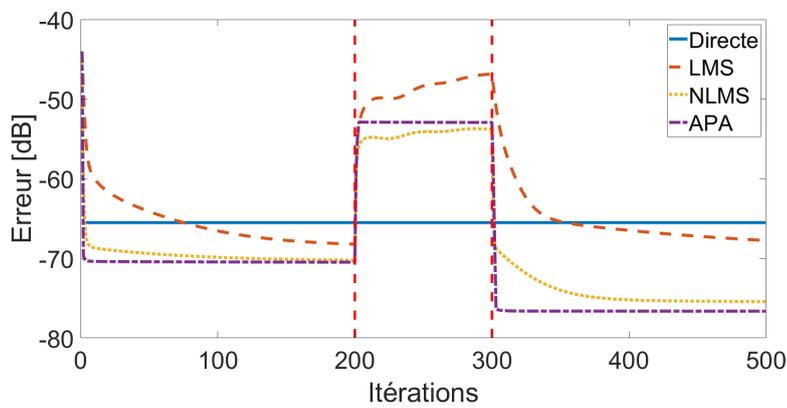


Figure 3.14 – Erreur quadratique moyenne dans la *Bright zone* pour les différents algorithmes pour un auditeur se déplaçant selon l'axe -Y entre les itérations 200 et 300. Les lignes verticales pointillées rouges repèrent le début et la fin du déplacement.

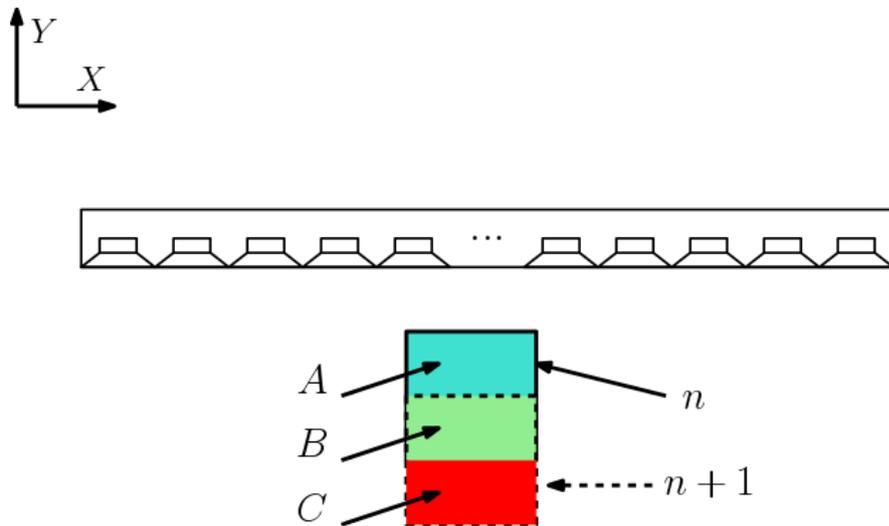


Figure 3.15 – Représentation des différentes zones de recouvrement du champ à reproduire entre deux itérations. Trois zones coexistent. La zone A correspond au champ reproduit à l'itération n , la zone B correspond à la partie commune entre l'itération n et l'itération $n+1$, la zone C correspond à la zone qui doit être reproduite à l'itération $n+1$ et qui est caractérisée par un champ nul jusqu'à l'itération n .

Pour ce qui est de l'effort de réseau (**Figure 3.16**), on remarque qu'il est différent entre les positions initiale et finale de l'auditeur, y compris pour la méthode Directe, ce qui indique que l'effort de reproduction nécessaire à un champ sonore donné dépend de la position de l'auditeur.

Les champs sonores pour l'ensemble des méthodes sont présentés sur les **Figures 3.17** à 3.19 pour les itérations 200, avant le déplacement, 250 pendant le déplacement et 500 à la fin du déplacement et après nouvelle convergence des algorithmes. On vérifie que la zone de reproduction du champ sonore cible, correspondant à la *Bright zone* et repérée par le carré en pointillés blancs, suit bien le déplacement de l'auditeur. On note que la focalisation de la reproduction de la *Bright zone* est faible pour les solutions NLMS et APA. Cette absence de focalisation s'explique par l'effort plus important de ces méthodes, lié à une position de l'espace différente de l'évaluation en **Figure 3.4**. Cependant, le champ sonore reproduit par la méthode LMS est distordu pendant son déplacement (3.18). De fait, aucune des méthodes itératives proposées ne permet de reproduire fidèlement les résultats de la méthode PM lors des déplacements. Cependant, malgré une distorsion du champ sonore, la méthode LMS permet de se rapprocher des résultats de la méthode Directe.

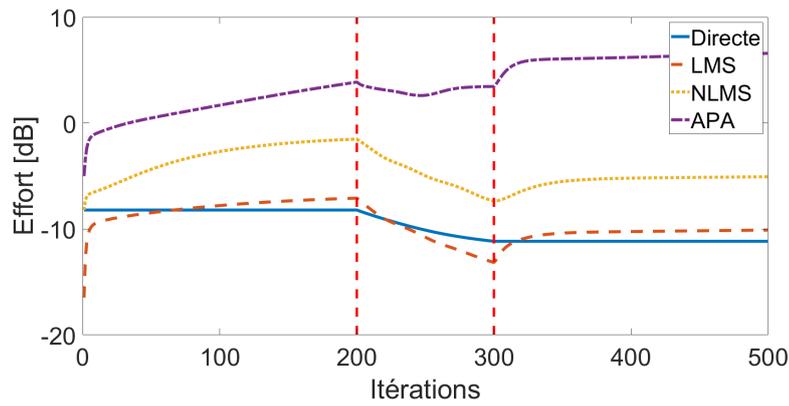


Figure 3.16 – Effort de réseau dans la *Bright zone* pour différents algorithmes pour un auditeur mobile entre les itérations 200 et 300 pour un auditeur se déplaçant selon l'axe $-Y$. Les lignes verticales pointillées rouges représentent le début et la fin du déplacement. Pour le calcul de l'effort, la source de référence est choisie de sorte à ce qu'elle reproduise un niveau sonore équivalent dans la zone avec une distance source/micro équivalente (Eq. (2.37)).

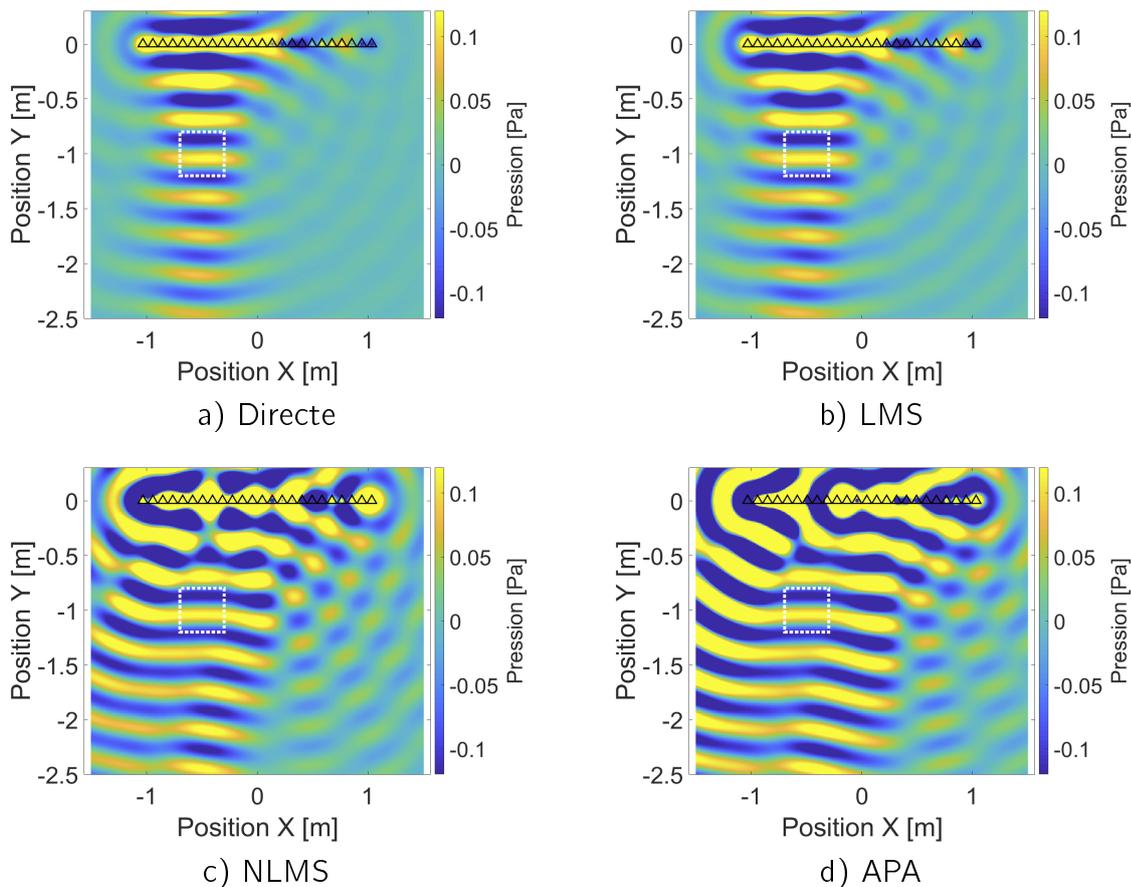


Figure 3.17 – Champ sonore pour les méthodes Directe a), LMS b), NLMS c) et APA d) à l'itération 200 avant le déplacement selon l'axe $-Y$. La position de la *Bright zone* est définie par le carré pointillé blanc.

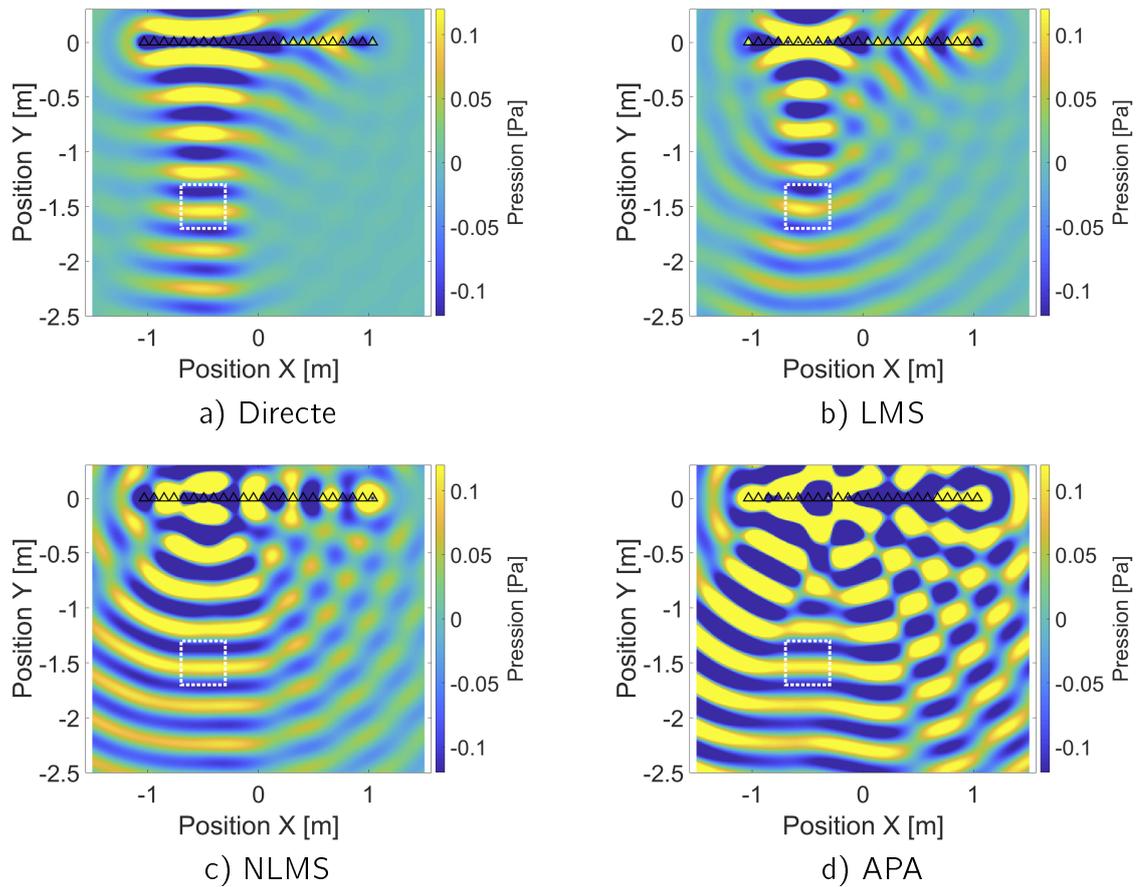


Figure 3.18 – Champ sonore pour les méthodes Directe a), LMS b), NLMS c) et APA d) à l'itération 250 durant le déplacement selon l'axe $-Y$. La position de la *Bright zone* est définie par le carré pointillé blanc.

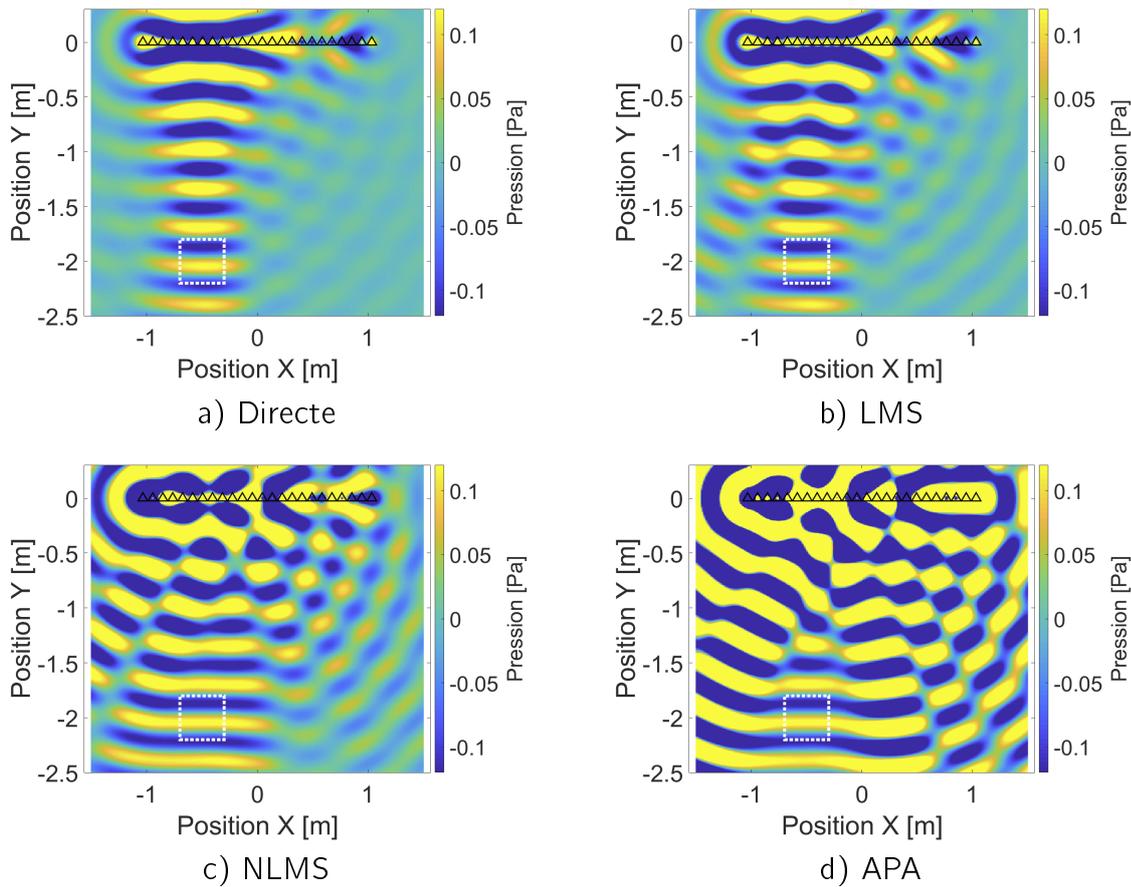


Figure 3.19 – Champ sonore pour les méthodes Directe a), LMS b), NLMS c) et APA d) à l'itération 500 après déplacement selon l'axe $-Y$ et convergence des algorithmes. La position de la *Bright zone* est définie par le carré pointillé blanc.

3.2.2 Déplacement selon l'axe X

La seconde partie de l'étude propose de réaliser un déplacement similaire mais selon l'axe X . De la même façon que précédemment, la trajectoire de la *Bright zone* pour cette étude est illustrée en **Figure** 3.20. Les paramètres des algorithmes restent inchangés.

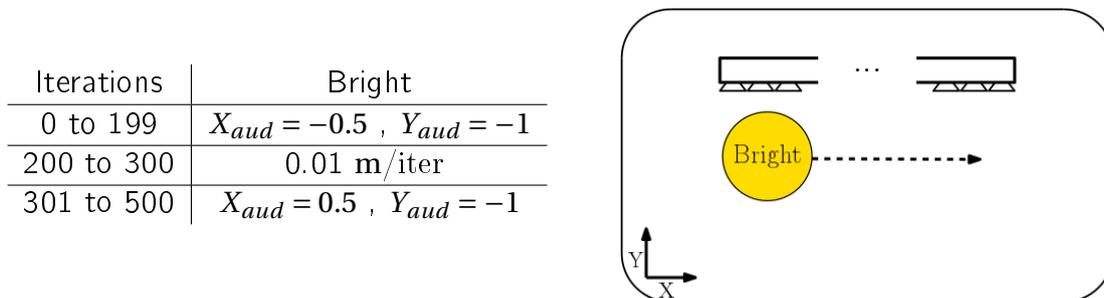


Figure 3.20 – Déplacement de l'auditeur selon l'axe X . Le déplacement s'effectue entre la 200^{ème} et la 300^{ème} itération.

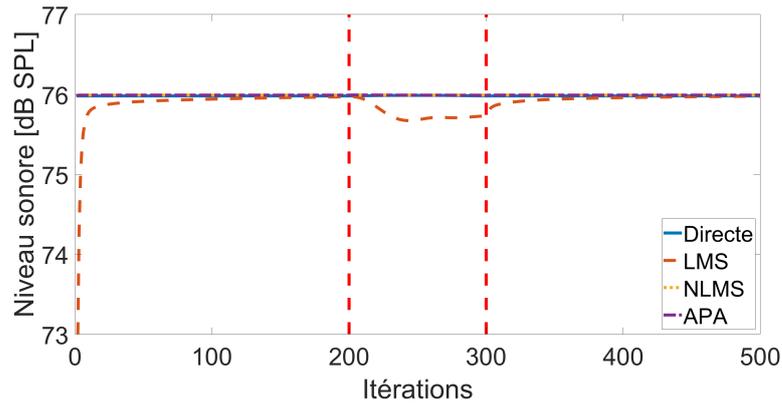


Figure 3.21 – Niveau sonore dans la *Bright zone* pour différents algorithmes pour un auditeur mobile entre les itérations 200 et 300 pour un auditeur se déplaçant selon l'axe X . Les lignes verticales pointillées rouges représentent le début et la fin du déplacement.

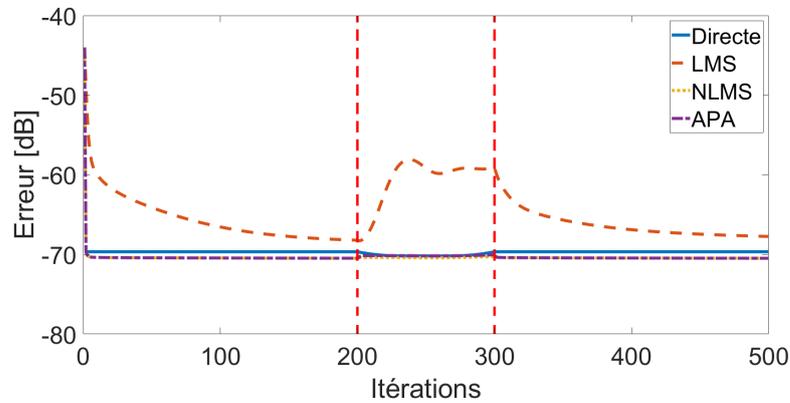


Figure 3.22 – Erreur quadratique moyenne dans la *Bright zone* pour différents algorithmes pour un auditeur mobile entre les itérations 200 et 300 pour un auditeur se déplaçant selon l'axe X . Les lignes verticales pointillées rouges représentent le début et la fin du déplacement.

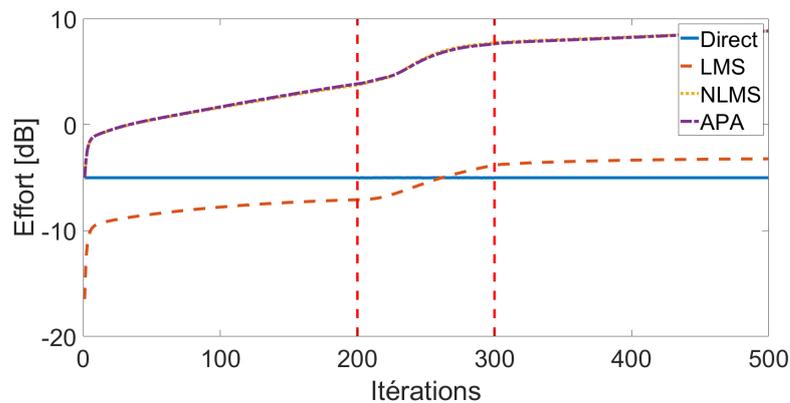


Figure 3.23 – Effort de réseau dans la *Bright zone* pour différents algorithmes pour un auditeur mobile entre les itérations 200 et 300 pour un auditeur se déplaçant selon l'axe X . Les lignes verticales pointillées rouges représentent le début et la fin du déplacement. La source de référence est choisie de sorte à ce qu'elle reproduise un niveau sonore équivalent dans la zone avec une distance source/micro équivalente (**Eq.** (2.37)).

De la même façon que pour le déplacement selon l'axe $-Y$, le mouvement induit une variation de niveau sonore (**Figure 3.21**), il en est de même pour l'erreur de reproduction (**Figure 3.22**). Cependant, cette variation est plus faible ici qu'en §3.2.1. Ce phénomène est différent de ce qui a pu être observé lors du déplacement selon l'axe Y car la direction de déplacement de l'auditeur et la direction de propagation de l'onde sont différentes. Etant donné que le mouvement de l'auditeur est orthogonal à la direction de propagation de l'onde, les variations du champ sonore cible entre deux itérations sont minorées par rapport au cas précédent, les champs sonores concernés étant plus similaires dans la *Bright zone* (**Figure 3.15**). L'augmentation de l'erreur est beaucoup plus faible dans le cas des filtres NLMS et APA puisque leur temps de convergence est nettement plus faible que pour la méthode LMS. On remarque aussi que l'effort (**Figure 3.23**) de tous les algorithmes itératifs augmente de façon significative entre le début et la fin du déplacement, phénomène que l'on ne retrouve pas dans le cas de la résolution Directe. Dans ce cas, il ne semble pas que l'augmentation de l'effort soit induite par une nouvelle position, comme c'était le cas pour le déplacement selon l'axe $-Y$, et ce pour plusieurs raisons. D'une part, l'effort de reproduction de la méthode Directe ne varie pas entre ces deux itérations. D'autre part, la position de départ est symétrique par rapport à la position d'arrivée, on s'attend donc à obtenir un comportement similaire. De plus, le déplacement semble affecter tous les algorithmes itératifs de la même façon. Cette amplification de l'effort suite au déplacement de la *Bright zone* résulte d'un effet mémoire qui est étudié plus en détail dans la section suivante.

3.3 L'effet mémoire

Les méthodes itératives utilisent les poids calculés à l'itération n pour calculer les nouveaux poids de l'itération $n+1$. Le filtre dispose donc d'une mémoire des événements précédents. Lors du déplacement de l'auditeur, le champ sonore qui a été reproduit à l'ancienne position et qui est absent de la nouvelle position (Zone *A* sur la **Figure 3.15**) ne fait plus partie du problème d'optimisation. Il n'est donc plus contraint et la partie des poids permettant cette ancienne reproduction n'est plus utile pour minimiser l'erreur à la nouvelle position, même si elle reste en mémoire dans l'algorithme. L'effet mémoire est donc le résidu du calcul des filtres précédents qui ne dégrade pas l'erreur de reproduction à l'itération suivante. Le champ sonore n'étant défini que sur une partie de l'espace correspondant à la *Bright zone*, seule cette zone est affectée par les modifications. La **Figure 3.24** représente le champ sonore reproduit avant, pendant et après le déplacement de l'auditeur, par la méthode Directe (insensible à l'effet mémoire) (**Figure 3.24 a**), **c**), **e**)) et par la méthode LMS (**Figure 3.24 b**), **d**), **f**)). On observe que dans le cas de la méthode LMS, il existe une "zone de trainée" sur le passage du réseau de microphones. L'augmentation de l'effort de réseau se traduit par la "non extinction" de haut-parleurs actifs aux itérations précédentes, mais qui ne sont plus utiles à la reproduction du champ sonore à l'itération courante. Pour pallier ce problème, nous proposons d'utiliser une méthode de type Leaky LMS ou VLLMS, solution que nous allons tester dans la section suivante.

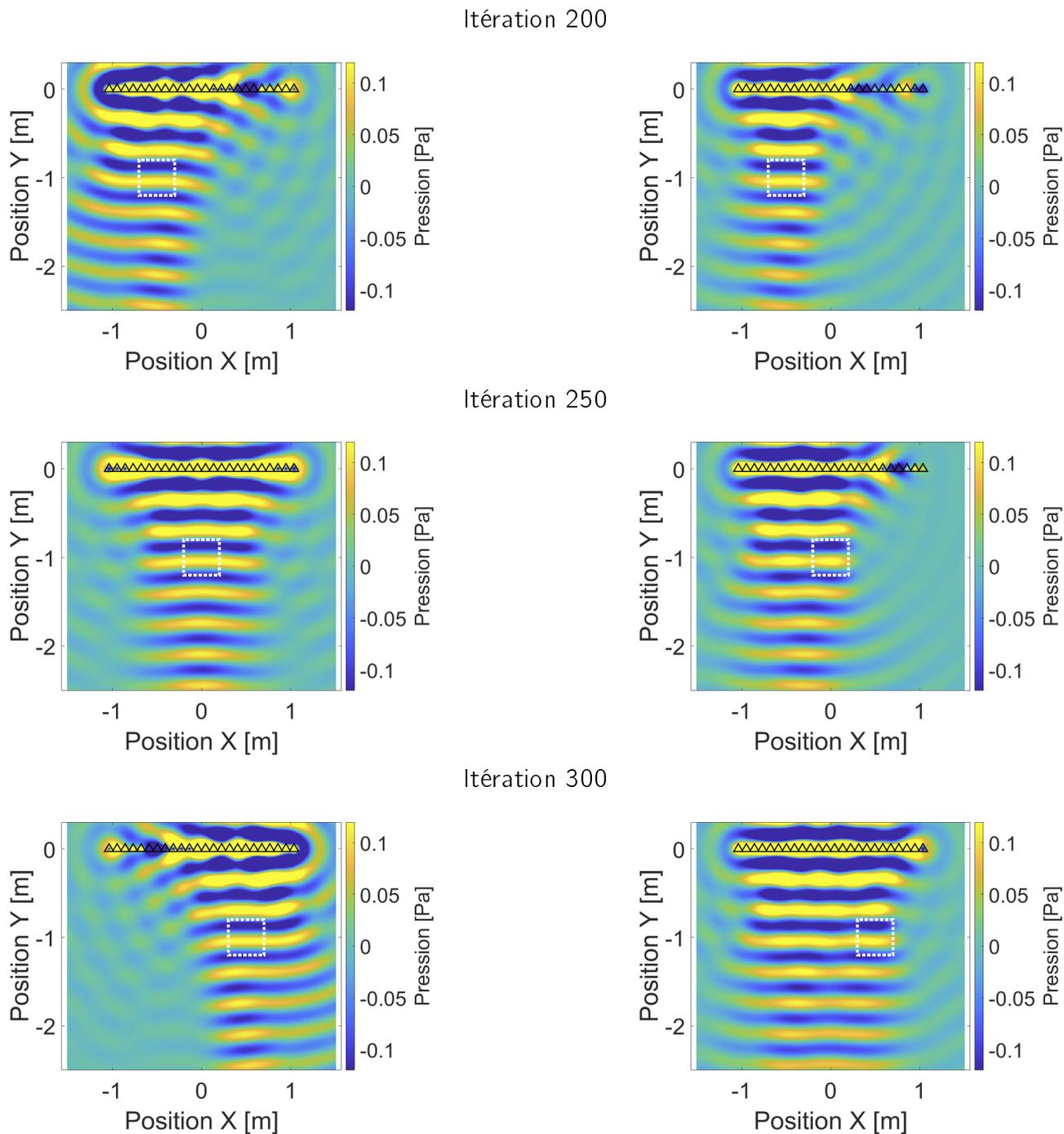


Figure 3.24 – Champ sonore pour les méthodes Directe (gauche) et LMS (droite) pour l'itération 200, c'est à dire avant le déplacement, l'itération 250 c'est à dire pendant le déplacement et pour l'itération 500 c'est à dire après déplacement e) et f) selon l'axe X. La position de la *Bright zone* est définie par le carré pointillé blanc.

3.3.1 Comparaison des performances des méthodes LMS et VLLMS

Les méthodes de type Leaky LMS introduisent une procédure d'oubli (ou de "fuite") afin de n'utiliser qu'une partie des poids calculés à l'itération précédente (§2.2.4), à la différence du filtre LMS qui garde l'intégralité de ces poids. Nous souhaitons ici évaluer l'intérêt de la méthode Leaky LMS et de sa variante VLLMS dans le but de réduire l'effet mémoire observé pour la solution LMS. Les paramètres d'étude proposés sont identiques à ceux du déplacement selon l'axe X (§3.3), puisque c'est dans cette

configuration que l'effet est le plus notable. Même si cet effet est également présent pour les algorithmes NLMS et APA, nous nous contenterons de l'appliquer à la méthode LMS. Pour rappel, le filtre VLLMS dispose de paramètres supplémentaires permettant de régler la vitesse de transition entre le filtre LMS et le filtre Leaky LMS (§2.2.4). Pour cette évaluation, on choisit $lu = 0.1$, $ld = 0.1$ et $\alpha = 3$. L'influence de ces paramètres est proposée en **Annexe D**. Seule l'influence du facteur d'oubli maximal γ_{max} est évaluée ici. Les performances obtenues avec les filtres Leaky LMS et VLLMS sont représentées sur les figures 3.25-3.27 pour γ_{max} variant de 0 à 0.05.

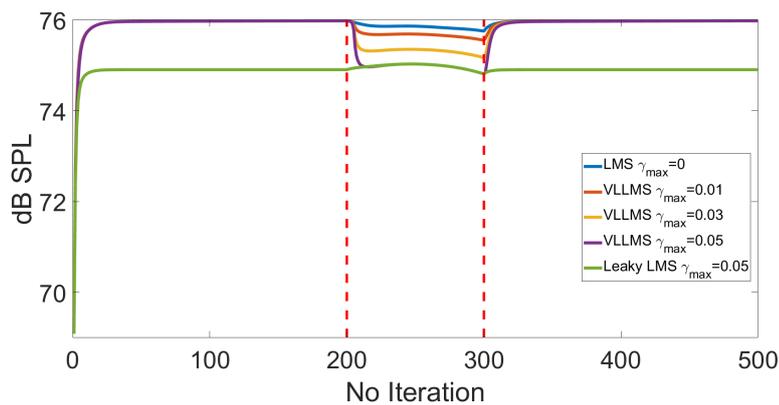


Figure 3.25 – Niveau sonore dans la *Bright zone* pour différentes valeurs du paramètre d'oubli γ_{max} pour un auditeur se déplaçant selon l'axe X pour les méthodes LMS, Leaky LMS et VLLMS. Les lignes verticales pointillées rouges représentent le début et la fin du déplacement.

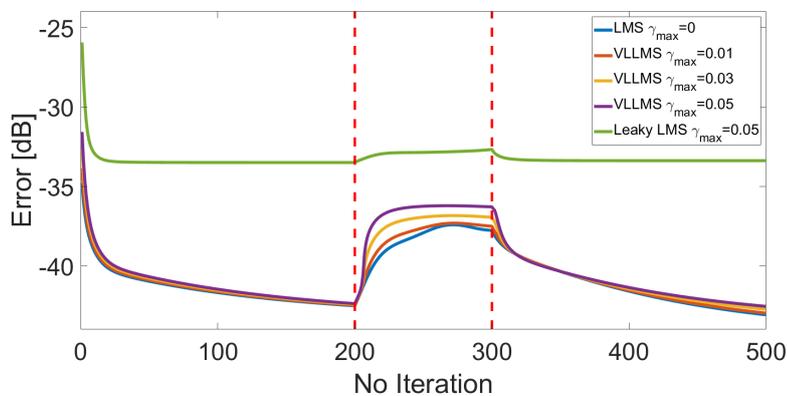


Figure 3.26 – Erreur quadratique moyenne dans la *Bright zone* pour différentes valeurs du paramètre d'oubli γ_{max} pour un auditeur se déplaçant selon l'axe X pour les méthodes LMS, Leaky LMS et VLLMS. Les lignes verticales pointillées rouges représentent le début et la fin du déplacement.

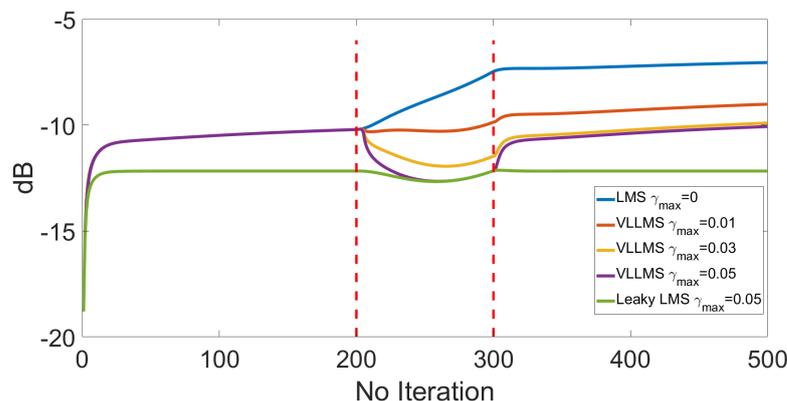


Figure 3.27 – Effort de réseau pour différentes valeurs du paramètre d'oubli γ_{max} pour un auditeur se déplaçant selon l'axe X pour les méthodes LMS, Leaky LMS et VLLMS. Les lignes verticales pointillées rouges représentent le début et la fin du déplacement. La source de référence est choisie de sorte à ce qu'elle reproduise un niveau sonore équivalent dans la zone avec une distance source/micro équivalente (**Eq.** (2.37)).

La diminution du niveau sonore (**Figure 3.25**) lors du déplacement de l'auditeur est d'autant plus grande que la valeur de γ_{max} est élevée mais ne dépasse pas la valeur minimale représentée par l'algorithme Leaky LMS correspondant à la valeur la plus élevée du paramètre d'oubli. Le rôle de la méthode VLLMS étant de passer du mode LMS au mode Leaky LMS. Lors du déplacement, le facteur d'oubli du filtre VLLMS augmente jusqu'à atteindre la valeur maximale γ_{max} . Plus cette valeur augmente, plus l'oubli est important et donc plus l'erreur sur le niveau sonore et sur l'erreur de reproduction (**Figure 3.26**) augmente.

On remarque par ailleurs que l'erreur de reproduction de la méthode VLLMS reste très inférieure à l'erreur de reproduction de la méthode Leaky LMS, quelle que soit la valeur du facteur d'oubli. Cependant, elle s'avère toujours supérieure à celle de la méthode LMS. Pour ce qui est de l'effort de réseau (**Figure 3.27**), l'augmentation du paramètre d'oubli permet à la méthode VLLMS de modérer l'accroissement de l'effort, lié à l'effet mémoire, présent sur la méthode LMS après le déplacement de l'auditeur. Ces résultats s'observent également sur le champ sonore reproduit (**Figure 3.28** et **Figure 3.29**). On vérifie que le champ résiduel est efficacement réduit par la méthode Leaky LMS dès l'application d'un facteur d'oubli de **0.01** (**Figure 3.28**). En revanche, l'effet de traînée persiste, quoique atténué, dans le cas de la méthode VLLMS, et ce même après convergence de l'algorithme à l'itération 500. Plusieurs méthodes permettent d'éliminer ce champ résiduel. Par exemple, on peut augmenter le facteur d'oubli maximal. Le champ sonore obtenu par la méthode VLLMS en fin de convergence se rapproche ainsi de celui reproduit par la méthode Leaky LMS, comme l'illustre la **Figure 3.29**. On peut également prolonger la mise en œuvre du processus d'oubli quelques itérations supplémentaires après arrêt du mouvement en jouant sur les paramètres l_u, l_d ou α .

La méthode Leaky LMS réussit à supprimer l'effet mémoire, même avec un facteur d'oubli faible, mais au prix d'une élévation de l'erreur de reproduction, et d'un biais sur le niveau sonore. La méthode VLLMS permet de trouver un meilleur compromis en réduisant le champ résiduel sans trop dégrader l'erreur. Son principal inconvénient est qu'elle nécessite le réglage de 3 paramètres supplémentaires (l_u, l_d et α) qui dépendent du type de déplacement (vitesse et durée du déplacement). Ces paramètres influent sur le facteur d'oubli $\gamma(n)$ au cours des itérations, plus particulièrement sur la vitesse de

passage entre la le filtre LMS et le filtre Leaky LMS. De plus, ils influent également sur le temps de convergence en début et en fin de déplacement (**Annexe D**).

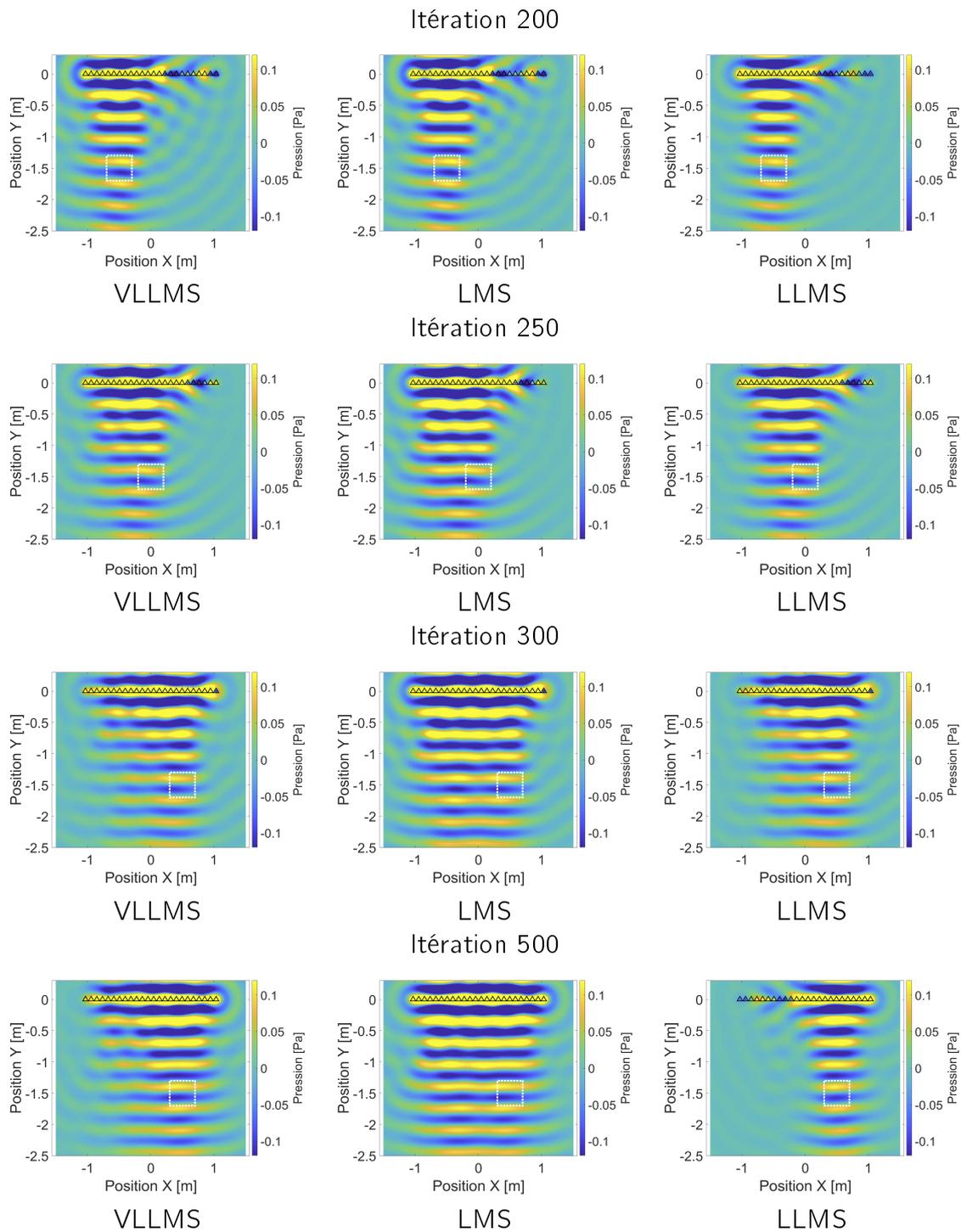


Figure 3.28 – Champ sonore reproduit par les méthodes VLLMS (gauche), LMS (milieu) et Leaky LMS (droite). Dans le cas des algorithmes Leaky LMS et VLLMS, un facteur d'oubli de 0.01 est appliqué.

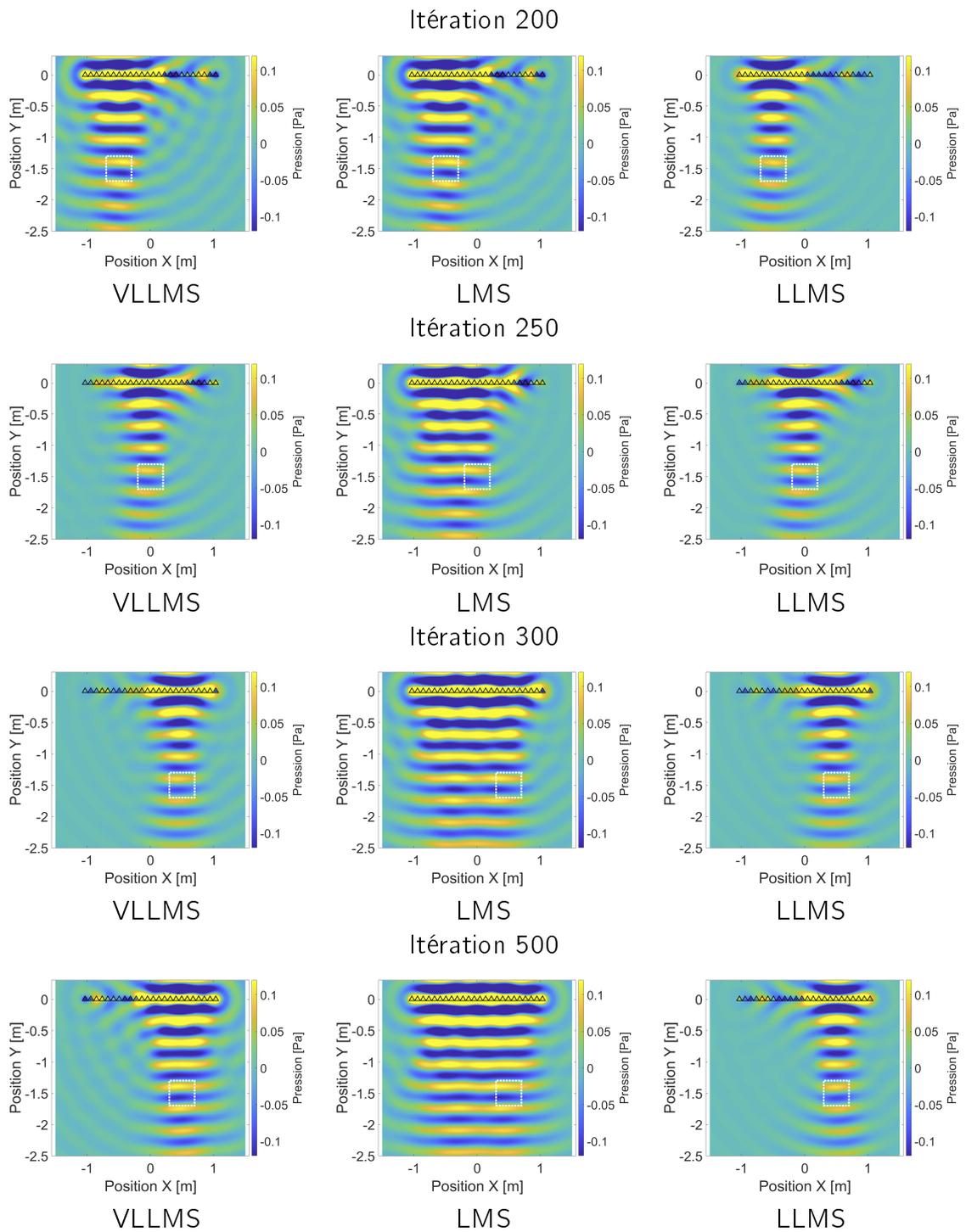


Figure 3.29 – Champ sonore reproduit par les méthodes VLLMS (gauche), LMS (milieu) et Leaky LMS (droite). Dans le cas des algorithmes Leaky LMS et VLLMS, un facteur d'oubli de 0.05 est appliqué.

3.4 Conclusion

La question traitée dans ce chapitre porte sur la production d'une *Bright zone* pour un auditeur unique en mouvement. On se focalise donc sur la reproduction d'une *Bright zone* sans chercher à réduire le champ sonore induit pour un second auditeur (la *Dark zone*). La spécificité de ce problème réside dans l'adaptation de la *Bright zone* aux déplacements de l'auditeur. Une des contraintes que doit respecter l'algorithme retenu est sa réactivité, afin de pouvoir suivre le mouvement de l'auditeur sans induire de latence dans la reproduction sonore (ce qui implique notamment une convergence rapide et un coût de calcul faible). Dans ce chapitre, nous avons évalué les performances des méthodes itératives, introduites au chapitre précédent, pour produire une *Sound Zones* en mouvement. Les algorithmes itératifs LMS, NLMS et APA ont été évalués dans le cas d'une *Bright zone* statique, afin de tester leur capacité à résoudre le problème indépendamment du déplacement de l'auditeur. Ces algorithmes demandent un certain nombre d'itérations (appelé temps de convergence) avant de converger vers la solution obtenue par la méthode Directe. Il a été observé que les algorithmes NLMS et APA ont une convergence très rapide en comparaison du filtre LMS. Cependant, ces algorithmes nécessitent l'utilisation d'un paramètre de régularisation β afin de stabiliser le filtre, ce qui tend à augmenter le temps de calcul. Par ailleurs, à la différence de la méthode Directe, le paramètre de régularisation β ne contrôle pas l'effort de réseau, ce qui conduit à des efforts de réseau très élevés à long terme et s'avère incompatible avec une implémentation réaliste.

Le filtre LMS est quant à lui caractérisé par un coût de calcul plus faible, d'autant plus qu'il ne nécessite pas le calcul d'un paramètre de régularisation supplémentaire. Cependant, pour assurer sa stabilité, le pas μ_L , qui dépend des valeurs propres des matrices de corrélation spatiales $\mathbf{G}^H\mathbf{G}$, doit être réglé correctement. La méthode retenue consiste à le calculer pour chaque itération. Cette solution ne constitue pas la méthode la plus rapide, mais assure à la fois stabilité et bonne rapidité de convergence. D'autres solutions, discutées dans le chapitre 5, sont possibles.

Par ailleurs, les méthodes LMS, NLMS et APA n'introduisent aucune contrainte sur l'effort, conduisant à son augmentation sur le long terme. L'utilisation des méthodes de type Leaky LMS (Leaky LMS et VLLMS) résout partiellement ce problème en introduisant un facteur d'oubli qui réduit l'influence des poids calculés aux itérations précédentes. La méthode Leaky LMS, ainsi que sa version variable dans le temps, la méthode VLLMS, permet ainsi de réguler l'effort de réseau si nécessaire. Cette méthode pourrait a priori s'utiliser également dans le cas des filtres NLMS et APA, ce qui n'a pas été exploré ici.

Après avoir évalué les performances des méthodes itératives en conditions statiques, ces dernières ont été testées dans le cas d'une zone de reproduction en mouvement, soit parallèlement soit orthogonalement au réseau de haut-parleurs. Une baisse significative des performances a été observée pour l'ensemble des algorithmes itératifs pendant le déplacement. Cette baisse est principalement due à la variation entre le champ sonore déjà existant et le champ sonore à reproduire à la nouvelle itération. Une zone commune à deux itérations successives implique une adaptation des algorithmes dépendant du champ sonore déjà présent dans la zone de reproduction. Lors d'un déplacement selon un axe perpendiculaire au sens de propagation de l'onde à reproduire, la conservation des poids calculés aux itérations précédentes introduit un effet mémoire qui se caractérise par une zone de traînée et une augmentation de l'effort de réseau.

Les méthodes Leaky LMS et VLLMS offrent une solution pour réduire cet effet

mémoire, puisqu'elles mettent en œuvre un processus d'oubli des itérations précédentes. De fait, les artefacts de trainée sont réduits, mais au prix d'une augmentation de l'erreur de reproduction.

Pour conclure, la méthode itérative la plus adaptée à notre problème est la méthode VLLMS. Même si les méthodes NLMS et APA présentent une convergence très rapide, leur coût de calcul est au moins aussi élevé que celui de la méthode Directe et est de surcroît alourdi par la détermination du paramètre de régularisation.

La méthode VLLMS combine quant à elle les avantages de la méthode LMS et ceux de la méthode LLMS. La méthode LMS dispose d'un coût de calcul relativement faible, même lorsque le pas μ_L est calculé à chaque itération, et fournit une bonne précision de reproduction du champ sonore. Le filtre Leaky LMS permet quant à lui de réduire l'effort de réseau ainsi que l'effet mémoire. C'est donc la méthode VLLMS qui sera utilisée dans le chapitre suivant où nous allons considérer la synthèse conjointe d'une *Bright zone* et d'une *Dark zone* pour des auditeurs en mouvement.

Principales avancées

- **Evaluation des méthodes développées** en chapitre 2 pour la reproduction focalisée dans une zone par la méthode PM. Les algorithmes **LMS, NLMS et APA** sont **comparés** à la **méthode Directe** en condition **statique** et dans le cas d'un **déplacement de l'auditeur**.
- **Mise en évidence de l'effet mémoire** lors des **déplacements** : effet mémoire résultant des **contributions des itérations précédentes** correspondant à un problème légèrement différent associé au **déplacement de l'utilisateur**.
- **Evaluation** des méthodes **VLLMS** et **LLMS** introduisant un **facteur d'oubli** pour **réduire l'effet mémoire**.
- **Choix** de la méthode **VLLMS** pour la **suite** dans le cas **multi-auditeurs** car elle **combine** les avantages de **LMS** et **LLMS**.

Chapitre 4

Vers des Sound Zones en mouvement

Dans le chapitre précédent, nous nous sommes focalisés sur le problème de restitution du champ sonore dans la *Bright zone*, la nouveauté introduite par rapport à l'état de l'art résidant dans le fait que l'auditeur cible est en mouvement. Dans ce chapitre, nous allons nous intéresser au problème de *Sound Zones* dans son intégralité, en considérant conjointement la reproduction d'un champ sonore cible dans *Bright zone* et dans la *Dark zone*, ces deux dernières pouvant être en mouvement. Il ne s'agit pas à proprement parler d'un nouveau problème, mais plutôt d'une reformulation du problème résolu pour la *Bright zone*, en considérant cette fois-ci un champ cible à reproduire combinant celui de la *Bright zone* et celui de la *Dark zone*. Pour résoudre ce problème, nous allons nous appuyer sur les solutions mises en oeuvre dans le cas d'une *Bright zone* mobile.

Pour rappel, l'évaluation menée dans le chapitre précédent d'une sélection d'algorithmes itératifs adaptés à ce problème nous a permis d'identifier l'algorithme VLLMS comme la solution répondant au mieux aux contraintes de la reproduction de *Sound Zones* mobiles. Cet algorithme dispose en effet d'une faible complexité de calcul, associée à un effet mémoire réduit, ce qui le rend réactif au déplacement de l'auditeur. De plus, il peut être aisément comparé à la méthode de résolution Directe qui constitue la solution de référence parmi les solutions de l'état de l'art. Dans ce chapitre, l'algorithme VLLMS est donc évalué pour les trois configurations de *Dark zones* proposées dans la partie 2.4. La première configuration est définie par la mise en oeuvre conventionnelle (champ de pression nul en *Dark zone*). La seconde configuration correspond à l'adaptation de la méthode WLS pour le filtre VLLMS dans laquelle la fonction de coût intègre conjointement la minimisation de la pression en *Dark zone* et la qualité de reproduction du champ sonore en *Bright zone*. Enfin, la troisième méthode propose de définir le champ sonore en *Dark zone* à chaque itération comme une reproduction atténuée de la pression qui est produite dans cette zone à l'itération précédente. Cette dernière méthode est composée de deux variantes. La première notée DZIG s'intéresse à l'atténuation sur l'ensemble de la *Dark zone*, tandis que la seconde appelée DZII réalise l'atténuation sur chacun des microphones de la *Dark zone* de façon indépendante comme défini en §2.4.2. Dans tous les cas, les résultats sont comparés à ceux obtenus par la méthode Directe. L'évaluation est d'abord menée dans le cas de deux auditeurs fixes §4.2.3, avant de considérer deux auditeurs en mouvement §4.3. Un déplacement simple est étudié dans un premier temps, puis on s'intéresse ensuite aux cas dits « d'occlusion », définis lorsqu'un auditeur passe devant un autre auditeur. Au préalable les notations utilisées par la suite sont rappelées (§4.1).

4.1 Généralisation du problème pour inclure une *Dark zone*

La méthode VLLMS est ici appliquée à la situation où plusieurs auditeurs sont présents simultanément et souhaitent écouter deux programmes audio distincts. Il convient alors de reproduire chaque programme dans la zone pour laquelle il est destiné (*Bright zone*) et de l'atténuer dans l'autre zone (*Dark zone*), ce qui signifie qu'on met donc en place des *Sound Zones*. Le problème est posé de façon équivalente au cas mono-auditeur. On note \mathbf{G}_B les fonctions de transfert entre les haut-parleurs et les microphones dans la *Bright zone* (là où un contenu audio doit être reproduit) et \mathbf{G}_D , les fonctions de transfert pour la *Dark zone* (là où le contenu audio doit être rendu inaudible). De la même façon, les pressions $\mathbf{p}_B = \mathbf{G}_B \mathbf{q}$ et $\mathbf{p}_D = \mathbf{G}_D \mathbf{q}$ correspondent respectivement aux pressions dans la *Bright zone* et la *Dark zone*. Il en est de même pour le champ cible \mathbf{p}_{T_B} et \mathbf{p}_{T_D} . Dans la suite, le champ sonore est considéré comme nul en *Dark zone* $\mathbf{p}_{T_D} = \mathbf{0}$ lorsqu'il n'est pas défini explicitement par la méthode utilisée. Pour retrouver la formulation du problème VLLMS présentée précédemment (§2.2.4), il suffit d'assembler toutes ces notations sous la forme :

$$\mathbf{G} = \begin{bmatrix} \mathbf{G}_B \\ \mathbf{G}_D \end{bmatrix} \quad \mathbf{p} = \begin{bmatrix} \mathbf{p}_B \\ \mathbf{p}_D \end{bmatrix} \quad \mathbf{p}_T = \begin{bmatrix} \mathbf{p}_{T_B} \\ \mathbf{p}_{T_D} \end{bmatrix} \quad (4.1)$$

Même si l'objectif est de rendre inaudible en *Dark zone* le champ sonore associé à la *Bright zone*, en pratique, il n'est pas possible d'obtenir un niveau sonore nul en *Dark zone*. Afin d'évaluer la qualité de séparation entre les deux zones, le paramètre de contraste est utilisé comme défini en équation (2.36). Ce paramètre permet de rendre compte de la différence de niveau sonore entre les deux zones et fournit donc un paramètre objectif de "confort" d'écoute en *Bright zone*, relativement aux interférences des autres zones. Comme vu dans la partie 2.4, il existe différentes façons de définir le champ sonore dans la *Dark zone* pour améliorer les *Sound Zones*. Une évaluation des différentes méthodes de calcul est donc proposée.

4.2 Influence de la stratégie de contrôle de la *Dark zone*

La méthode PM permet d'être flexible quant au choix du champ sonore dans la *Dark zone*. Comme présenté en 2.4, quatre solutions sont proposées pour le contrôle champ sonore en *Dark zone*. Deux d'entre elles sont issues de l'état de l'art et adaptées à la résolution par la méthode VLLMS. Les deux autres développées dans le travail de thèse s'appuient sur la formulation itérative du problème. La première consiste à définir un champ sonore nul dans la *Dark zone* en donnant autant d'importance dans l'optimisation des deux zones. La seconde consiste à utiliser la méthode *Weighted Least Squares* (WLS) de sorte à choisir l'importance donnée à la minimisation du champ sonore dans la *Dark zone* ou à la minimisation de l'erreur de reproduction dans la *Bright zone*. Les deux autres méthodes utilisent le champ sonore directement présent dans la *Dark zone* qui est atténué, soit de façon globale, constituant (méthode DZIG), soit de façon homogène sur la zone (méthode DZII [Roussel et Nicol, 2018b]).

Dans un premier temps, les méthodes de *Dark zone* sont évaluées dans le cas de *Sound Zones* statiques pour les méthodes itératives utilisant la résolution VLLMS en comparaison avec la méthode Directe.

Pour réaliser cette évaluation de façon optimale, elles sont évaluées à effort identique. Cet effort cible E_C est déterminé de façon arbitraire et défini égal à $E_C = -10$ dB.

Notons qu'à la différence du chapitre précédent, où le but était de minimiser l'erreur dans la *Bright zone*, ici le critère de l'erreur n'est plus le plus pertinent pour contraindre le problème, car le but n'est pas uniquement de reproduire un champ sonore dans la *Bright zone*, mais aussi de minimiser la pression dans la *Dark zone*. L'erreur de reproduction est alors dépendante de la priorité donnée soit à la minimisation de la pression dans la *Dark zone*, soit à la minimisation de l'erreur dans la *Bright zone*. Au contraire, dans ce problème, il est plus intéressant de limiter l'énergie délivrée par les haut-parleurs pour chaque méthode, et de comparer leurs performances en termes d'erreur et de contraste. De plus, dans le chapitre précédent, toutes les méthodes itératives ne permettaient pas de réguler l'effort, et il n'était donc pas possible d'utiliser l'effort comme paramètre limitant pour leur comparaison.

Pour rappel la **Figure 2.2**, définissant le problème général des *Sound Zones*, est reportée en **Figure 4.1** afin de faciliter la lecture de ce chapitre.

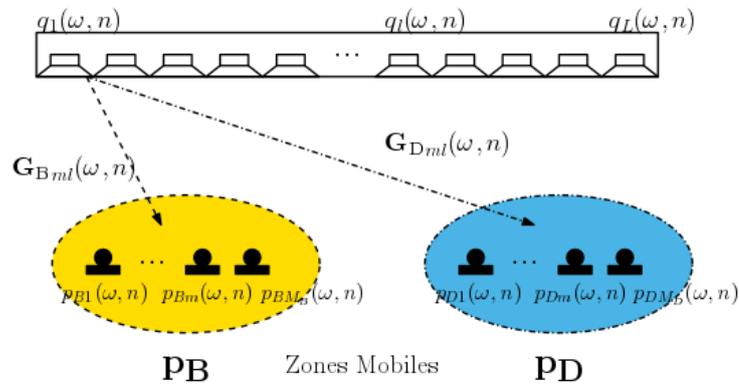


Figure 4.1 – Définition du problème acoustique pour la génération d'une sound zone dans un contexte de plusieurs auditeurs mobiles combinant une *Bright zone* (jaune) et une *Dark zone* (bleu). La matrice de fonctions de transfert $\mathbf{G}_B(\omega, n)$ définit les trajets acoustiques entre les M microphones et les L_{HP} haut-parleurs de la *Bright zone*, de la même façon, la matrice de fonctions de transfert $\mathbf{G}_D(\omega, n)$ définit les trajets acoustiques pour la *Dark zone*. Chaque microphone m délivre une pression $p_m(\omega, n)$, avec $p_{Bm}(\omega, n)$ pour les pressions en *Bright zone* et $p_{Dm}(\omega, n)$ pour les pressions en *Dark zone*. Les haut-parleurs sont pilotés par des poids $q_l(\omega, n)$.

4.2.1 Paramètres expérimentaux

Le problème consiste à reproduire dans la *Bright zone* une onde plane à 76 dB SPL, en provenance de la direction *Y* pour des zones positionnées comme proposé en **Figure 4.2**. Le but est d'évaluer les performances de chaque méthode de *Sound Zones* pour différentes stratégies de contrôle du champ sonore en *Dark zone*.

Iterations	Bright	Dark
0 à 500	$X_{aud} = -0.5$ $Y_{aud} = -2$	$X_{aud} = 0.5$ $Y_{aud} = -2$

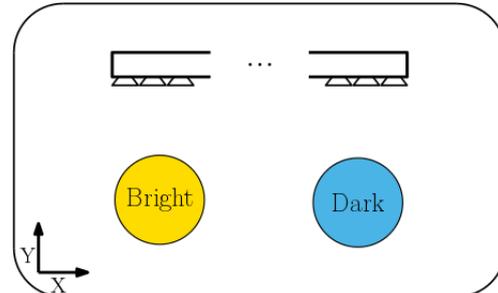


Figure 4.2 – Positions de la *Bright zone* et de la *Dark zone* pour l'évaluation des méthodes de *Sound Zones* dans le cas statique.

Il est choisi de laisser l'algorithme fonctionner pour 500 itérations afin de rendre compte d'un cas statique pour ces algorithmes et donc d'un nombre élevé d'itérations par rapport au temps de convergence défini pour une zone seule.

Pour rappel, les méthodes DZII et DZIG utilisent une version atténuée du champ sonore présent dans la *Dark zone* à l'itération précédente à l'aide d'un ou plusieurs coefficients d'atténuation ξ calculés à l'aide de (2.44).

La méthode DZII calcule les coefficients ξ indépendamment pour chaque microphone. La méthode DZIG quant à elle, calcule ξ pour l'ensemble de la zone. Ces deux méthodes nécessitent la définition d'un contraste cible à obtenir entre les zones. Dans un premier temps, pour les méthodes DZII et DZIG, le contraste cible est défini comme étant $L_C = 60$ dB, ce qui permet d'avoir un bon contraste entre les zones [Poletti, 2008]. Par la suite, l'influence de ce contraste est évalué et reporté dans le nom de la méthode. Par exemple dans le cas $L_C = 60$ dB, la méthode est appelée DZII 60.

Par ailleurs, la méthode WLS nécessite elle aussi le choix du paramètre η définissant un ratio entre les performances de la *Dark zone* et de la *Bright zone*. Pour la première partie de l'évaluation, $\eta = 0.8$ ce qui donne plus d'importance à la minimisation de la *Dark zone*. La méthode VLLMS avec la *Dark zone* définie comme nulle est notée VLLMS 0 dans ce chapitre.

4.2.2 Résultats

Dans cette partie, nous évaluons les performances des différentes méthodes testées, en termes de contraste et d'erreur de reproduction dans la *Bright zone*. La **Figure 4.3** représente le contraste obtenu pour l'ensemble des fréquences pour la dernière itération considérée. Dans cette configuration, les méthodes disposent d'un contraste maximal entre 700 Hz et 6.5 kHz. Par ailleurs, très peu de différences existent entre la méthode Directe et les méthodes testées. La proximité des résultats des méthodes Directe et VLLMS 0 provient du fait que les *Dark zones* sont définies de façon équivalente et donnent donc des résultats similaires. Pour ce qui est des méthodes DZII et DZIG, le contraste cible étant très élevé par rapport à la valeur optimale au sens des moindres

carrés¹, il n'est pas possible pour ces algorithmes d'être plus performants en termes de contraste que la méthode de référence. La méthode VLLMS WLS dont le ratio entre la *Bright zone* et la *Dark zone* est différent fait exception. Par ailleurs, les mêmes conclusions s'imposent pour l'erreur de reproduction quadratique dans la *Bright zone* proposée en **Figure 4.4**. Les méthodes se comportent de la même façon car elles sont proches dans leur fonctionnement. Pour ce qui est de la méthode VLLMS WLS, elle offre globalement un contraste légèrement supérieur aux autres méthodes au détriment d'une erreur de reproduction plus élevée dans la *Bright zone*. L'évaluation des méthodes DZII et DZIG avec la résolution VLLMS nécessite une analyse complémentaire pour voir l'influence du paramètre L_C associé au contrôle du contraste. Cet aspect est traité en §4.2.3.

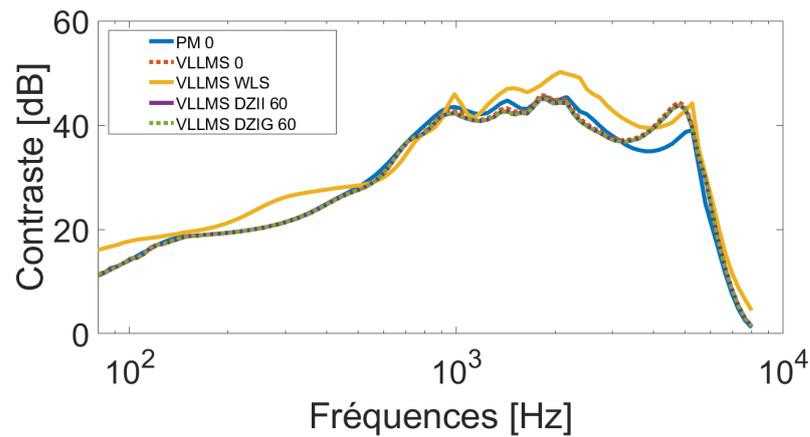


Figure 4.3 – Contraste pour les différentes méthodes de *Dark zone* dans le cas d'utilisateurs statiques.

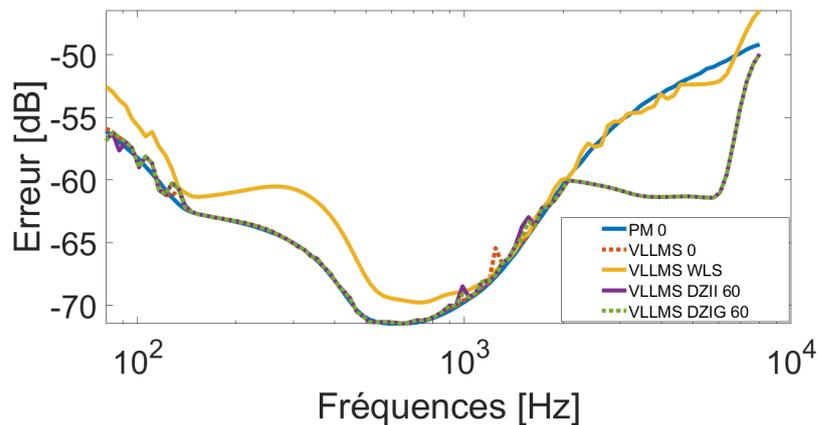


Figure 4.4 – Erreur quadratique moyenne dans la *Bright zone* pour les différentes méthodes de *Dark zone* dans le cas d'utilisateurs statiques.

1. Cette valeur optimale au sens des moindres carrés est obtenue par la méthode Directe, les autres méthodes étant des approximations

4.2.3 Influence du contraste pour les méthodes DZIG et DZII

La partie précédente n'a pas permis de mettre en évidence de réelles différences entre les méthodes DZII, DZIG et la méthode Directe. Pour rappel, les méthodes DZII et DZIG nécessitent le choix d'un paramètre de contraste cible L_C pour fonctionner. Or, ce paramètre modifie le niveau sonore dans la *Dark zone* et peut alors s'adapter à différents besoins. Dans cette partie, l'évaluation est réalisée pour trois valeurs cibles de L_C : 20 dB, 30 dB et 50 dB, donnant lieu aux méthodes DZII 20, DZII 30, DZII 50 ainsi que DZIG 20, DZIG 30 et DZIG 50.

La **Figure 4.5** présente le contraste pour chacune des méthodes et la **Figure 4.6** présente l'erreur quadratique de reproduction dans la *Bright zone*. La méthode DZIG est représentée en trait plein et la méthode DZII en traits pointillés. Pour les contrastes cibles peu élevés, correspondant à $L_C = 20\text{dB}$ ou $L_C = 30\text{dB}$, le contraste est constant sur une bande fréquentielle allant d'environ 600 Hz à 6.5 kHz. Pour ce qui est du contraste cible $L_C = 50$ dB, on observe que d'une part le contraste cible n'est pas atteint et que d'autre part, celui-ci n'est pas constant sur l'ensemble de la bande fréquentielle, comme pour la **Figure 4.3** (contraste cible $L_C = 60$ dB). Étant donné les capacités limitées de l'algorithme d'optimisation à résoudre le problème, lorsqu'un contraste très élevé lui est imposé, cela revient à une situation similaire à s'il lui était imposé de reproduire une pression nulle dans la *Dark zone*. Il reste donc à sa valeur maximale de contraste (au sens d'optimal du problème) sans pouvoir la dépasser.

La qualité de reproduction de l'onde plane dans la *Bright zone* n'est quant à elle pas affectée par cette différence de contraste. La valeur maximale de contraste n'est donc pas définie par la valeur cible de contraste mais par la capacité de l'algorithme à reproduire un niveau de contraste maximal définissant une limite, ce de façon indépendante pour chaque fréquence. Lorsque le contraste cible est inférieur à cette limite sur une certaine bande de fréquences, il est alors possible d'obtenir un contraste constant sur cette même bande de fréquences.

A noter qu'il serait intéressant d'évaluer la qualité perçue des *Sound Zones* lorsque celles-ci ne disposent pas d'un contraste maximal, mais d'un contraste plus faible et constant sur une bande de fréquences. Certains cas pourraient ne pas nécessiter un contraste maximal en s'appuyant sur ces paramètres perceptifs. Ici, les méthodes DZII et DZIG ne se différencient pas sur les performances de contraste ou d'erreur de reproduction en *Bright zone* mais plutôt sur leur capacité à générer un contraste constant sur une bande de fréquences.

Le contraste entre deux zones est défini comme la différence de niveau sonore moyen entre les deux zones. Cependant, le champ cible dans la *Dark zone* n'étant pas toujours défini explicitement par les méthodes (la méthode WLS par exemple), celui-ci peut correspondre à différentes distributions spatiales. Généralement, il est observé que le niveau sonore dans la *Dark zone* n'est pas homogène et seul le niveau sonore moyen est alors considéré [Coleman, 2014]. A noter que l'influence de cette non-homogénéité spatiale n'a été que très peu étudiée jusqu'à présent. Ici, l'intérêt est porté sur la répartition spatiale du niveau sonore dans la *Dark zone*. La **Figure 4.7** présente l'écart type sur le niveau sonore dans la *Dark zone* pour les méthodes DZII et DZIG au cours des itérations. C'est à dire l'évolution au cours des itérations de la dispersion du niveau sonore au sein de la *Dark zone*. Plus elle est élevée, plus le champ sonore au sein de la *Dark zone* est inhomogène.

Dans l'ensemble, pour le même niveau de contraste cible, la méthode DZII a un écart-type plus faible que son équivalent pour la méthode DZIG. Ce résultat est pré-

visible car la méthode DZII, en ajustant individuellement le niveau sonore en chaque point réduit l'écart-type de niveau sonore dans la zone, et améliore donc l'homogénéité. Cela est particulièrement sensible pour les faibles contrastes cibles, particulièrement $L_C = 20$ dB et $L_C = 30$ dB. Pour $L_C = 50$ dB, cette différence existe mais est plus faible. Pour l'ensemble des méthodes évaluées, l'écart type observe une première phase d'augmentation rapide puis une phase de diminution avant d'augmenter légèrement à nouveau. Ce phénomène est lié au compromis réalisé entre la qualité de reproduction de la *Bright zone* et le niveau de contraste à obtenir. La méthode DZII peut être vue comme une méthode WLS avec un paramètre η variable au sein de l'espace. C'est à dire qu'il permet de moduler le champ sonore cible en *Dark zone* en fonction du contraste observé à chaque position microphonique, en donnant une importance variable à la reproduction du champ dans une zone donnée de l'espace.

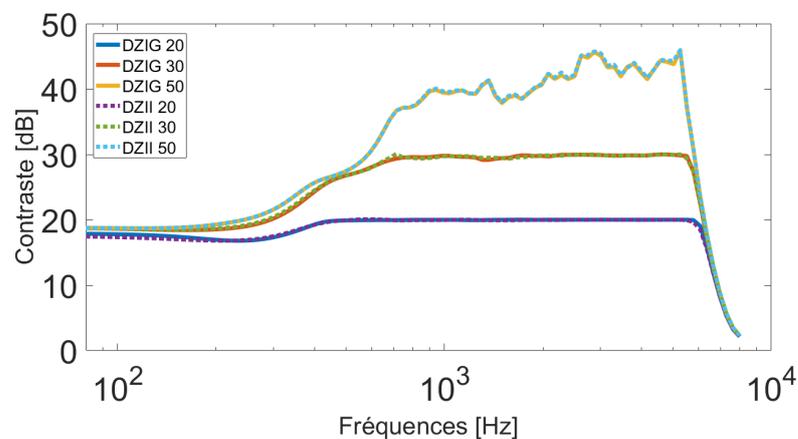


Figure 4.5 – Contraste entre la *Bright zone* et la *Dark zone* pour les méthodes DZII et DZIG pour différents contrastes cibles L_C .

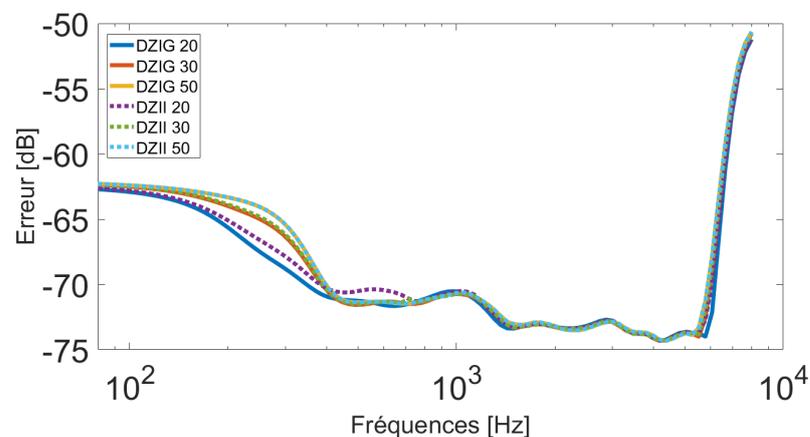


Figure 4.6 – Erreur de quadratique de reproduction dans la *Bright zone* pour les méthodes DZII et DZIG pour différents contrastes cibles L_C .

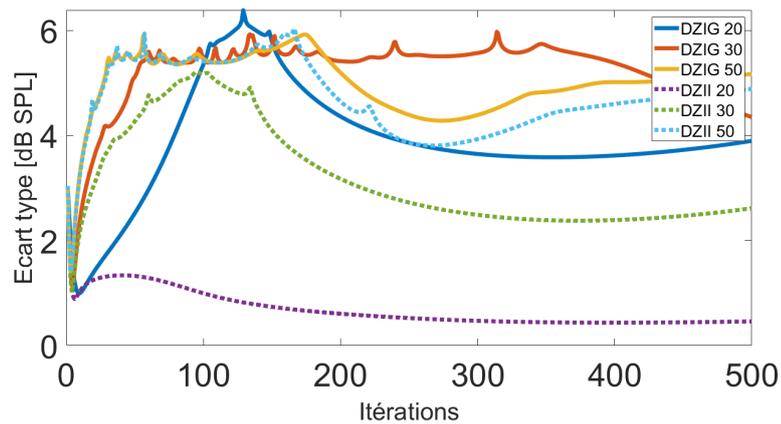


Figure 4.7 – Ecart-type du niveau sonore pour les microphones dans la *Dark zone*, pour les méthodes DZII et DZIG, pour différents niveaux de contraste en fonction des itérations.

Afin de mieux évaluer l'impact de l'écart-type de niveau sonore dans la *Dark zone*, le champ sonore est reproduit pour chacune des méthodes à l'itération 500. La **Figure 4.8** présente le niveau sonore sur l'ensemble de l'espace considéré pour la méthode DZII 50. On observe un niveau sonore constant en *Bright zone* (carré blanc) correspondant à la reproduction de l'onde plane. Dans la *Dark zone* (carré noir), le niveau sonore est plus faible et n'est pas constant. Les **Figures 4.9** représentent un zoom de la **Figure 4.8** en *Dark zone* (carré noir) pour les différentes méthodes étudiées, en termes d'écart au niveau sonore cible.

Lorsque cette différence est nulle, le niveau sonore correspond au niveau sonore souhaité. Lorsque la différence est positive, le niveau sonore est plus élevé que le niveau cible, ce qui signifie que pour cette position, le contraste est trop faible par rapport au contraste cible L_C . Si la différence est négative, le niveau sonore est trop faible. Pour cette position le contraste est trop élevé par rapport au contraste cible. Les **Figures 4.9 a), c) et e)** représentent la méthode DZIG pour les différents niveaux de contraste, alors que les **Figures 4.9 b), d) et f)** représentent la méthode DZII pour ces mêmes niveaux de contraste. On remarque que la *Dark zone* de la méthode DZIG n'est homogène pour aucune des valeurs de L_C (le niveau sonore n'est pas constant sur l'ensemble de la zone). Cela est dû au fait que lorsque le champ sonore cible p_{TD} de la *Dark zone* est calculé, l'algorithme cherche à minimiser l'erreur quadratique sur l'ensemble de la zone en donnant autant d'importance à tous les points, donc sans favoriser de zone spécifique, tant que le niveau sonore cible sur l'ensemble de la zone est respecté. Pour ce qui est de la méthode DZII, pour un contraste cible de $L_C = 20$ dB (**Figure 4.9 b)**), l'ensemble de la zone est très proche du niveau sonore cible, ce de façon homogène. Pour un contraste cible $L_C = 30$ dB, le champ sonore (**Figure 4.9 b)**) reste plus homogène que celui de la méthode DZIG 30 (**Figure 4.9 c)**), même s'il n'est pas complètement homogène sur l'ensemble de la zone. La méthode DZII favorise donc une homogénéité du niveau sonore dans la *Dark zone* en donnant plus de poids aux positions pour lesquelles le niveau sonore cible n'est pas atteint. Cependant, ceci dans les limites des performances de l'algorithme : pour un contraste cible $L_C = 50$ dB, la méthode DZII ne parvient pas à dépasser les performances de la méthode DZIG.

En conclusion, si le niveau sonore moyen de la *Dark zone* de la méthode DZIG est identique au niveau sonore de la méthode DZII, les stratégies diffèrent pour obtenir ce niveau sonore moyen, la méthode DZIG diminue fortement le niveau sonore sur certains

points de l'espace pour compenser les niveaux sonores plus élevés dans le reste de la zone. Pour un fort contraste cible, les méthodes DZII et DZIG se comportent de la même manière malgré le poids supplémentaire donné à certains points de l'espace par la méthode DZII. Si l'homogénéité de la *Dark zone* n'est aujourd'hui reliée à aucun critère de confort d'écoute, il serait néanmoins judicieux d'évaluer l'effet des hétérogénéités en termes perceptifs. Le coût de calcul supplémentaire associé à l'utilisation de cette méthode reste cependant faible et semble être envisageable pour une utilisation en temps réel.

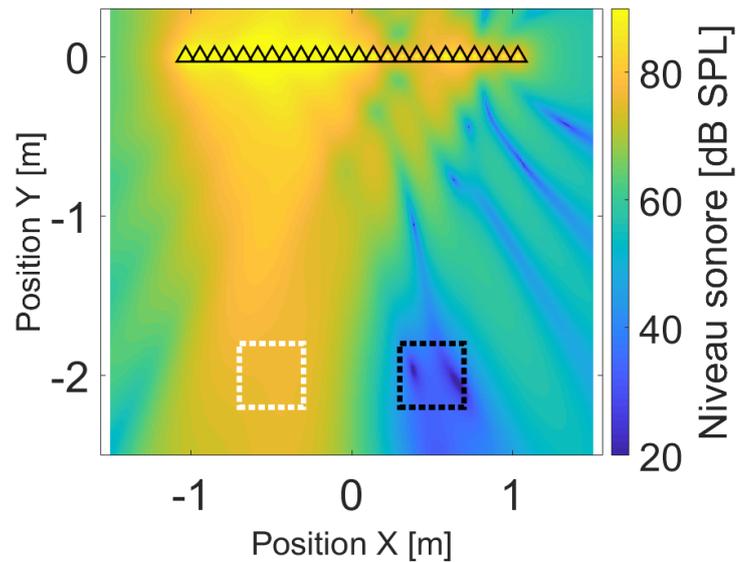


Figure 4.8 – Niveau Sonore sur l'ensemble du champ pour la méthode DZIG 50. Le cadre blanc correspond à l'emplacement de la *Bright zone* et le cadre noir à celui de la *Dark zone*.

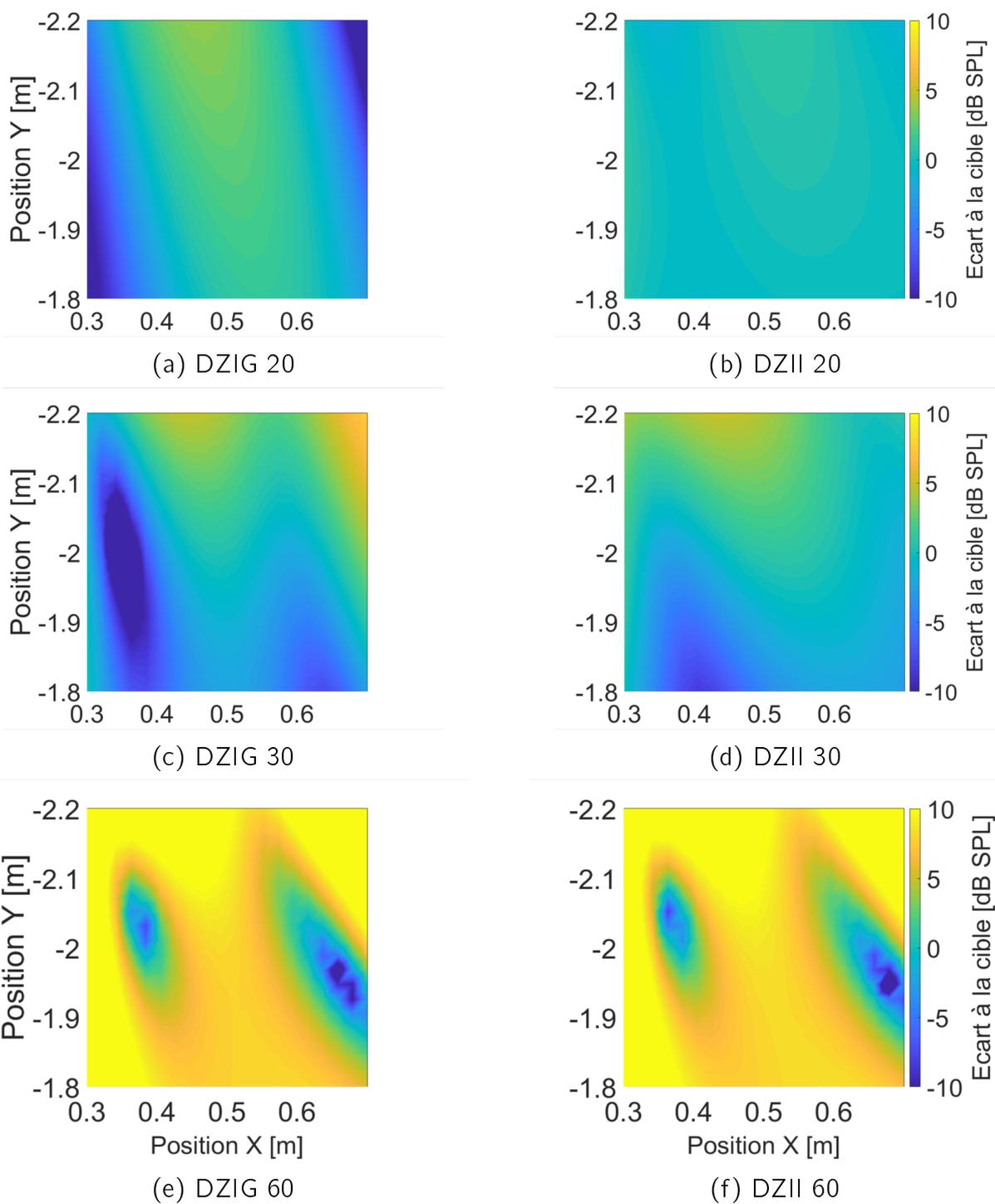


Figure 4.9 – Ecart de niveau sonore au champ cible à la position de la *Dark zone* pour l'ensemble des microphones pour différentes valeurs de contraste, pour les méthodes DZII et DZIG. La partie du champ sonore représentée correspond à un zoom sur la *Dark zone* de la **Figure 4.8**.

4.3 Déplacement de l'utilisateur

Dans la partie précédente, les méthodes de *Dark zone* ont été évaluées pour des utilisateurs statiques. Dans cette partie, nous nous intéressons à l'utilisation des méthodes de *Sound Zones* dans le cas d'auditeurs mobiles, et plus particulièrement à la répartition spatiale du champ sonore sur l'ensemble du domaine de reproduction (et non plus seulement au sein des zones). Cela permet de mettre en évidence deux aspects essentiels pour la suite : d'une part la compensation de l'effet mémoire par l'utilisation de l'algorithme VLLMS dans le cas des *Sound Zones*, et d'autre part l'influence que les deux zones ont l'une sur l'autre.

Par simplicité, seules deux méthodes sont comparées ici : la méthode Directe et la méthode VLLMS 0 (le champ sonore en *Dark zone* est choisi nul). Dans le contexte de *Sound Zones* pour des auditeurs mobiles, il est par ailleurs important de mesurer l'influence du déplacement sur la qualité des *Sound Zones* proposées. Un cas particulier à étudier propre aux *Sound Zones* mobiles est le problème d'occlusion [Poletti, 2008, Betlehem *et al.*, 2015]. Ce phénomène se produit lorsqu'une zone est dans le même axe qu'une autre. Il existe alors deux zones concurrentes. Ce problème est d'autant plus sensible dans le cas d'un réseau linéaire, car à la différence du réseau circulaire, il n'est pas possible d'utiliser les haut-parleurs situés derrière l'un des deux utilisateurs [Zhang *et al.*, 2016, Coleman, 2014]. De plus, la méthode PM n'est pas la plus robuste au problème d'occlusion car elle contraint fortement la direction du champ sonore dans la *Bright zone*. Une méthode comme le *Planarity Control* [Coleman *et al.*, 2014b] semble plus adaptée dans ce cas car la contrainte sur le champ sonore n'est pas définie sur sa direction d'arrivée, mais ce cas n'est pas traité ici.

L'évaluation des méthodes de *Sound Zones* mobiles est réalisée pour trois cas particuliers dans la configuration d'un réseau linéaire de haut-parleurs. Le premier cas correspond au déplacement de la *Bright zone* selon un axe ne gênant pas l'utilisateur de la *Dark zone*. Les deux autres cas mettent en avant le phénomène d'occlusion. L'étude des problèmes d'occlusion permet de vérifier la stabilité de l'algorithme dans des "conditions extrêmes" d'utilisation et son comportement après une forte perturbation. Les cas d'occlusion sont considérés comme des cas extrêmes car l'algorithme doit reproduire deux champs sonores très différents dans des zones de l'espace très proches et pointés dans la même direction. Il n'y a donc aucune garantie d'obtenir un résultat stable.

Le premier cas d'étude correspond à une occlusion simple dite asymétrique. L'auditeur de la *Bright zone* passe devant le second et s'arrête devant. Le second cas correspond à une occlusion dite symétrique. L'auditeur de la *Bright zone* passe devant l'autre auditeur sans s'arrêter et finit par s'arrêter à une position symétrique de la position de départ. Les différentes configurations sont définies respectivement en **Figure 4.10**, 4.14 et 4.18.

Pour l'ensemble des simulations, le protocole est défini comme suit :

- La *Dark zone* est fixe pour toute la durée de la simulation
- De l'itération 0 à l'itération 199, la *Bright zone* est à sa position initiale.
- De l'itération 200 à l'itération 300, la *Bright zone* se déplace de 0.01 m/iter dans une direction donnée.
- De l'itération 301 à l'itération 500, la *Bright zone* est à sa position finale et ne bouge plus.

Les positions X_{aud} et Y_{aud} correspondent à la position du centre du réseau de microphones par rapport au centre du réseau de haut-parleurs. Les 200 premières itérations ainsi que les dernières permettent au filtre VLLMS de se stabiliser avant et après le mouvement de l'auditeur de la *Bright zone*. Le début et la fin du déplacement des zones sont représentés par des lignes pointillées rouges verticales sur l'ensemble des figures. Le pas μ_L (Eq. (2.10)) est calculé à chaque itération comme étant la valeur propre maximale $\kappa_{max}/4$ de $\mathbf{G}_B^H \mathbf{G}_B$ ce qui correspond à un filtre VLLMS convergeant rapidement, mais proche de l'instabilité. Les paramètres du facteur d'oubli sont : $lu = 0.1$, $ld = 0.1$ et $\gamma_{max} = 0.2$, correspondant à un oubli maximal de 20% des poids précédents.

Pour la méthode Directe, le paramètre de régularisation est recalculé pour chaque itération du VLLMS, de sorte à ce que l'effort de réseau ne dépasse jamais -10 dB.

4.3.1 Déplacement de la *Bright zone* selon l'axe Y

Le premier cas de déplacement de l'auditeur consiste en un déplacement selon l'axe Y d'une distance de 1m réparti sur 100 itérations. Les deux utilisateurs restent sur deux axes distincts sans se croiser comme précisé en **Figure 4.10**.

Iterations	Bright	Dark
0 à 199	$X_{aud} = -0.5,$ $Y_{aud} = -0.5$	$X_{aud} = 0.5,$
200 à 300	0.01 m/iter	
301 à 500	$X_{aud} = -0.5,$ $Y_{aud} = -1.5$	$Y_{aud} = -1.5$

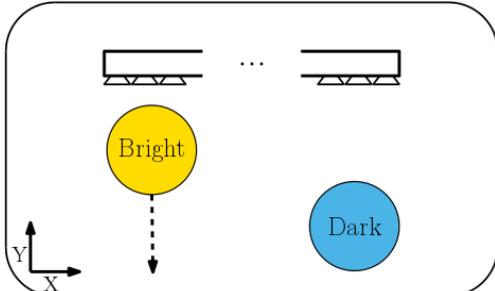


Figure 4.10 – Positions de la *Bright zone* et de la *Dark zone* pour l'évaluation des performances en fonction du déplacement de l'auditeur, pour différentes méthodes de calcul de la *Dark zone*. L'utilisateur se déplace selon l'axe Y. Le déplacement de la *Bright zone* se produit entre l'itération 200 et l'itération 300.

On remarque aussi bien sur le contraste (**Figure 4.11**) que sur l'erreur quadratique de reproduction (**Figure 4.12**) une différence de performance entre les positions de départ et d'arrivée. Cette différence est particulièrement visible sur le contraste, plus élevé et s'étendant sur une bande de fréquences plus large à la position finale qu'à la position initiale. Ces différences de performances entre la position initiale et la position finale peuvent s'expliquer par la position relative des deux zones comme suggéré dans [Bo Moller et Olsen, 2011]. En effet, la distance entre la *Bright zone* et la *Dark zone* ainsi que leur position relative ont une grande influence sur les performances des *Sound Zones*. Un des problèmes potentiels de l'utilisation des *Sound Zones* pour les utilisateurs mobiles est donc une grande disparité de contraste entre les différentes zones de l'espace. Une étude approfondie sur l'influence des positions relatives et absolues des auditeurs par rapport au réseau de haut-parleurs pourrait permettre de mieux comprendre comment les performances évoluent au sein de l'espace.

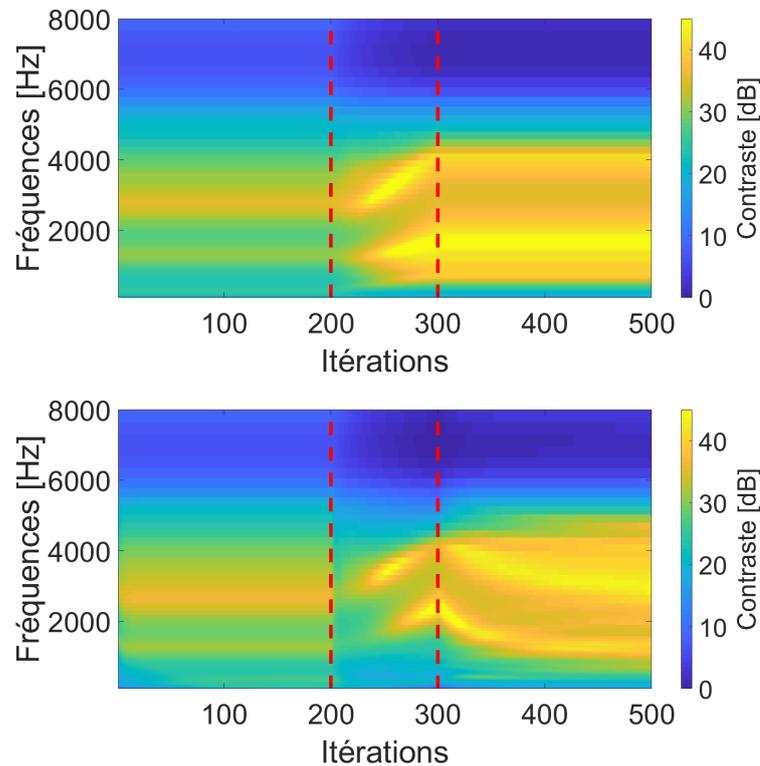


Figure 4.11 – Contraste au cours des itérations pour la méthode Directe (Haut) et la méthode VLLMS (Bas) pour un déplacement le long de l'axe Y. Le déplacement de la *Bright zone* se produit entre l'itération 200 et l'itération 300.

D'autre part, lors du déplacement de la *Bright zone*, la méthode VLLMS offre de moins bonnes performances que la méthode Directe en termes de qualité de reproduction du champ sonore (diminution du contraste plus importante au début de déplacement et erreur de reproduction plus élevée sur l'ensemble du déplacement). Cela est dû à plusieurs facteurs. D'une part, la méthode VLLMS ne converge pas en une seule itération, et ne peut donc pas s'adapter complètement pendant toute la durée du déplacement au changement. D'autre part, de la même façon que dans le cas d'un auditeur unique (§3.3), la qualité du champ sonore est dégradée lors du déplacement (**Figure 3.14**). En effet, à la différence de la méthode Directe qui recommence le calcul d'un nouveau problème indépendant à chaque fois, la méthode VLLMS calcule le champ sonore à l'itération $n+1$, à l'aide des poids déjà existant à l'itération n . Dans le cas d'un déplacement, ce champ n'est que partiellement connu (Voir **Figure 3.14**).

Il est à noter que même si les valeurs entre les algorithmes Direct et VLLMS sont différentes, les tendances en termes de performance restent similaires. Ces différences de performance entre les deux algorithmes peuvent aussi s'expliquer par le temps de convergence de la méthode VLLMS. Les champs sonores reproduits par les méthodes Directe et VLLMS pour une fréquence de 1kHz sont présentés en **Figure 4.13** à titre d'exemple. Il est possible de remarquer une très grande similitude dans les champs sonores reproduits par les deux méthodes.

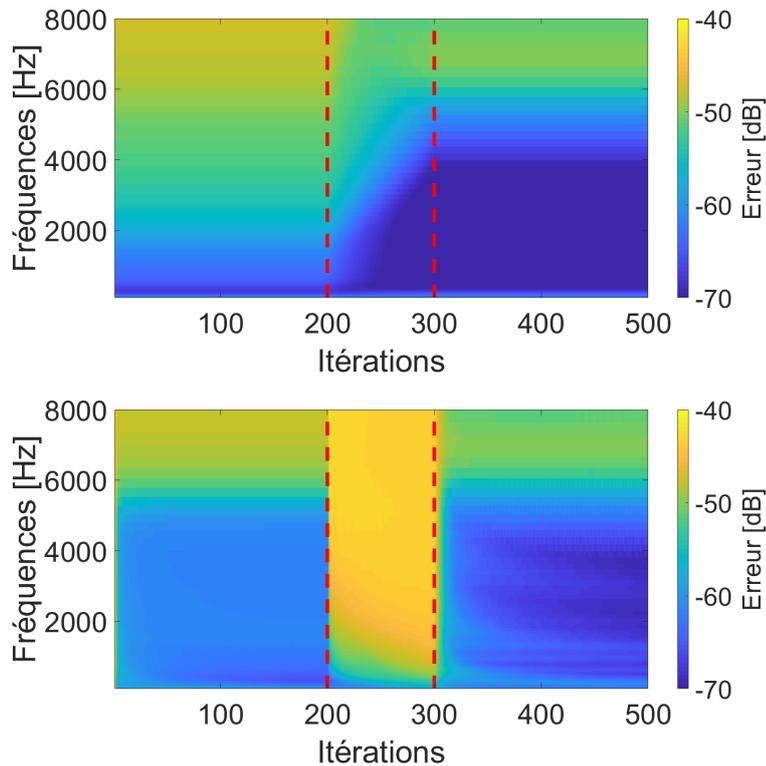


Figure 4.12 – Erreur quadratique moyenne au cours des itérations pour la méthode Directe (Haut) et la méthode VLLMS (Bas) pour un déplacement le long de l'axe Y. Le déplacement de la *Bright zone* se produit entre l'itération 200 et l'itération 300.

Les méthodes Directe et VLLMS offrent un comportement similaire lors du déplacement d'un utilisateur au sein de l'espace. Cependant, les variations de contraste en fonction des positions relatives des zones risquent de dégrader l'expérience sur l'ensemble de l'espace. Cela implique qu'il n'est pas possible de bénéficier d'un contraste maximal sur l'ensemble des positions de l'espace de reproduction. L'utilisation des méthodes DZII ou DZIG pourrait potentiellement permettre d'obtenir des performances homogènes tout au long du déplacement en définissant un contraste cible plus faible sur l'ensemble de l'espace. En effet, en regard de la **Figure 4.5**, il est possible de moduler le niveau de contraste cible. Certaines zones de l'espace de reproduction disposent d'un contraste plus élevé que d'autres. En diminuant le contraste maximal sur l'ensemble de l'espace de reproduction à l'aide des méthodes DZII ou DZIG, la dispersion des valeurs de contraste sur l'ensemble de l'espace pourrait être réduite. Cependant, pour définir quelle valeur de contraste cible L_C donner aux méthodes DZIG ou DZII, il est nécessaire de réaliser une évaluation complète des performances en fonction des positions relatives ainsi que de définir un contraste minimum pour garantir une séparation de contenus satisfaisante. Ceci nécessite de prendre en compte un grand nombre de paramètres et oblige la création d'un nouvel indicateur de la répartition du contraste dans l'espace. Ces performances dépendent notamment de la configuration choisie, de l'espace de reproduction souhaité ainsi que des couples de position *Bright zone*/*Dark zone*. En outre, la construction d'un tel indicateur est loin d'être trivial.

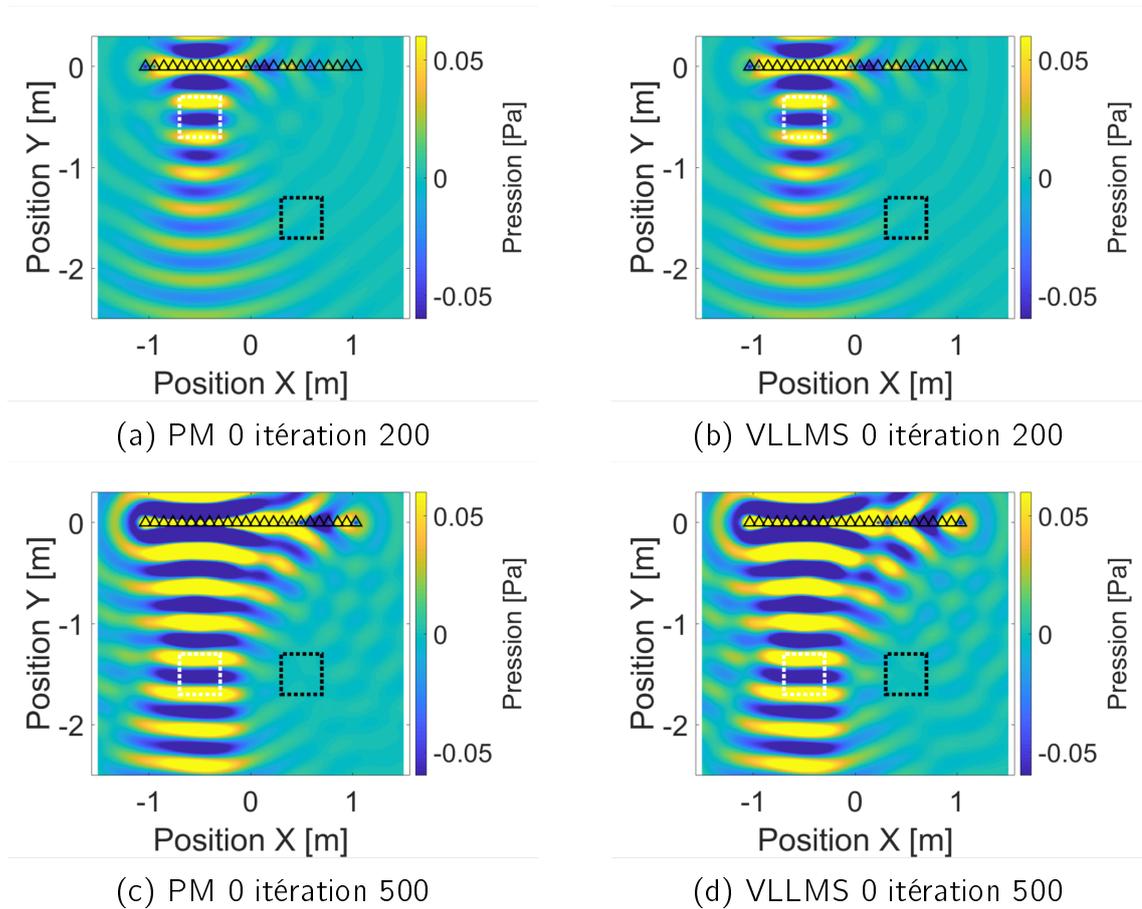


Figure 4.13 – Champ sonore reproduit par les méthodes Directe (Gauche) et VLLMS (Droite) pour différentes positions de l'utilisateur en *Bright zone*. L'itération 200 (Haut) correspond à l'itération juste avant le déplacement de l'utilisateur, l'itération 500 (Bas) correspond à la fin du déplacement de l'utilisateur.

4.3.2 Problème d'occlusion asymétrique

Le second cas d'étude pour les *Sound Zones* dans le cas d'auditeurs mobiles est le problème d'occlusion asymétrique, qui correspond à la première évaluation dans des conditions dites "extrêmes". L'intérêt est particulièrement porté sur la stabilité de l'algorithme. L'auditeur de la *Bright zone* se déplace depuis sa position initiale pour se placer devant le second auditeur lui-même positionné en *Dark zone* **Figure** (4.14). De la même façon que précédemment, l'auditeur se déplace de 1 m en 100 itérations. La **Figure** 4.15 compare les performances en termes de contraste entre la *Bright zone* et la *Dark zone* pour la méthode Directe et la méthode VLLMS.

Itérations	Bright	Dark
0 à 199	$X_{aud} = -0.5,$ $Y_{aud} = -1.5$	$X_{aud} = 0.5,$
200 à 300	0.01 m/iter	
301 à 500	$X_{aud} = -0.5,$ $Y_{aud} = -1$	$Y_{aud} = -1.5$

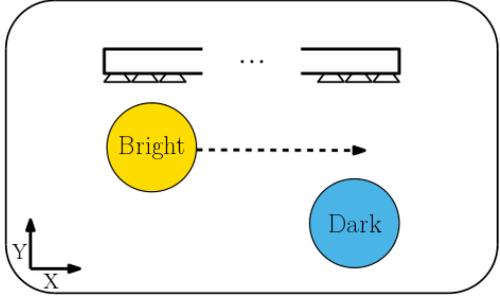


Figure 4.14 – Positions de la *Bright zone* et de la *Dark zone* pour l'évaluation de l'évolution des performances en fonction du déplacement de l'auditeur, des algorithmes de *Sound Zones* pour différentes méthodes de calcul de la *Dark zone*. L'utilisateur se déplace selon l'axe X . Le déplacement de la *Bright zone* se produit entre l'itération 200 et l'itération 300. La *Bright zone* passe devant la *Dark zone* et s'y arrête.

Les deux méthodes ont un comportement similaire au cours des itérations. On note une baisse conséquente de contraste lorsque la *Bright zone* se rapproche puis passe devant la *Dark zone*. Ce phénomène connu pour la méthode Directe reste présent dans les mêmes proportions dans le cas de la méthode VLLMS, avec des performances similaires à celles de la méthode Directe. Ce résultat n'est pas surprenant car les fonctions de coût des deux algorithmes sont identiques et donnent donc des résultats similaires pour ces conditions spécifiques. Cependant, il est à noter que l'algorithme VLLMS reste stable dans ces conditions et ne nécessite pas de régularisation supplémentaire.

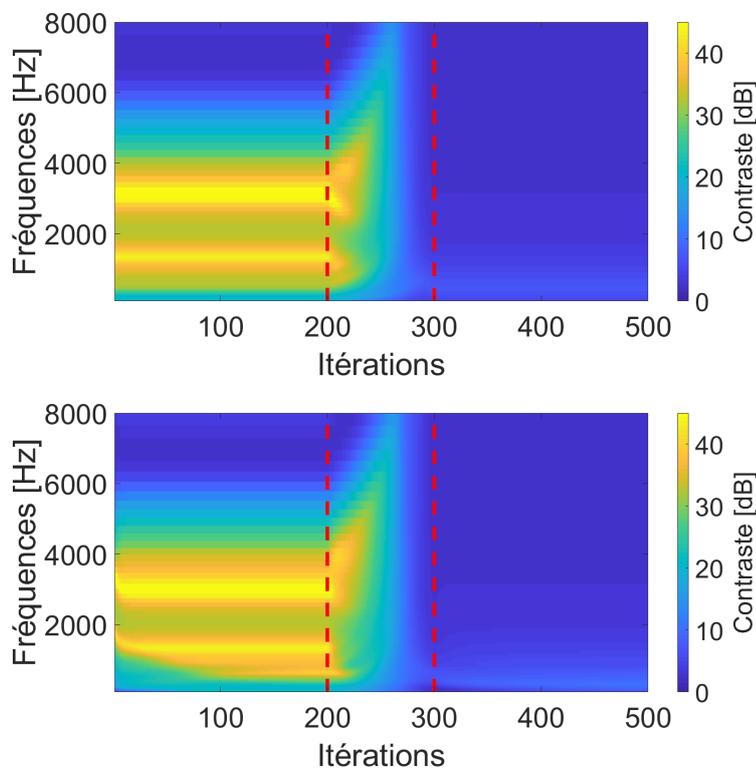


Figure 4.15 – Contraste au cours des itérations pour la méthode Directe (Haut) et la méthode VLLMS (Bas) pour un déplacement le long de l'axe X avec une occlusion asymétrique. Le déplacement de la *Bright zone* se produit entre l'itération 200 et l'itération 300.

Pour ce qui est de l'erreur de reproduction (**Figure 4.16**), on peut faire le même constat que pour le contraste. Dans ce cas, l'augmentation de l'erreur de reproduction à la position finale est due à la déformation du champ sonore dans la *Bright zone* induite par l'algorithme dans le but d'obtenir le meilleur compromis possible entre le contraste et la qualité de reproduction dans la *Bright zone*. La **Figure 4.17** montre le champ sonore reproduit par les deux algorithmes pour différentes positions de l'utilisateur pour une fréquence de 1 kHz. Puisque les performances en termes de contraste dans ces conditions sont très mauvaises et que la qualité de reproduction de champ sonore s'en ressent, on pourrait raisonnablement vouloir éliminer la *Dark zone* afin d'obtenir une meilleure reproduction du champ sonore et traiter le problème comme celui d'un utilisateur unique en négligeant le second utilisateur. Cependant, ce type de situation reste marginal et ne semble pas nécessiter un traitement spécifique. Il semble plus probable que cette situation soit transitoire lorsqu'un utilisateur passe devant l'autre, correspondant au cas de l'occlusion symétrique.

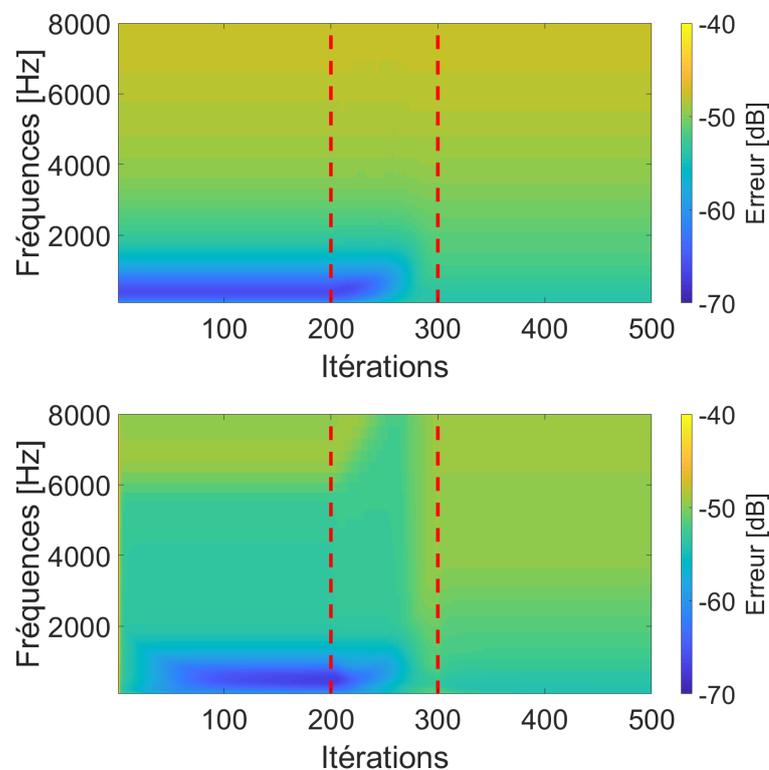


Figure 4.16 – Erreur quadratique moyenne au cours des itérations pour la méthode Directe (Haut) et la méthode VLLMS (Bas) pour un déplacement le long de l'axe X avec une occlusion asymétrique. Le déplacement de la *Bright zone* se produit entre l'itération 200 et l'itération 300.

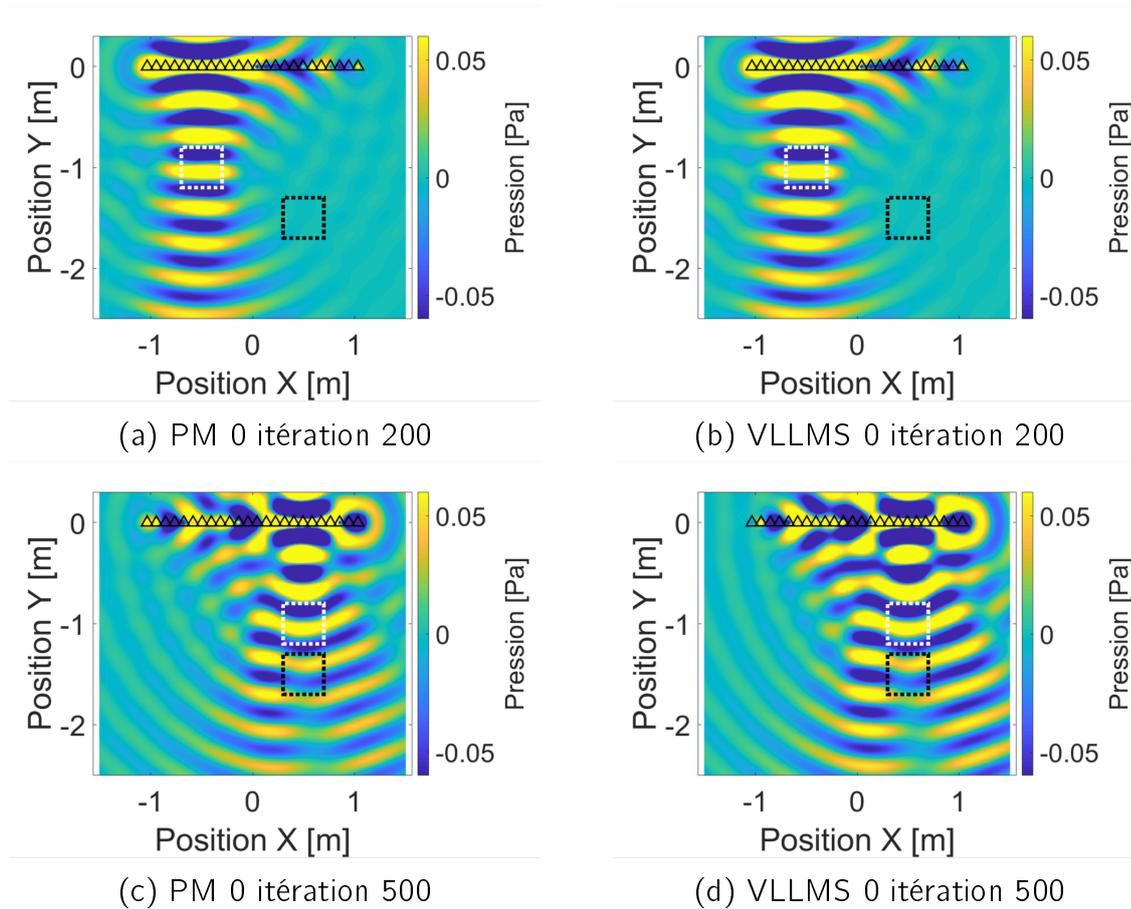


Figure 4.17 – Champ sonore reproduit par les méthodes Directe (Gauche) et VLLMS (Droite) pour différentes positions de l'utilisateur en *Bright zone*. L'itération 200 (Haut) correspond à l'itération juste avant le déplacement de l'utilisateur, l'itération 500 (Bas) correspond à la fin du déplacement de l'utilisateur.

4.3.3 Problème d'occlusion symétrique

Dans ce dernier cas d'étude, l'auditeur de la *Bright zone* se déplace de sa position initiale pour passer devant le second auditeur en *Dark zone* puis s'arrête dans une position symétrique à la position initiale (**Figure 4.18**). L'auditeur se déplace toujours de 1 m en 100 itérations.

La symétrie entre position initiale et position finale de l'utilisateur permet de mesurer les performances de l'algorithme face à une forte perturbation. La **Figure 4.19** compare les performances en termes de contraste entre la *Bright zone* et la *Dark zone* pour la méthode Directe et la méthode VLLMS. Une fois de plus, le comportement entre les deux méthodes reste similaire. Lorsque l'auditeur de la *Bright zone* passe devant la *Dark zone*, le contraste diminue fortement et la qualité du champ sonore en *Bright zone* est fortement dégradée, augmentant donc l'erreur de reproduction avant de se rétablir après le passage de l'utilisateur (**Figure 4.20**). Le cas de l'occlusion symétrique permet de rendre compte que, pour des positions symétriques, les performances sont identiques. Cet effet est d'autant plus intéressant dans le cas de la méthode VLLMS, qui, après avoir été perturbée se rétablit à l'identique et reste comparable à la méthode Directe, ce qui s'observe particulièrement sur les champs sonores à 1 kHz en **Figure 4.21**.

Itérations	Bright	Dark
0 à 199	$X_{aud} = -0.5,$ $Y_{aud} = -1$	$X_{aud} = 0,$
200 à 300	0.01 m/iter	
301 à 500	$X_{aud} = -0.5,$ $Y_{aud} = -1$	$Y_{aud} = -1.5$

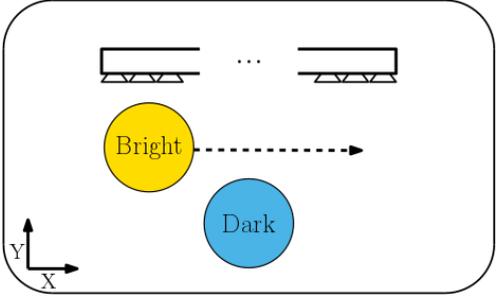


Figure 4.18 – Positions de la *Bright zone* et de la *Dark zone* pour l'évaluation des algorithmes de *Sound Zones*, pour différentes méthodes de calcul de la *Dark zone*. L'utilisateur se déplace selon l'axe X. Le déplacement de la *Bright zone* se produit entre l'itération 200 et l'itération 300. La *Bright zone* passe devant la *Dark zone* et s'arrête à une position symétrique de celle de départ.

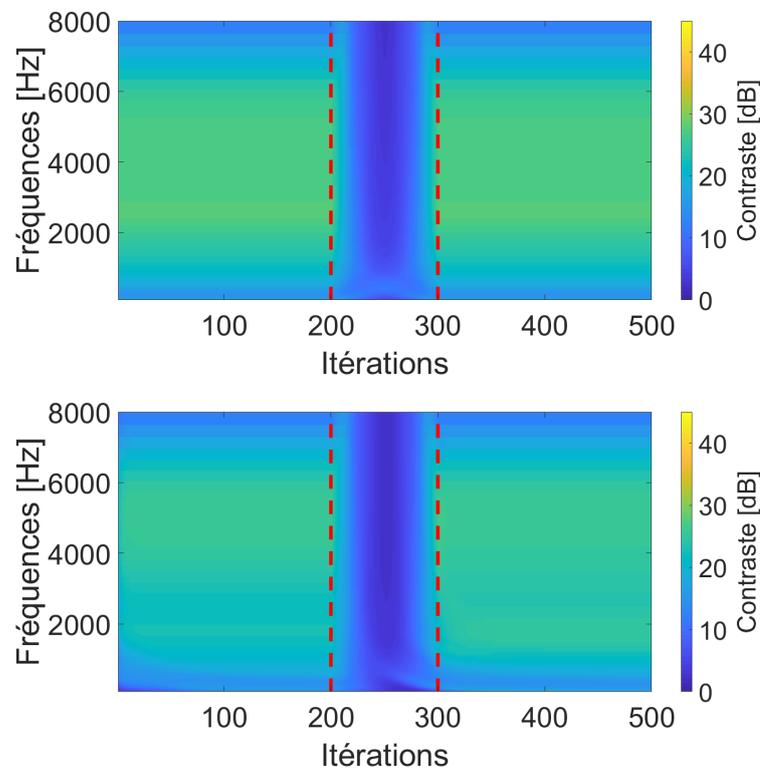


Figure 4.19 – Contraste au cours des itérations pour la méthode Directe (Haut) et la méthode VLLMS (Bas) pour un déplacement le long de l'axe X avec une occlusion symétrique. Le déplacement de la *Bright zone* se produit entre l'itération 200 et l'itération 300. Les positions initiale et finale de la *Bright zone* sont symétriques par rapport au centre du réseau de haut-parleurs.

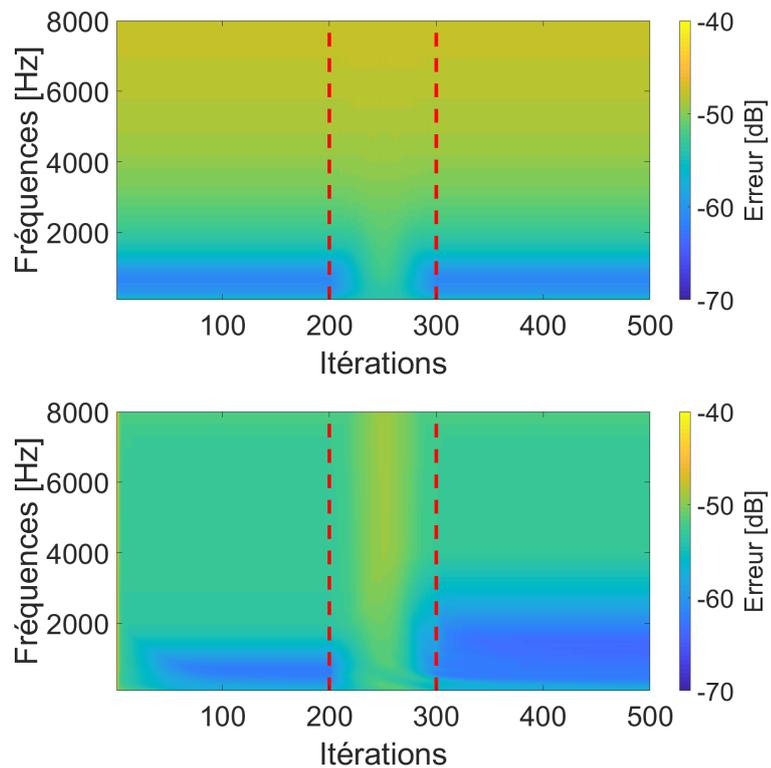


Figure 4.20 – Erreur quadratique moyenne au cours des itérations pour la méthode Directe (Haut) et la méthode VLLMS (Bas) pour un déplacement le long de l'axe X avec une occlusion symétrique. Le déplacement de la *Bright zone* se produit entre l'itération 200 et l'itération 300. Les positions initiale et finale de la *Bright zone* sont symétriques par rapport au centre du réseau de haut-parleurs.

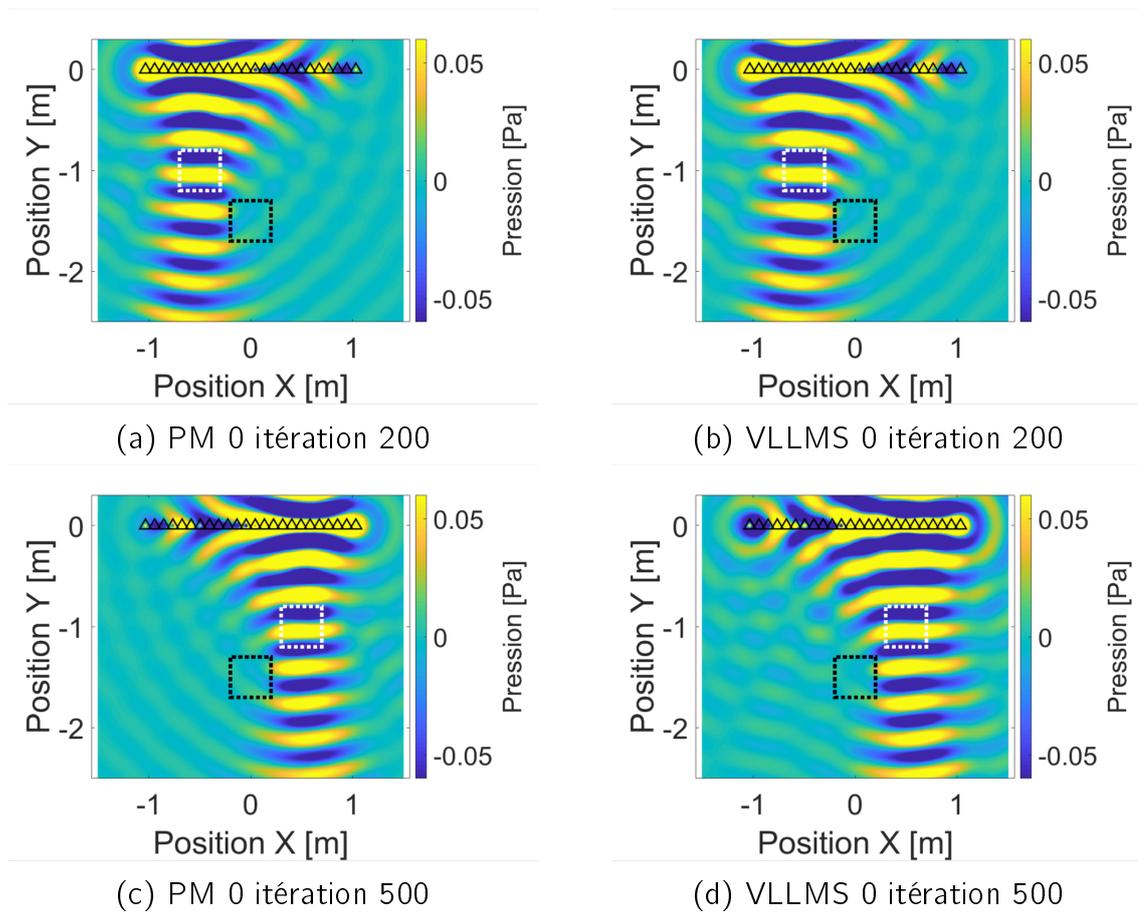


Figure 4.21 – Champ sonore reproduit par les méthodes Directe (Gauche) et VLLMS (Droite) pour différentes positions de l'utilisateur en *Bright zone*. L'itération 200 (Haut) correspond à l'itération juste avant le déplacement de l'utilisateur, l'itération 500 (Bas) correspond à la fin du déplacement de l'utilisateur.

Conclusion

Les méthodes de résolution VLLMS proposées dans le chapitre précédent ont pu être mises en place dans le cas du problème global de *Sound Zones* pour des auditeurs en mouvement et leurs performances ont été comparées dans différentes conditions.

Dans un premier temps, différentes stratégies de contrôle de la *Dark zone* ont été évaluées en conditions statiques. La méthode PM définit la *Dark zone* comme un champ sonore nul. La méthode WLS s'appuie sur un compromis entre qualité de reproduction et contraste. Enfin, les méthodes avec *Dark zone* itératives DZIG et DZII utilisent une version atténuée du champ sonore déjà présent en *Dark zone*. Ces méthodes sont flexibles et peuvent s'adapter à différents problèmes de *Sound Zones* de façon à améliorer les performances par rapport à la méthode PM.

Les méthodes DZII et DZIG présentent l'avantage de pouvoir régler le niveau de contraste entre les zones, de sorte à obtenir un contraste constant sur une certaine bande de fréquences grâce à la limitation du contraste cible. Par ailleurs, l'utilisation de la méthode DZII permet de définir le niveau sonore indépendamment pour chaque microphone du réseau de la *Dark zone*, ce qui tend à renforcer l'homogénéité de la *Dark zone*. Malgré les avantages des méthodes DZII et DZIG, lorsque le contraste cible L_C est trop élevé, elles se comportent de la même façon que les méthodes conventionnelles

définies par un champ nul en *Dark zone*.

Par la suite, une évaluation de la méthode VLLMS dans le cas d'utilisateurs mobiles a été réalisée. On a vu qu'elles obtiennent des résultats similaires à la méthode Directe de référence mais avec un coût de calcul réduit. Un des principaux problèmes induits par le déplacement des utilisateurs est la différence de performance des *Sound Zones* en fonction de la position relative des utilisateurs. Cette différence de performance ne permet pas d'obtenir un contraste homogène sur l'ensemble de l'espace de reproduction. Cependant, l'utilisation des méthodes de type DZII ou DZIG pourrait homogénéiser ce contraste sur l'ensemble de l'espace de reproduction en définissant un contraste cible plus faible mais constant sur l'ensemble de l'espace de reproduction. Cette utilisation des méthodes DZII et DZIG réduirait alors l'impact de la variation de contraste liée à la variation de position des utilisateurs. Cependant, la détermination de la valeur de contraste cible pour l'ensemble de l'espace de reproduction nécessite la définition d'un critère d'évaluation approprié, dépendant notamment de la configuration du système, des positions relatives des zones entre elles et par rapport au réseau de haut-parleurs. Cette hypothèse n'a pas été vérifiée et nécessiterait une étude à part entière.

Par ailleurs, il a été observé que la méthode Directe et la méthode VLLMS donnent des performances similaires y compris dans les cas d'occlusion dits symétriques ou asymétriques. Ces cas d'occlusion définissent des cas d'utilisations extrêmes qui confirment la stabilité de l'algorithme VLLMS dans ces conditions.

La méthode de résolution VLLMS permet donc bien de réduire le coût de calcul lié à l'utilisation des *Sound Zones* dans le cas d'utilisateurs mobiles, tout en disposant de nombreuses possibilités d'amélioration et avec des performances similaires à la méthode Directe. Des études complémentaires restent nécessaires pour non seulement pour améliorer, mais surtout évaluer les performances sur un plus large panel de configurations, représentatives des conditions réelles d'utilisation des *Sound Zones* pour les utilisateurs en mouvement. Des pistes à ce sujet sont proposées dans le chapitre suivant.

Principales avancées

- **Evaluation des stratégies de contrôle de Dark zone développées** au chapitre 2 pour les méthodes **WLS, DZIG et DZII** associées à l'algorithme **VLLMS**.
- Mise en évidence d'une **amélioration de l'homogénéité de la Dark zone** par les méthodes **DZIG et DZII**.
- Mise en évidence de l'intérêt de la méthode **VLLMS pour suivre les auditeurs** dans le cas de *Sound Zones*.
- Mise en évidence de la stabilité de la méthode **VLLMS pour les cas d'occlusion**

Chapitre 5

Discussions et Perspectives

Les chapitres précédents ont présenté le travail de thèse portant sur la mise au point de méthodes adaptatives pour la reproduction de champs sonores multizones pour des auditeurs en mouvement. Avant de conclure, ce chapitre de synthèse d'une part, récapitule les solutions proposées et d'autre part discute des perspectives en vue de les améliorer. Plusieurs aspects sont abordés : d'une part des solutions alternatives de résolutions adaptatives, et d'autre part, un certain nombre de questions est soulevé quant à la réalisation pratique de telles méthodes. Nous apportons dans ce chapitre des éléments de réflexion portant à la fois sur leur mise en place dans un contexte d'interaction avec un effet de salle ainsi que sur les aspects perceptifs.

5.1 Bilan des différentes solutions explorées dans les travaux de thèse

La méthode de résolution Directe du problème PM (*Pressure Matching*) de *Sound Zones* pour des auditeurs mobiles, présentée dans les chapitres 3 et 4, ne permet pas une utilisation dans le cadre d'un problème variant en temps-réel, à cause de sa complexité de calcul. D'une part, la reproduction du champ sonore nécessite un grand nombre de haut-parleurs et de microphones pour obtenir une reproduction satisfaisante du champ sonore. D'autre part, l'ajustement du paramètre de régularisation peut être long et doit être calculé indépendamment pour chaque fréquence, ce qui est incompatible avec une utilisation en temps réel. Par ailleurs, dans le cas d'une utilisation pour un ou plusieurs auditeurs en mouvement, l'ensemble du calcul doit être réalisé à nouveau pour chaque nouvelle position.

Les méthodes adaptatives, à la différence de la méthode Directe, utilisent quant à elles le résultat des calculs des itérations précédentes pour calculer la solution du nouveau problème de l'itération en cours, ce qui tend à alléger le coût de calcul. Les méthodes NLMS (*Normalized Least Mean Square*) et APA (*Affine Projection Adaptive*), malgré leur convergence rapide, ne permettent cependant pas de réduire ce coût de calcul en étant au minimum équivalent à celui de la méthode de résolution Directe. On notera d'ailleurs le cas du filtre APA dont la complexité algorithmique est même supérieure à la méthode Directe (§2.2). Par ailleurs, dans le cas des méthodes NLMS et APA, il est également nécessaire de calculer un paramètre de régularisation de la même façon que pour la méthode Directe.

Pour ce qui est de la méthode LMS et de ses dérivées, elles tirent parti de l'utilisation du calcul réalisé aux itérations précédentes et dispose d'une complexité algorithmique

plus faible que la méthode Directe. De plus, cette méthode ne nécessitant pas de régularisation en première approximation, elle est donc beaucoup mieux adaptée à une utilisation temps réel. Cependant, cette méthode ne résout pas tous les problèmes. D'une part, elle nécessite le réglage du pas μ_L de l'algorithme qui assure sa stabilité et régule sa vitesse de convergence. Ce paramètre, directement relié à la valeur propre maximale de la matrice de corrélation spatiale $\mathbf{G}^H\mathbf{G}$, nécessite par ailleurs d'être calculé à chaque position, de sorte à rendre l'algorithme stable et tout en conservant une bonne convergence sur l'ensemble de l'espace. L'ensemble des études menées dans les chapitres précédents considère que ce paramètre est choisi de façon quasi optimale pour chaque pas (la valeur propre de la matrice $\mathbf{G}^H\mathbf{G}$ est calculée à chaque itération). Le coût de calcul lié à l'estimation de ce paramètre n'est pas négligeable et constitue donc un point bloquant à résoudre. Plusieurs solutions sont proposées en §5.2.

D'autre part, cette méthode ne dispose pas d'une limitation de l'effort et pose de sérieux problèmes quant à l'effet mémoire induit par le déplacement de l'auditeur. Nous avons vu que ces défauts pouvaient être réduits en introduisant un facteur d'oubli (méthode Leaky LMS), conduisant à la méthode VLLMS qui s'avère être une solution satisfaisante à notre problème. Dans ce qui suit, nous allons nous focaliser sur cette dernière. De plus, il est à noter que la méthode proposée est une méthode de résolution qui peut aussi être utilisée pour des méthodes de *Sound Zones* autres que la méthode PM.

5.2 Amélioration de la solution de la méthode PM associée à une résolution VLLMS

5.2.1 Convergence de l'algorithme VLLMS

A l'instar de la méthode LMS, la méthode VLLMS nécessite l'ajustement d'un paramètre de pas μ_L qui doit être calculé de sorte à assurer la stabilité du filtre et obtenir une vitesse de convergence acceptable. Pour rappel, ce paramètre détermine à la fois la vitesse de convergence de l'algorithme ainsi que la stabilité du filtre. Dans l'ensemble des études réalisées, il est conditionné par la plus grande valeur propre κ_{max} de la matrice de corrélation spatiale $\mathbf{G}^H\mathbf{G}$ et doit vérifier

$$\mu_{Lmax} < \frac{2}{\kappa_{max}}. \quad (5.1)$$

Le calcul de la corrélation spatiale peut être coûteux en comparaison du reste de l'algorithme, d'autant plus qu'il doit être réalisé pour chaque pas et pour chaque fréquence dans le cas de l'utilisation dans un contexte mobile.

Une première solution pour réduire ce coût consiste à choisir un paramètre unique pour l'ensemble de la zone de reproduction, ce qui permet de ne le calculer qu'une seule fois, lors de l'initialisation de la méthode. Il faut cependant que cette valeur assure la stabilité du calcul sur l'ensemble de la zone de reproduction. Pour cela il faut choisir la valeur propre maximale de la matrice $\mathbf{G}^H\mathbf{G}$ appropriée sur l'ensemble de la zone de reproduction. Dans notre configuration, cette valeur est maximale à proximité du réseau de haut-parleurs. Le pas de l'algorithme est très faible pour l'ensemble de la zone de reproduction, ce qui induit une faible vitesse de convergence. Cependant, en s'éloignant du réseau de haut-parleurs l'algorithme reste toujours stable.

Une autre solution consiste à réaliser un dictionnaire de ces valeurs propres pour différentes positions du réseau de microphones. On calcule donc au préalable les valeurs propres pour un ensemble de positions. Il n'est pas nécessaire que ce dictionnaire soit réalisé sur l'intégralité des positions occupées par le réseau de microphones. En effet, seules quelques valeurs prélevées judicieusement sur la zone à reproduire peuvent être suffisantes pour obtenir un compromis entre stabilité et vitesse de convergence.

Une dernière solution consiste à calculer les valeurs propres de façon itérative. Entre deux positions proches de l'espace, les valeurs propres changent a priori assez peu. Il est donc possible de calculer la valeur propre à la nouvelle position à partir de la connaissance de celle de l'ancienne position. Il existe de nombreuses méthodes de calcul de valeurs propres de façon itérative [Chatelin, 2012, Bhaya et Kaszkurewicz, 2006] telle que la méthode de la puissance itérée [Golub et Van Loan, 2012], la méthode d'itération du quotient de Rayleigh [Batterson et Smillie, 1989], méthodes dont la pertinence pour la résolution de notre problème doit cependant être évaluée.

5.2.2 Sélection des paramètres VLLMS

Le réglage de la méthode VLLMS ne se limite pas à la sélection du paramètre μ_L de l'algorithme. En effet, plusieurs paramètres supplémentaires sont à ajuster : d'une part le facteur d'oubli γ_{max} , qui permet de réduire l'effet mémoire de l'algorithme, et d'autre part les paramètres l_u , l_d et α qui définissent la vitesse à laquelle la méthode VLLMS passe du mode LMS au mode Leaky LMS et inversement par le biais de la variation de $\gamma(n)$. Comme montré dans le chapitre 3, ces valeurs ont une influence sur le champ sonore reproduit ainsi que sur l'effort de réseau. Pour rappel, l_u définit le pas de montée vers le filtre Leaky LMS et l_d le pas de retour vers le filtre LMS. Le paramètre α influe sur la vitesse de passage de l'un à l'autre. Lors de fortes accélérations, il est intéressant d'augmenter l_u ou α pour que les poids des itérations précédentes soient oubliés rapidement puisque le problème à résoudre change rapidement d'une itération à l'autre. D'autre part, lorsque le mouvement de l'auditeur dure peu de temps, il est intéressant d'augmenter la valeur de l_d et de diminuer la valeur de α , car il est possible que les poids en mémoire n'aient pas eu le temps d'être oubliés lors du court déplacement. Il faut donc laisser le temps à l'algorithme de les oublier. Pour l'instant aucune méthode n'existe pour régler les paramètres l_u , l_d et α de façon optimale.

La solution proposée dans le chapitre 2 lie les paramètres α , l_u et l_d au déplacement de l'utilisateur. Cependant, cette approche reste incomplète et nécessite la prise en compte de la vitesse de déplacement de l'utilisateur. Afin de déterminer les valeurs de l'ensemble des paramètres de la méthode VLLMS, il serait utile de définir le lien entre la qualité de reproduction du champ sonore, les paramètres de la méthode VLLMS et la vitesse de déplacement de l'utilisateur. Cet aspect nécessite une étude approfondie afin d'établir ce lien, nécessitant potentiellement la prise en compte d'effets perceptifs.

Une autre solution permettant d'éviter le réglage de ces paramètres consisterait à utiliser uniquement la méthode Leaky LMS au prix d'un biais sur la reproduction du champ sonore. Cependant, ce biais pourrait être réduit par l'utilisation de méthodes différentes du filtre VLLMS, comme *Circular Leaky LMS* ou *Max-Tap Leaky LMS*, avec un coût de calcul supplémentaire relativement faible [Nascimento et Sayed, 1996]. Ces solutions semblent être une bonne alternative à la sélection complexe des paramètres de la méthode VLLMS mais n'a pas été évaluée dans le cadre de ce travail.

5.3 Vers d'autres solutions

Si les solutions proposées dans ce travail de thèse constituent un choix assumé, d'autres solutions auraient pu être envisagées. Cette partie présente un ensemble de solutions alternatives à la fois sur la méthode de reproduction de *Sound Zones*, et sur la méthode de résolution du problème pour une application en temps réel.

Par ailleurs, tout au long du travail de thèse, l'hypothèse d'une propagation en champ libre a été retenue. Il faut cependant avoir conscience que les *Sound Zones* sont destinées à être mises en œuvre dans des salles où l'utilisation de méthodes itératives, telle que la méthode VLLMS, peut s'avérer problématique en raison de l'interaction avec l'effet de salle. Il est donc intéressant de voir comment il est possible d'aborder ce problème en utilisant les Réponses Impulsionnelles (RI) de la salle ou en cherchant des solutions alternatives.

De plus, même si les méthodes de résolution de type VLLMS constituent une réduction conséquente du coût de calcul, la méthode PM nécessite un grand nombre de microphones pour contrôler la reproduction du champ sonore dans une zone. Plusieurs solutions sont proposées pour réduire ce nombre.

5.3.1 Hybridation avec d'autres méthodes de Sound Zones

La méthode PM de reproduction de champ sonore, choisie dans cette thèse, offre une bonne qualité de reproduction de champ sonore et reste performante dans le cas des *Sound Zones*. Cependant, ce n'est pas la seule méthode disponible (Voir chapitre 1). D'autres méthodes peuvent s'avérer avantageuses par certains aspects spécifiques dans le cas des *Sound Zones* mobiles. C'est par exemple cas de la méthode de *Planarity Control* (PC), qui consiste à reproduire le mieux possible une onde plane en choisissant un ensemble de directions de propagation, et qui présente l'avantage de pouvoir reproduire un contenu audio sur une zone de l'espace sans définir explicitement un champ sonore, mais en définissant un ensemble de directions d'arrivée. Cette définition plus flexible permet de réduire les effets de l'occlusion [Coleman, 2014]. Dans le cas d'un réseau linéaire, il est particulièrement intéressant d'avoir une flexibilité de la direction d'arrivée de l'onde car les cas d'occlusion sont plus fréquents que pour des réseaux de haut-parleurs circulaires.

5.3.2 Autres méthodes itératives

Les méthodes adaptatives pour résoudre un problème des moindres carrés, telle que la méthode PM, sont nombreuses [Haykin, 2002]. Leurs caractéristiques en termes de complexité de calcul, de convergence et de stabilité peuvent être très différentes. Les méthodes proposées dans ce document sont a priori celles qui disposent de la plus faible complexité algorithmique et ont été choisies pour cette raison. Cependant, ce ne sont ni les méthodes les plus stables ni les méthodes ayant la meilleure convergence. Il pourrait être intéressant de formuler le problème en utilisant par exemple une résolution par la méthode RLS (Recursive Least Squares) [Haykin, 2002] ou encore au moyen d'un filtre de Kalman [Haykin, 2002]. Malgré leur coût de calcul supérieur aux méthodes utilisées ici, ces méthodes restent très intéressantes dans des applications temps réel de problèmes complexes. Il semble aussi possible d'utiliser des méthodes de compensation de salle telle que la méthode WDAF (Wave Domain Adaptive Filtering) [Gauthier et Berry, 2006, Spors *et al.*, 2007] afin de les adapter au problème des *Sound Zones* mobiles en

intégrant une *Dark zone* dans le problème de reproduction de champ sonore. Cette méthode décompose le problème de sorte à réduire sa complexité en découplant le système total de dimension $M \times L$ en un ensemble de $\min(M, L)$ problèmes indépendants. Un des problèmes inhérents à cette méthode est la nécessité de calcul d'une décomposition aux valeurs propres (*Singular Value Decomposition*) qui peut elle aussi être coûteuse en temps de calcul [Gauthier et Berry, 2006].

5.4 Interaction des méthodes dynamiques avec l'effet de salle

L'ensemble du travail de thèse s'est focalisé sur l'évaluation des méthodes de *Sound Zones* en champ libre. Cependant, lorsque les *Sound Zones* définies en champ libre sont appliquées directement à la diffusion dans une salle, leurs performances sont altérées. Ceci est particulièrement visible sur le contraste [Olik *et al.*, 2013]. L'utilisation des méthodes de *Sound Zones* pour des auditeurs en mouvement et en interaction avec un effet de salle pose donc un ensemble de problèmes supplémentaires.

Afin de résoudre ce problème, une solution souvent utilisée est la mesure des Réponses Impulsionnelles (RI) aux positions d'écoute dans la salle de diffusion [Coleman, 2014, Betlehem *et al.*, 2015], ce qui permet de compenser en partie les effets de la salle. Cependant, les conditions acoustiques de la salle peuvent varier au cours du temps (température, hygrométrie, etc.) et il est donc nécessaire de réaliser régulièrement les mesures. Il existe un grand nombre de méthodes de compensation de salle utilisées pour les *Sound Zones* statiques, comme par exemple les méthodes de remodelage de réponses impulsionnelles (*Impulse Response Reshaping*) [Jungmann *et al.*, 2012] ou encore la méthode WDAF [Gauthier et Berry, 2006, Spors *et al.*, 2007].

Néanmoins, il n'est pas possible d'utiliser directement ces méthodes dans le cas d'utilisateurs mobiles, et ce, pour plusieurs raisons. Premièrement les zones de reproduction étant variables au sein de l'espace, il est nécessaire de réaliser une cartographie précise des RI de l'ensemble de la salle, ce qui peut être fastidieux. De plus, de la même façon que pour les *Sound Zones* statiques, lors d'un changement dans la salle, ces mesures sont à refaire, même si les méthodes présentent une certaine robustesse [Coleman *et al.*, 2013]. Par ailleurs, il n'est pas non plus envisageable de réaliser la mesure en continu aux emplacements des utilisateurs car cela nécessiterait un dispositif microphonique directement placé sur l'auditeur, ce qui risque de limiter la liberté de mouvement de l'auditeur et de réduire par conséquent l'intérêt des *Sound Zones* mobiles. Plusieurs solutions sont néanmoins possibles. Par exemple, les méthodes basées sur les fonctions plénacoustiques [Ajdlar *et al.*, 2006, Mignot *et al.*, 2012] permettent la cartographie complète des RI d'une salle à l'aide d'un nombre réduit de mesures grâce à des méthodes d'interpolation proche des approches de *compressive sensing*. Une fois ces mesures obtenues, il devient alors possible d'utiliser les méthodes usuelles de compensation.

Une autre solution résiderait dans l'utilisation de microphones de contrôle déportés. Cette méthode a été mise en œuvre dans le cas de zones de silence pour des environnements peu réfléchissants [Boulandet *et al.*, 2017]. Elle consiste à définir directement le problème d'optimisation en prenant en compte les RI à une position différente de la zone de reproduction et donc d'intégrer la correction de position à la méthode de *Sound Zones* ainsi reformulée. L'avantage de cette méthode réside dans le fait que son

intégration est directement prise en charge dans le calcul des *Sound Zones*, mais elle nécessite la connaissance des fonctions de transfert entre les microphones de contrôle et les zones de reproduction. Dans le cas d'une salle, ces fonctions de transfert ne sont pas triviales et doivent être mesurées ou modélisées.

5.5 Dispositif microphonique pour le contrôle du champ sonore

La méthode de *Sound Zones* PM nécessite un grand nombre de microphones pour reproduire correctement le champ sonore dans la *Bright zone*. Il importe donc de réfléchir au dispositif microphonique nécessaire au contrôle du champ reproduit (nombre de capteurs nécessairement limité, mise en place discrète et transparente pour les utilisateurs, etc.). Or, le temps de calcul de la méthode VLLMS est directement relié au nombre de microphones. Le nombre de multiplications à réaliser à chaque itération est proportionnel au nombre de microphones et de haut parleurs utilisés (**Annexe C**). Avec un nombre trop élevé de microphones, on risque de perdre les bénéfices de la faible complexité algorithmique de la méthode VLLMS. Il semble donc judicieux d'en réduire leur nombre afin de diminuer le coût de calcul et de permettre le fonctionnement de la méthode avec des ressources limitées. Le nombre de microphones peut être réduit par différentes méthodes. Une méthode triviale consiste à ramener le nombre de microphones à son strict nécessaire, c'est-à-dire un seul microphone par zone, comme ce qui a été proposé par [Simón Gálvez et Fazi, 2016]. Cette solution est d'ailleurs suffisamment peu coûteuse pour utiliser la méthode Directe qui n'a pas de besoin important en termes de régularisation dans cette configuration. Cependant, elle ne permet pas le contrôle d'un champ sonore sur l'ensemble d'une zone mais uniquement en une seule position de l'espace, ce qui peut être suffisant dans certains cas.

La réduction du nombre de microphones peut aussi passer par l'utilisation de méthodes plus complexes. Il existe par exemple des méthodes de *compressive sensing* [Jin et Kleijn, 2014] qui réduisent le nombre d'informations nécessaires à connaître sur le champ sonore tout en gardant une bonne résolution sur ce dernier grâce à des méthodes d'interpolation.

Une autre possibilité est de s'appuyer, dans la *Bright zone*, sur les méthodes de reproduction de champ sonore. Puisqu'il est possible de contrôler un champ sonore à l'intérieur d'un volume uniquement en connaissant ce qui se passe sur sa frontière (§1.1.4), il semble possible de contrôler le champ sonore dans une zone en ne le contrôlant qu'avec des micros en périphérie de la zone. Les méthodes de type HOA ou WFS pourraient être utilisées en l'occurrence. Plus précisément, pour un champ sonore donné dans la zone, on cherche dans un premier temps à connaître les valeurs de pression à obtenir sur les microphones situés en périphérie de celle-ci par des méthodes de type WFS ou HOA. Ces valeurs définissent alors les pressions P_t à reproduire par la méthode de *Sound Zones*. Les haut-parleurs sont alors contrôlés, par exemple, à l'aide de la méthode VLLMS. Cependant, même si ces méthodes réduisent le nombre de microphones utilisés, le coût de calcul ajouté doit être pris en compte. Une comparaison entre les différentes approches doit être réalisée afin de déterminer un compromis entre le nombre de microphones nécessaires, la qualité de reproduction du champ sonore et le coût de calcul.

5.6 Quid de la perception ?

Pour finir, il ne faut pas oublier que le champ sonore reproduit est destiné à être écouté. Au-delà de la caractérisation des ondes acoustiques reproduites, il importe d'évaluer comment elles sont perçues, d'autant que l'expérience d'écoute est assez particulière (focalisation dans une bulle sonore, contenu stable malgré le déplacement, présence potentielle d'artefacts).

L'ensemble du document s'est uniquement intéressé à la reproduction physique du champ sonore sans prendre en compte l'aspect perceptif des méthodes évaluées. Les systèmes dynamiques de reproduction audio et plus particulièrement les *Sound Zones* pour les utilisateurs mobiles changent rapidement de filtre lors du déplacement de l'utilisateur. Ce changement de filtre peut être brutal dans le cas de la méthode Directe et produire des artefacts audibles. Afin de limiter ces artefacts, [Simón Gálvez et Fazi, 2016] proposent d'utiliser une méthode *overlap-save* modifiée de sorte à réduire les artefacts et rendre le changement de filtre plus progressif [Wefers et Vorländer, 2014]. Dans le cas des méthodes itératives, si le changement de filtre est plus progressif grâce au temps de convergence de l'algorithme, il est possible que les artefacts soient moins perceptibles que dans le cas de la méthode Directe. Il est donc nécessaire d'évaluer cet aspect afin de garantir une écoute confortable.

Par ailleurs, une étude perceptive sur le fonctionnement en temps réel de la méthode VLLMS pourrait mettre en évidence un ensemble d'artefacts perceptibles liés à la reproduction itérative et qui n'auraient pas été relevés lors de l'étude des problèmes du champ sonore. De plus, la prise en compte de la perception peut permettre de revisiter, à l'aune des seuils de perception, les contraintes appliquées au problème de reproduction de *Sound Zones* mobiles, de manière similaire aux travaux réalisés par [Francombe *et al.*, 2012, Francombe *et al.*, 2013b, Francombe *et al.*, 2013a] dans le cas des méthodes de *Sound Zones* statiques et qui ont permis de déterminer des seuils d'acceptabilité de gêne entre deux programmes audio.

Enfin, il a été suggéré dans le Chapitre 4 que les méthodes permettant un contraste constant sur l'ensemble de la zone de reproduction telles que les méthodes DZII et DZIG sont potentiellement plus confortables que les méthodes usuelles. Cette hypothèse se base sur le fait que s'il existe une variation dans la gêne occasionnée par le contenu concurrent, lié au déplacement de l'auditeur, l'expérience globale peut être altérée. Cette hypothèse mérite d'être vérifiée pour conclure sur l'intérêt de ces méthodes.

Chapitre 6

Conclusion

Les méthodes de reproduction de champ sonore sont à même de reproduire des champs sonores complexes à l'aide de réseaux de haut-parleurs et peuvent être adaptées à différents cas de figures. Cette thèse s'est intéressée en particulier à l'utilisation de ces méthodes dans le cas de *Sound Zones*. Les méthodes de *Sound Zones* rendent possible l'écoute d'un contenu audio indépendant pour chaque auditeur en maximisant le niveau sonore dans une zone de l'espace tout en réduisant l'impact de ce dernier dans d'autres zones de l'espace. Ces méthodes ont pour effet de maximiser le contraste de niveau sonore entre plusieurs zones de l'espace. Un grand nombre de ces méthodes a été développé aux cours de la dernière décennie et leur implémentation en conditions réelles est déjà bien maîtrisée pour les auditeurs statiques. La méthode sélectionnée pour calculer les poids des *Sound Zones* est la méthode PM permettant d'avoir un bon compromis entre la qualité de reproduction du champ sonore et le contraste entre les zones.

Cette thèse a cherché à étendre les possibilités des *Sound Zones* au cas d'auditeurs se déplaçant. Cependant, le coût de calcul élevé des méthodes de résolution usuelles rend difficile leur utilisation dans un contexte temps réel, nécessaire à terme à l'implémentation de *Sound Zones* pour des auditeurs en mouvement. La résolution usuelle s'appuie sur le calcul direct de l'inversion des matrices de corrélation spatiale $\mathbf{G}^H\mathbf{G}$ et nécessite le calcul d'un paramètre de régularisation. Quelques implémentations en temps réel ont tout de même été réalisées mais avec de fortes restrictions sur le champ sonore reproduit.

La principale contribution de cette thèse est d'introduire et d'évaluer, à l'aide de simulations en conditions anéchoïques, un ensemble d'algorithmes itératifs, dans le but de réduire le coût de calcul du problème afin de l'appliquer à l'utilisation de *Sound Zones* pour des auditeurs mobiles dans un contexte temps réel. Ces travaux ont donné lieu au dépôt de deux brevets.

Dans un premier temps, trois algorithmes adaptatifs, LMS, NLMS et APA, ont été reformulés pour permettre la reproduction d'un champ sonore destiné à un unique utilisateur mobile. A la différence de la méthode Directe habituellement utilisée, ces algorithmes résolvent le problème en s'appuyant sur le résultat des itérations précédentes. Nous avons mis en avant la capacité de tous ces algorithmes à résoudre le problème PM. A l'instar de la méthode Directe, les algorithmes NLMS et APA nécessitent une inversion de matrice qui doit être régularisée ce qui ne leur confère aucun avantage sur cette dernière. La méthode LMS quant à elle permet de réduire le coût de calcul par rapport à la méthode Directe. Elle nécessite cependant le réglage d'un paramètre de pas μ_L dépendant des valeurs propres de la matrice de corrélation spatiale $\mathbf{G}^H\mathbf{G}$, influençant

conjointement la vitesse de convergence et la stabilité de l'algorithme. Malgré son coût de calcul relativement faible par rapport à la méthode Directe, la méthode LMS occasionne une augmentation de l'erreur de reproduction lors du déplacement de l'auditeur. Par ailleurs, il a été mis en évidence la présence d'un effet mémoire de ces algorithmes ayant pour conséquence une augmentation de l'effort de reproduction. De plus, la stabilité sur le long terme de ces algorithmes dans le contexte de la reproduction de champs sonores pour un unique utilisateur mobile est également limitée. Afin de résoudre les problèmes liés à l'effet mémoire ainsi qu'à la stabilité à long terme des algorithmes, nous avons introduit les méthodes Leaky LMS et VLLMS qui sont caractérisées par leur capacité à oublier une partie des poids de l'itération précédente, mais qui nécessitent le réglage de paramètres supplémentaires.

Dans un second temps, nous avons appliqué la méthode VLLMS dans le contexte des *Sound Zones* pour plusieurs auditeurs en déplacement, ce qui a fait l'objet d'un premier brevet. La comparaison des performances entre la méthode Directe et la méthode VLLMS a montré la capacité de la méthode VLLMS à résoudre le problème PM avec des performances similaires à la méthode Directe avec un coût de calcul réduit. De plus, il a été mis en évidence une variabilité des performances de la méthode PM en fonction de la position relative des utilisateurs par rapport au réseau de haut-parleurs. D'autre part, les méthodes itératives rendent possible l'utilisation des pressions présentes dans la *Dark zone* pour son calcul à l'itération suivante. Cela permet alors d'adapter le champ sonore en *Dark zone* de sorte à ne le contraindre qu'en amplitude et ainsi régler le niveau de contraste entre la *Bright zone* et la *Dark zone*. Cette solution a fait l'objet d'un second brevet. Cette méthode semble pouvoir résoudre le problème de la variabilité des performances liées à la position des utilisateurs en contraignant un contraste plus faible mais constant sur l'ensemble de l'espace. Cela nécessite cependant d'être évalué aussi bien sur l'aspect des performances que sur l'aspect perceptif.

Pour conclure, l'algorithme VLLMS tel que proposé dans ce manuscrit est un bon candidat pour l'utilisation de *Sound Zones* dans le cas d'utilisateurs mobiles malgré le nombre élevé de paramètres à régler pour son bon fonctionnement. Cependant, de nombreux aspects doivent encore être améliorés pour assurer l'implémentation temps réel d'une telle méthode. D'une part définir une méthode de calcul du paramètre de pas μ_L de sorte à réduire le coût induit par le calcul des valeurs propres de la matrice de corrélation spatiale et, d'autre part, il est nécessaire de permettre une implémentation de ces méthodes dans des conditions non anéchoïques.

Chapitre 7

Annexes

A Algorithmes pour les Sound Zones

Chaque méthode de *Sound Zones* définit un algorithme pour le calcul des poids \mathbf{q} . Les algorithmes des principales méthodes proposées dans les chapitres 1 et 2 sont présentés dans cette Annexe. Les notations sont les mêmes que celles utilisées dans le corps du document. Les algorithmes des méthodes présentées sont BC, ACC, PM, en résolution Directe et les méthodes LMS, NLMS, APA et VLLMS pour la méthode PM ainsi que la méthode VLLMS DZII sont présentées.

Augmentation du paramètre β pour la régularisation

Les algorithmes Directe, NLMS et APA, comme présenté en chapitre 2 doivent être régularisés. La méthode choisie consiste à augmenter le paramètre de régularisation jusqu'à ce qu'un critère de performance soit respecté. Au cours de cette thèse il a été choisi de ne pas augmenter linéairement le paramètre de régularisation β au cours des itérations de régularisation n_r , de sorte à couvrir rapidement une large gamme de valeurs. Dans ce but β est défini tel que

$$\beta(n_r + 1) = \beta(n_r) + \beta_{step}(n_r). \quad (1.1)$$

Dans l'ensemble du document on choisit :

$$\beta_{step}(n_r) = a_r * b_r^{n_r}, \quad (1.2)$$

avec $a_r = 1e^{-5}$ et $b_r = 1.025$. Le choix a été réalisé empiriquement de sorte à obtenir un compromis entre vitesse de convergence du paramètre de régularisation et respect des contraintes imposées.

A.1 Algorithme Brightness Control

L'algorithme *Brightness Control* (BC) calcule les poids pour un système avec un auditeur unique défini dans une *Bright zone* en choisissant le niveau sonore dans la zone à reproduire. L'effort de réseau et le niveau sonore sont liés.

Algorithme 1 Brightness control

```

1: procedure BC
2:   for all  $\omega$  frequencies do
3:     Calcul de  $\mathbf{G}_B$  pour toute la Bright zone
4:      $\hat{\mathbf{q}} \leftarrow$  vec. propre de la val. propre max de  $[\mathbf{G}_B^H \mathbf{G}_B]$ 
5:      $\hat{\mathbf{p}}_B^H \hat{\mathbf{p}}_B \leftarrow \hat{\mathbf{q}}^H \mathbf{G}_B^H \mathbf{G}_B \hat{\mathbf{q}}$ 
6:      $a \leftarrow \sqrt{\frac{M_B |p_r|^2 \times 10^{\frac{T_B}{10}}}{\hat{\mathbf{p}}_B^H \hat{\mathbf{p}}_B}}$ 
7:      $\mathbf{q}_{\text{opt}}(\omega) \leftarrow a \hat{\mathbf{q}}$ 
8:   end for
9: end procedure

```

A.2 Algorithme ACC

L'algorithme *Acoustic Contrast Control* (ACC) calcule les poids pour une *Bright zone* et une *Dark zone* pour deux auditeurs en choisissant le niveau sonore moyen dans la *Bright zone* et un contraste maximal entre les deux zones en respectant une contrainte sur l'effort de réseau $\mathbf{q}^H \mathbf{q}$ et sur le contraste maximal entre les deux zones. La contrainte de contraste maximal C_{max} permet d'obtenir une réponse fréquentielle plus plate que si l'on permettait à l'algorithme d'optimiser de façon plus performante. Cette valeur est généralement choisie de sorte à ce que $C_{max} = 60dB$.

Algorithme 2 Acoustic Contrast Control

```

1: procedure ACC
2:   for all  $\omega$  frequencies do
3:     Calcul de  $\mathbf{G}_B$  pour toute la Bright zone
4:     Calcul de  $\mathbf{G}_D$  pour toute la Dark zone
5:      $\beta \leftarrow 0$ 
6:     while  $E > E_C$  or  $C_{B \rightarrow D} > C_{max}$  or  $\text{cond}(\mathbf{G}_D^H \mathbf{G}_D + \beta \mathbf{I}) > 10^{10}$  do
7:        $\hat{\mathbf{q}} \leftarrow \text{vec. propre de la val. propre max de } [\mathbf{G}_D^H \mathbf{G}_D + \beta \mathbf{I}]^{-1} [\mathbf{G}_B^H \mathbf{G}_B]$ 
8:        $a \leftarrow \sqrt{\frac{M_B |p_r|^2 \times 10^{\frac{T_B}{10}}}{\hat{\mathbf{q}}^H \mathbf{G}_B^H \mathbf{G}_B \hat{\mathbf{q}}}}$ 
9:        $\mathbf{q} \leftarrow a \hat{\mathbf{q}}$ 
10:       $\mathbf{p}_B^H \mathbf{p}_B \leftarrow \mathbf{q}^H \mathbf{G}_B^H \mathbf{G}_B \mathbf{q}$ 
11:       $\mathbf{p}_D^H \mathbf{p}_D \leftarrow \mathbf{q}^H \mathbf{G}_D^H \mathbf{G}_D \mathbf{q}$ 
12:       $C_{B \rightarrow D} \leftarrow 10 \log \left( \frac{M_D \mathbf{p}_B^H \mathbf{p}_B}{M_B \mathbf{p}_D^H \mathbf{p}_D} \right)$ 
13:       $\beta \leftarrow \beta + \beta_{step}$ 
14:       $E \leftarrow 10 \log_{10} \left( \frac{\mathbf{q}^H(n) \mathbf{q}(n)}{|q_r(n)|^2} \right)$ 
15:     end while
16:      $\mathbf{q}_{opt}(\omega) \leftarrow \mathbf{q}$ 
17:   end for
18: end procedure

```

A.3 Algorithmes PM

L'algorithme *Pressure Matching* (PM) calcule les poids pour une *Bright zone* et une *Dark zone* pour deux auditeurs en choisissant un champ sonore cible à reproduire dans la *Bright zone*. La *Dark zone* peut être définie de façon différente. Soit en définissant un champ sonore nul en *Dark zone*, $\mathbf{p}_T = \mathbf{0}$ (Pressure Matching 1), soit en choisissant un contraste maximal dans la *Dark zone* en définissant un niveau sonore moyen dans la zone (Pressure Matching 2, (PM/ACC)), soit en utilisant la méthode *Weighted Least Squares* (Pressure Matching 3 (WLS)).

Algorithme 3 Pressure Matching 1

```

1: procedure PM
2:   for all  $\omega$  frequencies do
3:     Définition de  $\mathbf{p}_T$  pour toute la zone
4:     Calcul des  $\mathbf{G}$  pour toute la zone
5:      $\beta \leftarrow 0$ 
6:     while  $E > E_C$  or  $\text{cond}([\mathbf{G}^H \mathbf{G} + \beta \mathbf{I}]) > 10^{10}$  do
7:        $\mathbf{q} \leftarrow [\mathbf{G}^H \mathbf{G} + \beta \mathbf{I}]^{-1} [\mathbf{G}^H \mathbf{p}_T]$ 
8:        $\beta \leftarrow \beta + \beta_{step}$ 
9:        $E \leftarrow 10 \log_{10} \left( \frac{\mathbf{q}^H(n) \mathbf{q}(n)}{|q_r(n)|^2} \right)$ 
10:    end while
11:     $\mathbf{q}_{opt}(\omega) \leftarrow [\mathbf{G}^H \mathbf{G} + \beta \mathbf{I}]^{-1} [\mathbf{G}^H \mathbf{p}_T]$ 
12:  end for
13: end procedure

```

Algorithme 4 Pressure Matching 2

```

1: procedure ACC/PM
2:   for all  $\omega$  frequencies do
3:     Définition de  $\mathbf{p}_{T_B}$  pour la Bright zone
4:     Calcul de  $\mathbf{G}_B$  pour la Bright zone
5:     Calcul de  $\mathbf{G}_D$  pour la Dark zone
6:      $\beta \leftarrow 0$ 
7:     while  $E > E_C$  or  $\text{cond}([\mathbf{G}_B^H \mathbf{G}_B + \mathbf{G}_D^H \mathbf{G}_D + \beta \mathbf{I}]) > 10^{10}$  do
8:        $\mathbf{q} \leftarrow [\mathbf{G}_B^H \mathbf{G}_B + \mathbf{G}_D^H \mathbf{G}_D + \beta \mathbf{I}]^{-1} [\mathbf{G}_B^H \mathbf{p}_{T_B}]$ 
9:        $\beta \leftarrow \beta + \beta_{step}$ 
10:       $E \leftarrow 10 \log_{10} \left( \frac{\mathbf{q}^H(n) \mathbf{q}(n)}{|q_r(n)|^2} \right)$ 
11:    end while
12:     $\mathbf{q}_{opt}(\omega) \leftarrow [\mathbf{G}_B^H \mathbf{G}_B + \mathbf{G}_D^H \mathbf{G}_D + \beta \mathbf{I}]^{-1} [\mathbf{G}_B^H \mathbf{p}_{T_B}]$ 
13:  end for
14: end procedure

```

Algorithme 5 Pressure Matching 3

```

1: procedure WLS
2:   Définition  $0 > \eta > 1$ 
3:   for all  $\omega$  frequencies do
4:     Définition de  $\mathbf{p}_{T_B}$  pour la Bright zone
5:     Calcul de  $\mathbf{G}_B$  pour la Bright zone
6:     Calcul des  $\mathbf{G}_D$  pour la Dark zone
7:      $\beta \leftarrow 0$ 
8:     while  $E > E_C$  or  $\text{cond}([\eta \mathbf{G}_B^H \mathbf{G}_B + (1 - \eta) \mathbf{G}_D^H \mathbf{G}_D + \beta \mathbf{I}]) > 10^{10}$  do
9:        $\mathbf{q} \leftarrow [\eta \mathbf{G}_B^H \mathbf{G}_B + (1 - \eta) \mathbf{G}_D^H \mathbf{G}_D + \beta \mathbf{I}]^{-1} [\mathbf{G}_B^H \mathbf{p}_{T_B}]$ 
10:       $\beta \leftarrow \beta + \beta_{step}$ 
11:       $E \leftarrow 10 \log_{10} \left( \frac{\mathbf{q}^H(n) \mathbf{q}(n)}{|\mathbf{q}_r(n)|^2} \right)$ 
12:    end while
13:     $\mathbf{q}_{opt}(\omega) \leftarrow [\eta \mathbf{G}_B^H \mathbf{G}_B + (1 - \eta) \mathbf{G}_D^H \mathbf{G}_D + \beta \mathbf{I}]^{-1} [\mathbf{G}_B^H \mathbf{p}_{T_B}]$ 
14:  end for
15: end procedure

```

A.4 Algorithme LMS

L'algorithme LMS calcule les poids pour reproduire un champ sonore cible de façon itérative. Ce champ sonore peut être composé d'une *Bright zone* seule ou d'une *Bright zone* et d'une *Dark zone*, définissant ainsi un problème de *Sound Zones*.

Algorithme 6 LMS

```

1: procedure LMS
2:   Initialisation :
3:    $\mathbf{q}(0) = \mathbf{0}$ 
4:   for all  $n$  iterations do
5:     for all  $\omega$  frequencies do
6:        $\mu_L$  val. propre max de  $\mathbf{G}^H \mathbf{G}$  divisée par 4 (valeur arbitraire)
7:       Définition de  $\mathbf{p}_T$ 
8:       Calcul des  $\mathbf{G}(n)$  pour toute la zone
9:        $\mathbf{q}(n, \omega) = \mathbf{q}(n - 1, \omega) + \mu_L \mathbf{G}^H(n) (\mathbf{G}(n) \mathbf{q}(n, \omega) - \mathbf{p}_T(n))$ 
10:    end for
11:  end for
12: end procedure

```

A.5 Algorithme NLMS

L'algorithme NLMS calcule les poids pour reproduire un champ sonore cible de façon itérative avec la normalisation du pas $\tilde{\mu}$ à chaque étape. Le champ sonore à reproduire peut être composé d'une *Bright zone* seule ou d'une *Bright zone* et d'une *Dark zone*, définissant ainsi un problème de *Sound Zones*.

Algorithme 7 NLMS

```

1: procedure NLMS
2:   Initialisation :
3:    $\mathbf{q}(0) = \mathbf{0}$ 
4:   for all  $n$  iterations do
5:     for all  $\omega$  frequencies do
6:       Calcul des  $\mathbf{G}(n)$  pour toute la zone
7:       Définition de  $\mathbf{p}_T$ 
8:       while  $E_\psi > E_C$  or  $\text{cond}([\mathbf{G}^H(n)\mathbf{G}(n) + \beta\mathbf{I}]) > 10^{10}$  do
9:          $\beta \leftarrow \beta + \beta_{step}$ 
10:         $\mathbf{q}_\psi \leftarrow [\mathbf{G}^H(\mathbf{G}(n) + \beta\mathbf{I})^{-1}[\mathbf{G}(n)^H\mathbf{p}_T]]$   $\triangleright$  Permet le calcul de  $E$ 
11:         $E \leftarrow 10\log_{10}\left(\frac{\mathbf{q}_\psi^H(n)\mathbf{q}_\psi(n)}{|q_r(n)|^2}\right)$ 
12:      end while
13:       $\mathbf{q}(n, \omega) \leftarrow \mathbf{q}(n-1, \omega) + \mu_N[\mathbf{G}^H(n)\mathbf{G}(n) + \beta\mathbf{I}]^{-1}[\mathbf{G}^H(n)(\mathbf{G}(n)\mathbf{q}(n, \omega) - \mathbf{p}_T(n))]$ 
14:    end for
15:  end for
16: end procedure

```

A.6 Algorithme APA

L'algorithme APA calcule les poids pour reproduire un champ sonore cible de façon itérative avec la normalisation du pas $\tilde{\mu}$ à chaque étape en prenant en compte les itérations précédentes. Le champ sonore à reproduire peut être composé d'une *Bright zone* seule ou d'une *Bright zone* et d'une *Dark zone*, définissant ainsi un problème de *Sound Zones*.

Algorithme 8 APA

```

1: procedure APA
2:   Initialisation :
3:   Set  $K$ 
4:    $\mathbf{q}(0) \leftarrow \mathbf{0}$ 
5:    $\mathbf{G}_F \leftarrow \mathbf{0}$ 
6:   for all  $n$  iterations do
7:     for all  $\omega$  frequencies do
8:       Calcul des  $\mathbf{G}(n)$  pour toute la zone
9:        $\mathbf{G}_F(n) \leftarrow \mathbf{G}(n)$   $\triangleright$   $\mathbf{G}_F(n)$  correspond à une partie de la matrice  $\mathbf{G}_F$ 
10:      Définition de  $\mathbf{p}_{T_F}$ 
11:      while  $E > E_C$  or  $\text{cond}([\mathbf{G}_F^H\mathbf{G}_F + \beta\mathbf{I}]) > 10^{10}$  do  $\triangleright$  régularisation APA
12:         $\beta \leftarrow \beta + \beta_{step}$ 
13:         $\mathbf{q}_\psi \leftarrow [\mathbf{G}_F^H\mathbf{G}_F + \beta\mathbf{I}]^{-1}[\mathbf{G}_F^H\mathbf{p}_T]$   $\triangleright$  Permet le calcul de  $E$ 
14:         $E \leftarrow 10\log_{10}\left(\frac{\mathbf{q}_\psi^H(n)\mathbf{q}_\psi(n)}{|q_r(n)|^2}\right)$ 
15:      end while
16:       $\mathbf{q}(n, \omega) \leftarrow \mathbf{q}(n-1, \omega) + \mu_N[\mathbf{G}_F^H\mathbf{G}_F + \beta\mathbf{I}]^{-1}[\mathbf{G}_F^H(n)(\mathbf{G}_F(n)\mathbf{q}(n, \omega) - \mathbf{p}_T(n))]$ 
17:    end for
18:  end for
19: end procedure

```

A.7 Algorithme VLLMS

L'algorithme VLLMS calcule les poids de façon itérative pour reproduire un champ sonore cible, qui peut être une *Bright zone* et une *Dark zone*. Cette méthode présente l'avantage de limiter l'effet mémoire. Une seconde méthode est également présentée : la méthode DZII, qui atténue le champ sonore en *Dark zone* individuellement pour chaque microphone.

Algorithme 9 VLLMS

```

1: procedure VLLMS
2:   Initialisation :
3:    $\mathbf{q}(0) = \mathbf{0}$ 
4:    $\gamma = 0$ 
5:   for all  $n$  iterations do
6:     for all  $\omega$  frequencies do
7:        $\mu_L$  val. propre max de  $\mathbf{G}^H \mathbf{G}$  divisée par 4 (valeur arbitraire)
8:       Définition de  $\mathbf{p}_T$ 
9:       Calcul des  $\mathbf{G}(n)$  pour toute la zone
10:       $\mathbf{q}(n, \omega) = (1 - \mu_L \gamma) \mathbf{q}(n-1, \omega) + \mu_L \mathbf{G}^H(n) (\mathbf{G}(n) \mathbf{q}(n, \omega) - \mathbf{p}_T(n))$ 
11:      if  $x_{aud}(n) \neq x_{aud}(n-1)$  then
12:         $m \leftarrow \min(m + l_u, 1)$ 
13:      else  $m \leftarrow \max(m - l_d, 0)$ 
14:      end if
15:       $\gamma \leftarrow \gamma_{max} \left( \frac{m}{\chi} \right)^\alpha$ 
16:    end for
17:  end for
18: end procedure

```

Algorithme 10 VLLMS DZII

```

1: procedure VLLMS DZII
2:   Initialisation :
3:    $\mathbf{q}(0) = \mathbf{0}$ 
4:    $\gamma = 0$ 
5:   for all  $n$  iterations do
6:     for all  $\omega$  frequencies do
7:        $\mu_L$  val. propre max de  $\mathbf{G}^H \mathbf{G}$  divisée par 4 (valeur arbitraire)
8:       Définition de  $\mathbf{p}_{T_B}$ 
9:       Calcul de  $\mathbf{G}_B$  pour la Bright zone
10:      Calcul de  $\mathbf{G}_D$  pour la Dark zone
11:       $\mathbf{p}_D \leftarrow \mathbf{G}_D \mathbf{q}$ 
12:       $\mathbf{p}_{T_D} \leftarrow \circ \times \mathbf{p}_D$ 
13:       $\mathbf{G} \leftarrow$  Concaténation de  $\mathbf{G}_B$  et  $\mathbf{G}_D$ 
14:       $\mathbf{q}(n, \omega) = (1 - \mu_L \gamma) \mathbf{q}(n-1, \omega) + \mu_L \mathbf{G}^H(n) (\mathbf{G}(n) \mathbf{q}(n, \omega) - \mathbf{p}_T(n))$ 
15:      if  $x_{aud}(n) \neq x_{aud}(n-1)$  then
16:         $m \leftarrow \min(m + l_u, 1)$ 
17:      else
18:         $m \leftarrow \max(m - l_d, 0)$ 
19:      end if
20:       $\gamma \leftarrow \gamma_{max} \left( \frac{m}{\chi} \right)^\alpha$ 
21:    end for
22:  end for
23: end procedure

```

B Théorie des algorithmes adaptatifs

L'ensemble de la théorie sur les filtres adaptatifs présentée dans cette partie s'inspire fortement du livre "*Adaptive Filter Theory*" de Simon Haykin [Haykin, 2002].

B.1 Problème des moindres carrés

Le principe du filtrage optimal repose sur la notion de trouver la "meilleure solution" possible à un problème donné. Dans le cas présent, le but est de trouver la "meilleure valeur possible" du filtre $h(n)$ connaissant $x(n)$, l'entrée du filtre, et une sortie $y(n)$ non connue dans l'absolue, mais de laquelle on cherche à se rapprocher. La relation entrée/sortie d'un système linéaire peut être définie par son produit de convolution

$$y(n) = \sum_0^{\infty} h(n)x(n-k). \quad (1.3)$$

On dispose d'un signal $d(n)$ correspondant à la valeur désirée du signal de sortie et de la sortie réelle du filtre $y(n)$. Il est aisé de définir une erreur $e(n)$ entre la valeur désirée $d(n)$ et la valeur réelle de sortie du système $y(n)$:

$$e(n) = d(n) - y(n). \quad (1.4)$$

Dans le meilleur des cas, on cherche à avoir $y(n) = d(n)$ donc $e(n) = 0$. Une solution consiste à chercher la distance minimale qui sépare en moyenne les deux signaux. Il est donc possible d'écrire cette distance moyenne comme

$$J = E[\|e(n)\|^2], \quad (1.5)$$

ce qui correspond à l'erreur quadratique moyenne entre les deux signaux. Le but est alors de minimiser J . Une solution consiste à trouver les points stationnaires de J par dérivation, puisque J est une fonction quadratique, elle ne dispose que d'un point stationnaire qui correspond aussi à son minimum. Sachant que le filtre h peut être défini par k coefficients complexes avec $k \in \mathbb{N}^+$ on a :

$$h_k = a_k + j b_k. \quad (1.6)$$

On cherche alors à dériver l'équation (1.5) en définissant l'opérateur

$$\nabla_k J = \frac{\partial J}{\partial a_k} + j \frac{\partial J}{\partial b_k}. \quad (1.7)$$

L'équation (1.5) peut alors être écrite,

$$\nabla_k J = E \left[\frac{\partial e(n)}{\partial a_k} e^*(n) + \frac{\partial e^*(n)}{\partial a_k} e(n) + \frac{\partial e(n)}{\partial b_k} j e^*(n) + \frac{\partial e^*(n)}{\partial b_k} j e(n) \right] \quad (1.8)$$

avec * correspondant au complexe conjugué et

$$\frac{\partial e(n)}{\partial a_k} = -x(n-k) \quad \frac{\partial e(n)}{\partial b_k} = j x(n-k) \quad (1.9)$$

$$\frac{\partial e^*(n)}{\partial a_k} = -x^*(n-k) \quad \frac{\partial e^*(n)}{\partial b_k} = -j x^*(n-k) \quad (1.10)$$

correspondant aux dérivées partielles de $\nabla_k J$.

En remplaçant les équations (1.9) et (1.10) dans l'équation (1.8) et après simplification, il est alors possible d'écrire la dérivée de J

$$\nabla_k J = -2E[x(n-k)e^*(n)]. \quad (1.11)$$

Pour obtenir la valeur optimale des coefficients h_k on cherche alors $\nabla_k J = 0$, ce qui donne

$$E[x(n-k)e^*(n)] = 0, \quad (1.12)$$

qui remplit la condition nécessaire et suffisante pour minimiser J .

Une méthode, plusieurs utilisations

Dans le cas d'un filtre adaptatif, il existe toujours un signal désiré que l'on cherche à reproduire. En fonction de comment est choisi ce signal $d(n)$, il permet par exemple de réaliser l'identification ou encore la compensation de système. Seule la méthode d'égalisation (ou compensation) est succinctement présentée ici. La relation entrée/sortie du système inconnu est définie comme

$$\gamma(n) = \sum_0^{\infty} \eta(n)\chi(n-k), \quad (1.13)$$

avec $\gamma(n)$ la sortie du système, $\eta(n)$ les coefficients du filtre et χ le signal d'entrée.

Egalisation

Dans le cas de l'égalisation d'un système, on cherche à compenser ce qu'a modifié le filtre $\gamma(n)$ par le filtre $h(n)$. On fera alors en sorte de reproduire le signal d'entrée $\chi(n)$ du système inconnu ce qui correspond à

$$d(n) = \chi(n), \quad (1.14)$$

avec pour entrée du filtre adaptatif

$$x(n) = \gamma(n). \quad (1.15)$$

L'erreur à minimiser d'après l'équation (1.4) s'écrit alors

$$e(n) = \chi(n) - y(n). \quad (1.16)$$

La **Figure A.2.1** présente la relation entrées/sorties des filtres pour réaliser la compensation de la réponse du filtre inconnu par le filtre adaptatif.

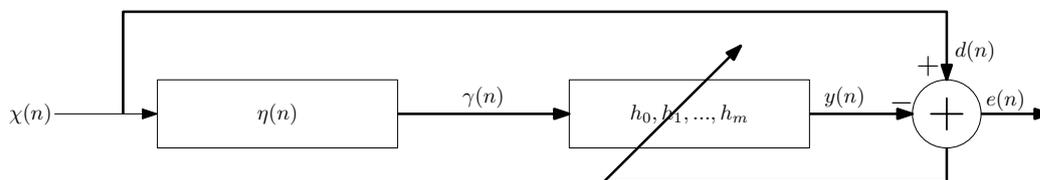


Figure A.2.1 – Schéma représentant le fonctionnement d'un filtre adaptatif en mode égalisation. Le filtre à égaliser se trouve avant le filtre adaptatif, qui devra compenser la sortie $\gamma(n)$ du filtre $\eta(n)$ en le comparant avec le signal d'entrée $\chi(n)$.

Filtrage de Wiener

Le filtrage de Wiener-Hopf permet de trouver les coefficients du filtre optimal h_{oi} de façon analytique, défini par la minimisation de $\nabla_k J$. En remplaçant $e^*(n)$ et $y(n)$ par leur expression respective définie dans les équations (1.5) et (1.3), il est possible d'écrire l'équation (1.12) sous la forme :

$$\mathbb{E} \left[x(n-k) \left(d^*(n) - \sum_{i=0}^{\infty} h_{oi} x^*(n-i) \right) \right] = 0. \quad (1.17)$$

En développant cette équation, on obtient

$$\sum_{i=0}^{\infty} h_{oi} \mathbb{E} [x(n-k)x^*(n-i)] = \mathbb{E} [x(n-k)d^*(n)]. \quad (1.18)$$

On reconnaît alors l'autocorrélation du signal d'entrée $x(n)$ définie par

$$r(i-k) = \mathbb{E} [x(n-k)x^*(n-i)], \quad (1.19)$$

ainsi que l'intercorrélacion entre le signal d'entrée $x(n)$ et le signal désiré $d^*(n)$

$$p(-k) = \mathbb{E} [x(n-k)d^*(n)]. \quad (1.20)$$

Il est alors possible de réécrire l'équation (1.17) sous la forme

$$\sum_{i=0}^{\infty} h_{oi} r(i-k) = p(-k), \quad (1.21)$$

ce qui permet de définir les coefficients optimaux h_{oi} en fonction du signal d'entrée $x(n)$ dans le but d'obtenir une sortie désirée $d(n)$.

Par ailleurs, il est également possible de décrire le problème sous forme matricielle en posant

$$\mathbf{R} = \mathbb{E} [\mathbf{x}(n)\mathbf{x}^H(n)] = \begin{bmatrix} r(0) & r(1) \dots & r(M-1) \\ r^*(1) & r(0) & \dots & r(M-2) \\ \vdots & \vdots & \ddots & \vdots \\ r^*(M-1) & r^*(M-2) & \dots & r(0) \end{bmatrix} \quad (1.22)$$

la matrice d'autocorrélation, avec H la transposée hermitienne et $\mathbf{x}(n) = [x(n), x(n-1), \dots, x(n-M+1)]$ dans le cas d'un signal avec M échantillons. De la même façon, il est possible de définir l'intercorrélacion entre le signal d'entrée $x(n)$ et le signal désiré $d(n)$,

$$\mathbf{p} = \mathbb{E} [\mathbf{x}(n)d^*(n)]. \quad (1.23)$$

Cette formulation permet alors d'écrire l'équation (1.21) sous la forme

$$\mathbf{R}\mathbf{h}_o = \mathbf{p}. \quad (1.24)$$

On en déduit la valeur des coefficients optimaux d'un système connaissant la matrice d'autocorrélacion du signal d'entrée ainsi que l'intercorrélacion du signal d'entrée et du signal désiré, suivant

$$\mathbf{h}_o = \mathbf{R}^{-1}\mathbf{p}. \quad (1.25)$$

Filtre LMS

Le filtrage adaptatif permet de trouver le filtre optimal au sens des moindres carrés de façon itérative plutôt que de calculer directement la réponse par la méthode de Wiener-Hopf. Le principe est de calculer une estimation du gradient de l'erreur pour obtenir la solution optimale au problème de façon itérative. En partant de la définition de l'erreur (**Eq.** (1.4)) entre le signal désiré et le signal de sortie du système défini par son produit de convolution (**Eq.** (1.3))

$$e(n) = d(n) - \sum_{k=0}^{M-1} h_k^* x(n-k), \quad (1.26)$$

Il est possible d'écrire la fonction de coût comme étant l'erreur quadratique moyenne :

$$J = \mathbb{E}[e(n)e^*(n)] \quad (1.27)$$

$$= \mathbb{E}[|d(n)|^2] - \sum_{k=0}^{M-1} h_k^* \mathbb{E}[x(n-k)d^*(n)] - \sum_{k=0}^{M-1} h_k \mathbb{E}[x^*(n-k)d(n)] + \sum_{k=0}^{M-1} \sum_{i=0}^{M-1} h_k^* h_i \mathbb{E}[x(n-k)x^*(n-i)]. \quad (1.28)$$

Il est possible de reconnaître quatre termes distincts qui sont

$$\sigma_d^2 = \mathbb{E}[|d(n)|^2] \quad p(-k) = \mathbb{E}[x(n-k)d^*(n)] \quad (1.29)$$

$$p^*(-k) = \mathbb{E}[x^*(n-k)d(n)] \quad r(i-k) = \mathbb{E}[x(n-k)x^*(n-i)]. \quad (1.30)$$

σ^2 représente la variance du signal désiré, $p(-k)$ et $p^*(-k)$ correspondent à l'inter-corrélation entre $x(n)$ et $d(n)$ et $r(i-k)$ l'autocorrélation de $x(n)$. En remplaçant ces termes dans l'équation (1.28), on obtient

$$J = \sigma_d^2 - \sum_{k=0}^{M-1} h_k^* p(-k) - \sum_{k=0}^{M-1} h_k p^*(-k) + \sum_{k=0}^{M-1} \sum_{i=0}^{M-1} h_k^* h_i r(i-k). \quad (1.31)$$

L'équation (1.31) peut aussi être écrite sous forme matricielle comme présenté dans la partie sur le filtrage de Wiener, en équations (1.22) et (1.23), ce qui s'écrit

$$J(\mathbf{h}) = \sigma_d^2 - \mathbf{h}^H \mathbf{p} - \mathbf{p}^H \mathbf{h} + \mathbf{w}^H \mathbf{R} \mathbf{w}. \quad (1.32)$$

Cette formulation permet d'obtenir la fonction de coût d'un problème des moindres carrés posé sous forme matricielle pour un ensemble des coefficients de filtre \mathbf{h} . Pour obtenir le minimum de cette fonction de façon itérative, il est nécessaire de calculer son gradient instantané $\nabla_k J(n)$. Lorsque celui-ci s'annule pour $\nabla_k J(n) = \mathbf{0}$, le filtre correspond à sa valeur optimale. On cherche alors la dérivée de cette fonction de coût, ce qui donne dans le cas d'un filtre $h_k = a_k + j b_k$ à coefficients complexes

$$\nabla_k J = -2\mathbf{p} + 2\mathbf{R}\mathbf{h}(n) \quad (1.33)$$

$$= -2\mathbf{x}(n)d^*(n) + 2\mathbf{x}(n)\mathbf{x}^H(n)\mathbf{h}(n) \quad (1.34)$$

$$= -2\mathbf{x}(n)[d^*(n) - \mathbf{x}^H(n)\mathbf{h}(n)] \quad (1.35)$$

$$= -2\mathbf{x}(n)e^*(n) \quad (1.36)$$

avec

$$\mathbf{R}(n) = \mathbf{x}(n)\mathbf{x}^H(n) \qquad \mathbf{p}(n) = \mathbf{x}(n)d^*(n). \qquad (1.37)$$

et

$$y(n) = \mathbf{h}^H(n)\mathbf{x}(n) \qquad e(n) = d(n) - y(n). \qquad (1.38)$$

Cet estimateur est alors appliqué à la méthode de descente de gradient et permet de définir les poids du filtre pour chaque étape comme étant :

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \mu\mathbf{x}(n)e^*(n), \qquad (1.39)$$

avec μ correspondant au pas d'adaptation permettant de choisir la vitesse de descente. Il doit par ailleurs être choisi de façon judicieuse pour assurer la stabilité du système ainsi que pour avoir une erreur de minimale proche de la valeur optimale du filtre. En effet, le pas d'adaptation de la méthode LMS est proportionnel à l'amplitude du signal d'entrée. Lorsque la solution est proche de l'optimal, il continue d'exister une faible variation de ce filtre correspondant à un bruit d'adaptation. Lorsque le signal est de forte amplitude, il est possible d'avoir une erreur d'estimation importante, ce qui peut mener à une instabilité du filtre.

La méthode LMS permet donc de façon itérative de trouver les coefficients d'un filtre, ce qui permet d'éviter l'inversion de la matrice d'autocorrélation nécessaire au calcul du filtre optimal par la méthode de Wiener.

Pour pallier la contrainte du choix du coefficient μ , il est possible d'utiliser d'autres méthodes basées sur le même principe et permettant d'adapter le pas μ en fonction de différents critères. Il existe par exemple le NLMS (Normalized Least Square) qui réalise la normalisation par rapport aux données d'entrées, la méthode APA qui est une généralisation de la méthode NLMS¹ ou encore le RLS qui choisi le pas d'adaptation en fonction de l'inverse de la matrice de corrélation et de l'entrée du filtre. D'autres méthodes adaptatives existent comme par exemple les réseaux de neurones qui permettent aussi d'adapter les coefficients d'un système complexe d'un ensemble de filtres en fonction des données d'entrée du filtre. Chacune de ces méthodes ont des caractéristiques de convergence et de complexité de calcul² qui les rendent plus ou moins adaptées à un certain nombre de problèmes.

1. Un filtre NLMS correspond à un filtre APA d'ordre 1. Les filtres APA permettent de prendre en compte les données d'entrée des K itérations précédentes.

2. Les méthodes sont présentées dans l'ordre de complexité croissante.

Filtre NLMS

Le principe du filtre NLMS (Normalized LMS) est d'automatiser le choix du paramètre de pas μ de sorte, d'une part, à ne pas subir l'influence de l'amplitude des données d'entrée $\mathbf{x}(n)$ causant une amplification du bruit de gradient (lorsque le filtre a convergé), et d'autre part, une vitesse de convergence optimale grâce au choix de μ variable à chaque pas. Le principe est de normaliser les données d'entrée du filtre pour atteindre ces caractéristiques. Pour réaliser cette normalisation, le filtre se base sur le principe des faibles perturbation. C'est à dire que la variation $\delta\mathbf{h}(n+1)$ entre le coefficient $\mathbf{h}(n)$ à l'itération n et le coefficient $\mathbf{h}(n+1)$ à l'itération $n+1$ doit être minimale, ce qui implique

$$\delta\mathbf{h}(n+1) = \mathbf{h}(n+1) - \mathbf{h}(n), \quad (1.40)$$

tout en respectant la contrainte sur la valeur désirée du filtre

$$\mathbf{h}^H(n+1)\mathbf{x}(n) = d(n). \quad (1.41)$$

Pour résoudre ce problème d'optimisation sous contraintes, il est possible d'utiliser la méthode des multiplicateurs de Lagrange. La fonction de coût peut alors s'écrire sous la forme

$$J(n) = \|\delta\mathbf{h}(n+1)\|^2 + \Re[\lambda^*(d(n) - \mathbf{h}(n+1)\mathbf{x}(n))], \quad (1.42)$$

avec λ le multiplicateur de Lagrange complexe.

Pour obtenir la valeur optimale de ce problème qui minimise la cette fonction de coût, on dérive dans un premier temps l'équation (1.42) :

$$\frac{\partial J(n)}{\partial \mathbf{h}^*(n+1)} = 2(\mathbf{h}(n+1) - \mathbf{h}(n)) - \lambda^* \mathbf{x}(n). \quad (1.43)$$

En annulant cette dérivée, on obtient

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \frac{1}{2}\lambda^* \mathbf{x}(n). \quad (1.44)$$

Dans un second temps, on cherche à trouver la valeur du multiplicateur de Lagrange λ en remplaçant la valeur de $\mathbf{h}(n+1)$ (Eq. (1.44)) dans la contrainte définie en équation (1.41), ce qui donne

$$d(n) = \mathbf{h}^H(n+1)\mathbf{x}(n) \quad (1.45)$$

$$= \left(\mathbf{h}(n) + \frac{1}{2}\lambda^* \mathbf{x}(n) \right)^H \mathbf{x}(n) \quad (1.46)$$

$$= \mathbf{h}^H(n)\mathbf{x}(n) + \frac{1}{2}\lambda \mathbf{x}^H(n)\mathbf{x}^H(n) \quad (1.47)$$

$$= \mathbf{h}^H(n)\mathbf{x}(n) + \frac{1}{2}\lambda \|\mathbf{x}(n)\|^2. \quad (1.48)$$

On déduit alors la valeur de λ :

$$\lambda = \frac{2e(n)}{\|\mathbf{x}(n)\|^2}, \quad (1.49)$$

avec

$$e(n) = d(n) - \mathbf{h}^H(n)\mathbf{x}(n), \quad (1.50)$$

correspondant à l'erreur de reproduction.

En combinant les résultats des équations (1.44) et (1.49), il vient

$$\delta\mathbf{h}(n+1) = \mathbf{h}(n+1) - \mathbf{h}(n) \quad (1.51)$$

$$= \frac{1}{\|\mathbf{x}(n)\|} \mathbf{x}(n) e^*(n). \quad (1.52)$$

Un facteur positif réel $\tilde{\mu}$ est alors introduit pour contrôler la quantité de changement entre deux itérations. On écrit alors les équations 1.51 et 1.52 sous la forme :

$$\delta\mathbf{h}(n+1) = \mathbf{h}(n+1) - \mathbf{h}(n) \quad (1.53)$$

$$= \frac{\tilde{\mu}}{\|\mathbf{x}(n)\|} \mathbf{x}(n) e^*(n). \quad (1.54)$$

A partir des équations 1.53 et 1.54, il est équivalent d'écrire

$$\mathbf{h}(n+1) - \mathbf{h}(n) = \frac{\tilde{\mu}}{\|\mathbf{x}(n)\|} \mathbf{x}(n) e^*(n), \quad (1.55)$$

on en déduit alors la formulation récursive correspondant au filtrage NLMS :

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \frac{\tilde{\mu}}{\|\mathbf{x}(n)\|} \mathbf{x}(n) e^*(n). \quad (1.56)$$

Si on compare l'équation 1.56 avec celle du filtre LMS (**Eq.** (1.39)), on note que le filtre NLMS correspond à un filtre LMS avec

$$\mu(n) = \frac{\tilde{\mu}}{\|\mathbf{x}(n)\|^2}, \quad (1.57)$$

c'est à dire, que le filtre NLMS est une version du filtre LMS normalisée par la puissance des données d'entrée du filtre.

Filtre APA d'ordre K

De la même façon que le filtre NLMS, le filtre APA cherche à minimiser la variation des coefficients du filtre entre deux itérations, c'est à dire

$$\delta\mathbf{h}(n+1) = \mathbf{h}(n+1) - \mathbf{h}(n), \quad (1.58)$$

tout en respectant une contrainte sur la valeur désirée $d(n)$ du filtre. A la différence du filtre NLMS, qui ne dispose d'une contrainte que sur une unique valeur désirée, le filtre APA cherche à respecter cette contrainte sur un ensemble de K valeurs correspondant aux K itérations précédentes. Cet ensemble de contraintes s'écrit :

$$d(n-k) = \mathbf{h}^H(n+1)\mathbf{x}(n-k) \quad \text{avec} \quad k = 0, 1, \dots, K-1. \quad (1.59)$$

de la même façon que pour le filtre NLMS, le problème peut être résolu sous la forme des multiplicateurs de Lagrange, ce qui donne :

$$J(n) = \|\mathbf{h}(n+1) - \mathbf{h}(n)\|^2 + \sum_{k=0}^{K-1} \Re [\lambda_k^* (d(n-k) - \mathbf{h}^H(n+1)\mathbf{x}(n-k))]. \quad (1.60)$$

Afin de faciliter la lecture, un ensemble de matrices peuvent être introduites :

- Une matrice de dimension $K \times M$, avec M le nombre de coefficients du filtre, correspondant à la matrice des données d'entrée :

$$\mathbf{A}^H(n) = [\mathbf{x}(n), \mathbf{x}(n-1), \dots, \mathbf{x}(n-K+1)]. \quad (1.61)$$

- Un vecteur des valeurs désirées de dimensions $K \times 1$:

$$\mathbf{d}^H(n) = [d(n), d(n-1), \dots, d(n-K+1)]. \quad (1.62)$$

- Un vecteur de Lagrange contenant tous les multiplicateurs de Lagrange, de dimension $K \times 1$:

$$\boldsymbol{\lambda}^H = [\lambda_0, \lambda_1, \dots, \lambda_{K-1}]. \quad (1.63)$$

En utilisant ces définitions, on peut alors réécrire la fonction de coût (**Eq.** (1.60)) sous la forme

$$J(n) = \|\mathbf{h}(n+1) - \mathbf{h}(n)\|^2 + \Re [(\mathbf{d}(n) - \mathbf{A}(n)\mathbf{h}(n+1))^H \boldsymbol{\lambda}]. \quad (1.64)$$

Pour minimiser cette fonction de coût, elle est dérivée, ce qui donne :

$$\frac{\partial J(n)}{\partial \mathbf{h}^*(n+1)} = 2(\mathbf{h}(n+1) - \mathbf{h}(n)) - \mathbf{A}^H(n)\boldsymbol{\lambda}. \quad (1.65)$$

En annulant cette dérivée, il vient

$$\delta \mathbf{h}(n+1) = \frac{1}{2} \mathbf{A}^H(n)\boldsymbol{\lambda}. \quad (1.66)$$

Pour éliminer le vecteur de Lagrange, de l'équation (1.66), on utilise les définitions des équations (1.61) et (1.62) pour réécrire l'équation (1.59) sous la forme :

$$\mathbf{d}(n) = \mathbf{A}(n)\mathbf{h}(n+1). \quad (1.67)$$

En multipliant des deux côtés l'équation (1.66) par la matrice des entrées $\mathbf{A}(n)$ et en utilisant les équations (1.58) et (1.67) pour éliminer le vecteur $\mathbf{h}(n+1)$, il est possible d'écrire :

$$\mathbf{d}(n) = \mathbf{A}(n)\mathbf{h}(n) + \frac{1}{2} \mathbf{A}(n)\mathbf{A}^H(n)\boldsymbol{\lambda}. \quad (1.68)$$

On peut alors en déduire le vecteur d'erreur

$$\mathbf{e}(n) = \mathbf{d}(n) - \mathbf{A}(n)\mathbf{h}(n). \quad (1.69)$$

En utilisant l'équation (1.68), on obtient le vecteur des multiplicateurs de Lagrange sous la forme :

$$\boldsymbol{\lambda} = 2(\mathbf{A}(n)\mathbf{A}^H(n))^{-1} \mathbf{e}(n), \quad (1.70)$$

avec $^{-1}$ une inversion de matrice.

En remplaçant la solution de cette équation dans l'équation (1.66), on obtient

$$\delta \mathbf{h}(n+1) = \mathbf{A}^H (\mathbf{A}(n) (\mathbf{A}^H(n)))^{-1} \mathbf{e}(n). \quad (1.71)$$

De la même façon que pour le filtre NLMS, un facteur positif réel $\tilde{\mu}$ est introduit pour contrôler la quantité de changement entre deux itérations, ce qui donne :

$$\delta \mathbf{h}(n+1) = \tilde{\mu} \mathbf{A}^H (\mathbf{A}(n) (\mathbf{A}^H(n)))^{-1} \mathbf{e}(n). \quad (1.72)$$

De façon équivalente, on peut écrire la mise à jour des coefficients sous la forme

$$\mathbf{h}(n+1) = \mathbf{h}(n) + \tilde{\mu} \mathbf{A}^H(n) (\mathbf{A}(n) \mathbf{A}^H(n))^{-1} \mathbf{e}(n). \quad (1.73)$$

C Notes sur la complexité algorithmique

La vitesse d'exécution d'un algorithme peut être estimée par sa complexité algorithmique. Celle-ci correspond au nombre d'opérations nécessaires pour obtenir un résultat donné. Plus un algorithme est complexe, plus celui-ci met de temps à s'exécuter. De manière générale, la complexité algorithmique est donnée par l'élément le plus lourd à calculer et en fonction des dimensions du problème [Golub et Van Loan, 2012, Cormen, 2004]. Pour le calcul matriciel, les complexités algorithmiques sont connues pour différents problèmes. Dans notre cas, deux opérations sont réalisées : la multiplication matricielle et l'inversion. Ces deux opérations ont une complexité identique, car il peut être montré qu'une inversion peut être traitée de la même façon qu'une multiplication matricielle [Cormen, 2004].

Dans le cas de la multiplication matricielle de deux matrices carrées de dimension n , la complexité est en n^3 , correspondant au nombre de multiplications réalisées³. Le problème augmente donc avec le cube de la dimension du problème. Dans le cas de la méthode Directe, LMS, NLMS et APA, toutes les matrices ne sont pas carrées, la complexité est donc calculée en fonction du nombre de haut-parleurs L_{HP} , du nombre de microphones M et de l'ordre du filtre K , pour le filtre APA. Les complexités sont données dans le tableau suivant.

Méthode	Complexité
Méthode LMS et VLLMS	ML_{HP}
Méthode Directe et NLMS	ML_{HP}^2 pour $L_{HP} < \sqrt{M}$ et L_{HP}^3 pour $0 < \sqrt{M} < L_{HP}$
Méthode APA	$M(KL_{HP})^2$ pour $KL_{HP} < \sqrt{M}$ et KL_{HP}^3 pour $0 < \sqrt{M} < KL_{HP}$

Le coût de calcul induit par ces différentes méthodes met donc en évidence que la méthode LMS dispose de la plus faible complexité algorithmique, et augmente linéairement avec la dimension du problème. Les méthodes NLMS et Directe disposent quant à elles d'une augmentation au moins avec le carré du nombre de haut-parleurs et au plus avec le cube du nombre de haut-parleurs. La méthode APA de son côté est encore plus lourde en prenant en compte l'ordre du filtre.

3. Le problème de multiplication matricielle peut être résolu de façon plus efficace par certains algorithmes, sous un certain nombre de contraintes. Nous ne nous intéressons ici qu'à la méthode naïve qui fait référence [Cormen, 2004].

D Influence des paramètres de la méthode VLLMS

D.1 Paramètres de pas de transition : ld et lu

Les paramètres lu et ld correspondent aux pas de montée et de descente du paramètre $\gamma(n)$. Lorsque leur valeur est faible, le filtre VLLMS, met plus de temps à passer du "mode LMS" au mode "Leaky LMS". L'ensemble des figures présentées correspond à un déplacement selon l'axe X entre les itérations 200 et 300, dans les mêmes conditions que celles proposées dans la partie 3.3 liée à l'effet mémoire.

Le paramètre ld permet de passer de Leaky LMS vers LMS et lu de passer de LMS vers Leaky LMS, ce de façon indépendante. L'augmentation de lu induit un passage rapide vers le mode Leaky LMS (lors d'un déplacement par exemple). La diminution de ld diminue la vitesse de retour vers le filtre LMS, utile pour diminuer l'effet mémoire après l'arrêt, mais au prix d'une vitesse de convergence plus faible. Il faut donc trouver un compromis entre la vitesse de convergence, et l'oubli. Les **Figures** A.4.2 et A.4.4 illustrent les paramètres d'erreur, d'effort et de niveau sonore pour un ld variable et un $lu = 0.2$. Les **Figures** A.4.8 et A.4.9 illustrent quant à elles ces paramètres pour un lu variable et un $ld = 0.2$. On observe sur la **Figure** A.4.8, avec $lu = 0.01$, que si la transition n'est pas réalisée assez vite, le filtre ne désactive pas tous les haut-parleurs. L'effort final est donc plus élevé que dans le cas des deux autres valeurs de lu .

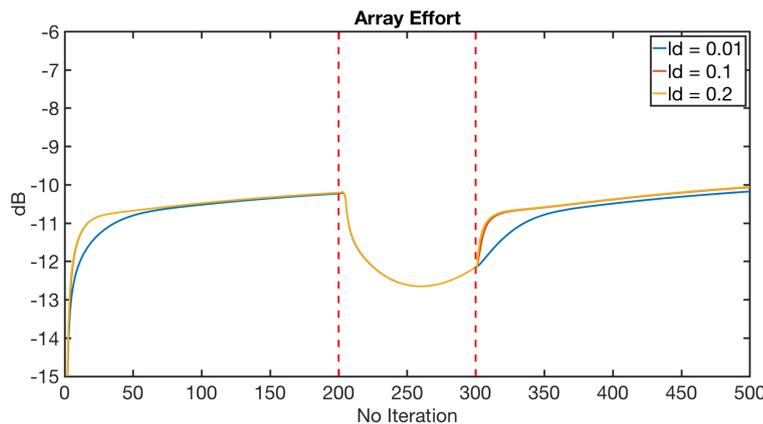


Figure A.4.2 – Effort de réseau pour différents ld pour un auditeur se déplaçant à un d'un centimètre par itération pendant 100 itérations entre les itérations 200 et 300. Le déplacement est réalisé selon l'axe X . Les lignes verticales pointillées rouges représentent le début et la fin du déplacement. La source de référence est choisie de sorte à ce qu'elle reproduise un niveau sonore équivalent dans la zone avec une distance source/micro équivalente (**Eq.** (2.37)).

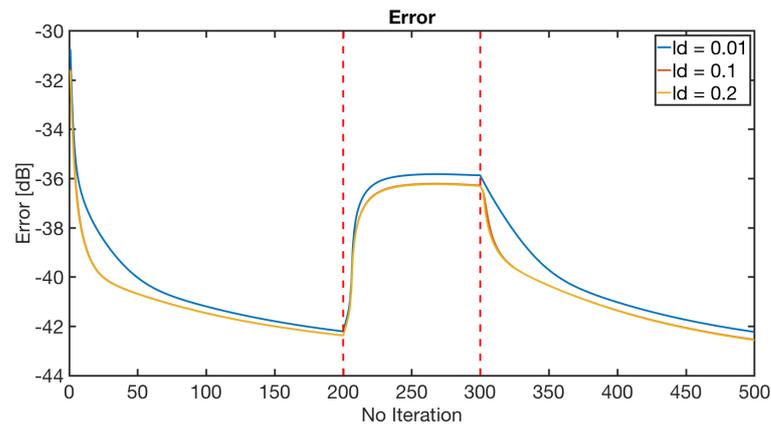


Figure A.4.3 – Erreur de reproduction pour différents ld pour un auditeur se déplaçant à un d d'un centimètre par itération pendant 100 itérations entre les itérations 200 et 300. Le déplacement est réalisé selon l'axe X. Les lignes verticales pointillées rouges représentent le début et la fin du déplacement.

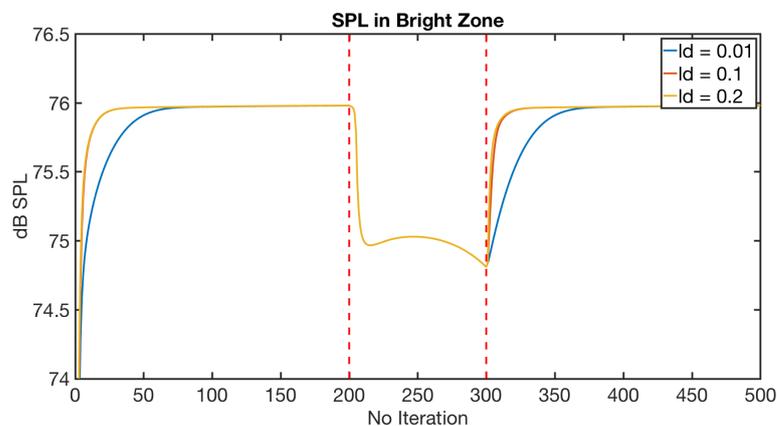


Figure A.4.4 – Niveau sonore moyen pour différents ld pour un auditeur se déplaçant à un d d'un centimètre par itération pendant 100 itérations entre les itérations 200 et 300. Le déplacement est réalisé selon l'axe X. Les lignes verticales pointillées rouges représentent le début et la fin du déplacement.

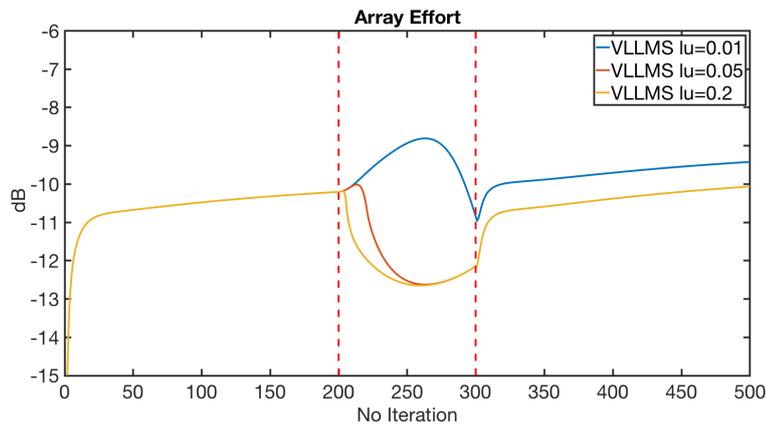


Figure A.4.5 – Effort de réseau pour différents lu pour un auditeur se déplaçant à un d'un centimètre par itération pendant 100 itérations entre les itérations 200 et 300. Le déplacement est réalisé selon l'axe X. Les lignes verticales pointillées rouges représentent le début et la fin du déplacement. La source de référence est choisie de sorte à ce qu'elle reproduise un niveau sonore équivalent dans la zone avec une distance source/micro équivalente (**Eq. (2.37)**).

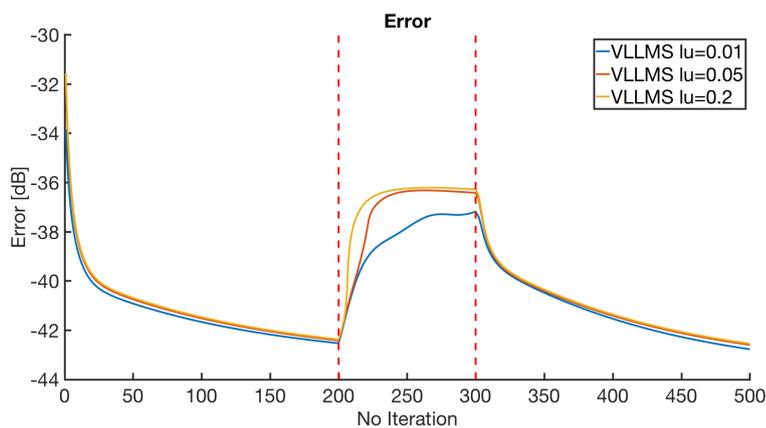


Figure A.4.6 – Erreur de reproduction pour différents lu pour un auditeur se déplaçant à un d'un centimètre par itération pendant 100 itérations entre les itérations 200 et 300. Le déplacement est réalisé selon l'axe X. Les lignes verticales pointillées rouges représentent le début et la fin du déplacement.

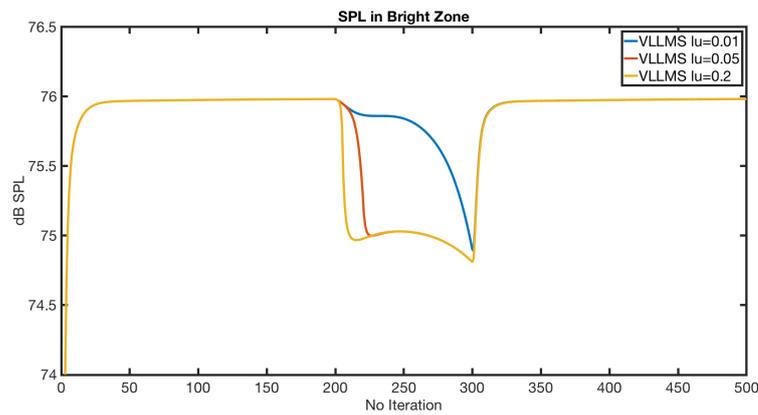


Figure A.4.7 – Niveau sonore pour différents lu pour un auditeur se déplaçant à un d'un centimètre par itération pendant 100 itérations entre les itérations 200 et 300. Le déplacement est réalisé selon l'axe X. Les lignes verticales pointillées rouges représentent le début et la fin du déplacement.

D.2 Paramètre de vitesse de transition : α

Le paramètre alpha permet de choisir la vitesse à laquelle on atteint γ_{max} , lorsqu'il s'agit de se comporter comme un filtre Leaky LMS (pendant les transitions), et la vitesse à laquelle on atteint 0 lorsqu'il s'agit de se comporter comme un filtre LMS (pendant les phases stationnaires). Son influence est principalement exercée sur la vitesse de convergence l'algorithme en début et en fin de période transitoire. L'ensemble des **Figures** A.4.8, A.4.9 et A.4.10 correspond à un déplacement selon l'axe X entre les itérations 200 et 300, dans les mêmes conditions que celles proposées dans la partie 3.3 liée à l'effet mémoire.

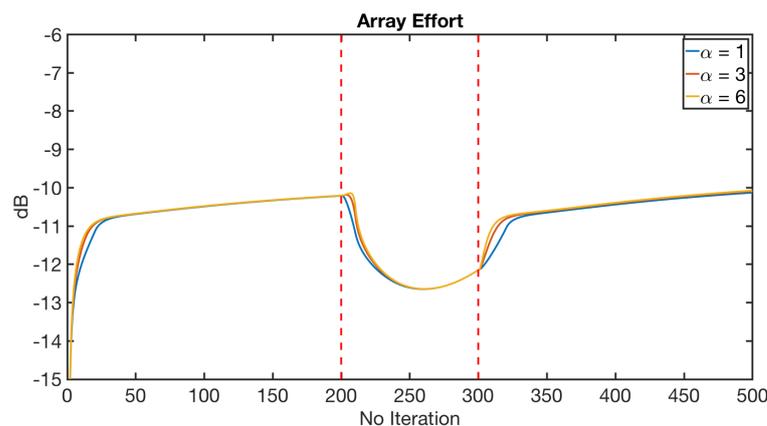


Figure A.4.8 – Effort de réseau pour différents α pour un auditeur se déplaçant à un d'un centimètre par itération pendant 100 itérations entre les itérations 200 et 300. Le déplacement est réalisé selon l'axe X. Les lignes verticales pointillées rouges représentent le début et la fin du déplacement. La source de référence est choisie de sorte à ce qu'elle reproduise un niveau sonore équivalent dans la zone avec une distance source/micro équivalente (**Eq.** (2.37)).

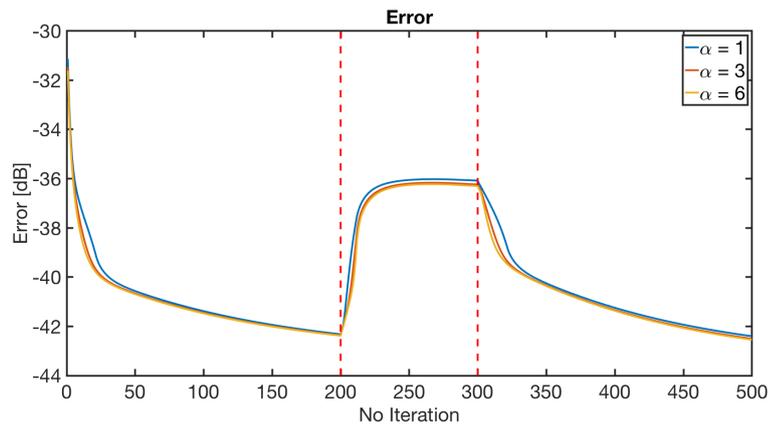


Figure A.4.9 – Erreur de reproduction pour différents α pour un auditeur se déplaçant à un d'un centimètre par itération pendant 100 itérations entre les itérations 200 et 300. Le déplacement est réalisé selon l'axe X. Les lignes verticales pointillées rouges représentent le début et la fin du déplacement.

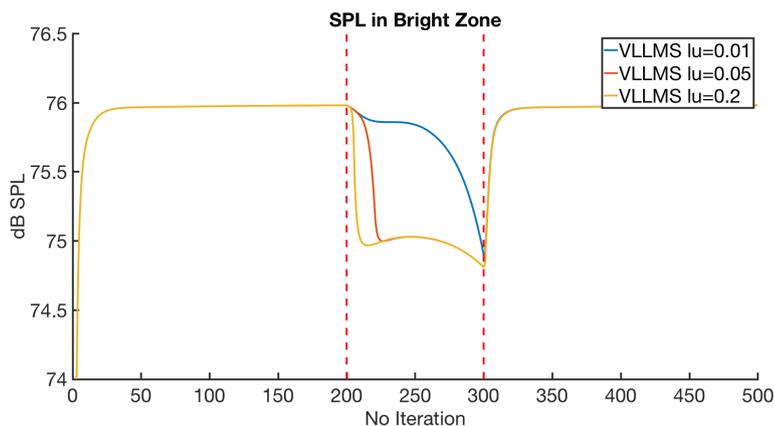


Figure A.4.10 – Niveau sonore pour différents α pour un auditeur se déplaçant à un d'un centimètre par itération pendant 100 itérations entre les itérations 200 et 300. Le déplacement est réalisé selon l'axe X. Les lignes verticales pointillées rouges représentent le début et la fin du déplacement.

E Champs sonores NLMS et APA

Cette annexe présente les figures complémentaires pour l'augmentation de l'effort en fonction de la fréquence pour les méthodes NLMS et APA. Les conditions de comparaison sont identiques à celles de la partie 3.1.4. On observe que l'effet mémoire est présent de façon équivalente dans le cas des filtre NLMS et APA.

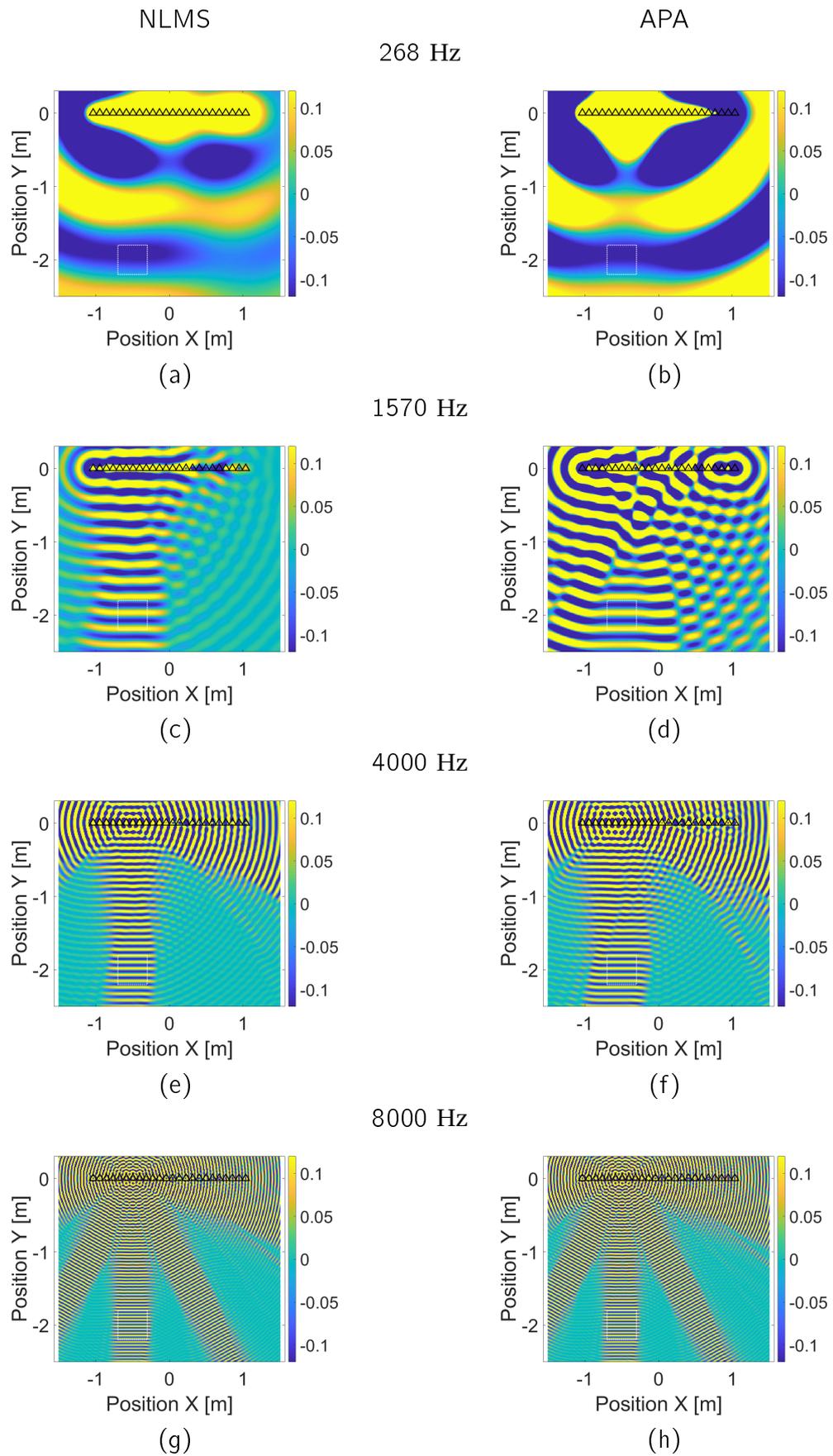


Figure A.5.11 – Champ sonore pour une seule *Bright zone* pour différentes fréquences, dans le cas des méthodes NLMS (a),(c),(e),(g) et APA (b),(d),(f),(h)

F Effet mémoire des méthodes NLMS et APA

Cette annexe présente les champs sonores pour les méthodes NLMS et APA dans les conditions proposées en §3.3 dans le but de mettre en évidence l'effet mémoire. Il est à noter que dans le cas du filtre APA, l'effet mémoire devrait pouvoir être réduit en imposant aux itérations $n - k$ $k \in [1, K - 1]$ d'avoir une pression cible $\mathbf{p}_T(n - k) = \mathbf{0}$ (Ce qui n'est pas possible dans le cas du filtre NLMS). Cet aspect n'a pas été exploré dans la thèse car peu d'intérêt a été porté à l'amélioration du filtre APA à cause de son coût de calcul élevé. Cet aspect pourrait faire l'objet d'évaluations complémentaires.

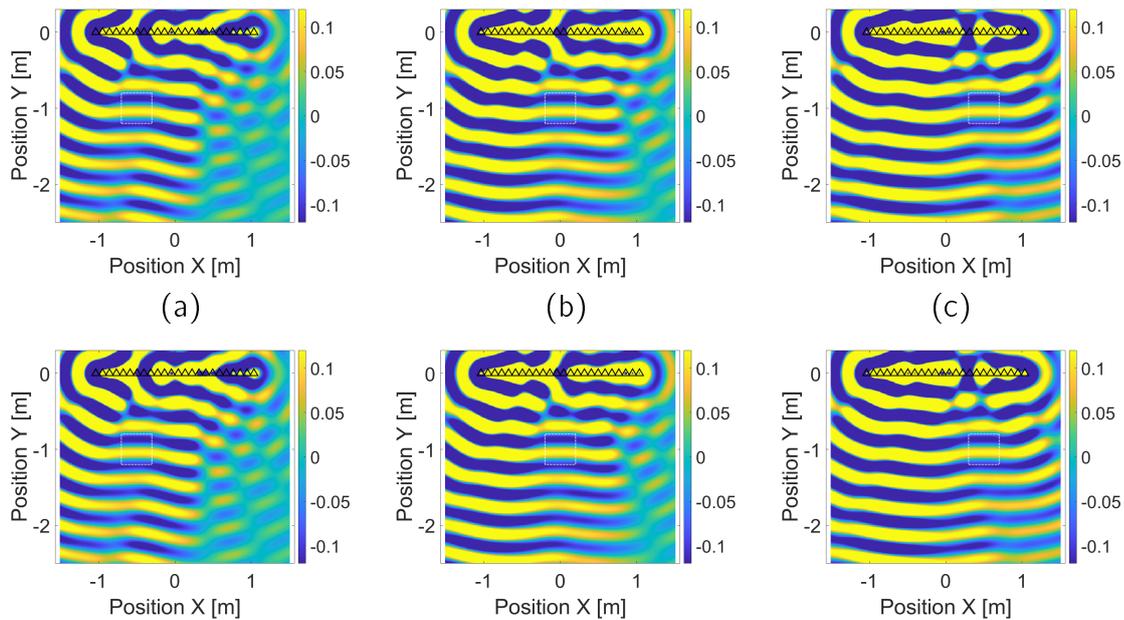


Figure A.6.12 – Champ sonore pour les méthodes NLMS (Haut) et APA (Bas) pour les itérations 200 (a), 250 (b) et 300 (c) pour un déplacement selon l'axe X.

Bibliographie

- [Ahrens et Spors, 2010] Ahrens, J. et Spors, S. (2010). Sound field reproduction using planar and linear arrays of loudspeakers. *IEEE transactions on audio, speech, and language processing*, 18(8):2038–2050.
- [Ajdler *et al.*, 2006] Ajdler, T., Sbaiz, L. et Vetterli, M. (2006). The plenacoustic function and its sampling. *IEEE transactions on Signal Processing*, 54(10):3790–3804.
- [Albert et Gardner, 1967] Albert, A. E. et Gardner, L. A. (1967). Stochastic approximation and nonlinear regression, research monograph no. 42.
- [Atal et Schroeder, 1966] Atal, B. S. et Schroeder, M. R. (1966). Apparent sound source translator. US Patent 3,236,949.
- [Batterson et Smillie, 1989] Batterson, S. et Smillie, J. (1989). The dynamics of rayleigh quotient iteration. *SIAM journal on numerical analysis*, 26(3):624–636.
- [Baykaner *et al.*, 2015] Baykaner, K., Coleman, P., Mason, R., Jackson, P. J., Francome, J., Olik, M. et Bech, S. (2015). The relationship between target quality and interference in sound zone. *Journal of the Audio Engineering Society*, 63(1/2):78–89.
- [Berkhout, 1988] Berkhout, A. J. (1988). A holographic approach to acoustic control. *Journal of the audio engineering society*, 36(12):977–995.
- [Betlehem et Abhayapala, 2005] Bethlehem, T. et Abhayapala, T. D. (2005). Theory and design of sound field reproduction in reverberant rooms. *The Journal of the Acoustical Society of America*, 117(4):2100–2111.
- [Betlehem *et al.*, 2010] Bethlehem, T., Anderson, C. et Poletti, M. (2010). A directional loudspeaker array for surround sound in reverberant rooms. *In Proc. Int. Conf. Acoust.(ICA'10)*, pages 1–6.
- [Betlehem *et al.*, 2015] Bethlehem, T., Zhang, W., Poletti, M. A. et Abhayapala, T. D. (2015). Personal sound zones : Delivering interface-free audio to multiple listeners. *IEEE Signal Processing Magazine*, 32(2):81–91.
- [Bhaya et Kaszkurewicz, 2006] Bhaya, A. et Kaszkurewicz, E. (2006). *Control perspectives on numerical algorithms and matrix problems*, volume 10. SIAM.
- [Bitmead et Anderson, 1980] Bitmead, R. et Anderson, B. (1980). Performance of adaptive estimation algorithms in dependent random environments. *IEEE Transactions on Automatic Control*, 25(4):788–794.
- [Blauert, 1983] Blauert, J. (1983). Spatial hearing.
- [Bo Moller et Olsen, 2011] Bo Moller, M. et Olsen, M. (2011). Sound zones. Master thesis, DTU, Dannemark.
- [Bohannon, 1997] Bohannon, R. W. (1997). Comfortable and maximum walking speed of adults aged 20—79 years : reference values and determinants. *Age and ageing*, 26(1):15–19.

- [Bose et Carter, 1984] Bose, A. G. et Carter, J. (1984). Headphoning. US Patent 4,455,675.
- [Boulandet *et al.*, 2017] Boulandet, R., Laurence, T. et Lissek, H. (2017). Design of remote quiet zones using spot-type sound reducers. *Acta Acustica united with Acustica*, 103(5):858–868.
- [Chang et Jacobsen, 2012] Chang, J.-H. et Jacobsen, F. (2012). Sound field control with a circular double-layer array of loudspeakers. *The Journal of the Acoustical Society of America*, 131(6):4518–4525.
- [Chatelin, 2012] Chatelin, F. (2012). *Eigenvalues of Matrices : Revised Edition*. SIAM.
- [Cheer *et al.*, 2013] Cheer, J., Elliott, S. J. et Gálvez, M. F. S. (2013). Design and implementation of a car cabin personal audio system. *Journal of the Audio Engineering Society*, 61(6):412–424.
- [Choi *et al.*, 2008] Choi, J.-W., Kim, Y., Ko, S. et Kim, J.-H. (2008). Super-directive loudspeaker array for the generation of personal sound zone. *In Audio Engineering Society Convention 125*. Audio Engineering Society.
- [Choi et Kim, 2002] Choi, J.-W. et Kim, Y.-H. (2002). Generation of an acoustically bright zone with an illuminated region using multiple sources. *The Journal of the Acoustical Society of America*, 111(4):1695–1700.
- [Coleman, 2014] Coleman, P. (2014). *Loudspeaker array processing for personal sound zone reproduction*. Thèse de doctorat, University of Surrey.
- [Coleman *et al.*, 2014a] Coleman, P., Jackson, P. J., Olik, M. et Abildgaard Pedersen, J. (2014a). Personal audio with a planar bright zone. *The Journal of the Acoustical Society of America*, 136(4):1725–1735.
- [Coleman *et al.*, 2014b] Coleman, P., Jackson, P. J., Olik, M., Møller, M., Olsen, M. et Abildgaard Pedersen, J. (2014b). Acoustic contrast, planarity and robustness of sound zone methods using a circular loudspeaker array. *The Journal of the Acoustical Society of America*, 135(4):1929–1940.
- [Coleman *et al.*, 2013] Coleman, P., Jackson, P. J., Olik, M., Olsen, M., Møller, M. et Pedersen, J. A. (2013). The influence of regularization on anechoic performance and robustness of sound zone methods. *In Proceedings of Meetings on Acoustics ICA2013*, volume 19, page 055055. ASA.
- [Conference, 2016] Conference, S. F. C. (2016). Sound Field Control Conference, Audio Engineering Society. <http://www.aes.org/conferences/2016/sfc/>.
- [Cormen, 2004] Cormen, T. (2004). *Introduction à l'algorithmique*. Dunod.
- [Donley *et al.*, 2016] Donley, J., Ritz, C. et Kleijn, W. B. (2016). Improving speech privacy in personal sound zones. *In Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on*, pages 311–315. IEEE.
- [Donley et Ritz, 2015] Donley, J. et Ritz, C. H. (2015). An efficient approach to dynamically weighted multizone wideband reproduction of speech soundfields. *In ChinaSIP 2015*. IEEE.
- [Elliott *et al.*, 2012] Elliott, S. J., Cheer, J., Choi, J.-W. et Kim, Y. (2012). Robustness and regularization of personal audio systems. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(7):2123–2133.

- [Elliott *et al.*, 2010] Elliott, S. J., Cheer, J., Murfet, H. et Holland, K. R. (2010). Minimally radiating sources for personal audio. *The Journal of the Acoustical Society of America*, 128(4):1721–1728.
- [Elliott et Jones, 2006] Elliott, S. J. et Jones, M. (2006). An active headrest for personal audio. *The Journal of the Acoustical Society of America*, 119(5):2702–2709.
- [Francombe *et al.*, 2013a] Francombe, J., Coleman, P., Olik, M., Baykaner, K., Jackson, P. J., Mason, R., Dewhurst, M., Bech, S. et Pederson, J. A. (2013a). Perceptually optimized loudspeaker selection for the creation of personal sound zones. In *Audio Engineering Society Conference : 52nd International Conference : Sound Field Control-Engineering and Perception*. Audio Engineering Society.
- [Francombe *et al.*, 2012] Francombe, J., Mason, R., Dewhurst, M. et Bech, S. (2012). Determining the threshold of acceptability for an interfering audio programme. In *Audio Engineering Society Convention 132*. Audio Engineering Society.
- [Francombe *et al.*, 2013b] Francombe, J., Mason, R., Dewhurst, M. et Bech, S. (2013b). Modelling listener distraction resulting from audio-on-audio interference. In *Proceedings of Meetings on Acoustics ICA2013*, volume 19, page 035036. ASA.
- [Gauthier et Berry, 2006] Gauthier, P.-A. et Berry, A. (2006). Adaptive wave field synthesis with independent radiation mode control for active sound field reproduction : Theory. *The Journal of the Acoustical Society of America*, 119(5):2721–2737.
- [Gauthier et Berry, 2015] Gauthier, P.-A. et Berry, A. (2015). Sound field reproduction of microphone array recordings using the lasso and the elastic-net : Theory, application examples and artistic potentials. *Proceedings of inSONIC2015, Aesthetics of Spatial Audio in Sound, Music, and Sound Art*.
- [Gauthier *et al.*, 2017] Gauthier, P.-A., Lecomte, P. et Berry, A. (2017). Source sparsity control of sound field reproduction using the elastic-net and the lasso minimizers. *The Journal of the Acoustical Society of America*, 141(4):2315–2326.
- [Gauthier *et al.*, 2016] Gauthier, P.-A., Pasco, Y. et Berry, A. (2016). Generalized singular value decomposition for personalized audio using loudspeaker array. In *Audio Engineering Society Conference : 2016 AES International Conference on Sound Field Control*. Audio Engineering Society.
- [Gerzon, 1973] Gerzon, M. A. (1973). Periphony : With-height sound reproduction. *Journal of the Audio Engineering Society*, 21(1):2–10.
- [Golub et Van Loan, 2012] Golub, G. H. et Van Loan, C. F. (2012). *Matrix computations*, volume 3. JHU press.
- [Gupta *et al.*, 2004] Gupta, N., Barreto, A. et Choudhury, M. (2004). Modeling head-related transfer functions based on pinna anthropometry. In *Proc. of the Second International Latin American and Caribbean Conference for Engineering and Technology (LACCEI)*.
- [Haykin, 2002] Haykin, S. S. (2002). *Adaptive filter theory*. Pearson Education.
- [Jin et Kleijn, 2014] Jin, W. et Kleijn, W. B. (2014). Multizone soundfield reproduction in reverberant rooms using compressed sensing techniques. In *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 4728–4732. IEEE.
- [Jungmann *et al.*, 2012] Jungmann, J. O., Mazur, R., Kallinger, M., Mei, T. et Mertins, A. (2012). Combined acoustic mimo channel crosstalk cancellation and room impulse

- response reshaping. *IEEE Transactions on Audio, Speech, and Language Processing*, 20(6):1829–1842.
- [Kamenetsky et Widrow, 2004] Kamenetsky, M. et Widrow, B. (2004). A variable leaky lms adaptive algorithm. In *Signals, Systems and Computers, 2004. Conference Record of the Thirty-Eighth Asilomar Conference on*, volume 1, pages 125–128. IEEE.
- [Kennedy *et al.*, 2007] Kennedy, R. A., Sadeghi, P., Abhayapala, T. D. et Jones, H. M. (2007). Intrinsic limits of dimensionality and richness in random multipath fields. *IEEE Transactions on Signal processing*, 55(6):2542–2556.
- [Kirkeby *et al.*, 1997] Kirkeby, O., Nelson, P. A. et Hamada, H. (1997). The "stereo dipole" : Binaural sound reproduction using two closely spaced loudspeakers. In *Audio Engineering Society Convention 102*. Audio Engineering Society.
- [Kirkeby *et al.*, 1996] Kirkeby, O., Nelson, P. A., Orduna-Bustamante, F. et Hamada, H. (1996). Local sound field reproduction using digital signal processing. *The Journal of the Acoustical Society of America*, 100(3):1584–1593.
- [Lilis *et al.*, 2010] Lilis, G. N., Angelosante, D. et Giannakis, G. B. (2010). Sound field reproduction using the lasso. *IEEE Transactions on Audio, Speech, and Language Processing*, 18(8):1902–1912.
- [Lueg, 1936] Lueg, P. (1936). Process of silencing sound oscillations. US Patent 2,043,416.
- [Mignot *et al.*, 2012] Mignot, R., Daudet, L. et Ollivier, F. (2012). Interpolation of room impulse responses in 3d using compressed sensing. In *Acoustics 2012*.
- [Møller, 1992] Møller, H. (1992). Fundamentals of binaural technology. *Applied acoustics*, 36(3-4):171–218.
- [Møller et Olsen, 2016] Møller, M. B. et Olsen, M. (2016). Sound zones : On performance prediction of contrast control methods. In *Audio Engineering Society Conference : 2016 AES International Conference on Sound Field Control*. Audio Engineering Society.
- [Møller *et al.*, 2012] Møller, M. B., Olsen, M. et Jacobsen, F. (2012). A hybrid method combining synthesis of a sound field and control of acoustic contrast. In *Audio Engineering Society Convention 132*. Audio Engineering Society.
- [Moreau, 2006] Moreau, S. (2006). *Étude et réalisation d'outils avancés d'encodage spatial pour la technique de spatialisation sonore Higher Order Ambisonics : microphone 3D et contrôle de distance*. Thèse de doctorat.
- [Nagumo et Noda, 1967] Nagumo, J.-I. et Noda, A. (1967). A learning method for system identification. *IEEE Transactions on Automatic Control*, 12(3):282–287.
- [Nascimento et Sayed, 1996] Nascimento, V. H. et Sayed, A. H. (1996). An unbiased and cost-effective leaky-lms filter. In *Conference Record of the Thirtieth Asilomar Conference on Signals, Systems and Computers*, pages 1078–1082. IEEE.
- [Nicol, 1999] Nicol, R. (1999). *Restitution sonore spatialisée sur une zone étendue : application à la téléprésence*. Thèse de doctorat, Université du Maine.
- [Nicol, 2010] Nicol, R. (2010). Binaural technology. Audio Engineering Society.
- [Nicol *et al.*, 2018] Nicol, R., Emerit, M. et Gros, L. (2018). Hrtf « prêt-à-porter » pour le son binaural dans les futurs contenus d'orange. In *Congrès Français d'Acoustique*. Société Française d'Acoustique.

- [Olik *et al.*, 2013] Olik, M., Jackson, P. J. et Coleman, P. (2013). Influence of low-order room reflections on sound zone system performance. In *Proceedings of Meetings on Acoustics ICA2013*, volume 19, page 015058. ASA.
- [Ozeki et Umeda, 1984] Ozeki, K. et Umeda, T. (1984). An adaptive filtering algorithm using an orthogonal projection to an affine subspace and its properties. *Electronics and Communications in Japan (Part I : Communications)*, 67(5):19–27.
- [Poletti, 2008] Poletti, M. (2008). An investigation of 2-d multizone surround sound systems. In *Audio Engineering Society Convention 125*. Audio Engineering Society.
- [Poletti et Abhayapala, 2011] Poletti, M. et Abhayapala, T. (2011). Interior and exterior sound field control using general two-dimensional first-order sources. *The Journal of the Acoustical Society of America*, 129(1):234–244.
- [Poletti *et al.*, 2010] Poletti, M., Fazi, F. et Nelson, P. (2010). Sound-field reproduction systems using fixed-directivity loudspeakers. *The Journal of the Acoustical Society of America*, 127(6):3590–3601.
- [Pulkki, 1997] Pulkki, V. (1997). Virtual sound source positioning using vector base amplitude panning. *Journal of the audio engineering society*, 45(6):456–466.
- [Radmanesh et Burnett, 2013] Radmanesh, N. et Burnett, I. S. (2013). Generation of isolated wideband sound fields using a combined two-stage lasso-ls algorithm. *IEEE Trans. Audio, Speech & Language Processing*, 21(2):378–387.
- [Rayleigh, 1896] Rayleigh, J. W. S. B. (1896). *The theory of sound*, volume 2. Macmillan.
- [Rice, 1929] Rice, C. W. (1929). Loud-speaker. US Patent 1,707,570.
- [Robin et Kaddouch, 2018] Robin, P. et Kaddouch, I. (2018). Unidirectional loudspeaker enclosure. US Patent App. 10/129,657.
- [Roussel et Nicol, 2018a] Roussel, G. et Nicol, R. (2018a). Méthode itérative de contrôle de champ sonore avec correction de l'erreur induite par le déplacement de l'auditeur. Brevet, App. 201659FR01.
- [Roussel et Nicol, 2018b] Roussel, G. et Nicol, R. (2018b). Méthode itérative de contrôle de champ sonore multizone avec contrôle de contraste homogène. Brevet, App. 201556FR01.
- [Schroeder *et al.*, 1962] Schroeder, M., Atal, B. et Kuttruff, K. (1962). Perception of coloration in filtered gaussian noise—short-time spectral analysis by the ear. *The Journal of the Acoustical Society of America*, 34(5):738–738.
- [Simon Galvez et Fazi, 2017] Simon Galvez, M. et Fazi, F. (2017). A loudspeaker array for 2 people transaural reproduction. In *International Congress in Sound and Vibration (ICSV), 2017*.
- [Simón Gálvez et Fazi, 2016] Simón Gálvez, M. F. et Fazi, F. M. (2016). Listener adaptive filtering strategies for personal audio reproduction over loudspeaker arrays. In *Audio Engineering Society Conference : 2016 AES International Conference on Sound Field Control*. Audio Engineering Society.
- [Smagowska et Pawlaczyk-Łuszczynska, 2013] Smagowska, B. et Pawlaczyk-Łuszczynska, M. (2013). Effects of ultrasonic noise on the human body—a bibliographic review. *International Journal of Occupational Safety and Ergonomics*, 19(2):195–202.

- [Song *et al.*, 2010] Song, M.-S., Zhang, C., Florencio, D. et Kang, H.-G. (2010). Personal 3d audio system with loudspeakers. *In Multimedia and Expo (ICME), 2010 IEEE International Conference on*, pages 1600–1605. IEEE.
- [Spors et Ahrens, 2008] Spors, S. et Ahrens, J. (2008). A comparison of wave field synthesis and higher-order ambisonics with respect to physical properties and spatial sampling. *In Audio Engineering Society Convention 125*. Audio Engineering Society.
- [Spors *et al.*, 2007] Spors, S., Buchner, H., Rabenstein, R. et Herbordt, W. (2007). Active listening room compensation for massive multichannel sound reproduction systems using wave-domain adaptive filtering. *The Journal of the Acoustical Society of America*, 122(1):354–369.
- [Steinberg et Snow, 1934] Steinberg, J. C. et Snow, W. B. (1934). Auditory perspective-physical factors. *Transactions of the American Institute of Electrical Engineers*, 53(1):12–17.
- [Sullenberger *et al.*, 2019] Sullenberger, R. M., Kaushik, S. et Wynn, C. M. (2019). Photoacoustic communications : delivering audible signals via absorption of light by atmospheric h₂o. *Optics letters*, 44(3):622–625.
- [Talagala *et al.*, 2014] Talagala, D. S., Zhang, W. et Abhayapala, T. D. (2014). Efficient multi-channel adaptive room compensation for spatial soundfield reproduction using a modal decomposition. *IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP)*, 22(10):1522–1532.
- [Tokuno *et al.*, 1996] Tokuno, H., Hamada, H., Kirkeby, O. et Nelson, P. A. (1996). *Binaural sound reproduction in a stereo dipole system*. Thèse de doctorat, ASA.
- [Trevino *et al.*, 2016] Trevino, J., Sakamoto, S. et Suzuki, Y. (2016). Presenting spatial sound to moving listeners using high-order ambisonics. *In Audio Engineering Society Conference : 2016 AES International Conference on Sound Field Control*. Audio Engineering Society.
- [Van Veen et Buckley, 1988] Van Veen, B. D. et Buckley, K. M. (1988). Beamforming : A versatile approach to spatial filtering. *IEEE assp magazine*, 5(2):4–24.
- [Wefers et Vorländer, 2014] Wefers, F. et Vorländer, M. (2014). Efficient time-varying fir filtering using crossfading implemented in the dft domain. *In Proceedings of the 2014 7th Medical and Physics Conference Forum Acusticum, Cracow, Poland*, pages 7–12.
- [Wen *et al.*, 2005] Wen, Y., Yang, J. et Gan, W.-S. (2005). Strategies for an acoustical-hotspot generation. *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 88(7):1739–1746.
- [Westervelt, 1963] Westervelt, P. J. (1963). Parametric acoustic array. *The Journal of the Acoustical Society of America*, 35(4):535–537.
- [Widrow et Hoff, 1960] Widrow, B. et Hoff, M. E. (1960). Adaptive switching circuits. Rapport technique, STANFORD UNIV CA STANFORD ELECTRONICS LABS.
- [Williams, 1999] Williams, G. (1999). *Fourier Acoustics : Sound Radiation and Nearfield Acoustical Holography*. Elsevier Science.
- [Wu et Abhayapala, 2009] Wu, Y. J. et Abhayapala, T. D. (2009). Theory and design of soundfield reproduction using continuous loudspeaker concept. *IEEE Transactions on Audio, Speech, and Language Processing*, 17(1):107–116.

- [Wu et Abhayapala, 2011] Wu, Y. J. et Abhayapala, T. D. (2011). Spatial multizone soundfield reproduction : Theory and design. *IEEE Transactions on audio, speech, and language processing*, 19(6):1711–1720.
- [Zhang et al., 2016] Zhang, W., Abhayapala, T. D., Betlehem, T. et Fazi, F. M. (2016). Analysis and control of multi-zone sound field reproduction using modal-domain approach. *The Journal of the Acoustical Society of America*, 140(3):2134–2144.

Titre : Contributions à la mise au point de méthodes adaptatives pour la reproduction de champs sonores multizone pour des auditeurs en mouvement

Mots clés : Sound Zones, Auditeurs en mouvement, Filtrage adaptatif, Audio spatialisé

Résumé : Le nombre croissant d'appareils de diffusion de contenus audio pose le problème de partager le même espace physique sans partager le même espace sonore. Les *Sound Zones* rendent possible la reproduction de programmes audio indépendants et spatialement séparés, à partir d'un ensemble de haut-parleurs et de méthodes de reproduction de champs sonores. Le problème est alors décomposé en deux zones : la *Bright zone*, où le contenu doit être reproduit et la *Dark zone*, où il doit être annulé. De nombreuses méthodes existent pour résoudre ce problème, mais la plupart ne traite que le cas d'auditeurs en position statique. Elles s'appuient sur la résolution directe des méthodes d'optimisation adaptative, telle que la méthode de *Pressure Matching* (PM). Or, pour des utilisateurs en mouvement, ces méthodes ont un coût de calcul trop élevé, rendant impossible leur application à un problème dynamique. Le but de cette thèse est de développer une solution présentant une complexité compatible avec un contrôle dynamique des *Sound*

Zones, tout en conservant les performances des méthodes conventionnelles. Sous l'hypothèse que les déplacements sont lents, une résolution itérative du problème PM est proposée et évaluée. Les algorithmes LMS, NLMS et APA sont comparés sur la base de simulations en champ libre. On s'est d'abord intéressé au problème restreint à la *Bright zone*. La méthode LMS s'avère la plus avantageuse en termes de complexité, mais elle souffre d'une erreur de reproduction. Un effet mémoire limitant la réactivité des algorithmes est aussi mis en évidence. Il est corrigé en implémentant une variante introduisant un facteur d'oubli (*Variable Leaky LMS* ou VLLMS). Cette solution est ensuite appliquée avec succès au problème global incluant la *Dark zone*. Deux améliorations, qui tirent parti de la formulation itérative, sont apportées d'une part sur les spécifications dans la *Dark zone*, et d'autre part au facteur d'oubli contrôlé par le déplacement de l'auditeur.

Title : Contributions to the development of adaptive methods for the reproduction of multizone sound fields for moving listeners

Keywords : Sound zones, Moving listeners, Adaptive filtering, Spatial audio

Abstract : The growing number of audio devices raises the problem of sharing the same physical space without sharing the same sound space. Sound Zones make it possible to play independent and spatially separated audio programs by loudspeaker array in combination with sound field reproduction methods. The problem is then split into two zones: the Bright zone, where the audio content must be reproduced and the Dark zone, where it must be cancelled. There are many methods available to solve this problem, but most only deal with auditors in a static position. They are based on the direct resolution of adaptive optimization methods, such as the Pressure Matching (PM) method. However, for moving users, these methods have a too high computation cost, making it impossible to apply them to a dynamic problem. The aim of this thesis is to develop a solution offering a level of complexity compatible with a dynamic control of Sound Zones,

while maintaining the performance of conventional methods. Under the assumption that displacements are slow, an iterative resolution of the PM problem is proposed and assessed. The LMS, NLMS and APA algorithms are compared on the basis of free field simulations. First, the problem restricted to the Bright zone was addressed. The LMS method is the most advantageous in terms of complexity, but it suffers from a reproduction error. A memory effect limiting the reactivity of the algorithms is also highlighted. It is corrected by implementing a leaky variant (Variable Leaky LMS or VLLMS) introducing a forgetting factor. This solution is then successfully applied to the global problem including the Dark zone. Two improvements, which take advantage of the iterative formulation, are made with respect to the specifications in the Dark zone, and on the forgetting factor controlled by the listener's movement.