



**HAL**  
open science

# Réseaux à grand nombre de microphones : applicabilité et mise en œuvre

Charles Vanwynsberghe

► **To cite this version:**

Charles Vanwynsberghe. Réseaux à grand nombre de microphones : applicabilité et mise en œuvre. Acoustique [physics.class-ph]. Université Pierre et Marie Curie - Paris VI, 2016. Français. NNT : 2016PA066474 . tel-01508547

**HAL Id: tel-01508547**

**<https://theses.hal.science/tel-01508547>**

Submitted on 14 Apr 2017

**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.

# École Doctorale SMAER

Institut d'Alembert

## Thèse de doctorat

Présentée par

**Charles VANWYNSBERGHE**

pour obtenir le grade de

DOCTEUR DE L'UNIVERSITÉ PARIS VI

---

# Réseaux à grand nombre de microphones : applicabilité et mise en œuvre

---

Spécialité : Acoustique

Soutenue le 12/12/2016 devant un jury composé de :

---

Rapporteur	<b>M. Jérôme Antoni</b>	Professeur, INSA Lyon
Rapporteur	<b>M. Manuel Melon</b>	Professeur, Université du Maine
Examineur	<b>M. Frank Poisson</b>	Ingénieur, SNCF
Président du jury	<b>M. Jean-Luc Zarader</b>	Professeur, UPMC Paris 6
Directeur de thèse	<b>M. Pascal Challande</b>	Professeur, UPMC Paris 6
Co-encadrant de thèse	<b>M. François Ollivier</b>	Maître de conférences, UPMC Paris 6
Co-encadrant de thèse	<b>M. Jacques Marchal</b>	Maître de conférences, UPMC Paris 6
Co-encadrant de thèse	<b>M. Régis Marchiano</b>	Professeur, UPMC Paris 6

# Résumé

La localisation de sources acoustiques consiste à retrouver la position et la signature spectrale des différentes sources dans l'espace. Ce problème inverse nécessite la mesure de la pression acoustique en différents points de l'espace avec une antenne de microphones, et s'appuie sur un modèle de propagation dans le milieu. L'ensemble des techniques connues pour la localisation des sources requiert de façon cruciale la connaissance précise de la géométrie de l'antenne. Pour améliorer la précision de la localisation, une solution consiste à augmenter le nombre de microphones et l'envergure de l'antenne. Mais cette solution fait face à de nombreux défis technologiques et scientifiques.

Récemment l'apparition de microphones numériques MEMS (*Micro Electro Mechanical System*) a ouvert de nouvelles perspectives. L'une d'elles est le développement de systèmes d'acquisition massivement multi-canaux de grande envergure. De tels systèmes répondent aux conditions requises pour localiser les sources. En revanche, de nouvelles contraintes se posent. La première est le flux élevé de données issues de l'antenne, qui doit être traité en un temps raisonnable. La deuxième contrainte est de connaître la position des nombreux microphones déployés *in situ*. Ce manuscrit propose des méthodes répondant à ces deux contraintes.

La première partie contient la description du système d'acquisition développé au cours de cette thèse. Deux points sont abordés : la caractérisation des microphones MEMS, et l'imagerie en temps réel. Dans un premier temps, on montre que les microphones MEMS ont les caractéristiques adaptées aux applications d'antennerie. Ensuite, le traitement en temps réel des signaux acquis est accompli grâce à une implémentation sur carte graphique. Cette stratégie est une solution au problème de flux de données. On peut ainsi construire un outil d'imagerie temps réel de sources large bande, permettant d'établir un diagnostic dynamique de la scène sonore sur une durée arbitraire.

La deuxième partie présente différentes méthodes robustes de calibration géométrique pour retrouver la position de microphones à partir de la mesure des signaux microphoniques. En effet, dans des conditions réelles d'utilisation, les méthodes actuelles sont inefficaces pour des antennes étendues et à grand nombre de microphones. Ce manuscrit propose des techniques qui permettent de garantir la robustesse du processus de calibration. Les méthodes proposées couvrent les différents environnements acoustiques existants, du champ libre au champ réverbérant. Leur efficacité est prouvée par différentes campagnes expérimentales.

## Mots-clé :

Grand réseau de microphones, MEMS, problèmes inverses acoustiques, localisation de sources, imagerie acoustique, calibration géométrique, GPU, parcimonie.

# **Implementation and applicability of very large microphone arrays**

# Abstract

Acoustic source localization consists in identifying the position and spectrum of sources in the environment. This inverse problem requires measuring the acoustical pressure at different points in space with a microphone array, and relies on a wave propagation model. Most methods need the accurate knowledge of the microphone array geometry. In order to enhance the localization performance, one solution consists in increasing the number of microphones and the array aperture. Yet this solution implies several technological and scientific challenges.

Recently, digital MEMS (*Micro Electro Mechanical System*) microphones came out and have opened new perspectives. One of them is the design of large-aperture and massively multichannel acquisition systems. Such systems meet the good requirements for source localization. However, new problems arise. First, an important data flow comes from the array, and must be processed fast enough. Second, if the large array is set up *in situ*, retrieving the position of numerous microphones becomes a challenging task. This thesis proposes methods dealing with these two problems.

The first part exhibits the description of the acquisition system, which has been developed during the thesis. Two points are dealt with : the MEMS microphone characterization, as well as real-time acoustic imaging. First, we show that MEMS microphone characteristics are suitable for array processing applications. Then, real-time processing of channel signals is achieved by a graphics card implementation. This strategy is one solution to the heavy data flow processing issue. In this way, a real-time acoustic imaging tool can be developed, and enables a dynamic wide-band diagnosis, for an arbitrary duration.

The second part presents several robust geometric calibration methods : they retrieve microphone positions, based only on the array acoustic signals. Indeed, in real-life conditions, the state of the art methods are inefficient with large arrays. This thesis proposes techniques that guarantee the robustness of the the calibration process. The proposed methods allow calibration in the different existing soundscapes, from free field to reverberant field. Various experimental scenarios prove the efficiency of the methods.

## **Keywords :**

Large-aperture microphone array, very large array, MEMS microphone, array processing, inverse problems, acoustic source localization, acoustic imaging, geometric calibration, GPU, sparsity.

# Remerciements

Avant tout, je remercie chaleureusement mes encadrants Pascal, Régis, Jacques et François, de m'avoir donné la chance de faire une thèse dans leur équipe. Travailler avec eux a toujours été un plaisir, on a fait ensemble un véritable travail d'équipe sur le projet *Megamicros* de réseaux de microphones MEMS. Il y a toujours eu une dynamique excellente, une motivation de la part de tous qui a permis de construire beaucoup sur un projet relativement neuf. Mes encadrants m'ont toujours laissé aller au bout de mes idées et choix, tout en réussissant à m'orienter quand il le fallait pour garder une bonne cohésion globale de mon travail final. Grâce à eux, j'ai pu être véritablement passionné par mon sujet de thèse. Je tiens particulièrement à remercier et féliciter Hélène, Christian et Dominique, des équipes techniques d'électronique et de mécanique à Saint-Cyr. Ils ont conçus les systèmes, construits tout le matériel que j'ai utilisé lors des expériences de ma thèse. Cela représente un travail colossal mais très précieux, à partir duquel j'ai pu construire mes trois ans de travaux de recherche, et obtenir mon savoir-faire actuel en réseaux à grand nombre de microphones.

Je tiens également à remercier les membres du Jury : Manuel Melon et Jérôme Antoni pour avoir accepté d'être rapporteur de mon manuscrit, et aussi Frank Poisson et Jean-Luc Zarader pour être venu assister à ma soutenance en tant qu'examinateur et Président. J'ai été ravi de voir, notamment pendant les questions, leur grand intérêt envers mes travaux de thèse.

Pendant mon doctorat, j'ai aussi eu la chance de participer à plusieurs workshops avec différents chercheurs de l'Inria et de l'EPFL. Cela a été l'occasion de faire des collaborations scientifiques extrêmement enrichissantes, de rencontrer une communauté mathématique qui m'a beaucoup appris. Grâce à eux j'ai pris de la hauteur sur différents outils théoriques et algorithmiques que j'ai employé durant la thèse. Pour cela je remercie tous les chercheurs réunis lors de ces workshops, et particulièrement Antoine Liutkus, Mathieu Fontaine, Laurent Daudet, Thibault Nowakowski et Gilles Puy.

En parallèle de ma thèse, j'ai pu vivre dans un institut avec une ambiance qui, je pense, est exceptionnelle. La bonne humeur et la cohésion entre les non-permanents a été une motivation humaine forte pour venir au labo (qui, on ne le dira jamais assez, devient particulièrement importante dès qu'il faut plus d'une semaine pour arriver à bout d'une erreur dans un code). Je remercie pour cela Chia, Valentin, Joseph, Claire, David *le maléfique*, Thomas, Barend, Geoffroy, JB, Arthur, Etienne, Estelle, Elisabeth, Benjamin, Paul, Marguerite, Mehdi, Andrés, Walid, Christian, Antoine, Simona. La liste est quand même sacrément longue... désolé pour les oubliés! Merci

également aux co-bureaux Nicolas, Virgile, Omar et Hugo, avec qui il nous arrivait de discuter parfois de science, puis parfois d'autre chose que science. Et enfin, un grand merci à Raphaël Leiba, à la fois collègue proche de l'équipe d'acoustique, mais aussi ami débordant d'énergie et d'une générosité rare.

Plus personnellement je remercie Adrien et Rémy, colocataires durant ma thèse et collègues de promotion à Cachan, pour leur amitié remarquable. Et surtout, un grand merci à mes parents et ma famille, pour leur soutien et leur attention à tout moment.

# Table des matières

<b>Introduction</b>	<b>1</b>
<b>1 Le système d'acquisition</b>	<b>11</b>
1.1 L'architecture matérielle . . . . .	12
1.1.1 Des microphones MEMS pour construire une antenne protéiforme . . . .	12
1.1.2 L'interface maître . . . . .	13
1.2 Caractérisation acoustique de l'antenne de microphones . . . . .	15
1.2.1 L'expérience en chambre anéchoïque . . . . .	15
1.2.2 Résultats . . . . .	17
1.3 Quelques illustrations du système <i>in situ</i> . . . . .	20
<b>2 Imagerie acoustique large bande</b>	<b>25</b>
2.1 Introduction . . . . .	26
2.2 Localisation de sources en bande étroite . . . . .	27
2.2.1 Problème direct en bande étroite . . . . .	27
2.2.2 Reconstruction par formation de voies . . . . .	29
2.3 Localisation de sources large bande . . . . .	31
2.3.1 Problème direct . . . . .	31
2.3.2 Formation de voies large bande . . . . .	31
2.3.3 Imagerie par bande d'octave . . . . .	32
2.3.4 Comment garantir un traitement en temps réel . . . . .	34
2.4 Expériences d'imagerie . . . . .	36
2.4.1 Scénario 1 : expérience de calibrage . . . . .	37



2.4.2	Scénario 2 : l'aspirateur . . . . .	38
2.5	Discussion : impact de l'erreur en position des microphones . . . . .	42
2.5.1	Le montage manuel de l'antenne induit une erreur géométrique . . . . .	42
2.5.2	Méthodes de correction des dispersions de mesure . . . . .	43
2.6	Conclusion . . . . .	45
<b>3</b>	<b>Calibration géométrique en champ diffus</b>	<b>47</b>
3.1	Introduction . . . . .	48
3.2	Méthodologie . . . . .	51
3.2.1	Description du problème . . . . .	51
3.2.2	Calibration géométrique : description fonctionnelle globale . . . . .	51
3.3	Estimation des distances par paire en champ diffus . . . . .	52
3.3.1	La cohérence en champ diffus . . . . .	52
3.3.2	Estimation des distances par paire entre microphones . . . . .	54
3.3.3	Limitations de l'estimation des distances par régression . . . . .	54
3.4	Estimation robuste de la géométrie du réseau . . . . .	57
3.4.1	Le MDS pour la localisation de capteurs . . . . .	57
3.4.2	Le RMDS garantit la robustesse de l'estimation . . . . .	61
3.4.3	Le LRMDs, une approche locale du RMDS adaptée aux grands réseaux . . . . .	62
3.4.4	Algorithme de résolution . . . . .	62
3.5	La réalisation expérimentale . . . . .	65
3.5.1	Scène et dispositif expérimentaux . . . . .	65
3.5.2	Estimation des cohérences et des distances . . . . .	67
3.5.3	Le LRMDs en pratique : la configuration paramétrique . . . . .	69
3.6	Résultats expérimentaux . . . . .	73
3.6.1	Performances de la méthode proposée . . . . .	73
3.6.2	Comparaison de résultats avec d'autres approches de la littérature . . . . .	76
3.7	Discussions . . . . .	81
3.7.1	Applicabilité à l'imagerie acoustique . . . . .	81

3.7.2	Au delà du milieu réverbérant . . . . .	84
3.7.3	Matrice de distances et loi de distribution . . . . .	88
3.8	Conclusion . . . . .	89
<b>4</b>	<b>Calibration géométrique par temps de vol</b>	<b>91</b>
4.1	Introduction . . . . .	92
4.2	Calibration géométrique par temps de vol . . . . .	94
4.2.1	Le protocole classique . . . . .	94
4.2.2	Une approche par RMDS robuste aux phénomènes de réflexion . . . . .	99
4.2.3	Expérience numérique et comparaison . . . . .	100
4.3	Calibration géométrique par différence de temps de vol . . . . .	103
4.3.1	Le protocole classique . . . . .	103
4.3.2	Première approche robuste : débruitage préalable de l'observable . . . . .	106
4.3.3	Deuxième approche robuste : calibration conjointe à l'identification des observables bruités . . . . .	109
4.3.4	Expérience numérique et comparaison . . . . .	112
4.4	Analyse globale des approches dans une expérience <i>in situ</i> . . . . .	116
4.5	Conclusion . . . . .	118
	<b>Conclusions</b>	<b>121</b>
<b>A</b>	<b>Le modèle de cohérence en champ diffus</b>	<b>125</b>
A.1	Cas du champ isotrope 3D . . . . .	125
A.1.1	Cas d'un champ diffus harmonique, à $\omega = \omega_0$ . . . . .	125
A.1.2	Cas d'un champ diffus large bande . . . . .	127
A.2	Cas particulier du champ isotrope plan . . . . .	128
A.3	Estimation de la fonction de Green du milieu par l'écoute du champ diffus . . . . .	129
<b>B</b>	<b>LRMDS : majorante et conditions de pondération</b>	<b>131</b>
B.1	Construction de la fonction majorante . . . . .	131
B.1.1	Construction de $g$ : cas du RMDS, sans pondération . . . . .	132

B.1.2	Construction de $g$ : cas du RMDS pondéré . . . . .	134
B.2	Propriétés du RMDS . . . . .	135
	<b>Bibliographie</b>	<b>137</b>

# Nomenclature

## Notations générales

$f_s$	Fréquence d'échantillonnage
$c_0$	Célérité du son
$p_n(t)$	Signal temporel de pression issu du microphone $n$
$\hat{p}_n(f)$	Transformée de Fourier à court terme du signal $p_n$
$\hat{\mathbf{p}}$	Vecteur de pressions mesurée à une fréquence donnée
$f, f_l$	Fréquence dans le domaine continu, pointe en fréquence discretisée
$r_{nm}(\tau)$	Intercorrélation temporelle du couple de signaux $(p_n, p_m)$
$S_{nm}(f)$	densité interspectrale de puissance du couple de signaux $(p_n, p_m)$ à la fréquence $f$
$\Gamma_{nm}(f)$	Cohérence des signaux issu du couple de signaux $(p_n, p_m)$ , à la fréquence $f$
$\tilde{\Gamma}_{nm}(f)$	Estimation de $\Gamma_{nm}(f)$

## Indices et dimensions

$n, m$	Indices de microphone
$i$	Indice de source
$l$	Indice de fréquence
$M$	Nombre de microphones, taille du vecteur de mesure
$I$	Nombre de sources, taille du vecteur source
$L$	Nombre de fréquences
$D$	Dimension spatiale du réseau de microphones

## Paramètres géométriques

$\mathbf{x}_n, \mathbf{X}$	Position du microphone $n$ , puis forme matricielle de taille $D \times M$ , pour $M$ microphones
$\mathbf{s}_i, \mathbf{S}$	Position de la source $i$ , puis forme matricielle de taille $D \times I$ , pour $I$ sources

$d_{nm}$ , $\mathbf{D}$	Distance théorique entre les microphones $n$ et $m$ , puis forme matricielle
$\delta_{nm}$ , $\mathbf{\Delta}$	Distance estimée entre les microphones $n$ et $m$ par régression sur les cohérences, puis forme matricielle
$t_{im}$	Temps de vol théorique entre la source $i$ et le microphone $m$
$\tau_{im}$	Temps de vol estimé par GCC-PHAT, entre la source $i$ et le microphone $m$
$t_{imn}$ , $\mathbf{T}_i$	Différence de temps de vol théorique entre les microphones $n$ et $m$ issu de la source $i$ , puis forme matricielle telle que $[\mathbf{T}_i]_{mn} = t_{imn}$
$\tau_{imn}$ , $\mathbf{T}_i$	Différence de temps de vol estimé par GCC-PHAT entre les microphones $n$ et $m$ , issu de la source $i$ , puis forme matricielle telle que $[\mathbf{T}_i]_{mn} = \tau_{imn}$

### Notations dans les problèmes de calibration géométrique

$\epsilon_{nm}$	Terme de bruit aléatoire iid centré associé à l'observable $\delta_{nm}$ ou $\tau_{im}$
$\epsilon_{imn}$ , $\mathbf{E}_i$	Terme de bruit aléatoire iid centré associé à l'observable $\tau_{imn}$ , puis forme matricielle telle que $[\mathbf{E}_i]_{mn} = \epsilon_{imn}$
$o_{nm}$ , $\mathbf{O}$	Terme de valeur aberrante associé à l'observable $\delta_{nm}$ ou $\tilde{\tau}_{nm}$ , puis forme matricielle
$o_{imn}$ , $\mathbf{O}_i$	Terme de valeur aberrante associé à l'observable $\tau_{imn}$ , puis forme matricielle telle que $[\mathbf{O}_i]_{mn} = o_{imn}$
$\tilde{\mathbf{X}}$	Estimation des positions de microphones $\mathbf{X}$
$\tilde{\mathbf{S}}$	Estimation des positions de sources $\mathbf{S}$
$\tilde{\mathbf{O}}$	Estimation des erreurs de valeurs aberrantes $\mathbf{O}$
$\nu$	Paramètre de régularisation de la norme $\ell_1$ sur $\mathbf{O}$ ou $\mathbf{O}_i$
$\kappa$	Critère d'arrêt des algorithmes itératifs

### Notations mathématiques

$\mathbb{A}_h^M$	Ensemble des matrices $M \times M$ antisymétriques à éléments diagonaux nuls
$\mathbb{S}_h^M$	Ensemble des matrices $M \times M$ symétriques à éléments diagonaux nuls
$\cdot^T$	Transposition
$\cdot^H$	Transposition hermitienne
$\text{tr } \mathbf{A}$	Trace de la matrice $\mathbf{A}$
$\text{diag}(v)$	Matrice diagonale carrée dont la diagonale contient les éléments du vecteur $v$
$\ \mathbf{A}\ _F$	Norme de Frobenius de $\mathbf{A}$ , $(\sum_{n,m} [\mathbf{A}]_{nm}^2)^{1/2}$
$\mathcal{H}_{2K}(\mathbf{A})$	Opérateur de <i>hard thresholding</i> , convertant les $2K$ éléments de plus grande valeur absolue dans la matrice $\mathbf{A}$
$\mathcal{S}_\nu(u)$	Opérateur de <i>soft thresholding</i> sur la variable $u$ , avec seuillage de valeur $\nu$

# Introduction

## Contexte

De nombreuses applications nécessitent l'utilisation d'une antenne de microphones [6]. La plus courante est la localisation de sources, qui consiste à retrouver la position spatiale et la signature acoustique (*i.e.* l'amplitude et la phase) des sources rayonnant dans le milieu. L'imagerie acoustique est une application particulière de la localisation pour le diagnostic. Elle cartographie les niveaux acoustiques sur une échelle de couleurs, permettant de situer et évaluer les sources par interprétation visuelle. Les applications intéressent essentiellement les industriels, en particulier les constructeurs de moyens de transports de tous les types : routier, ferroviaire ou aéronautique. En effet, le bruit engendré par le transport est soumis à des réglementations pour limiter la nuisance sonore. La localisation de sources devient alors un enjeu important lors du développement et de la conception d'un véhicule. Elle permet d'identifier les sources de bruits sur la structure, comprendre la cause physique, et ainsi réduire le bruit global. Il existe également des applications au multimédia ou aux systèmes de conférence, dans lesquels on cherche à réduire le bruit ambiant dans un enregistrement vocal.

Toutes ces applications sont fondées sur le principe de *filtrage* spatial. Il s'agit d'un problème inverse. Les sources acoustiques sont les inconnues recherchées, et l'antenne de microphones donne l'observable connu de la scène sonore. Deux types d'antennes de microphones existent :

- les réseaux synchrones, dont les canaux ont une synchronisation temporelle à l'échantillon près. Dans ce cas, tous les canaux sont contrôlés par une horloge commune. Un réseau synchrone nécessite donc un matériel dédié, où tous les canaux sont connectés à une interface maître ;
- les réseaux asynchrones : l'acquisition est issue de différents microphones connectés à un réseau sans fil, typiquement plusieurs téléphones portables. Les réseaux asynchrones sont intéressants du fait qu'ils ne nécessitent pas un matériel dédié. Cependant les décalages temporels entre les différents canaux sont inconnus, et peuvent varier au cours du temps. Ils doivent nécessairement être déterminés pour effectuer de la localisation de sources par les méthodes de l'état de l'art.

Par la suite, nous considérons uniquement la famille des réseaux synchrones, dont les canaux

sont supposés synchronisés dans le temps. Cette dernière condition est essentielle pour exploiter la plupart des techniques de la littérature.

La formation de voies est la première technique utilisée pour résoudre le problème inverse sur les sources acoustiques [16]. Elle rétropropage les mesures dans les directions de l'espace analysé, en inversant les effets de propagation décrits par la fonction de Green du milieu entre les sources et les microphones. En pratique, il s'agit de compenser les retards de propagation, et éventuellement la décroissance de l'amplitude. Suite à cette étape, les signaux compensés de tous les microphones sont sommés. La somme de ces signaux est constructive, et a théoriquement la valeur maximale aux points où se trouvent les sources [164]. Donc la formation de voies permet *a priori* de localiser les sources aux valeurs maximales du signal formé. Aujourd'hui, cette technique reste souvent privilégiée du fait de sa simplicité et de sa robustesse aux bruits et incertitudes de mesures des microphones. Cependant, elle est sensible au choix de la géométrie de l'antenne. En effet, la mesure des pressions, ponctuelles et en nombre fini, révèle le diagramme de directivité de l'antenne :

- d'une part le lobe principal au niveau de la source s'élargit quand la fréquence diminue. Cela a pour effet de dégrader la **résolution**, c'est à dire la précision de localisation spatiale. En revanche, à fréquence donnée ce lobe devient plus étroit en augmentant la taille de l'antenne ;
- d'autre part la présence de lobes secondaires peut masquer les sources acoustiques voisines, de niveau plus faible. Cela dégrade la **dynamique** de détection, qui est quantifiée par la différence de niveau entre le lobe principal et le lobe secondaire prépondérant ;
- pire encore, si la structure de l'antenne est régulière, l'échantillonnage spatial de la mesure engendre le phénomène de **recouvrement** en haute fréquence. Cela fait apparaître des lobes de réseau, qui sont une réplique du lobe principal mais dans des directions différentes. Dans ce cas, les lobes de réseau créent une ambiguïté pour localiser la vraie position de la source. Le problème de recouvrement semble *a priori* le plus gênant, mais il peut être simplement supprimé : il suffit de casser la régularité de la géométrie du réseau, en disposant les microphones de façon aléatoire.

Pour concevoir la géométrie optimale d'un réseau, il existe encore d'autres approches, par exemple fondées sur l'erreur d'interpolation du champ acoustique dans un volume [32]. Cependant, le diagramme de directivité reste le critère le plus répandu . Les méthodes existantes consistent généralement à trouver le meilleur compromis entre la résolution et la dynamique.

## Les réseaux de microphones : deux tendances opposées existent

Précédemment, nous avons identifié deux caractéristiques majeures pour la localisation de sources : la résolution et la dynamique. En améliorant ces deux points, le rapport signal bruit d'une source extraite d'une scène sonore par formation de voies augmente. Théoriquement, la manière naïve d'améliorer le rapport signal bruit requiert d'augmenter à la fois la taille globale

de l'antenne, ainsi que le nombre de microphones [124]. Mais jusqu'à récemment, une antenne (réseau synchrone) nécessitait une mise en œuvre complexe et coûteuse, ce qui limitait rapidement le système à quelques dizaines de voies.

### **1. Exploiter un réseau de petite taille et à faible nombre de microphones, avec une méthode haute résolution**

Pour compenser cette limitation technologique, la première tendance globale consiste à résoudre le problème inverse sur les sources avec une technique alternative à la formation de voies, pour réduire voire supprimer la réponse de l'antenne de petite taille. La première approche connue consiste à minimiser la contribution des lobes secondaires, tout en préservant un gain fixe dans la direction de la source. Cette approche, appelée CAPON [29] ou Minimum Variance Beamforming [91], nécessite le calcul statistique de la matrice de covariance des pressions mesurées, ce qui impose des sources stationnaires et fixes. Une seconde approche consiste à ôter la contribution du diagramme de directivité dans la formation de voies, par une optimisation globale avec DAMAS [24] ou de façon itérative sur chaque source avec CLEAN [127]. Enfin, si le nombre de sources acoustiques est petit relativement à la dimension de la scène sonore étudiée, il est possible de résoudre le problème inverse par approximation parcimonieuse des sources [115, 166, 167]. Globalement, les alternatives à la formation de voies sont un atout précieux pour utiliser des petites antennes, mais au prix d'une robustesse moindre aux bruits et/ou incertitudes de mesures, ou encore aux erreurs sur le modèle de propagation. Elles peuvent aussi imposer des contraintes de mise en œuvre, liées à des hypothèses fortes telles que la stationnarité des sources, ou encore liées au coût de calcul.

### **2. Exploiter un réseau de grande taille et à nombre élevé de microphones**

La deuxième tendance, plus rare, consiste à revenir sur la stratégie naïve d'augmenter le nombre de microphones jusqu'à plusieurs centaines, sur une envergure supérieure au mètre. Bien que le premier intérêt est d'améliorer la résolution et la dynamique de la formation de voies, il est également possible d'entourer une source, voire une scène sonore complète, par les microphones. Des applications nouvelles sont alors possibles, telles que l'identification du diagramme de directivité tridimensionnel d'une source [12, 115], ou encore la suppression du champ réverbérant d'une salle de la mesure [33]. Quelques équipes de recherche ont conçu des systèmes avec différentes architectures. Ces dernières sont plus ou moins optimales selon les différents critères inhérents à la multiplication du nombre de voies :

- l'encombrement de la connectique et du bâti;
- la complexité de la conception électronique et de la fabrication du système;
- la complexité de maintenance, pour remplacer les voies défectueuses;
- la difficulté et le temps nécessaire pour déployer le système *in situ*.

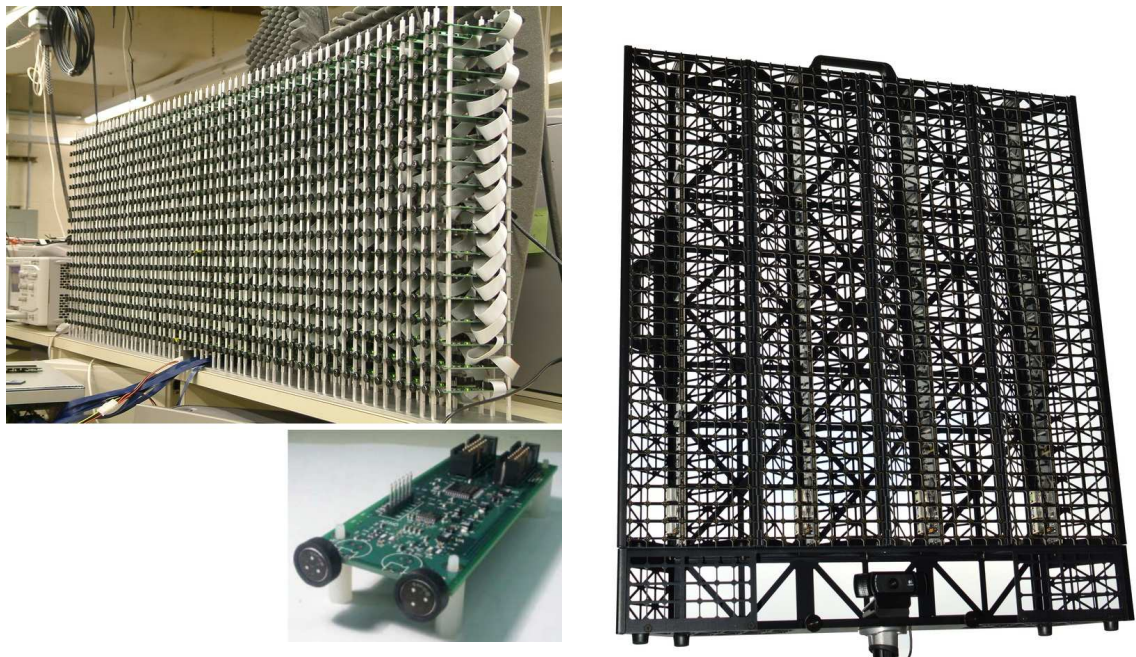


Le premier type d'architecture consiste à relier tous les microphones directement vers un étage d'amplification et de conversion analogique numérique, puis un étage de sérialisation des données. Cette méthode nécessite un câble par microphone, transférant du signal analogique. Elle est donc la moins optimale en terme de déploiement et de rapport signal bruit sur les signaux mesurés. Cependant, elle est la plus courante et la plus simple à mettre en œuvre puisqu'il s'agit de paralléliser une chaîne d'instrumentation classique. De ce fait, elle est présente à la fois dans le domaine de la recherche (120 [115], 128 [143, 144] ou 256 [126] voies) et industriel (par exemple Brüel & Kjaer ou 01dB).

L'équipe de Silverman *et al.*, à l'Université de Brown, a été pionnière sur la fabrication de grands réseaux à la fin des années 90. Le *Huge Microphone Array* est constitué de 512 microphones à électret [134, 135]. L'architecture est décomposée en modules de 16 microphones contenant la conversion analogique numérique, puis reliés par fibre optique à l'interface maître. Les données centralisées sont soit récupérées, soit envoyées dans un étage de traitement par formation de voies, inclus dans l'interface. Ce premier système est une prouesse technologique, et nécessite des ressources matérielles importantes. Il montre la seconde famille d'architectures : les microphones sont déportés en plusieurs sous-réseaux. L'avantage réside dans la possibilité de positionner les sous-réseaux sur une grande étendue. De plus cette architecture nécessite relativement peu de fils comparée à la précédente. Takagi *et al.* ont développé un système semblable, avec 7 sous-réseaux de 16 microphones (112 au total), et en outre la formation de voies est déportée dans chacun des sous-réseaux [142].

Enfin, le LOUD array, conçu au MIT en 2004, a longtemps été le plus grand réseau, muni de 1020 canaux [162]. Pour cela, l'équipe a conçu la structure la plus modulaire possible : la chaîne d'instrumentation est compactée au plus près de chaque microphone (cf figure 1a). Deux avantages importants sont la simplicité de fabrication et de maintenance : tous les modules sont identiques, et peuvent être facilement remplacés en cas de panne. De plus les signaux transmis sont entièrement numériques.

Dans l'ensemble de ces travaux, les validations expérimentales mettent en évidence la nette amélioration du rapport signal bruit d'une source acoustique extraite par la formation de voies, en fonction du nombre de microphones [133, 136, 163]. Toutefois ces réalisations sont principalement des défis technologiques. Elles montrent la difficulté que pose la multiplication du nombre de microphones à électret avec leur chaîne d'instrumentation. En cela l'utilisation de grandes antennes a toujours été difficile à démocratiser. Pour la suite, notons que l'architecture modulaire du LOUD array semble *a priori* la plus appropriée pour une augmentation en masse du nombre de canaux.



(a) En 2004 au MIT, le projet LOUD a accompli le défi technologique de construire une antenne modulaire de 1020 microphones, de  $1.8 \text{ m} \times 0.51 \text{ m}$  [162]. Un module de 2 microphones en bas à gauche.

(b) En 2014, l'antenne CAM<sup>1K</sup> de 1024 microphones MEMS de  $64 \text{ cm} \times 64 \text{ cm}$ , commercialisée par Sorama, a un coût de 13500 €.

**FIGURE 1** – Deux exemples illustrant l'évolution des grands réseaux de microphones ces dernières années.

## Les grands réseaux aujourd'hui : des verrous scientifiques et technologiques

Les réseaux à grand nombre de microphones relèvent trois défis majeurs :

1. la fabrication comme nous avons pu le voir précédemment. En effet, la complexité matérielle et l'encombrement des chaînes d'acquisition conventionnelles sont un frein à leur mise en œuvre à grande échelle ;
2. le flux de données élevé, globalement de plusieurs dizaines de MB/s. Il fait apparaître trois limites : le taux de transfert de données, le volume de stockage nécessaire à la sauvegarde des signaux acquis, et la vitesse de traitement des données ;
3. le déploiement du système sur un terrain d'expérience. Généralement, l'antenne est une structure rigide, facilement transportable tant que sa taille est petite, mais devient inadaptée si l'envergure atteint plusieurs mètres. Une solution est de considérer une structure souple, ou en plusieurs sous-réseaux rigides.

## 1. La fabrication du système

La contrainte de mise en œuvre, précédemment expliquée, s'est nettement réduite depuis récemment grâce à l'utilisation de microphones de type MicroElectroMechanical System (MEMS) à sortie numérique. Les premiers de ces microphones sont réalisés par Analog Device en 2011 [47], et sont aujourd'hui intégrés dans la plupart des téléphones portables. Popularisé dans les années 2000, le principe de la technologie MEMS est de réaliser la partie mécanique d'un capteur ou d'un actionneur de taille submillimétrique, et de l'intégrer dans un composant électronique assurant le conditionnement du signal [25]. La fabrication des composants MEMS suit le même procédé que celui des composants électroniques, entraînant :

- une densité d'intégration nettement supérieure à une chaîne d'instrumentation classique : les capteurs ou actionneurs ont la taille d'une puce électronique ;
- une forte homogénéité des composants car ils sont construits en échelle industrielle. Pour la même raison, le coût unitaire est nettement réduit.

Si les microphones MEMS sont initialement conçus pour des applications multimédia ou de communication, ils ont également été implémenté dans de récentes réalisations d'antennes, la première étant par Hafizovic *et al.* [63]. Sur le marché, l'entreprise Sorama vend le plus grand réseau existant (1024 voies, voir la figure 1b), et tient également le record du système avec le plus grand nombre de voies, 4096 voies simultanées [119].

La figure 1 montre l'intérêt de l'utilisation de microphones MEMS. Elle compare l'antenne de Sorama avec le *LOUD array*, deux réseaux avec le même nombre de canaux, et avec des architectures modulaires relativement proches. Les microphones MEMS apportent une compacité inégalée par rapport à la technologie classique.

Cependant au début de cette thèse, la littérature ne montre toujours pas d'usage d'une antenne de microphones MEMS avec une géométrie large. Mais le fait est que l'utilisation de tels composants à sortie numérique convient parfaitement pour une extension spatiale.

## 2. L'acquisition et son traitement

Le grand flux de données issu de la mesure acoustique multi-canal pose le problème du transfert de données du système vers un ordinateur, qui assure le stockage et/ou le traitement des signaux. Depuis une dizaine d'années, plusieurs bus de communications sont capables d'effectuer des taux de transfert de données suffisamment élevés :

- l'USB 3.0 a un taux de transfert maximal de 625 MB/s. C'est le protocole le plus répandu pour le transfert rapide de données ;
- le PCIe assure jusqu'à 15 GB/s de débit dans la version 3.0 (la plus récente actuellement) ;
- les récentes technologies Ethernet Gigabit sont conçues pour des débits supérieurs au Gb/s, c'est à dire supérieur à 120 MB/s<sup>1</sup> environ.

---

1. L'unité b correspond au bit ; le Byte (B), d'une taille de 8 bits, équivaut à l'octet.

Si l'on considère un réseau de 1000 canaux à un échantillonnage standard de 24 bits et de fréquence 44 100 Hz, il requiert un taux de transfert supérieur à 126 MB/s. Donc le problème du débit est aujourd'hui écarté.

Si les signaux sont sauvegardés pour un post-traitement, le volume devient rapidement très important. En gardant l'exemple précédent, un débit de 126 MB/s génère 7.4 GB de données en une minute. De nombreuses méthodes de compression sont employées pour réduire la taille d'un fichier audio, par exemple le célèbre format mp3. Mais elles sont dédiées à la compression d'un seul canal, et n'exploite pas les corrélations existantes entre les canaux. Koyano *et al.* proposent de représenter une acquisition multi-canal par une base d'ondes planes réduite en analyse en composantes principales, approchant la solution de Helmholtz [81]. Hormis cela, la compression multi-canal reste une piste rarement étudiée.

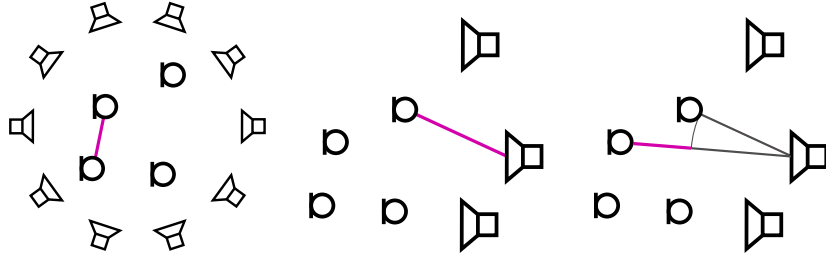
La dernière alternative reste le traitement à la volée. On ne sauvegarde alors qu'une partie sensiblement réduite des données, issue de l'algorithme de localisation de sources. Le problème de la rapidité de calcul s'impose pour localiser les sources en temps réel. D'abord, il peut être résolu par une implémentation optimisée. Les architectures de calcul parallèle montrent des performances intéressantes, et peuvent suffire pour effectuer la formation de voies en temps réel [11, 90]. Si l'optimisation de l'implémentation n'est pas suffisante, il faut réduire la quantité de données à traiter. Dans le contexte de localisation acoustique dans une salle, Lai *et al.* proposent de n'utiliser qu'une partie du réseau de microphones pour effectuer un traitement [83]. Le choix des sous-ensembles est déterminé en fonction de la position à analyser. Pour chaque point de l'espace, le critère optimise le diagramme de directivité sur une plage de fréquence, tout en minimisant le nombre de capteurs nécessaires.

Au début de la thèse, l'état de l'art ne montre pas de solution technologique, proposant à la fois une implémentation optimisée du traitement pour la localisation en temps réel, et avec des résultats expérimentaux sur une antenne à grand nombre de microphones.

### 3. Le déploiement de l'antenne

Comme il n'est plus possible de considérer l'antenne comme une unique structure rigide, son installation est faite *in situ*. Dans ce cas, sa géométrie n'est pas connue et doit être déterminée pour appliquer les méthodes de localisation de sources. Déterminer la position des capteurs dans l'espace, à l'aide d'un mètre ou télémètre laser, devient rapidement fastidieux si le nombre de microphones est grand. La détermination de la géométrie des réseaux par des méthodes acoustiques devient un avantage décisif dans leur mise en œuvre.

Les revues exhaustives de Plinge *et al.* [118] et Wang *et al.* [158] mettent en évidence trois approches principales de calibration géométrique, résumées dans la tableau 1. Les cas en champ libre (colonnes 2 et 3) sont les plus connus. Elles impliquent l'utilisation de sources contrôlées ou non contrôlées. Elles s'appuient sur la mesure de temps de vol des sources aux micros, ou des



Champ	réverbérant	libre	libre
Sources	bruit ambiant	fixes, de signaux connus et mesurés	fixes
Paramètres connus	distances entre microphones	"TOA" : distances sources-microphones	"TDOA" : différence des distances sources-microphones
Expérimenté sur grands réseaux	non	oui [76, 125]	oui [86, 115]

TABLE 1 – Positionnement de microphones : trois paradigmes possibles.

différences de temps de vol entre les micros. La résolution du problème de calibration détermine conjointement la position des microphones et celle des sources. Pour ces deux cas, des validations expérimentales prouvent la faisabilité sur des réseaux étendus, et de 120 [76, 86, 115] à 450 [125] microphones. Enfin, notons que dans le cas particulier où l'antenne déployée est constituée de plusieurs sous-réseaux de formes connues, des méthodes de calibrations géométriques ont été élaborées en s'appuyant sur l'information géométrique partiellement connue [117, 147].

Cependant ces techniques, supposant un champ libre, se dégradent en présence de réverbération. Il existe donc le deuxième cas de calibration géométrique (première colonne), consacré aux antennes déployées dans une salle réverbérante. Contrairement au cas du champ libre, il fonctionne par l'écoute du bruit ambiant. L'approche est intéressante car elle ne nécessite pas de protocole particulier. Cependant, les validations expérimentales montrent qu'elle est limitée qu'aux réseaux de petite taille [138].

Finalement dans l'ensemble, il n'existe donc pas d'approche unifiée pour résoudre le problème de calibration géométrique, pour les antennes de géométrie arbitraire, à grand nombre d'éléments répartis sur une grande étendue.

## Les objectifs de la thèse

Les travaux présentés dans ce manuscrit s'articulent autour d'une problématique commune : comment tirer bénéfice d'un grand réseau de microphones par une mise en œuvre aisée ? Le manuscrit présente avant tout des aspects méthodologiques, pour mettre expérimentalement en œuvre le déploiement et l'utilisation d'un réseau étendu sur plusieurs mètres, et comprenant

plusieurs centaines d'éléments. **Trois axes** concernant l'applicabilité et le déploiement d'un grand nombre de microphones sont présentés : l'implémentation matérielle d'une grande antenne (chapitre I), la localisation de sources par le calcul en temps réel (chapitre II), et enfin la calibration géométrique d'une grande antenne (chapitre III et IV).

En **chapitre I**, le système d'acquisition Megamicros est présenté. Il s'agit d'une antenne synchrone, constituée à ce jour de 256 microphones MEMS numériques. Elle est utilisée pour tous les travaux et validations expérimentales qui suivent. Elle est fabriquée au sein de l'Institut Jean le Rond d'Alembert, et nous discuterons de l'architecture matérielle et les choix de conceptions. Les microphones MEMS étant initialement conçus pour les téléphones portables, leur caractérisation acoustique est effectuée, pour analyser l'homogénéité des capteurs en vue d'une application à l'antennerie acoustique.

Le **chapitre II** est dédié à la localisation de sources en large bande par formation de voies. L'intérêt est avant tout de proposer une implémentation pour effectuer un diagnostic par imagerie en temps réel *in situ*. Le GPU (Graphic Processing Unit), une unité de calcul parallèle incluse dans les ordinateurs portables, est adaptée pour répondre à ce problème. Les performances de calculs sont analysées pour assurer le calcul en temps réel de la formation de voies sur le spectre audible, et avec un grand nombre de capteurs. Deux scénarios d'applications nous permettent à la fois (a) de quantifier la précision de l'outil diagnostique sur les niveaux acoustiques imagés, et (b) d'illustrer le fonctionnement et l'intérêt de l'outil développé.

Exposée dans le **chapitre III**, une première partie traitant de la calibration géométrique se place dans le contexte des milieux réverbérant et du champ diffus. Dans ce milieu l'observable connu, qui est la distance entre les microphones, augmente très rapidement (au carré du nombre de microphones) par rapport aux nombres d'inconnues sur la géométrie du réseau (proportionnel au nombre de microphones). Appliqué à une grande antenne, ce problème est donc sur-déterminé. Cependant une partie de l'observable géométrique est dégradé par des erreurs importantes, consistant en des valeurs aberrantes. La résolution du problème, sensible au bruit, est dégradée par la présence de ces valeurs aberrantes. Le chapitre propose donc une approche robuste, permettant conjointement de positionner les microphones et d'identifier les valeurs aberrantes de l'observable. La faisabilité sur les réseaux de grandes tailles est analysée expérimentalement.

Le **chapitre IV** propose d'étendre l'approche robuste employée précédemment, à la calibration en champ *quasi* libre. En effet, le problème résolu consiste toujours à conjointement retrouver la géométrie et les valeurs aberrantes de l'observable géométrique. Dans ce cas, ce sont les distances ou différences de distances entre les sources et les microphones qui sont mesurées. Nous verrons que les réflexions sur les obstacles présents dans le milieu engendrent des mesures de distance fausses. Ce dernier travail généralise donc les différentes conditions possibles pour la calibration géométrique d'un grand réseau, dans la perspective d'être appliqué à une antenne déployée *in situ*.



# Chapitre 1

## Présentation du système d'acquisition

### Sommaire

---

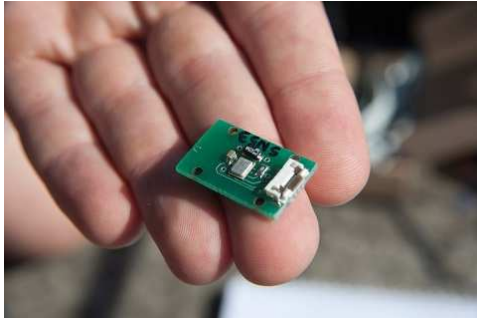
<b>1.1</b>	<b>L'architecture matérielle</b>	<b>12</b>
1.1.1	Des microphones MEMS pour construire une antenne protéiforme	12
1.1.2	L'interface maître	13
1.1.2.1	Première génération : 128 canaux	13
1.1.2.2	Deuxième génération : 256... ou 4x256 canaux	14
<b>1.2</b>	<b>Caractérisation acoustique de l'antenne de microphones</b>	<b>15</b>
1.2.1	L'expérience en chambre anéchoïque	15
1.2.2	Résultats	17
<b>1.3</b>	<b>Quelques illustrations du système <i>in situ</i></b>	<b>20</b>

---

Ce chapitre liminaire présente l'outil expérimental clé autour duquel se sont articulés mes travaux de thèse. C'est un système d'acquisition conçu à l'Institut  $\partial$ 'Alembert, qui est le fruit du projet *MegaMicros*. Son architecture globale reste semblable à celle d'un réseau de microphones conventionnel. La différence essentielle réside dans la nature des microphones implémentés : ils sont construits à la base d'une technologie MEMS à sortie digitale. Cela a permis de concevoir un système d'acquisition entièrement numérique. Grâce à ces particularités, nous disposons d'un outil expérimental (a) rapidement déployable *in situ*, (b) massivement multi-canal, (c) potentiellement à forte étendue spatiale, et (d) protéiforme, dans le sens où l'antenne de microphones est souple et donc la géométrie facilement adaptable.

Dans un premier temps, nous décrivons l'architecture matérielle globale, et ses spécificités dans sa conception. Ensuite nous présentons les caractéristiques acoustiques des microphones MEMS déterminées expérimentalement. Enfin, différentes mises en œuvre permettent d'illustrer l'utilisation du système.





(a) Un microphone MEMS : le composant a une taille  $5 \times 4 \times 1$  mm, ce qui facilite grandement l'intégration.



(b) Lignes de huit microphones MEMS : la connectique transmet uniquement des signaux numériques (horloge d'échantillonnage et signaux acoustiques multiplexés) par des câbles RJ-45.

**FIGURE 1.1** – Le choix des microphones MEMS réduit significativement l'encombrement et la complexité matérielle du réseau.

## 1.1 L'architecture matérielle

### 1.1.1 Des microphones MEMS pour construire une antenne protéiforme

Le capteur de pression acoustique utilisé est le composant ADMP441 fabriqué par Analog Device [47]. Il a été le premier microphone MEMS numérique commercialisé (2011). Le composant intègre toute la chaîne d'instrumentation :

1. le transducteur acoustique, qui assure *a priori* une acquisition omnidirectionnelle entre 60 Hz et 15 kHz ;
2. le convertisseur analogique-numérique de type  $\Sigma$ - $\Delta$  : son premier étage échantillonne le signal à une fréquence élevée, ce qui permet de repousser le bruit de quantification au-delà du spectre audible. Le second étage est un filtre numérique passe-bande, écartant les très basses fréquences ainsi que le repliement spectral en coupant les hautes fréquences au delà de  $f_s/2$ . Le dernier étage décime le signal à la fréquence audio de sortie  $f_s = 50$  kHz ;
3. Le signal numérique est transmis suivant le protocole I2S. Sur un même bus série, deux microphones envoient les échantillons codés sur 24 bits, à la fréquence  $f_s$  donnée par l'horloge de l'interface.

L'interface est maître sur tous les composants MEMS : elle cadence l'acquisition des canaux par une horloge commune définissant l'échantillonnage. Cela assure l'alignement temporel des mesures à l'échantillon près. On nomme un tel système réseau synchrone (*synchronous array* en anglais).

La forte miniaturisation des microphones par la technologie MEMS est le premier atout majeur pour un système multicanal (voir figure 1.1a). D'une taille de  $5 \times 4 \times 1$  mm, ils apportent une compacité inégalée par rapport aux microphones à électret. Ces derniers nécessitent une chaîne d'instrumentation (amplification + conversion analogique-numérique) logée dans des racks volu-

mineux, ainsi qu'une connectique plus lourde pour la bonne transmission du signal analogique. Notons que, pour réduire la connectique et simplifier le déploiement de l'antenne, un même et unique câble RJ-45 assure la connexion entre 8 microphones MEMS et l'interface (cf fig. 1.1b).

Le deuxième atout est la transmission entièrement numérique, qui est beaucoup plus robuste au bruit qu'une transmission analogique. Nous avons pu déporter les microphones jusqu'à une distance de 20 m par rapport à l'interface ; au-delà l'impédance de ligne du câble RJ-45 devient trop significative et dégrade le signal numérique. L'étendue du réseau de microphones est donc potentiellement très grande, contenue dans un disque de 40 m de diamètre.

Notons que le constructeur InvenSense a racheté la gamme de microphones MEMS d'Analog Device [71]. Aujourd'hui, il s'est spécialisé dans ces composants et en a élargi la gamme. Certains présentent des caractéristiques acoustiques améliorées (réponse en fréquence, rapport signal sur bruit, homogénéité), tandis que d'autres utilisent différents protocoles pour transmettre la sortie numérique. Un nouveau composant [72] sera particulièrement intéressant pour les systèmes multicanaux, car il permettra de multiplexer 16 voies sur un même fil avec le protocole TDM (contre 2 avec le protocole I2S).

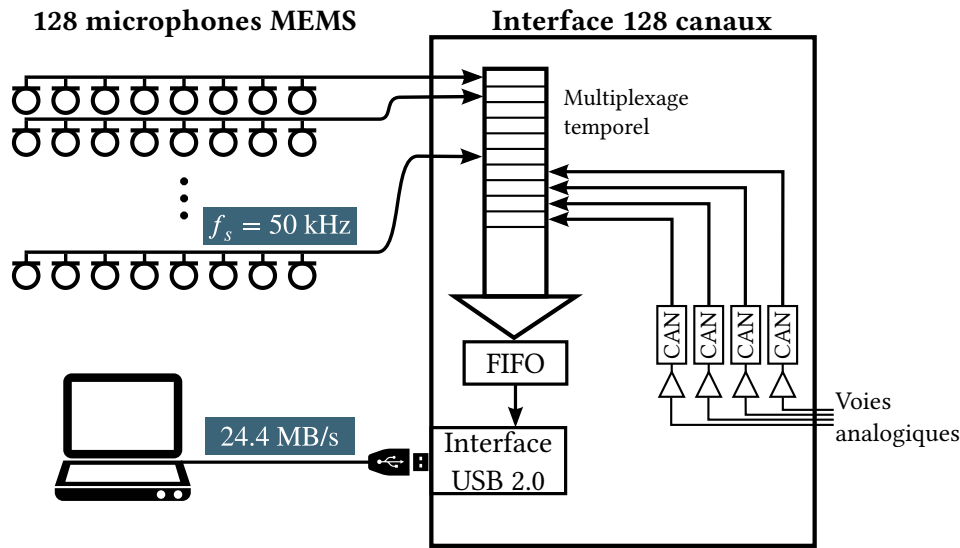
## 1.1.2 L'interface maître

### 1.1.2.1 Première génération : 128 canaux

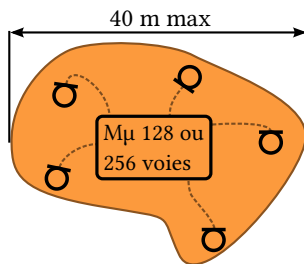
Le schéma de la première génération du système est donné en figure 1.2a. L'interface 128 canaux a pour rôle de centraliser les échantillons synchronisés des 128 microphones MEMS. Il est également muni de 4 entrées analogiques supplémentaires, qui sont notamment utilisées pour ajouter des microphones de référence. Toutes ces données sont ensuite sérialisées, et transmises au bus USB 2.0, qui assure la connexion entre l'interface et l'ordinateur. Quand le nombre maximal de canaux est utilisé et quand la fréquence d'échantillonnage est maximale, les données sont envoyées avec un flux de  $24.4 \text{ MB s}^{-1}$ . Au cœur de l'interface, toutes les tâches sont régies par un circuit logique programmable (ou FPGA, de l'anglais Field-Programmable Gate Array).

Du point de vue de l'utilisateur, la commande du système se résume aux tâches suivantes :

- le choix d'une des valeurs de fréquence d'échantillonnage suivantes :  $f_s = 50, 25, 20, 12.5, 10$  ou  $6.25 \text{ kHz}$  ;
- le choix des canaux à activer parmi les 128 (il en est de même pour les voies analogiques) ;
- le choix de la durée totale d'acquisition. Cette durée peut être non définie *a priori*, dans ce cas il faut envoyer une commande d'arrêt ultérieurement ;
- le lancement de l'acquisition : il se fait généralement par USB, après la série de commandes précédente. Mais il est possible de lancer l'acquisition par une commande logique externe connectée à l'interface maître : cela est utile pour synchroniser l'acquisition par rapport à un signal physique.



(a) Schéma fonctionnel du système à 128 canaux.



(b) Le réseau peut être étendu jusqu'à 20 m autour de l'interface quand il est branché directement en USB.

**FIGURE 1.2** – 1<sup>ère</sup> génération du système d'acquisition : 128 canaux d'entrée et 4 voies analogiques annexes. La liaison USB 2 permet la réception d'un flux maximal de  $24.4 \text{ MB s}^{-1}$ . Ce système est utilisé dans les expériences présentées aux chapitres 2 et 3.

Comme le montre le schéma 1.2b, l'étendue possible de la première génération du système *Megamicros* est de 40 m. Une telle envergure représente une progression notable par rapport aux réseaux synchrones existants. Cependant le nombre total de 128 voies reste du même ordre de grandeur que ce qui existe déjà parmi les systèmes commercialisés ou de laboratoire. Or la première génération du système communique par USB 2.0, ce qui bride le débit théorique maximal à  $35 \text{ MB s}^{-1}$ .

### 1.1.2.2 Deuxième génération : 256... ou 4x256 canaux

La nouvelle version du système a pour but d'augmenter le nombre de canaux, mais également l'étendue d'écoute. Le schéma de la deuxième génération du système est en figure 1.3a. L'interface est connectée à 256 voies MEMS, ce qui donne un débit de  $48.8 \text{ MB s}^{-1}$ . La liaison USB a été mise à niveau par la version 3.0. En plus des voies analogiques, deux voies d'entrée logique sont ajoutées, pour pouvoir enregistrer le signal d'une transmission série asynchrone, comme par exemple un bus classique RS-232. À cause de l'augmentation du nombre de canaux, donc du nombre total

d'entrées-sorties, nous utilisons un circuit FPGA intégré dans un boîtier de type BGA (de l'anglais Ball Gate Array).

De plus, il est maintenant possible de déporter le système par fibre optique. Un étage supplémentaire peut connecter jusqu'à quatre systèmes à 256 voies, et donc concentrer en tout 1024 signaux. Contrairement à une connexion cuivrée, la fibre optique a très peu de pertes de transmission. Nous utilisons actuellement une fibre optique de 10 m, mais théoriquement, les émetteurs-récepteurs utilisés garantissent une transmission jusqu'à 10 km de distance (cf fig. 1.3b). Expérimentalement, nous avons utilisé 256 voies déportées par fibre optique et le concentrateur. Des tests de débits avec le concentrateur ont validé la possibilité de recevoir un flux total de 1024 voies, c'est à dire  $195.2 \text{ MB s}^{-1}$ . La construction de trois interfaces maîtres supplémentaires est en cours, pour obtenir un système 1024 voies complet.

Avec cette architecture à deux étages, le système global est pionnier en terme d'étendue. Il peut ouvrir des perspectives nouvelles d'expériences acoustiques, notamment pour le monitoring du bruit urbain, qui devient possible à l'échelle d'un quartier.

## 1.2 Caractérisation acoustique de l'antenne de microphones

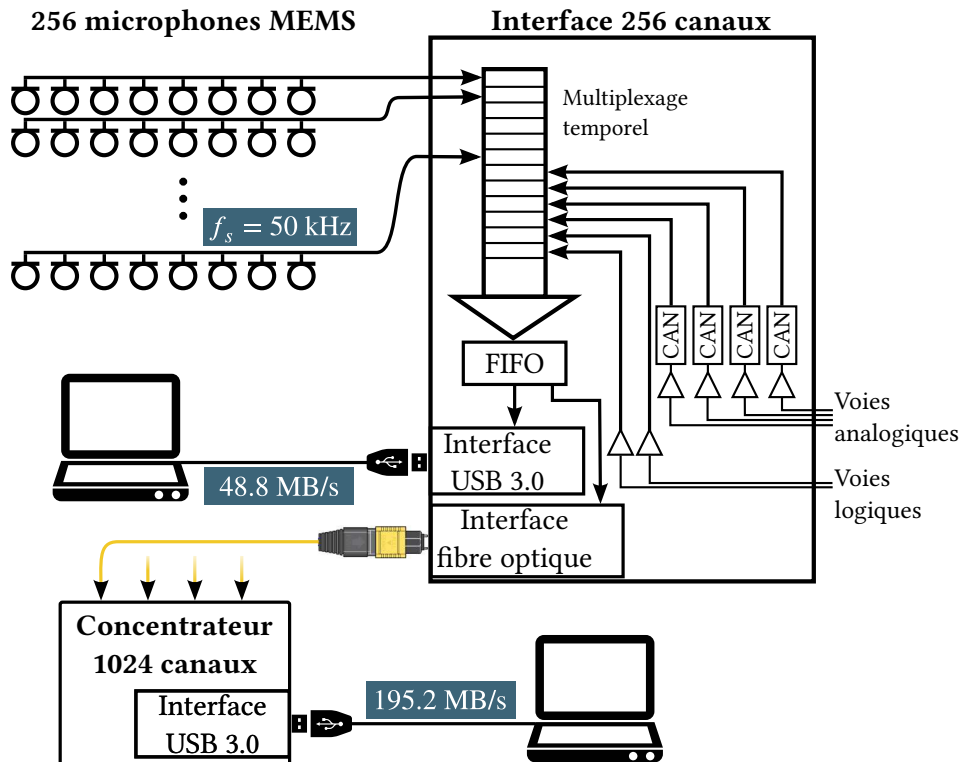
La documentation constructeur du microphone liste les caractéristiques statistiques du composant. Cependant, ces caractéristiques peuvent changer en fonction des choix de conception de la carte électronique (cf. photo 1.1a). De plus, l'intégration globale du circuit sur le support mécanique a un impact sur le comportement acoustique. Le constructeur mentionne notamment un éventuel impact sur la directivité et/ou un effet de résonance de Helmholtz [70], dus à la structure de maintien.

Nous souhaitons définir les performances de ces microphones en vue d'applications à l'antennerie acoustique. Dans ce cas précis, il faut avant tout quantifier la variabilité des capteurs. En effet, la plupart du temps, les hypothèses de base pour mettre en œuvre une technique d'antennerie considèrent des capteurs de mesure parfaitement homogènes. Les travaux d'Hafizovic *et al.* présentent la première création d'antenne de microphones MEMS [63], mais ne donne pas de caractérisation acoustique.

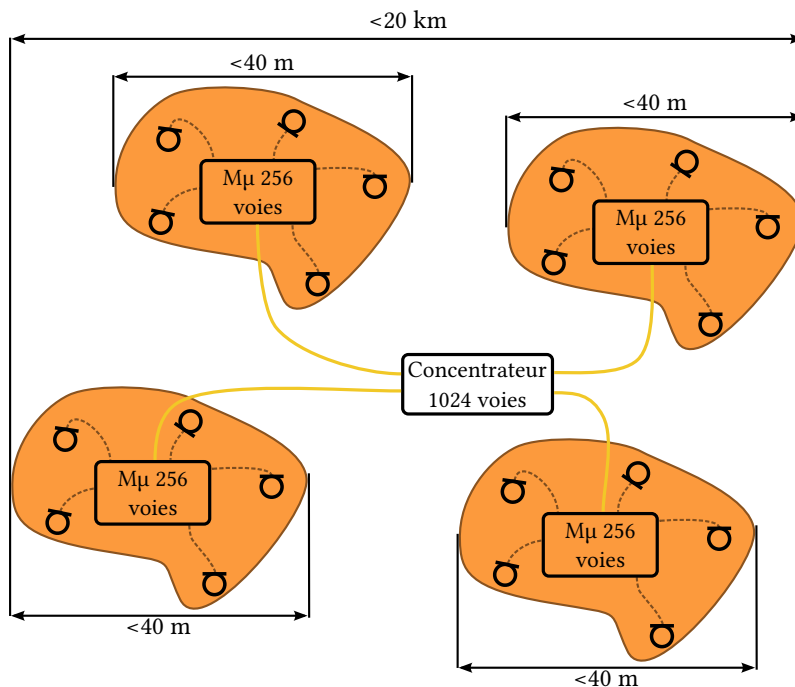
### 1.2.1 L'expérience en chambre anéchoïque

On souhaite déterminer cinq spécifications importantes pour les techniques d'antennerie :

- la sensibilité : dans l'idéal tous les capteurs ont le même niveau de sensibilité ;
- la directivité : dans l'idéal tous les capteurs sont omnidirectionnels ;
- la réponse en fréquence : elle détermine la bande passante du bruit mesurable ;
- la distorsion de phase : on espère obtenir la phase la plus linéaire possible ;



(a) Schéma fonctionnel du système à 256 canaux.



**FIGURE 1.3** – 2<sup>ème</sup> génération du système d'acquisition : 256 canaux d'entrée, et 4 voies analogiques + 2 voies logiques annexes. Soit un ordinateur est directement connecté au système par USB 3, soit le flux est envoyé vers un concentrateur 1024 voies par liaison optique. Ce système est utilisé en expérience au chapitre 4.

(b) Un réseau à 256 voies peut être étendu jusqu'à 20 m autour de l'interface. Mais l'utilisation du concentrateur 1024 voies permet de déporter jusqu'à 4 de ces réseaux dans un rayon de 10 km, par fibre optique.

– le bruit de fond : il détermine le seuil limite du bruit mesurable.

Avec la première version du système d'acquisition, des expériences de caractérisation sont effectuées au Laboratoire National de métrologie et d'Essai (LNE) [3], en chambre anéchoïque. 128 microphones MEMS sont assemblés sur un cadre de taille réduite (cf fig. 1.4). Une source acoustique monopolaire est positionnée à 5 m en face du cadre. Cette configuration géométrique permet de faire l'approximation qu'au niveau de l'antenne, l'onde est plane et donc perçue de manière identique par l'ensemble des microphones. L'intérêt est de dégager un comportement statistique des 128 microphones, pour analyser leur homogénéité. Un microphone *B&K 4955* est placé au niveau du cadre pour obtenir une mesure de référence. Il est relié à une des voies analogiques du système d'acquisition afin d'assurer la synchronisation temporelle avec les voies MEMS.

### 1.2.2 Résultats

**Sensibilité** L'histogramme 1.5 montre les niveaux SPL des 128 microphones, écoutant la même source de bruit blanc. Il met en évidence une variation en sensibilité faible des composants. La régression de l'histogramme par une loi normale est tracée en rouge. L'écart type correspondant est de 0.8 dB, ce qui place 99.7 % des sensibilités dans la plage  $\pm 2.4$  dB. Ce résultat est en adéquation avec la spécification constructeur, qui garantit une variation en sensibilité de  $\pm 3$  dB. Si nécessaire, ces variations en sensibilité peuvent éventuellement être corrigées par des coefficients d'égalisation.

**Directivité** La directivité est présentée en figure 1.6. Elle représente uniquement la partie avant des microphones : la directivité arrière n'est pas pertinente car elle dépend du support de fixation des circuits. La directivité mesurée est large bande car elle est issue d'une mesure de bruits blancs. La directivité moyenne, tracée en rouge, a une variation totale de 1.7 dB sur 180°. Nous pouvons donc considérer les microphones omnidirectionnels, d'autant plus que les directivités restent très homogènes d'un microphone à l'autre.

**Réponse en fréquence** La réponse fréquentielle est estimée par la méthode expliquée dans [13]. Elle propose une méthode d'estimation des fonctions de transfert par mesure de densités spectrales croisées. Cela permet d'écarter la contribution du bruit des chaînes d'acquisition. Pour obtenir les fonctions de transfert, la mesure de référence est fournie par le microphone *B&K 4955*.

La figure 1.7 montre une variation moyenne inférieure à 2 dB entre 80 et 2000 Hz. Dans cette plage, le niveau absolu est biaisé de 1 dB environ. Au-delà de 2 kHz, la valeur croît jusqu'à un pic à 9 dB. Cependant les réponses en fréquence sont homogènes parmi les différents microphones : si il faut une bonne justesse de mesure, il est possible d'appliquer une correction commune des signaux mesurés par une égalisation spectrale.

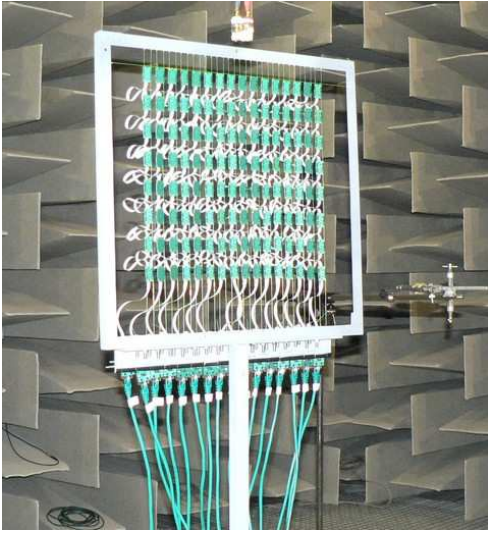


FIGURE 1.4 – Expérience au LNE pour la caractérisation des microphones. Dimension du cadre : 50 cm × 50 cm.

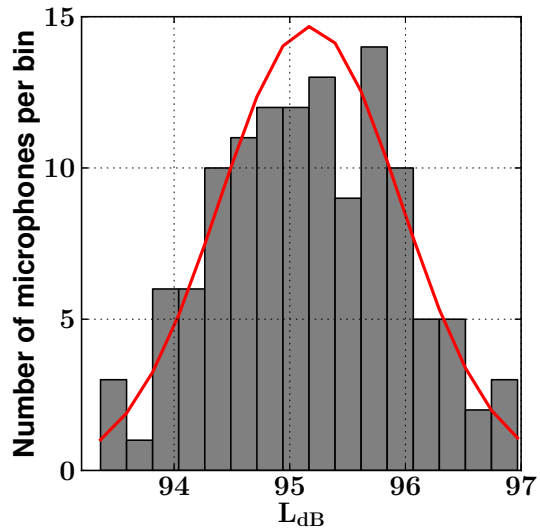


FIGURE 1.5 – Sensibilités : histogramme des 128 niveaux SPL, référés à 20  $\mu\text{Pa}$ ; régression normale tracée en rouge.

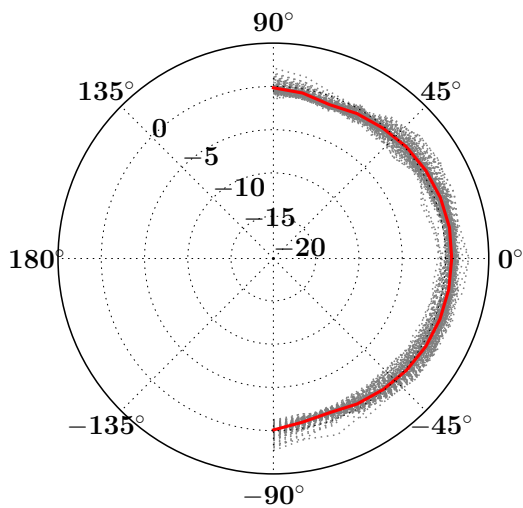


FIGURE 1.6 – Directivités en bruit blanc, en dB, normalisées par rapport à la valeur maximale; moyenne tracée en rouge.

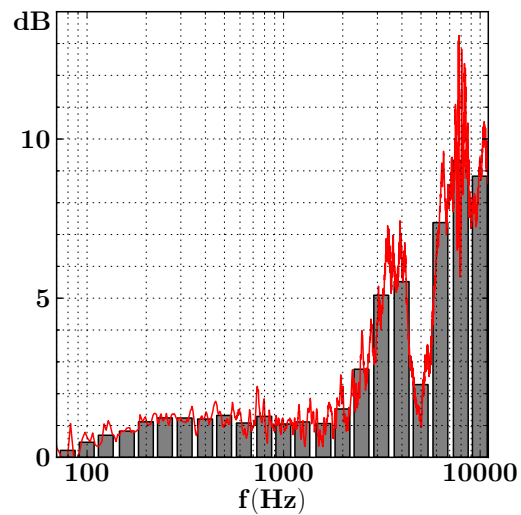
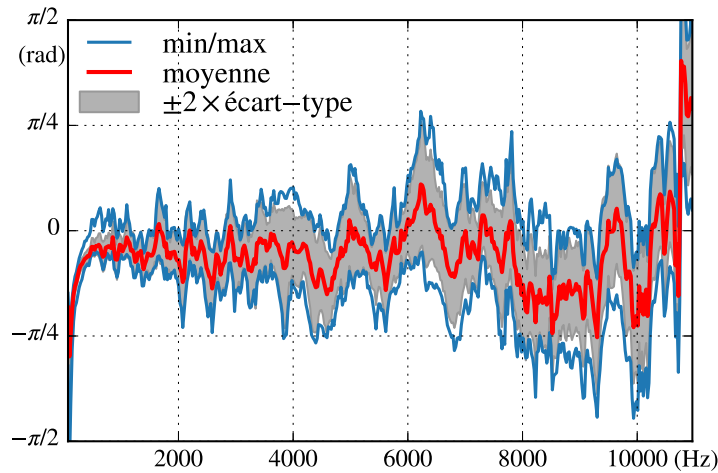


FIGURE 1.7 – Réponse en fréquence en dB (référéncée par rapport à la mesure du microphone *B&K*) tracée en rouge; niveaux par tiers d'octave en gris.



**FIGURE 1.8** – Variabilité de phase des microphones MEMS ; la phase est donnée par l'argument des fonctions de transfert entre les microphones MEMS et la voie de référence (microphone *B&K*).

Notons que l'évaluation de la réponse est restreinte sur l'intervalle fréquentiel 80 – 11000 Hz. Au-delà de cette plage, la cohérence entre les voies MEMS et la voie de référence est trop faible ; dans ce cas la mesure n'est plus pertinente. L'explication vient soit des microphones qui mesurent trop en marge de leur bande passante, soit de la bande passante réduite de la source.

**Variabilité de la phase** Les applications de la localisation de sources s'appuient sur la phase des signaux mesurés. C'est pour cela que la variabilité des phases issues des mesures doit être très faible. Deux causes physiques liées au système peuvent exister :

- les décalages temporels, qui engendrent une erreur de phase linéaire, proportionnelle à la fréquence. Puisque la même horloge synchronise tous les microphones MEMS, leurs signaux sont alignés à l'échantillon près ;
- une erreur de phase, non-linéaire, qui engendre une distorsion des signaux temporels.

Nous supposons que la variabilité en phase n'a pas d'impact tant qu'elle reste proche de 0 et qu'elle est très petite devant  $\pm\pi$ .

La figure 1.8 trace les phases des 128 signaux. J'ai utilisé les mêmes fonctions de transfert qui ont données les réponses en fréquence (fig. 1.7), pour en extraire les arguments. J'ai également ôté la contribution linéaire dans la phase, qui est uniquement due à l'erreur d'alignement entre le microphone de référence *B&K* et les voies MEMS. La moyenne de la phase (en rouge) garde une amplitude inférieure à  $\pi/4$  dans l'intervalle 80-10000 Hz. La variation moyenne est plus significative quand la fréquence augmente, notamment au-dessus de 4 kHz. L'écart-type des 128 phases mesurées montre un comportement bien homogène en basse fréquence ( $< \pi/20$  rad en dessous de 3 kHz), et augmente légèrement en fonction de la fréquence, avec un pic de 0.24 rad à 6 kHz.



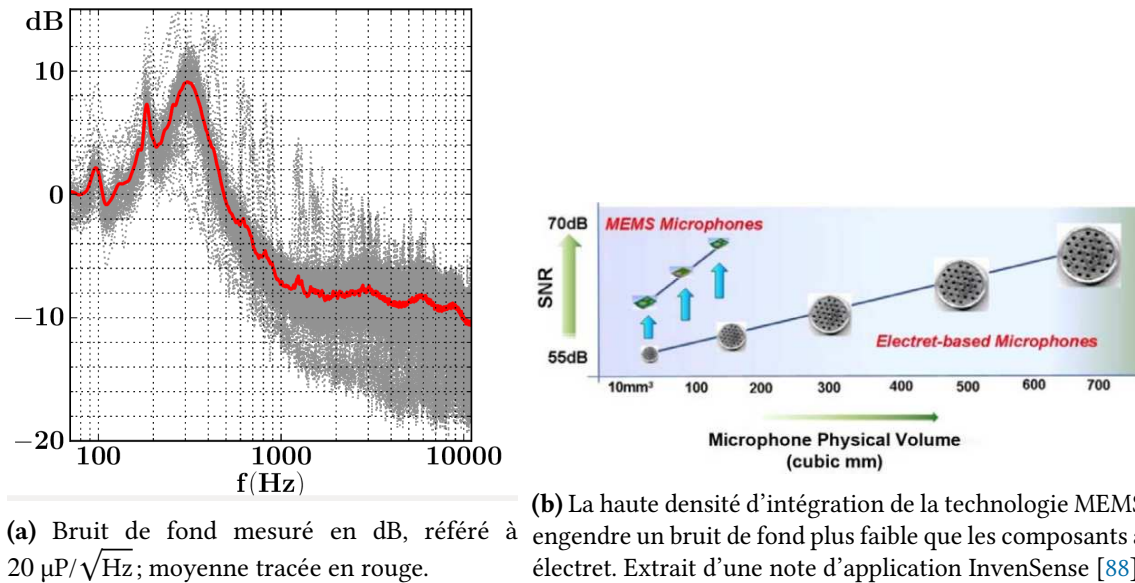


FIGURE 1.9 – Bruit de fond des microphones MEMS.

**Bruit de fond** Dans la chambre anéchoïque, la mesure des voies MEMS sans présence de source reflète le bruit inhérent au système et à la chaîne d'acquisition. Le niveau SPL global de bruit de fond est de 33.5 dB. La densité spectrale de puissance des mesures donne la répartition du bruit sur le spectre, donc son impact en fonction de la fréquence sur les mesures acoustiques en terme de rapport signal bruit. Comme l'illustre la figure 1.9a, la densité spectrale du bruit de fond reste en-deçà de  $10 \text{ dB}/\sqrt{\text{Hz}}$ . Pour les fréquences supérieures à 500 Hz, elle est négligeable car inférieure à  $0 \text{ dB}/\sqrt{\text{Hz}}$ ; le maximum de  $9 \text{ dB}/\sqrt{\text{Hz}}$  se situe à 300 Hz. Notons surtout que, comme le montre la figure 1.9b, les microphones MEMS ont un rapport signal sur bruit plus élevé que les microphones à électret du fait d'une densité d'intégration beaucoup plus élevée.

De manière générale, les caractéristiques acoustiques sont *a priori* bien adaptées pour une utilisation à l'antennerie acoustique. Un des points négatifs reste le pic dans la bande passante, qui n'est pas présent dans la documentation constructeur [47]. Cependant, l'homogénéité des composants est très bonne, donc une correction éventuelle, de réponse en fréquence ou en phase, peut être appliquée en commun sur tous les canaux.

### 1.3 Quelques illustrations du système *in situ*

Durant ma thèse, j'ai pu participer ou assister à diverses expériences impliquant l'utilisation des deux générations du système d'acquisition, et de multiples antennes de microphones MEMS. Les quelques exemples présentés en figure 1.10 sont décrits ci-dessous. Ils illustrent bien le caractère protéiforme du système *Megamicros*, ainsi que sa spécificité en terme d'étendue spatiale.

(a) Dans le cadre de la perception du bruit urbain, notre équipe a monté une antenne linéaire le long d'un axe routier des quais de la Seine (cf. figure 1.10a). Étendue de 21.6 m, la géométrie est régulière, avec un espacement de 17 cm entre les 128 microphones. Du fait de sa grande taille et de sa proximité à la route, elle a permis de faire de l'imagerie de champ proche, pour localiser les sources de bruits issus des véhicules. Grâce au couplage avec la vidéo prise par une webcam, on a été capable d'associer, à chaque véhicule, la ou les sources de bruits correspondantes, typiquement le bruit moteur et le bruit de roulement [87]. Concernant l'ergonomie du système, il est à noter le montage de l'antenne par 3 personnes a pris environ 30 min.

(b) Au cours de deux études sur la nuisance sonore, le système a été placé sur le site Jussieu de l'UPMC, au sommet de la tour Zamansky. La photo 1.10b montre les microphones alignés le long d'une arête de la tour. La géométrie employée est la même que dans l'étude précédente sur les quais de la Seine. La première étude a porté sur la perception du bruit urbain [97], avec une application au bruit de chantier. La seconde campagne de mesures a été effectuée durant le défilé aérien du 14 Juillet. Cette situation particulière implique des conditions intéressantes : les sources larges bandes (avions et hélicoptères) sont situées en champ lointain, et se déplacent avec une vitesse constante. Les sources pouvant être considérées ponctuelles et en faible nombre, cette expérience m'a permis d'étudier l'applicabilité de la localisation de sources large bande par une approche de parcimonie groupée [148].

(c) Dans le but d'analyser plus finement le bruit des véhicules, notre équipe a établi une campagne de mesure pour mener une étude paramétrique sur différents véhicules de transport. Pour cela, des mesures de bruit au passage de véhicule ont été effectuées sur une piste d'essai, face à une antenne rectangulaire de 256 microphones, d'une taille de 2.11 m × 19.2 m. La photo de la figure 1.10c montre l'envergure particulièrement grande du montage : le positionnement précis des microphones a nécessité une construction préalable de l'antenne en plusieurs blocs, qui ont finalement été assemblés sur la piste. Dans le chapitre 4, ce scénario sera le cœur expérimental de l'étude de la calibration géométrique par temps ou différences de temps de vol.

(d) Dans le cadre d'une étude sur le bruit aéroacoustique, le système a été employé dans une expérience dans la soufflerie anéchoïque BETI à l'Institut de PPrime (figure 1.10d). Le nombre élevé de microphones permet d'entourer la veine, dans laquelle se trouve un barreau ou une marche montante. Avec une mesure acoustique sur une surface fermée, il est possible de mettre à profit une technique de retournement temporel pour identifier les sources aéro-acoustiques [92].

(e) Pour finir, je mentionne la construction d'une antenne d'holographie de champ proche, de 16 × 16 éléments [96]. Contrairement aux exemples précédents, cette réalisation montre que tous les microphones MEMS peuvent être intégrés sur un seul circuit imprimé, et donc également être



(a) Mesure en environnement urbain, devant une route des quais de la Seine, Paris [87].



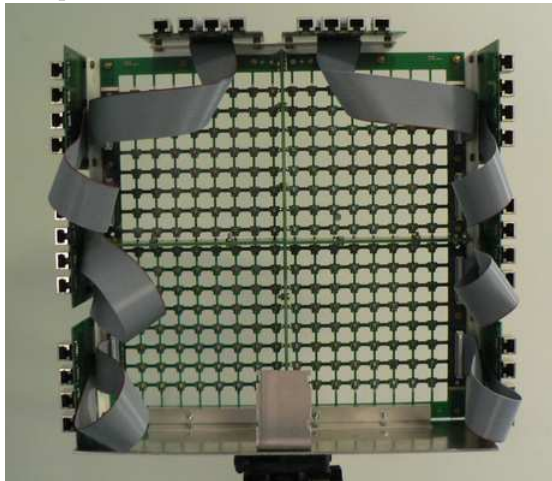
(b) Mesure en environnement urbain au sommet d'un immeuble, Paris [97].



(d) Étude du bruit d'écoulement en soufflerie anéchoïque [92].



(c) Série de mesures de bruit au passage de véhicule, sur une piste d'essai [87].



(e) Holographe temps réel à 256 canaux [96].

FIGURE 1.10 – Quelques exemples en images d'utilisations du système *Megamicros*.

concentrés sur une rétine de petite surface. Parmi les exemples décrits, il est à noter que cette géométrie est la seule qui n'est pas protéiforme, car les microphones sont soudés sur une structure rigide commune. De ce fait, cette géométrie est également la seule à être déterminée et connue *a priori* au millimètre près.

Au final, l'ensemble des illustrations révèle les problèmes énoncés en introduction générale, liés à l'utilisation d'un grand nombre de microphones. Premièrement, les grandes antennes déployées ont des géométries variées, spécifiques à chaque scénario expérimental. Or, assurer le montage géométriquement précis de ces antennes s'est avéré long et contraignant. Pour cela, l'intérêt de retrouver la position des microphones dans l'espace de façon automatique s'est révélé de façon naturelle. Le problème de calibration géométrique, traité en chapitre 3 et 4, sera le cœur de cette première thématique.

Deuxièmement se pose le problème du volume de données mesurées. En employant 256 capteurs, les mesures de pression engendrent un flux de  $48.8 \text{ MB s}^{-1}$ . En imaginant par exemple une expérience d'analyse de bruit urbain sur une heure, cela représente au total 172 GB de mesures brutes. Ce volume considérable de données motive l'intérêt de mettre en œuvre le traitement à la volée des mesures microphoniques. Le chapitre suivant se focalise sur ce sujet, et plus particulièrement sur la localisation de sources en temps réel.



## Chapitre 2

# Imagerie acoustique de sources large bande par un grand réseau

### Sommaire

---

<b>2.1</b>	<b>Introduction</b>	<b>26</b>
<b>2.2</b>	<b>Localisation de sources en bande étroite</b>	<b>27</b>
2.2.1	Problème direct en bande étroite	27
2.2.2	Reconstruction par formation de voies	29
<b>2.3</b>	<b>Localisation de sources large bande</b>	<b>31</b>
2.3.1	Problème direct	31
2.3.2	Formation de voies large bande	31
2.3.3	Imagerie par bande d'octave	32
2.3.4	Comment garantir un traitement en temps réel	34
2.3.4.1	Implémentation sur GPU	34
2.3.4.2	Étude paramétrique des performances	35
<b>2.4</b>	<b>Expériences d'imagerie</b>	<b>36</b>
2.4.1	Scénario 1 : expérience de calibrage	37
2.4.2	Scénario 2 : l'aspirateur	38
<b>2.5</b>	<b>Discussion : impact de l'erreur en position des microphones</b>	<b>42</b>
2.5.1	Le montage manuel de l'antenne induit une erreur géométrique	42
2.5.2	Méthodes de correction des dispersions de mesure	43
2.5.2.1	Dispersion issues de la mesure	43
2.5.2.2	Dispersion issues du modèle	44
<b>2.6</b>	<b>Conclusion</b>	<b>45</b>

---

## 2.1 Introduction

Le diagnostic acoustique est une pratique importante dans le milieu industriel. Typiquement la conception de véhicules de transports est contrainte par des réglementations européennes sur le bruit [49, 69]. En effet, la nuisance sonore a un impact sur les personnes dans l'environnement proche du véhicule, en plus de celui sur le confort des passagers [23]. La conformité à la plupart de ces réglementations est vérifiée par des protocoles de mesure de bruit au passage de véhicule (appelés tests *by-pass*), généralement via l'utilisation d'un sonomètre. C'est l'outil standardisé pour le diagnostic acoustique, donnant une description précise des niveaux de bruit. Une description énergétique globale est donnée par le niveau SPL (Sound Pressure Level), et une description spectrale affinée est fournie par les niveaux par bande d'octave ou de tiers d'octave. Une approche complémentaire, par antennerie acoustique, consiste à localiser la position des sources acoustiques sur la structure du véhicule. Pour les industriels, cet outil est précieux pour identifier l'origine physique des sources de bruit. Dans le cas des moyens de transport, il a été appliqué à l'étude des voitures [7], avions [132] et trains [95]. Cependant, la localisation de sources par réseau de microphones reste une méthode du domaine de la recherche. Malgré certaines propositions, il n'existe toujours pas de protocole standardisé d'imagerie acoustique [80]. Jusqu'à récemment, d'une part la localisation de sources était restreinte à une analyse en bande étroite à cause du temps de calcul nécessaire si l'antenne est munie d'un grand nombre de capteurs ; d'autre part, elle nécessitait une antenne de microphones analogiques, impliquant un matériel complexe et coûteux. Ces deux freins technologiques sont au cœur de la problématique que nous posons : que faudrait-il apporter aujourd'hui à la mise en œuvre de l'imagerie acoustique pour en faire un moyen de diagnostic plus répandu ?

Deux technologies récentes permettent d'apporter des solutions à ces deux problèmes de temps de calcul, et de complexité matérielle. D'abord, les cartes graphiques, ou GPUs (*Graphic Processing Unit*), sont devenues une solution très intéressante pour faire du calcul parallèle massif, tout en restant peu coûteuses et compactes. Or, la formation de voies est une approche de localisation de sources qui se prête facilement au calcul parallèle [103]. Malgré ses limitations connues liées à la directivité de l'antenne, elle reste encore une méthode utilisée dans le milieu industriel. L'utilisation du GPU pour l'imagerie a déjà fait ses preuves dans le domaine médical [5, 93, 120, 168], et de l'acoustique sous-marine [26, 27, 48]. Nous verrons que l'imagerie acoustique aérienne large bande sur GPU est bien plus performante que sur un processeur classique. De plus, le gain est suffisant pour faire de l'imagerie en temps réel. Deuxièmement, comme indiqué précédemment, nous utilisons des microphones MEMS numériques. En construisant une antenne acoustique avec de tels composants, la charge matérielle est considérablement réduite [63, 153].

En tirant profit de ces deux technologies, je propose dans ce chapitre un outil de diagnostic

acoustique temps réel, par imagerie large bande. Il comprend toute la chaîne, de l'acquisition au traitement des données. L'imagerie consiste à cartographier les niveaux acoustiques absolus par bande d'octave, ce qui permet d'obtenir une mesure comparable à celle d'un sonomètre. En première partie, ce chapitre énonce rapidement le formalisme de la formation de voies en bande étroite, ainsi que sa généralisation en large bande. Ensuite, j'explique l'approche choisie pour implémenter l'algorithme en calcul parallèle sur GPU, puis vérifie si le traitement en temps réel est possible. Enfin, deux expériences *in situ* sont effectuées pour valider l'intérêt de l'outil proposé. Je termine par une discussion positionnant les chapitres suivants du manuscrit : l'impact de l'erreur de la position des microphones sur la formation de voies.

## 2.2 Localisation de sources en bande étroite

### 2.2.1 Problème direct en bande étroite

Nous nous plaçons dans un problème de propagation exprimé dans le domaine des fréquences. Soit  $\hat{p}_m(\omega)$  la pression mesurée par le microphone  $m$  à la pulsation  $\omega$ . Avec un réseau de  $M$  microphones omnidirectionnels, nous obtenons, pour une pulsation donnée, le vecteur complexe de mesures  $\hat{\mathbf{p}} = [\hat{p}_1 \dots \hat{p}_m \dots \hat{p}_M]^\top$ . Les positions des microphones sont notées  $\mathbf{x}_1, \dots, \mathbf{x}_m, \dots, \mathbf{x}_M$ .

La formulation classique du problème de localisation de sources cherche à déterminer les amplitudes de sources monopolaires situées à des positions fictives sur une grille de  $I$  points dans l'espace [164]. Notons, toujours à la même pulsation, le vecteur inconnu complexe  $\hat{\boldsymbol{\beta}} = [\hat{\beta}_1 \dots \hat{\beta}_i \dots \hat{\beta}_I]^\top$  que nous souhaitons déterminer. Dans un milieu homogène et non-dispersif, la propagation de la source  $i$  vers les  $M$  microphones suit un modèle linéaire :

$$\hat{\mathbf{p}} = \mathbf{a}_i \hat{\beta}_i, \quad (2.1)$$

avec  $\mathbf{a}_i$  le vecteur de propagation<sup>1</sup> de la source  $i$  vers le vecteur de mesures. Dans le domaine des fréquences, il applique l'effet de la célérité du milieu, en ajoutant les retards adéquats sur les phases des signaux. Il dépend donc de la position des points de mesures et des sources. Comme le montre la figure 2.1, deux conditions géométriques existent :

- si les sources sont suffisamment éloignées par rapport à la taille caractéristique du réseau de microphones, elles peuvent être considérées à l'infini. Ceci se traduit par une propagation en ondes planes. Le vecteur d'ondes  $\mathbf{k}_i$  définit la direction et la vitesse de propagation, ainsi que la pulsation. On peut donc définir  $\mathbf{a}_i$  par :

1. dans la littérature le vecteur  $\mathbf{a}_i$  est souvent appelé *steering vector*.



$$\mathbf{a}_i = \begin{pmatrix} e^{-j\mathbf{k}_i \cdot \mathbf{x}_1} \\ \vdots \\ e^{-j\mathbf{k}_i \cdot \mathbf{x}_m} \\ \vdots \\ e^{-j\mathbf{k}_i \cdot \mathbf{x}_M} \end{pmatrix}; \quad (2.2)$$

- si la distance séparant les sources des microphones est du même ordre de grandeur que la taille caractéristique du réseau, la propagation d'onde est sphérique avec une décroissance de l'amplitude en fonction de la distance. Elle est définie par la fonction de Green de l'équation de Helmholtz, et permet d'écrire  $\mathbf{a}_i$  :

$$\mathbf{a}_i = \begin{pmatrix} \frac{1}{d_{1i}} e^{-j\frac{\omega}{c_0} d_{1i}} \\ \vdots \\ \frac{1}{d_{mi}} e^{-j\frac{\omega}{c_0} d_{mi}} \\ \vdots \\ \frac{1}{d_{Mi}} e^{-j\frac{\omega}{c_0} d_{Mi}} \end{pmatrix} \quad (2.3)$$

où  $d_{mi}$  est la distance entre la source  $i$  et le microphone  $m$  (cf fig. 2.1b).

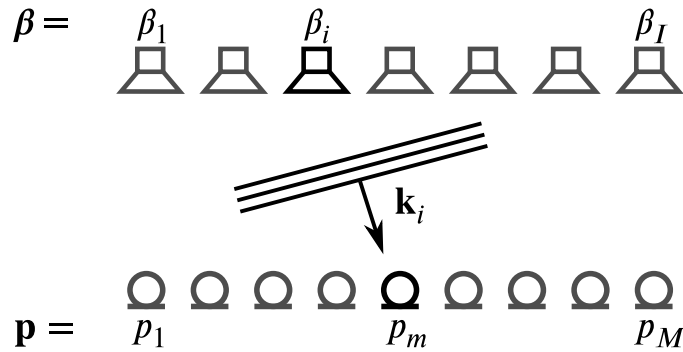
La frontière entre le champ proche et le champ lointain est fonction de la taille de l'antenne et de la fréquence considérée. Le champ proche se situe dans la zone de diffraction de Fresnel : la taille apparente de l'antenne  $y$  est suffisamment grande pour pouvoir *focaliser* sur une source de bruit. La frontière entre le champ lointain et le champ proche peut être quantifiée par la distance de Fraunhofer :

$$d_{\text{Fr}} = \frac{2d_{\text{ant}}^2 f}{c_0}, \quad (2.4)$$

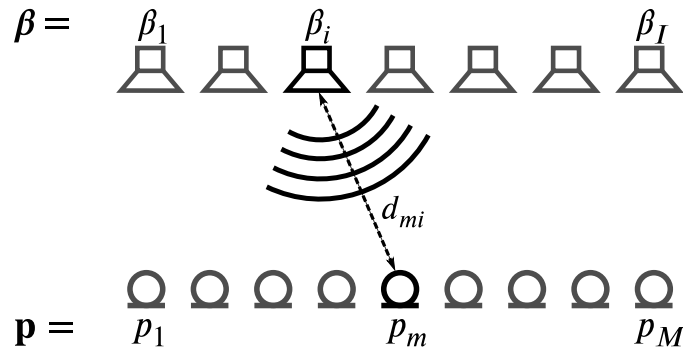
où  $d_{\text{ant}}$  est la taille caractéristique de l'antenne. Si l'espacement entre les microphones est homogène, on peut considérer  $d_{\text{ant}}$  égal à la valeur maximale des distances par paire entre les microphones.

Avec un tel modèle, le vecteur de mesure obtenu par les microphones est la contribution de chacune des sources situées sur les points de la grille. Il s'agit donc d'une combinaison linéaire des vecteurs de propagation  $\mathbf{a}_i$ , pondérés par les niveaux des sources :

$$\hat{\mathbf{p}} = \sum_{i=1}^I \hat{\beta}_i \mathbf{a}_i \quad (2.5)$$



(a) Condition de champ lointain : les microphones reçoivent des ondes planes de chaque source  $i$ , caractérisées par les vecteurs d'onde  $\mathbf{k}_i$ .



(b) Condition de champ proche : les microphones reçoivent des ondes sphériques de chaque source  $i$ .

**FIGURE 2.1** – Propagation acoustique des sources aux microphones en milieu homogène : deux conditions géométriques.

En notant la matrice de propagation  $\mathbf{A} = [\mathbf{a}_1 \dots \mathbf{a}_i \dots \mathbf{a}_I]$ , nous pouvons exprimer la propagation linéaire sous forme matricielle :

$$\hat{\mathbf{p}} = \mathbf{A}\hat{\boldsymbol{\beta}} \quad (2.6)$$

Avec ce modèle fréquentiel, la mesure peut être interprétée comme la convolution des sources par la matrice  $\mathbf{A}$ . Les paramètres qui régissent cette convolution sont essentiellement géométriques : la forme du réseau, et la position des points de la grille de sources. Nous verrons par la suite l'impact de ces caractéristiques géométriques sur la localisation de sources.

### 2.2.2 Reconstruction par formation de voies

Pour estimer les niveaux de pression des sources, la méthode la plus simple consiste à rétropropager les mesures obtenues selon le modèle de propagation. Géométriquement, cela consiste à

projeter les mesures dans les directions définies par les vecteurs  $\mathbf{a}_i$ . Cette opération est appelée formation de voies, ou *Delay & Sum Beamforming* (*DS Beamforming*).

Dans le cas du champ lointain, en notant  $\hat{b}_i$  le niveau reconstruit de la source  $\hat{s}_i$ , nous représentons la projection des mesures par le produit scalaire :

$$\hat{b}_i = \left\langle \frac{\mathbf{a}_i}{M}, \hat{\mathbf{p}} \right\rangle = \frac{1}{M} \mathbf{a}_i^H \hat{\mathbf{p}} \quad (2.7)$$

Comme les vecteurs  $\mathbf{a}_i$  définies par la relation (2.2) ont une norme égale à  $M$  (cf. eq (2.2)), il est nécessaire de normaliser le produit scalaire pour obtenir les bonnes amplitudes des sources. En notant  $\hat{\mathbf{b}} = [\hat{b}_1 \dots \hat{b}_i \dots \hat{b}_I]^T$  le vecteur de reconstruction des niveaux de sources, la formation de voies en champ lointain s'écrit :

$$\hat{\mathbf{b}} = \frac{1}{M} \mathbf{A}^H \hat{\mathbf{p}} \quad (2.8)$$

où  $^H$  est l'opérateur de transposition hermitienne. Physiquement, l'opération de *DS Beamforming* consiste donc en :

- l'étape *delay* : compenser les retards des mesures, dus à la propagation. Dans le domaine des fréquences, il s'agit d'appliquer les avances de phase adéquates ;
- l'étape *sum* : sommer les signaux compensés, qui seront constructifs à chacun des points sources.

Dans le cas du champ proche, si nous souhaitons reconstruire les niveaux d'amplitudes des sources, il faut également compenser la décroissance des amplitudes. Nous préférons alors la relation suivante pour appliquer la formation de voies en champ proche :

$$\hat{\mathbf{b}} = \frac{1}{M} \begin{bmatrix} d_{11} e^{\frac{j\omega}{c_0} d_{11}} & \dots & d_{m1} e^{\frac{j\omega}{c_0} d_{m1}} & \dots & d_{M1} e^{\frac{j\omega}{c_0} d_{M1}} \\ \vdots & & \vdots & & \vdots \\ d_{1i} e^{\frac{j\omega}{c_0} d_{1i}} & \dots & d_{mi} e^{\frac{j\omega}{c_0} d_{mi}} & \dots & d_{Mi} e^{\frac{j\omega}{c_0} d_{Mi}} \\ \vdots & & \vdots & & \vdots \\ d_{1I} e^{\frac{j\omega}{c_0} d_{1I}} & \dots & d_{mI} e^{\frac{j\omega}{c_0} d_{mI}} & \dots & d_{MI} e^{\frac{j\omega}{c_0} d_{MI}} \end{bmatrix} \hat{\mathbf{p}} \quad (2.9)$$

L'inconvénient majeur de la formation de voies classique est de faire apparaître le diagramme de directivité de l'antenne. Ceci est dû à la convolution des sources par la matrice de propagation, ce qui a pour effet de faire apparaître les lobes secondaires, et éventuellement les lobes de réseau. Ces lobes dégradent la localisation de sources et la dynamique des images acoustiques. Ce phénomène est largement expliqué dans la littérature : retenons juste qu'il est possible de minimiser l'amplitude des lobes selon la forme de l'antenne, et que la géométrie doit être suffisamment irrégulière pour diminuer les lobes de réseau.

## 2.3 Localisation de sources large bande

### 2.3.1 Problème direct

En pratique, le vecteur de mesure  $\hat{\mathbf{p}}$  est extrait d'une transformée de Fourier discrète des signaux acquis. La propriété d'orthogonalité de cette transformée est importante car elle permet une généralisation en large bande. En effet, dans le domaine fréquentiel, le problème de localisation de sources en bande étroite, à la pulsation  $\omega_l$ , est indépendant des autres pulsations.

Nous notons  $\{\omega_1, \dots, \omega_l, \dots, \omega_L\}$  les  $L$  pulsations sur lesquelles sont définies les transformées de Fourier à court terme des pressions mesurées. Il est alors possible de poser  $L$  problèmes directs de propagation :

$$\hat{\mathbf{p}}_l = \mathbf{A}_l \hat{\boldsymbol{\beta}}_l, l \in [1, L] \quad (2.10)$$

en construisant les matrices de propagation  $\mathbf{A}_l$  aux pulsations  $\omega_l$ , avec les vecteurs de propagations soit en champ lointain (eq. (2.2)), soit en champ proche (eq. (2.3)).

Grâce à la généralisation large bande, la reconstruction des  $L$  vecteurs inconnus  $\hat{\boldsymbol{\beta}}_l$  permettra donc non seulement d'identifier la position des sources acoustiques, mais aussi d'accéder à leur signature spectrale.

### 2.3.2 Formation de voies large bande

Comme les propagations sont indépendantes pour chacune des  $L$  fréquences, il en est de même pour les rétropropagations par formation de voies. Il suffit donc d'appliquer l'opération (2.8) aux  $L$  vecteurs de pressions issus de la transformée de Fourier discrète des signaux :

$$\hat{\mathbf{b}}_l = \mathbf{A}_l^H \hat{\mathbf{p}}_l, l \in [1, L] \quad (2.11)$$

La reconstruction conduit à la matrice  $\hat{\mathbf{B}} = [\hat{\mathbf{b}}_1 \dots \hat{\mathbf{b}}_l \dots \hat{\mathbf{b}}_L]$ , de taille  $I \times L$ . Chaque élément de la matrice  $\hat{\mathbf{B}}$  est calculé en fonction du point source et de la fréquence. En champ lointain, l'expression est :

$$\hat{b}_{il} = \langle \frac{\mathbf{a}_i}{M}, \hat{\mathbf{p}}_l \rangle, l \in [1, L] \quad (2.12)$$

En champ proche, toujours pour conserver la fidélité de reconstruction en amplitude,  $\hat{b}_{il}$  est donné par la relation :

$$\hat{b}_{il} = \frac{1}{M} \left[ d_{1i} e^{j \frac{\omega_l}{c_0} d_{1i}} \quad \dots \quad d_{mi} e^{j \frac{\omega_l}{c_0} d_{mi}} \quad \dots \quad d_{Mi} e^{j \frac{\omega_l}{c_0} d_{Mi}} \right] \hat{\mathbf{p}}_l. \quad (2.13)$$

Nous garderons cette dernière relation car dans les expériences présentées plus loin, la configuration géométrique sera en champ proche pour la plus grande partie du spectre. En effet, le plan de source sera distant de 5 m, parallèle à une antenne de diamètre 2 m (cf. figure 2.5a). En considérant l'intervalle de fréquence typique du bruit urbain [50 Hz; 5 kHz] (intervalle défini dans [2]), la distance de Fraunhofer  $d_{Fr}$  varie dans un intervalle de 1.2 à 120 mètres. Cette distance coïncide avec la position du plan source (c'est à dire quand  $d_{Fr} = 5$  m) à la fréquence de 210 Hz. Donc dans tous les cas, nous considérerons que la localisation de sources est effectuée dans une condition de champ proche.

### 2.3.3 Imagerie par bande d'octave

Dans le cadre d'un diagnostic par imagerie acoustique en bande étroite, on observe généralement la valeur absolue de  $\hat{\mathbf{b}}$  sur une échelle logarithmique, en dB. Si nous étendons la formation de voies sur  $L$  fréquences, il serait en pratique impossible d'observer par imagerie les  $L$  reconstructions indépendamment. À titre d'exemple, avec une fréquence d'échantillonnage de  $f_s = 50$  kHz, une transformée de Fourier à court terme sur 125 ms donne en tout 3125 fréquences. L'imagerie large bande nécessite donc une étape de réduction des données pour être observable.

Classiquement, un diagnostic large bande est obtenu à l'aide d'un sonomètre, qui donne des niveaux SPL par bande d'octave (éventuellement de tiers d'octave) en un point. Dans le cas d'un signal de pression continu quelconque  $p(t)$  de durée  $T$ , le niveau SPL de la bande d'octave centrée sur la fréquence  $f_0$  se calcule par la relation :

$$\text{SPL}(f_0) = 10 \log \left( \frac{1}{p_{\text{ref}}^2} \frac{2}{T} \int_{f_0/\sqrt{2}}^{\sqrt{2}f_0} |\hat{p}(f)|^2 df \right) \quad (2.14)$$

où  $p_{\text{ref}} = 20 \mu\text{Pa}$  est la pression de référence dans l'air. L'égalité de Parseval montre que cette relation donne la puissance d'une trame du signal  $p$  de durée  $T$ , restreinte à l'intervalle de fréquence  $[f_0/\sqrt{2}, \sqrt{2}f_0]$ . Le calcul de niveaux par bande d'octave permet de quantifier la répartition spectrale de l'énergie acoustique, avec uniquement 10 valeurs dans le domaine audible<sup>2</sup>. Notons que, dans le cas d'un signal échantillonné, l'équivalent discret de la relation (2.14) s'écrit :

$$\text{SPL}(f_0) = 10 \log \left( \frac{1}{p_{\text{ref}}^2} \frac{2}{N^2} \sum_{f_0/\sqrt{2} < f_l < \sqrt{2}f_0} |\hat{p}(f_l)|^2 \right) \quad (2.15)$$

où  $N$  est le nombre d'échantillons issus de la trame de pression monocanal  $p$ .

Généralement, le niveau SPL obtenu par la relation 2.15 correspond à celui perçu à 1 m de

2. nous rappelons les fréquences centrales normalisées des bandes d'octave : 31.5, 63, 125, 250 et 500 Hz, 1, 2, 4, 8, et 16 kHz.

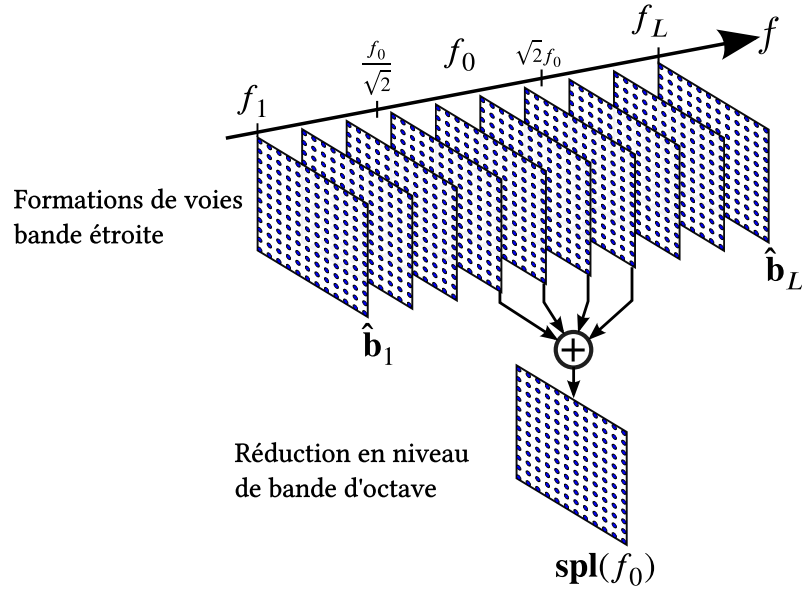


FIGURE 2.2 – Imagerie large bande : approche proposée.

distance de la source. On l'appelle alors le niveau ramené à 1 m. En pratique, la mesure par sonomètre est effectuée à une distance  $d$  connue par rapport à la source. En supposant une propagation en ondes sphériques, on ajoute  $10 \log(d^2)$  au niveau mesuré par sonomètre pour avoir le niveau équivalent à 1 m.

Pour faire un diagnostic global large bande par imagerie acoustique, l'idée proposée est de faire des images par "bande d'octave" : une image cartographie l'énergie acoustique contenue dans la bande d'octave correspondante. Si on considère une bande de fréquence centrale  $f_0$ , alors l'énergie dans cette bande est obtenue en s'inspirant de l'équation (2.15), en utilisant comme signal de pression la reconstruction par formation de voies  $\hat{\mathbf{B}}$ . Comme la décroissance de l'amplitude est prise en compte dans la rétropropagation (cf. éq. 2.13), la valeur obtenue en un point de la grille source s'apparente au niveau SPL ramené à 1 m. En notant  $\mathbf{SPL}(f_0)$  le vecteur de niveaux SPL des  $I$  points sources ramenés à 1 m, nous pouvons finalement calculer le  $i$ -ème élément par :

$$[\mathbf{SPL}(f_0)]_i = 10 \log \left( \frac{2}{N^2 p_{\text{ref}}^2} \sum_{f_0/\sqrt{2} < f_i < \sqrt{2}f_0} |\hat{b}_{ii}|^2 \right) \quad (2.16)$$

Notons qu'il est possible de spécialiser l'approche proposée selon le diagnostic. D'une part, il est possible d'appliquer une pondération sur ces niveaux (type A, B ou C). D'autre part, la bande de fréquence peut être une bande de tiers d'octave, ou n'importe quel ensemble de fréquences pertinent pour le diagnostic.

### 2.3.4 Comment garantir un traitement en temps réel

Le calcul en temps réel de la formation de voies en bande étroite est aujourd'hui facilement réalisable sur un processeur classique. C'est le cas de nombreuses solutions industrielles proposées. Comme le montre l'équation (2.8), le temps de calcul dépend de la taille de la grille  $I$  de reconstruction des sources. Il dépend également du nombre d'éléments du réseau  $M$ .

Cependant, étendre la formation de voies à une large bande devient très coûteux en temps de calcul. Sur des trames temporelles de 125 ms échantillonnées à 50 kHz, il faut effectuer le processus de formation de voies pour 3125 fréquences (2497 si on se limite à la bande audible), ce que est inaccessible avec un CPU pour garantir le temps réel.

#### 2.3.4.1 Implémentation sur GPU

La nature du calcul de formation de voies est identique quel que soit le point de la grille de reconstruction  $i$ , et la fréquence  $f_i$ . En effet, le calcul de chaque élément de la matrice  $\hat{\mathbf{B}}$  s'obtient par le produit scalaire donné en équation (2.7). Les différences sont uniquement associées aux valeurs des paramètres géométriques, de la fréquence, et des pressions mesurées  $\hat{\mathbf{P}}$ . En terme d'implémentation, cela se traduit par des instructions identiques, mais faisant appel à des données en mémoire différentes. Or le paradigme de programmation parallèle nommé Single Instruction Multiple Data (SIMD) convient précisément à cette situation.

Les architectures de calcul parallèle les plus courantes sont les clusters, qui sont des systèmes volumineux et chers. Rappelons que notre but est de développer un outil de diagnostic acoustique embarquable. Grâce à l'utilisation de microphones MEMS, nous avons pu fortement réduire la charge matérielle par rapport aux systèmes classiques. Pour rester fidèle à cette optique de mise en œuvre, nous préférons donc utiliser un GPU (Graphic Processing Unit) pour faire le calcul parallèle. Historiquement, les GPUs sont les processeurs des cartes graphiques utilisées en graphisme ou pour le jeu vidéo. Depuis une dizaine d'années, les constructeurs ont diversifié les applications du GPU dans le calcul scientifique parallèle [36, 99]. De plus, ils ont augmenté la simplicité de développement d'un code de calcul parallèle par rapport aux architectures type cluster ou FPGA [77]. En outre, cette architecture présente l'avantage d'être d'un coût du même ordre de grandeur que le système d'acquisition *Megamicros*. Enfin, comme il s'agit d'une carte graphique, toutes les données en mémoire peuvent être typées en texture pour faire de l'affichage performant, en utilisant par exemple la librairie *OpenGL*. Notons qu'une étude précédente a exploré l'intérêt de l'utilisation de GPUs dédiés à l'imagerie acoustique multi-fréquence en temps réel [109].

La carte graphique est un périphérique indépendant, branché sur le port PCIe du PC hôte. De ce fait, il contient sa propre mémoire, et n'exécute jamais de calcul sur la mémoire du PC (RAM). Toutes les données et paramètres doivent être d'abord transférées vers la mémoire globale du

GPU. Les données doivent être traitées à la volée pour obtenir les images en cinq étapes :

- Étape 1. Transférer les pressions des  $M$  canaux dans la mémoire globale du GPU, d'une durée de 125 ms par canal ;
- Étape 2. Calculer la matrice  $\hat{\mathbf{P}}$ . Une transformée de Fourier discrète est appliquée sur chaque canal de mesure.  $M$  transformées de Fourier rapides (FFT) sont effectuées en parallèle ;
- Étape 3. Calculer la matrice  $\hat{\mathbf{B}}$ . Par une approche SIMD, le calcul de chaque élément  $\hat{b}_{il}$  de  $\hat{\mathbf{B}}$  (eq. (2.12)) est affecté à un fil d'exécution (*thread*). Il y a donc au total  $I \times L$  threads ;
- Étape 4. Construire les images par bande d'octave. Toujours par SIMD, on parallélise le calcul de la valeur de chaque pixel des images par bande d'octave, donc l'équation (2.16) ;
- Étape 5. Afficher les images. De bonnes performances sont obtenues par la librairie *OpenGL*, car les données sont déjà en mémoire de la carte graphique. Il suffit de typer les vecteurs de niveaux acoustiques  $\mathbf{SPL}(f_0)$  dans des textures RGBA<sup>3</sup> puis de les afficher. Avec des performances quasi équivalentes, l'affichage peut être géré par la librairie *OpenCV*.

Le code global a été écrit en *python*. Comme nous utilisons des cartes graphiques *Nvidia*, la librairie *CUDA* a été utilisé pour coder les tâches de calcul parallèle. L'appel à des fonctions *CUDA* sous *python* est fait avec le module *pycuda* [78].

Pour les performances en temps réel, le CPU hôte n'est donc pas sollicité pour les calculs. Il gère la communication avec le GPU, et la récupération du flux de données par USB. Nous utilisons la librairie *libusb* pour gérer les transferts d'acquisition. Le protocole USB permet de faire du transfert asynchrone : la réception des données du système d'acquisition se fait par des tâches non-bloquantes, gérées par des interruptions. Ce point est essentiel pour libérer la charge du CPU, car le flux de données du système d'acquisition est dense et continu. Ainsi il reste du temps libre au CPU pour lancer les tâches de calcul au GPU (étapes 2, 3 et 4).

#### 2.3.4.2 Étude paramétrique des performances

Nous avons mesuré la vitesse de calcul d'une itération (Étapes 1 à 5 décrites en section 2.3.4.1) avec deux GPUs différents : (i) une *Nvidia Titan* muni de 2688 cœurs<sup>4</sup> *CUDA* et (ii) une *Nvidia Quadro 2000* contenant 192 cœurs. Pour quantifier le gain apporté par la parallélisation de la formation de voies, nous avons implémenté un code en C pour CPU (processeur Intel Xeon X5570) remplissant la même fonction que le code GPU. Le code CPU est exécuté sur un seul cœur.

Nous choisissons une configuration représentative d'un cas réaliste : la bande d'octave centrée sur 1 kHz est calculée sur une grille de  $100 \times 100$  points sources, à partir d'une trame de 125 ms

3. Pour décrire une texture, la couleur de chaque pixel est défini par le codage de couleurs RGBA : le rouge, vert, bleu, et la transparence Alpha.

4. Bien que la description physique donnée par le constructeur *Nvidia* d'un cœur *CUDA* soit peu précise, le nombre de cœurs dans un GPU peut être considéré comme le nombre de cœurs de CPU : il représente le nombre de threads exécutables en parallèle.



	CPU	GPU Quadro 2000	GPU Titan
Temps de calcul	8.12 s	31.1 ms	4.96 ms
Performance de calcul (pix/s)	1200	322000	2016000
Gain de vitesse GPU/CPU	/	261	1637

**TABLE 2.1** – Étude de performances : cas du calcul d’une image acoustique de bande d’octave, sur 125 ms de mesure.  $I = 10^4$ ,  $f_0 = 1$  kHz,  $f_s = 50$  kHz,  $M = 128$ ,  $N = 6144$ .

d’une acquisition sur 128 microphones, échantillonnée à  $f_s = 50$  kHz. Dans le cas présent il est trivial d’évaluer si le temps réel est possible : étant donné que le flux d’acquisition est continu, le temps de calcul d’une trame doit être inférieur à la durée de cette trame, c’est à dire 125 ms. Le tableau 2.1 montre les temps de calcul, obtenus via l’usage de timers pour garantir la meilleure précision de mesure.

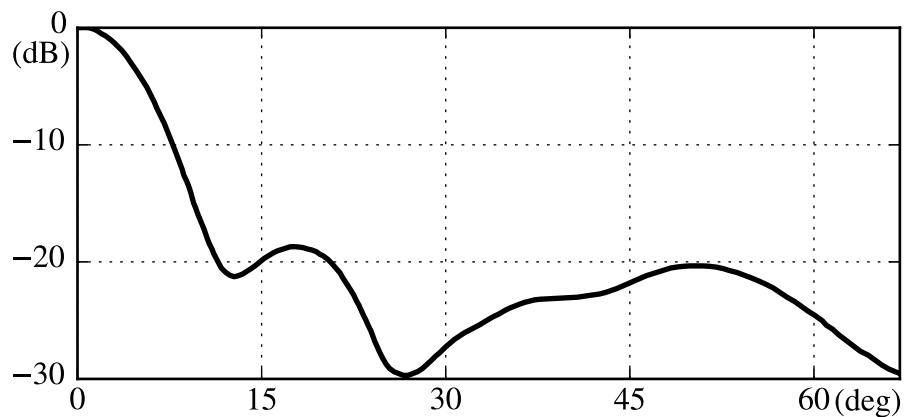
L’implémentation sur GPU montre un gain particulièrement important vis à vis de l’implémentation purement séquentielle. Pour un fonctionnement en temps réel de la formation de voies large bande, la parallélisation massive semble indispensable. En effet, sur un seul cœur CPU le calcul est de 8.12 s, ce qui représente 64 fois le temps de calcul limite acceptable. La parallélisation sur plusieurs dizaines de cœurs est une solution simple, d’autant plus que le calcul de la formation de voies par l’équation (2.12) est relativement élémentaire (produit scalaire). Ces raisons motivent d’autant plus le choix du GPU comme outil de calcul.

## 2.4 Expériences d’imagerie

Deux scénarios de localisation de sources sont mis en œuvre pour les expériences de validation :

- Le premier scénario s’apparente à une expérience de *calibrage* pour évaluer la précision des niveaux SPL ramenés à 1 m équivalent, par imagerie. La configuration est simple de manière à maîtriser correctement le niveau de deux sources harmoniques. Pour le calibrage, les niveaux évalués par la caméra acoustique sont comparés avec un sonomètre *Brüel & Kjaer 2236* de classe 1 (garantissant une incertitude inférieure à 1.5 dB) et offrant une résolution de 0.1 dB ;
- le deuxième scénario est une illustration d’un diagnostic temps réel avec un aspirateur immobile. Nous verrons l’intérêt de l’imagerie par bande d’octave pour obtenir conjointement deux caractéristiques d’une scène sonore : la localisation spatiale des sources et leur contenu spectral.

La caméra acoustique est configurée de manière à ressembler au plus à un sonomètre standard. D’une part, les images donnent des niveaux par bande d’octaves aux fréquences normalisées (cf. norme ISO 266). D’autre part, le taux de rafraîchissement des images est de 125 ms, ce qui correspond à la durée d’intégration normalisée des trames temporelles pour calculer le niveau de pression instantané.



**FIGURE 2.3** – Antenne *Soleil* (photo 2.5a) : diagramme de directivité théorique radial en champ lointain, obtenu par moyennage angulaire autour de l'axe central de l'antenne.

Le système d'acquisition comprend 128 microphones : la géométrie de l'antenne est circulaire (cf. *Soleil* en photo 2.5a). 16 barres de 8 microphones constituent les rayons d'un disque de 2.5 m de diamètre. Sur chaque barre, les microphones sont régulièrement espacés de 17 cm. Pour casser la régularité, les barres sont décalées aléatoirement par rapport au centre du disque. Le diagramme de directivité moyen de l'antenne *Soleil* est tracé en figure 2.3. La fréquence d'échantillonnage des signaux est  $f_s = 50$  kHz. L'imagerie est effectuée sur une grille carrée de  $I = 100 \times 100$  points. Les calculs sont effectués sur le GPU *Nvidia GTX Titan*, qui présente des performances amplement suffisantes pour le temps réel (cf. section précédente 2.3.4.2).

### 2.4.1 Scénario 1 : expérience de calibrage

Nous mettons en place une expérience pour quantifier la précision des niveaux SPL absolus cartographiés par formation de voies, en présence de deux sources contrôlées. Nous comparons ces niveaux avec ceux donnés par un sonomètre homologué, en présence de chaque source isolée.

Cette expérience s'apparente donc à un calibrage de l'outil de diagnostic, qui comprend à la fois le système d'acquisition et la partie algorithmique. Deux enceintes sont placées à 5 m devant le réseau circulaire. Elles émettent une sinusoïde de 3 kHz d'amplitude constante. Pour analyser la dynamique limite due aux lobes secondaires, nous faisons trois tests avec des amplitudes différentes :

- (a) l'enceinte droite/gauche émet à 69/69 dB ;
- (b) l'enceinte droite/gauche émet à 69/64 dB ;
- (c) l'enceinte droite/gauche émet à 69/59 dB.

Nous ne prenons pas en compte la directivités des enceintes. Les sources, situées dans le champ proche, sont supposées émettre des ondes sphériques. Pour s'assurer des bons niveaux acoustiques

Test	(a)	(b)	(c)
Niveau gauche	69.6 (69.0)	63.8 (64.0)	60.9 (59.0)
Niveau droite	69.8 (69.0)	68.2 (69.0)	69.3 (69.0)
Niveau relatif	0.2 (0)	4.5 (5.0)	8.4 (10.0)

**TABLE 2.2** – Calibrage de la caméra acoustique. Niveaux acoustiques des sources en dB donnés par l'imagerie ; niveaux  $L_{eq}$  du sonomètre entre parenthèses.

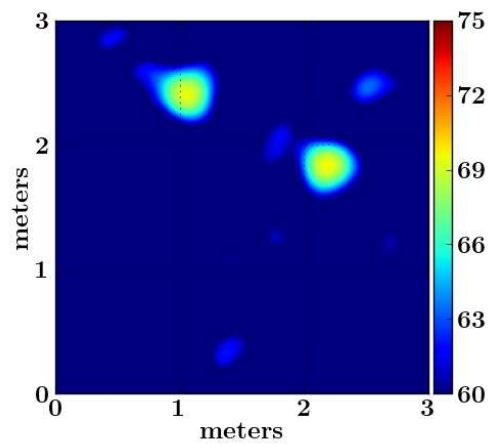
émis, nous plaçons un sonomètre à 1 m devant chacune des sources rayonnant séparément. Nous réglons les amplificateurs audio pour que le niveau équivalent continu ( $L_{eq}$ ), mesuré sur 30 s, soit celui souhaité. Compte tenu de la précision de mesure du sonomètre, les niveaux à 1 mètre équivalent des sources sont précis à 0.1 dB.

La figure 2.4 présente les images intégrées sur la bande d'octave 4 kHz pour les trois différents tests ; ces images permettent bien de localiser les deux enceintes. Le tableau 2.2 compare les niveaux ramenés à 1 m du sonomètre avec ceux de la formation de voie focalisant à la position des enceintes. Les niveaux SPL par bande d'octave absolus estimés par formation de voies ont un biais inférieur à  $0.8 \pm 0.1$  dB dans les tests (a) et (b). Dans le cas (c), le niveau de l'enceinte gauche est biaisé de  $1.9 \pm 0.1$  dB. Il s'agit du cas où les deux enceintes émettent avec la plus grande différence : 10 dB. De plus, la figure 2.4c laisse apparaître des lobes secondaires, d'amplitude équivalente à l'enceinte gauche. Cela montre que le cas (c) est proche de la dynamique limite du système imageur pour cette fréquence. Au-delà, si deux sources acoustiques ont une différence de niveau supérieure à environ 12 ou 13 dB, la source la plus faible se retrouve noyée dans les lobes secondaires issus de la source dominante. Cependant cette dynamique dépend de la directivité de l'antenne, qui est fonction de la géométrie et du nombre d'éléments. Il est possible d'aller au-delà de cette dynamique, soit en réduisant les lobes secondaires (Capon [91]), soit en passant des méthodes de déconvolution de la réponse de l'antenne (typiquement CLEAN [112, 115] ou DAMAS [24]).

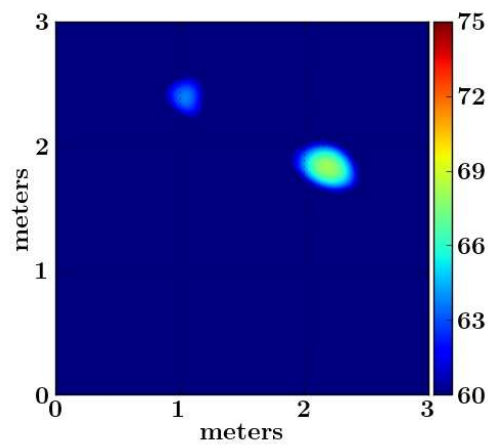
Avec l'antenne circulaire utilisée, la formation de voies intégrée sur bande d'octave produit des niveaux de source très comparables à ceux mesurés avec le sonomètre pour la source dominante. Le niveau de la source secondaire en revanche est mal évaluée s'il approche le niveau relatif des lobes secondaires de l'antenne. Il est alors surévalué du fait de l'énergie apportée par ces artefacts qui fixent la dynamique de détection des sources.

### 2.4.2 Scénario 2 : l'aspirateur

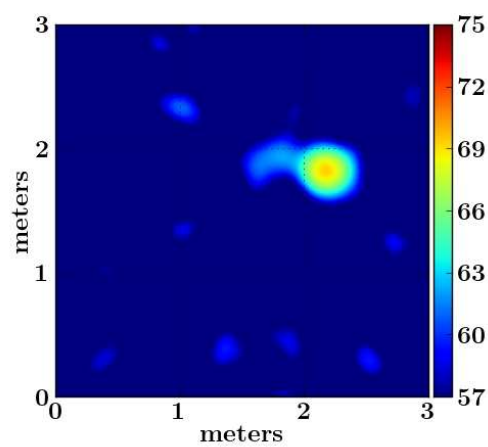
Ce deuxième scénario est une démonstration d'un diagnostic type industriel. L'objet étudié est un aspirateur placé à 4 m face à l'antenne. La figure 2.5c est représentative des images que l'on obtient en temps réel (à 8 images par seconde). La caméra cartographie le son sur une grille de  $3 \text{ m} \times 3 \text{ m}$ , discrétisée en  $100 \times 100$  points. Six bandes d'octaves, de  $f_0 = 250$  à 8000 Hz, sont



(a) Niveaux émis : 69 dB (gauche et droit).



(b) Niveaux émis : 64 dB (gauche), 69 dB (droit).



(c) Niveaux émis : 59 dB (gauche), 69 dB (droit).

FIGURE 2.4 – Calibrage de la caméra acoustique. Deux enceintes émettent un sinus de 3 kHz.

analysées simultanément.

Une observation globale de la figure valide trois points :

- à l’instar de l’imagerie monochromatique classique, le filtrage spatial par formation de voies localise et dissocie les différentes sources dans l’espace. Ici, les trois sources prépondérantes de l’aspirateur sont la partie moteur, la bouche d’aspiration et l’échappement ;
- l’imagerie par bande d’octave apporte bien plus d’information qu’en bande étroite. Les sources de bruit étudiées sont généralement large bande, et ici on obtient une définition espace-fréquence relativement riche, avec un nombre limité d’images. Dans un contexte industriel, cela permettrait d’orienter un diagnostic rapidement, et d’identifier la nature de chacune des sources. Par exemple, le bruit en basse fréquence indique le rayonnement du moteur. En moyenne et haute fréquences, le bruit aéroacoustique venant de la bouche d’aspiration rayonne sur une bande plus large. Notons également, dans la bande 2 et 4 kHz, un bruit près du moteur. Il s’agit du bruit aéroacoustique d’échappement d’air ;
- d’après l’expérience de calibrage, les niveaux acoustiques SPLs peuvent être correctement estimés pour chaque source de bruit. Mais contrairement au sonomètre qui donne un niveau SPL de la scène globale, la caméra par bande d’octave *isole* virtuellement chaque source de bruit grâce au filtrage spatial. Le niveau d’une source peut être interprété comme un sonomètre mesurant le bruit rayonné, issu uniquement de cette source. Cela dit, la formation de voies permet aussi d’extraire le spectre de chaque source, via la matrice  $\hat{\mathbf{B}}$ .

Malgré la grande taille de l’antenne, un inconvénient majeur lié à la formation de voies apparaît : la directivité de l’antenne varie en fonction de la fréquence. En conséquence, les images en basse fréquence ont une résolution faisant apparaître une tache large pour le bruit de moteur. À l’inverse, les images haute fréquence montre une résolution fine. L’image à 8 kHz l’illustre bien, en mettant en évidence deux taches séparées aux niveau de la bouche d’aspiration : il s’agit de l’entrée d’air principale et du clapet de régulation.

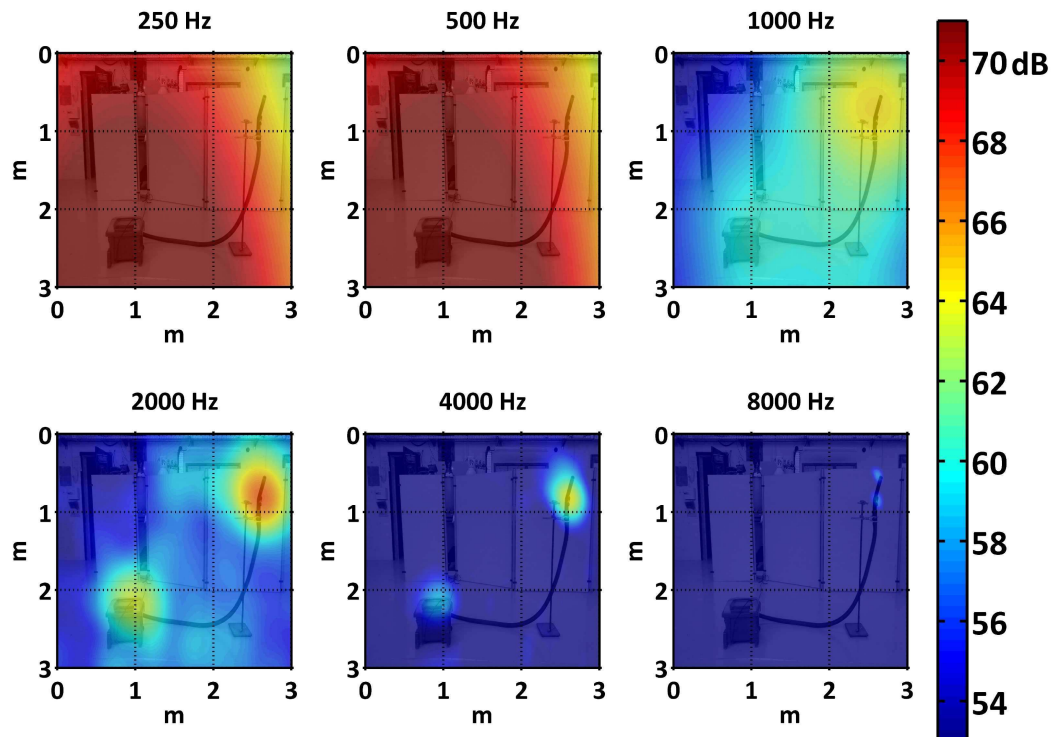
Enfin, le dernier point fondamental apporté par le temps réel est la possibilité d’un diagnostic dynamique. Elle permet d’observer des sources non stationnaires. Dans le cas de l’aspirateur, il est par exemple possible de suivre l’évolution des niveaux acoustiques pendant des changements de régime du moteur. Cette dimension dynamique serait d’une aide précieuse aux ingénieurs dans la mise au point des machines.



(a) L'antenne *Soleil* de 128 microphones MEMS, utilisée pour le calibrage et l'imagerie de l'aspirateur ; diagramme de directivité en figure 2.3.

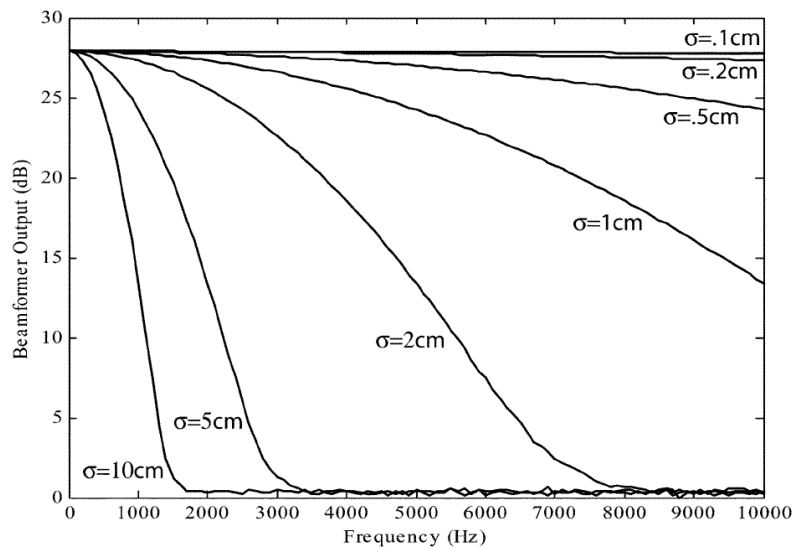


(b) L'aspirateur est placé 4 m devant l'antenne.



(c) Résultat de l'imagerie large bande d'un aspirateur sur 6 bandes d'octave. Les niveaux SPLs sont relatifs à la pression de référence  $p_{\text{ref}} = 20 \mu\text{Pa}$ .

FIGURE 2.5 – Scénario type industriel : diagnostic temps-réel d'un aspirateur.



**FIGURE 2.6** – Figure extraite de l'article de Sachar et al. [125]. La formation de voies large bande est construite pour identifier une source de bruit blanc avec une densité spectrale de puissance de  $28 \text{ dB}/\sqrt{\text{Hz}}$ . Une erreur est ajoutée sur la connaissance *a priori* de la géométrie du réseau, pour différents écarts types  $\sigma$  allant de 0.1 à 10 cm. Le réseau est muni de 448 microphones étendus sur  $6 \text{ m} \times 4 \text{ m}$ .

## 2.5 Discussion : impact de l'erreur en position des microphones

### 2.5.1 Le montage manuel de l'antenne induit une erreur géométrique

Les expériences précédentes mettent en valeur la capacité de l'outil *in situ*. Un atout majeur est sa transportabilité, et la rapidité d'installation de l'antenne. Durant ma thèse, j'ai pu assister ou participer à diverses expériences impliquant l'utilisation du réseau de 128 ou 256 microphones MEMS couplé à l'imageur acoustique temps réel. Comme le montrent les illustrations en figure 1.3 du chapitre 1, des configurations d'antenne variées ont été employées.

Cependant ces expériences successives ont révélé un problème inhérent à l'installation des réseaux *in situ* : il est difficile d'obtenir un montage de l'antenne de grande précision. En effet, les microphones des antennes classiques sont habituellement fixés sur une structure rigide. Une telle approche est difficile quand la géométrie globale s'étend sur plusieurs mètres. L'architecture matérielle du système *Megamicros* est décomposée en plusieurs sous-réseaux, chacun étant une barre de 8 microphones (identiques aux barres qui constituent l'antenne circulaire fig. 2.5a). Étant données les dimensions des réseaux ( $\approx 20 \text{ m}$ ) et le protocole de montage, la précision de positionnement des microphones est limitée. Comment cela se répercute-t-il sur l'application à la localisation de sources ?

Une erreur de position se traduit par un décalage temporel du signal acoustique, ce qui correspond à une erreur en phase linéaire, proportionnelle à la fréquence. Ce décalage temporel apparent est également fonction de la direction de la source de son. La figure 2.6, extraite de

l'article [125], montre bien la dégradation du gain de formation de voies augmenter quand la longueur d'onde diminue. Globalement, l'erreur n'est plus tolérable au-delà d'un biais moyen de 2 cm pour le domaine fréquentiel correspondant au bruit urbain. Dans le cas du filtrage spatial appliqué à la parole, l'erreur tolérable est de 1.5 cm [100] à 2 cm [68] selon le type de formation de voies. Pour plus de détails, une étude analytique établit une liste exhaustive des effets de l'erreur géométrique sur le *Beamforming* large bande [34].

## 2.5.2 Méthodes de correction des dispersions de mesure

L'erreur de positionnement est une des dispersions de mesure, et il en existe d'autres qui ne sont pas prises en compte par le modèle de propagation (équation (2.6)). Dans le cas de mesures réelles, deux types de dispersions doivent être considérés.

### 2.5.2.1 Dispersion issues de la mesure

Les caractéristiques intrinsèques des microphones engendrent deux types de bruits sur le vecteur de mesure  $\hat{\mathbf{p}}$  :

1. la dispersion de phase, venant de la réponse en fréquence, engendre une distorsion du signal temporel. La localisation de sources dans le domaine des fréquences est particulièrement sensible à l'erreur sur la phase ;
2. la dispersion en amplitude, qui est fonction de la fréquence à cause la variabilité de la réponse en fréquentielle de chaque microphones. La formation de voies est moins sensible aux erreurs de gain qu'aux erreurs de phases. En revanche, si les microphones sont directifs, alors la dispersion de gain varie en fonction de l'orientation des sources par rapport à l'antenne. Cette contrainte bien plus gênante est à éviter en expérience, car ces derniers sont considérés omnidirectionnels dans le modèle de propagation. Dans notre cas, nos expériences échappent à ce biais car les microphones MEMS du système *Megamicros* sont omnidirectionnels (cf. caractérisation acoustique au chapitre 1).

Pour pallier ces erreurs, la première approche possible consiste à rendre le *Beamforming* large bande robuste aux dispersions de phase et d'amplitude [51, 156]. Pour cela *Doclo et al.* proposent d'extraire les vecteurs de propagation optimaux  $\mathbf{a}_i$ , de manière à ce que les diagrammes de directivités de l'antenne apparents (issues du *Beamforming* des sources), correspondent au mieux aux diagrammes théoriques. Cette approche peut être envisagée selon deux méthodes différentes. La première proposée (*mean-case performance*) nécessite les densités de probabilité des dispersions de gain et de phase, pouvant être issues de la documentation constructeur ou d'une caractérisation des microphones (cf. figures 1.5-1.8 en chapitre 1). La deuxième méthode (*worst-case performance*) adapte le problème par une optimisation par *minimax*, et ne nécessite aucun *a priori* sur les microphones [35, 104].



Enfin, le deuxième type d'approche consiste à corriger, préalablement au *Beamforming*, les gains et phases avec une méthode de calibration. Il s'appuie généralement sur un *a priori* sur les sources mesurées. Par exemple, les calibrations fondées sur MUSIC nécessitent la connaissance du nombre de sources pour dissocier les espaces signal et bruit [101, 102]. Un autre exemple, s'appuyant sur l'hypothèse d'un faible nombre sources, permet d'utiliser une approche parcimonieuse pour identifier les erreurs de gain et/ou de phase conjointement à la position des sources [15, 62].

### 2.5.2.2 Dispersions issues du modèle

Le montage et les conditions expérimentales engendrent des erreurs sur le modèle de propagation  $\mathbf{A}$  :

1. la position des microphones, comme énoncé en section 2.5.1. Elle est analogue à une erreur de phase, fonction des positions géométriques et de la fréquence. C'est donc une erreur affectant rapidement la localisation de sources ;
2. la célérité du milieu, qui intervient dans la phase des vecteurs de propagation (cf. expression des vecteurs  $\mathbf{a}_i$ ). Même si elle varie sur une faible plage, elle peut engendrer une erreur géométrique élevée quand les dimensions de l'expérience deviennent grandes. Nous considérons l'erreur sur la célérité négligeable si la température du milieu est connue.

De la même manière qu'avec les erreurs de mesure, soit le *Beamforming* est rendu robuste aux erreurs géométriques, soit nous faisons appel à une méthode de calibration géométrique pour corriger la position des microphones. Il existe des variantes des *Beamforming* du type *worst/mean-case performance* énoncées section 2.5.2.1, robustes à une erreur en position, et dont la densité de probabilité est connue [104] ou non [8]. Autrement, il est possible de corriger l'erreur préalablement à l'application de la localisation de sources. Il existe une calibration géométrique pour affiner une géométrie, fondée sur MUSIC [57]. Cependant elle considère un domaine plan, dans lequel les sources sont positionnées à l'infini. Actuellement, je n'ai pas trouvé de calibration géométrique fondée sur la parcimonie des sources ; je pense qu'il serait intéressant de s'inspirer de [15] en adaptant la correction de phase à l'erreur géométrique, pour établir une méthode de calibration géométrique en présence d'un faible nombre de sources.

Nous avons énoncé des méthodes qui correspondent à une *correction* de la géométrie d'antenne. Cela dit, il existe des approches plutôt adaptées à l'*identification* d'une géométrie inconnue. Cette discussion positionne le problème de calibration géométrique à l'aveugle, qui sera le sujet des deux prochains chapitres. En effet, nous avons vu que la grande taille du réseau, déployé sur le terrain d'expérience, limite la précision de positionnement des microphones. Dans ce cas, l'intérêt de méthodes de calibration géométrique est :

- soit de corriger la géométrie si l'antenne est construite à partir d'un plan,
- soit d'identifier la géométrie inconnue si le réseau est déployé arbitrairement.

Enfin, cette discussion montre qu'il sera important de quantifier les performances d'une méthode

de calibration vis à vis de l'application qui découlera de l'antenne acoustique. Pour cela nous vérifierons l'applicabilité des résultats à l'imagerie acoustique.

## 2.6 Conclusion

Ce chapitre a présenté la première application dédiée à un grand réseau de microphones MEMS dans le cadre du projet *Megamicros*. Nous avons construit et mis en œuvre un instrument de mesure acoustique complet, de la partie matérielle au traitement de données. Il associe en temps réel les fonctions d'imageur acoustique et de sonomètre, et permet le diagnostic dynamique des sources de bruit. L'analyse quantifiée de la scène sonore s'effectue simultanément sur plusieurs bandes de fréquences. Nous avons montré qu'il est possible de détourner l'utilisation de microphones MEMS, normalement dédiés à l'acquisition de la parole, afin de construire une antenne pour la métrologie. Utilisés en grand nombre et couplés à la formation de voies, ces microphones sont bien adaptés puisqu'ils restituent fidèlement les niveaux acoustiques.

Un diagnostic en temps réel est bien plus ergonomique qu'un post-traitement, et permet d'obtenir une analyse *in situ*. Le calcul en temps réel a donc été une priorité, satisfaite par une implémentation sur GPU. L'exécution sur une architecture parallélisée réduit considérablement le temps de calcul vis à vis d'un CPU. Notons que la méthode de *Beamforming* par Capon se prête également au GPU avec des bon gains de performance [4, 5].

Le choix de la formation de voies s'est imposé car nous maîtrisons bien ses propriétés : ces premiers travaux de thèse ont avant tout comme but de valider l'applicabilité des microphones MEMS pour l'antennerie acoustique. Pour faire de l'imagerie par bande d'octaves, l'utilisation d'algorithmes de localisation plus performants pourrait constituer une prochaine étape. On pourrait envisager par exemple d'implémenter des méthodes par déconvolution (par exemple Clean [160], Damas [39]), avec régularisation parcimonieuse [21, 115], ou par approche bayésienne [38] qui ont été adaptées pour les sources large bande.



## Chapitre 3

# Calibration géométrique passive et robuste de grands réseaux de microphones en champ diffus

### Sommaire

---

<b>3.1</b>	<b>Introduction</b>	<b>48</b>
<b>3.2</b>	<b>Méthodologie</b>	<b>51</b>
3.2.1	Description du problème	51
3.2.2	Calibration géométrique : description fonctionnelle globale	51
<b>3.3</b>	<b>Estimation des distances par paire en champ diffus</b>	<b>52</b>
3.3.1	La cohérence en champ diffus	52
3.3.2	Estimation des distances par paire entre microphones	54
3.3.3	Limitations de l'estimation des distances par régression	54
3.3.3.1	Limitations numériques	55
3.3.3.2	Limitation physique : la diffusivité acoustique	56
<b>3.4</b>	<b>Estimation robuste de la géométrie du réseau</b>	<b>57</b>
3.4.1	Le MDS pour la localisation de capteurs	57
3.4.1.1	Approche par minimisation de la fonction <i>raw stress</i>	57
3.4.1.2	Approches locales de MDS pour la localisation de capteurs	59
3.4.2	Le RMDS garantit la robustesse de l'estimation	61
3.4.3	Le LRMDS, une approche locale du RMDS adaptée aux grands réseaux	62
3.4.4	Algorithme de résolution	62
3.4.4.1	Mise à jour de $\mathbf{O}$	63
3.4.4.2	Mise à jour de $\mathbf{X}$	63
3.4.4.3	Initialisation et convergence	64
<b>3.5</b>	<b>La réalisation expérimentale</b>	<b>65</b>

3.5.1	Scène et dispositif expérimentaux . . . . .	65
3.5.2	Estimation des cohérences et des distances . . . . .	67
3.5.3	Le LRMSD en pratique : la configuration paramétrique . . . . .	69
<b>3.6</b>	<b>Résultats expérimentaux . . . . .</b>	<b>73</b>
3.6.1	Performances de la méthode proposée . . . . .	73
3.6.2	Comparaison de résultats avec d'autres approches de la littérature . . . . .	76
3.6.2.1	Autre méthode de positionnement de microphones en champ diffus . . . . .	76
3.6.2.2	Deux MDS locaux utilisés pour le positionnement de capteurs . . . . .	78
<b>3.7</b>	<b>Discussions . . . . .</b>	<b>81</b>
3.7.1	Applicabilité à l'imagerie acoustique . . . . .	81
3.7.2	Au delà du milieu réverbérant . . . . .	84
3.7.2.1	Cocktail party en environnement intérieur . . . . .	85
3.7.2.2	Bruit urbain en environnement extérieur . . . . .	85
3.7.3	Matrice de distances et loi de distribution . . . . .	88
<b>3.8</b>	<b>Conclusion . . . . .</b>	<b>89</b>

## 3.1 Introduction

Dans de nombreuses situations réelles où il est fait appel aux techniques d'antennerie acoustique, l'environnement est réverbérant. Les problèmes inverses classiques tels que la localisation ou séparation de sources deviennent alors plus complexes du fait de cette réverbération. Différentes méthodes de filtrage spatial ont été proposées, qui permettent d'enlever la contribution de la réverbération. Un premier exemple est la séparation du champ direct et du champ réverbéré pour faire de la localisation de sources : l'intérêt est de supprimer les fausses détections dues aux réflexions sur les obstacles [33, 105]. Un second exemple est le suivi d'une source acoustique mobile robuste à la présence des réflexions [161]. Enfin, certaines approches préfèrent exploiter les échos comme une source d'information sur le milieu, pour par exemple retrouver la forme d'une salle [53], ou encore retrouver l'impédance des murs [14]. Dans ce chapitre, nous cherchons à tirer avantage de la réverbération pour résoudre le problème de la calibration géométrique des grands réseaux.

Les études théorique et expérimentale du champ acoustique diffus ont été formalisées à partir des années 50, et décrivent le champ caractéristique d'un milieu réverbérant. Les travaux de Schroeder restent historiquement les plus connus [128, 129]. D'autres études ont porté sur la corrélation spatiale de la pression acoustique mesurée en deux points distincts [37, 40, 73]. Elles montrent que l'intercorrélacion temporelle et la densité spectrale croisée sont fonctions de la distance séparant les deux points de mesure. En effet, en notant  $S_{nm}$  la densité spectrale croisée entre deux mesures microphoniques  $n$  et  $m$ , nous avons la relation caractéristique d'un champ diffus :

$$S_{nm} \propto \text{sinc}(k_0 d_{nm}) \quad (3.1)$$

où  $k_0$  est le nombre d'onde et  $d_{nm}$  la distance entre les microphones  $n$  et  $m$ . Il est également possible de calculer la cohérence  $\Gamma_{nm}$  qui correspond à la densité spectrale croisée normalisée entre  $-1$  et  $1$ . Dans ce cas, il est possible de montrer que :

$$\Gamma_{nm} = \text{sinc}(k_0 d_{nm}) \quad (3.2)$$

qui est une relation d'égalité, contrairement à l'équation 3.1 qui reste une relation de proportionnalité. En pratique, c'est donc la cohérence qui est mesurée pour obtenir la distance entre micros. Ce résultat est également retrouvé dans une chambre réverbérante électromagnétique, par la cohérence du champ électrique en deux points [67].

Dans un problème de calibration géométrique où la position des éléments est inconnue, l'écoute d'un champ diffus permet donc de tirer une connaissance géométrique élémentaire : la distance entre les microphones deux à deux. La première description d'une méthodologie de calibration géométrique en champ diffus aérien est exposée par McCowan *et al.* [94]. Elle s'appuie sur la régression d'une cohérence analytique sur les cohérences mesurées entre toutes les combinaisons possibles des couples de microphones, menant à l'ensemble des distances par paire du réseau. Pour apporter de la robustesse à cette procédure, plusieurs régressions sont faites sur différentes cohérences au cours du temps, puis une méthode de classification (*clustering*) de type K-mean en déduit la valeur de distance optimale (cette étape sera détaillée dans la section 3.6.2.1 de ce chapitre). Cela réduit le risque d'extraire une valeur aberrante issue d'une unique régression. Enfin, la dernière étape transforme l'ensemble des distances en positions des microphones, et fait appel à une méthode de Multidimensional Scaling (MDS). McCowan *et al.* parviennent à retrouver la géométrie d'une antenne à 8 microphones et d'une taille de 20 cm, avec une précision du centimètre. Notons également qu'ils génèrent le champ diffus artificiellement, avec des sources émettant un bruit blanc.

À partir de cette idée très novatrice, Taghizadeh *et al.* proposent des améliorations importantes à la méthode générique [139]. La première porte sur l'estimateur de la cohérence : son biais est considérablement réduit en faisant une moyenne au cours du temps. Ce changement diminue sensiblement l'erreur sur la régression des cohérences. Deuxièmement, la méthode de clustering par K-mean est remplacée par un histogramme 2D, qui est plus efficace pour écarter les valeurs aberrantes (voir également la section 3.6.2.1 pour les détails sur les méthodes de clustering). Enfin il est démontré que le champ diffus naturel, issu du bruit ambiant, peut suffire pour obtenir des mesures de cohérences exploitables. La validation expérimentale implique un réseau de 9 microphones de dimension 20 cm. Elle montre une réduction de l'erreur de calibration géométrique, qui est de 5 cm de moyenne avec la méthode proposée, contre 8.8 cm avec la méthode générique de McCowan. Cependant cette nouvelle approche ne montre pas d'amélioration concernant l'étendue du réseau (qui reste de 20 cm) et le nombre de microphones (toujours inférieur à 10).

Hennecke *et al.* s'inspirent également de la méthode générique, pour proposer une approche hybride, par mesure de champ diffus puis par mesure de différence de temps de vol. Il s'agit de

positionner un ensemble de petits sous-réseaux distribués dans une salle [64]. Deux expériences sont exposées, dans lesquelles chaque sous-réseau est constitué de 4 (resp. 8) microphones, d'une taille de 14 cm (resp. 20 cm); puis finalement 4 (resp. 2) sous-réseaux sont distribués dans une salle de 3.7 m × 6.8 m. La calibration géométrique est hiérarchisée en deux étapes : d'abord la forme de chaque sous-réseau est retrouvée par une calibration en champ diffus. La deuxième étape positionne les sous-réseaux par rapport à quelques sources ponctuelles dans la salle, ce qui donne la géométrie totale. La deuxième étape s'apparente donc à une calibration géométrique par différence de temps de vol (dont le principe est présenté dans le chapitre 4). Bien que l'antenne complète soit étendue sur plusieurs mètres, la calibration en champ diffus reste limitée aux petits sous-réseaux.

La littérature met en évidence un point commun : la validation expérimentale n'est jamais appliquée à un réseau étendu à grand nombre de microphones, et ce à cause de deux limitations. Premièrement, la mesure de la cohérence en champ diffus se dégrade si deux microphones deviennent trop éloignés. À titre d'exemple, *une* étude paramétrique expérimentale constate que la régression sur une cohérence devient critique quand la distance par paire est supérieure à 73 cm [140] (les raisons de cette limitation seront exposées dans ce chapitre). En pratique, cela signifie que si un réseau est étendu, alors certaines distances sont inconnues, et la matrice des distances est incomplète. De nombreuses techniques proposent d'obtenir les distances manquantes par complétion de la matrice de distances euclidiennes [17, 52, 138, 141]. Mais une seconde limitation s'ajoute : le nombre de distances par paire (et par conséquent celui des régressions de cohérences) augmente au carré du nombre de microphones. Bien que les méthodes de *clustering* par K-mean ou histogramme 2D peuvent réduire la présence de valeurs aberrantes, nous verrons qu'elles ne sont pas infaillibles et ne sont plus adaptées pour un grand nombre de microphones (cf. section 3.6.2.1). Par conséquent, il faudrait à la fois *compléter* la matrice de distances et *débruiter* les distances connues des valeurs aberrantes. Or la littérature montre que l'on peut (a) soit débruiter une matrice de distances complète contaminée par des valeurs aberrantes [30, 123], (b) soit compléter une matrice de distances contaminées par un bruit de faible amplitude [54]. L'analyse de cette littérature montre que les approches proposées ne résolvent pas conjointement la *complétion* et le *débruitage* des matrices de distances. Elles ne conviennent donc pas à notre problème.

Le but de ce chapitre est de proposer une méthode adaptée aux grands réseaux de microphones. Plutôt que de réduire la présence des valeurs aberrantes, nous considérons leur présence comme une inconnue, associée à celle de la géométrie du réseau. De plus, la méthode est locale : elle ne nécessite que les distances entre les microphones les plus proches. Le choix de la technique MDS sera la clé, fondée sur les travaux de Forero *et al.* [59]. D'abord, le chapitre expose la méthodologie globale de la calibration géométrique, et décrit le modèle de cohérence et son l'estimateur. La théorie de la méthode employée de résolution MDS sera détaillée en section 3.4. Ensuite, la première partie expérimentale 3.5 analysera la mise en pratique de la méthode MDS proposée. La deuxième partie expérimentale 3.6 quantifiera les performances de la méthode; elle comprendra une comparaison avec différentes approches de l'état de l'art, employées pour le positionnement

de microphones [139], ou de capteurs d'un réseau sans fil. Enfin nous discuterons 3.7 de l'impact de l'erreur de calibration géométrique sur l'imagerie acoustique, puis de l'applicabilité de la méthode *in situ* dans des environnements de terrain.

## 3.2 Méthodologie

### 3.2.1 Description du problème

Nous définissons  $p_n(t)$  la pression temporelle mesurée par le  $n$ -ème microphone, échantillonnée à la fréquence  $f_s$ .  $\hat{p}_n(f)$  est la transformée de Fourier à court terme associée à  $p_n(t)$ , calculée sur une trame de  $N$  échantillons temporels, donc de durée  $T = (N - 1)/f_s$ .

Nous considérons un réseau de  $M$  microphones, distribués dans l'espace aux positions définies par les  $M$  vecteurs colonnes  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M$  de taille  $D \times 1$ . Selon la forme de l'antenne, volumique, plane ou linéique,  $D = 3, 2$  ou  $1$  (nous supposons cette information connue *a priori*, sinon par défaut nous choisissons  $D = 3$ ). Nous posons la forme matricielle contenant toutes les coordonnées du réseau :  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_M]$ , définie dans l'ensemble  $\mathbb{R}^{D \times M}$ . Dans le cadre d'une calibration géométrique,  $\mathbf{X}$  est l'inconnue à déterminer. L'objectif est de trouver une estimation de la position des  $M$  microphones du réseau, appelée  $\tilde{\mathbf{X}} = [\tilde{\mathbf{x}}_1, \dots, \tilde{\mathbf{x}}_M]$ .

Enfin, notons  $d_{nm} = \|\mathbf{x}_n - \mathbf{x}_m\|$  la distance séparant les microphones  $n$  et  $m$ , ainsi que la matrice de distances euclidiennes  $\mathbf{D}$  de taille  $M \times M$ . Dans un espace euclidien,  $d_{nm} = d_{mn} \forall (n, m)$ ,  $\mathbf{D}$  appartient donc à l'ensemble des matrices symétriques dont les éléments diagonaux sont nuls :  $\mathbf{D} \in \mathbb{S}_h^M$ .

### 3.2.2 Calibration géométrique : description fonctionnelle globale

Dans une vue d'ensemble, la calibration géométrique en champ diffus se décline en trois étapes, que ce soit pour les méthodes présentées dans la littérature [64, 94, 139] ou pour celle proposée dans ce chapitre.

1. Le champ acoustique est mesuré par les microphones, sur une durée de plusieurs secondes ou minutes. La cohérence mesurée, notée  $\tilde{\Gamma}_{nm}$ , est obtenue de manière statistique grâce aux mesures, pour toutes les paires de microphones  $(n, m)$ ;
2. À partir d'un modèle du champ diffus, une fonction analytique décrit la cohérence. Dans le cas particulier du champ diffus, nous montrerons qu'elle dépend de la distance  $d_{nm}$ . Une régression du modèle sur la cohérence mesurée permet d'obtenir pour chaque couple  $(n, m)$  une estimation de la distance. Nous notons  $\delta_{nm}$  cette distance. La forme matricielle  $\mathbf{\Delta}$  réunit l'ensemble des distances par paire telle que  $[\mathbf{\Delta}]_{nm} = \delta_{nm}$ . Elle est de taille  $M \times M$  et appartient à  $\mathbb{S}_h^M$ ;



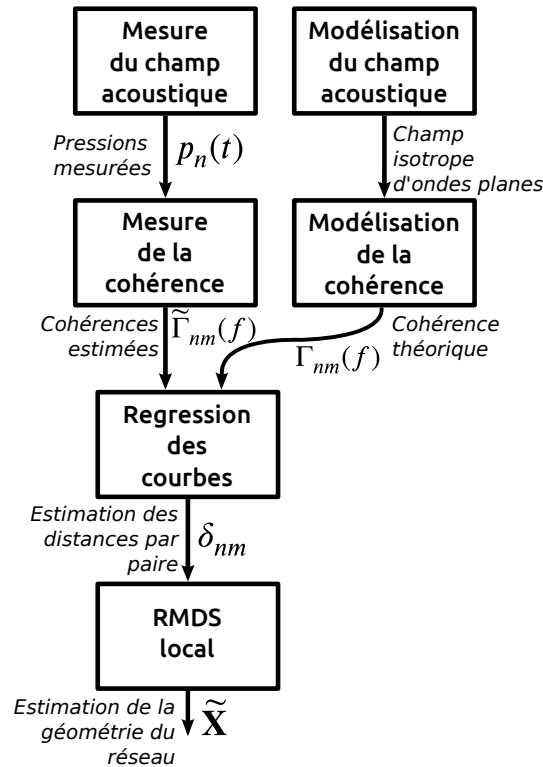


FIGURE 3.1 – Calibration géométrique en champ diffus : schéma de la méthode.

- La dernière étape consiste à tirer une estimée  $\tilde{\mathbf{X}}$  de la géométrie, à partir de la matrice de distances  $\mathbf{\Delta}$  issue des régressions. Cette étape fait appel à un algorithme de MDS.

L'ensemble du processus est également présenté dans le schéma de la figure 3.1.

### 3.3 Estimation des distances par paire en champ diffus

#### 3.3.1 La cohérence en champ diffus

L'interspectre (ou densité spectrale croisée de puissance) est une fonction d'analyse spectrale entre deux signaux  $p_n$  et  $p_m$ . Dans le cas d'un processus stationnaire, elle s'exprime par :

$$S_{nm}(f) = \mathbb{E} [\hat{p}_n(f) \hat{p}_m^*(f)] \quad (3.3)$$

où  $\hat{p}_n$  représente la transformée de Fourier discrète à court terme (TFCT) de  $p_n$ . L'interspectre d'un signal avec lui-même  $S_{nn}(f)$  est appelé autospectre (ou densité spectrale de puissance).

Par définition, la cohérence entre les signaux des microphones  $n$  et  $m$  est une fonction de la fréquence, à valeurs complexes telle que :

$$\Gamma_{nm}(f) = \frac{S_{nm}(f)}{\sqrt{S_{nn}(f)S_{mm}(f)}}, \quad (3.4)$$

La cohérence de deux signaux de microphones en milieu réverbérant a été étudiée à plusieurs reprises dans la littérature [40, 74, 108]. Le modèle de champ diffus couramment utilisé considère un champ isotrope d'ondes planes décorréliées en phase, à toutes les fréquences du spectre. Pour exprimer la pression temporelle mesurée par le microphone  $m$  à la fréquence  $f$ , le champ isotrope est représentée par l'intégrale sur l'angle solide d'ondes planes s'étant propagées jusqu'au microphone  $m$ . Ces ondes planes sont d'amplitude égales (pour conserver l'isotropie) et de phase  $\phi$  suivant une loi uniforme entre 0 et  $2\pi$ . La pression  $p_m(t)$  s'écrit alors :

$$p_m(t) = \iint_{4\pi} \cos(\mathbf{k}(\Theta_{m'}) \cdot \mathbf{x}_m - \omega t + \phi(\Theta_{m'})) d\Theta_{m'} \quad (3.5)$$

où  $\mathbf{k}$  est le vecteur d'onde, et  $\Theta_{m'}$  est l'angle polaire décrivant les directions de l'angle  $4\pi$ .

À partir de ce modèle, l'annexe A détaille la démonstration qui donne la cohérence analytique entre deux mesures microphoniques  $n$  et  $m$ . Elle montre que la cohérence est une fonction de la distance euclidienne entre les deux microphones  $n$  et  $m$ , noté  $d_{nm}$ . Elle s'exprime par :

$$\Gamma_{nm}(f) = \text{sinc}\left(\frac{2\pi f d_{nm}}{c_0}\right), \quad (3.6)$$

où  $c_0$  est la vitesse du son, constante dans le milieu. La partie imaginaire de la cohérence est donc *a priori* nulle dans le modèle de champ diffus considéré.

L'isotropie du champ acoustique est une hypothèse forte vis à vis de la qualité d'un champ diffus réel et des conditions expérimentales. D'autres modèles de champ diffus ont été étudiés. De façon générique, Blake *et al.* étudient une distribution d'ondes planes se propageant dans un angle solide fini [19]. L'expression analytique de la cohérence dépend alors de paramètres géométriques supplémentaires : l'ouverture et l'orientation de l'angle solide selon lequel se propagent les ondes planes. Dans [106] est étudiée la cohérence d'un champ diffus avec des microphones placés proches d'un mur : les ondes directes sont corrélées avec les ondes réfléchies, et en conséquence la cohérence est fonction de la distance des microphones par rapport au mur. Enfin, Kuster calcule analytiquement la cohérence entre les signaux enregistrés par des microphones directifs : l'expression analytique dépend de l'orientation relative des microphones dans l'espace [82].

De manière générale, il peut être utile d'obtenir une expression analytique alternative de la cohérence, car dans des conditions complexes le modèle en sinus cardinal n'est plus représentatif de la mesure. Cependant, elle sera fonction de paramètres géométriques supplémentaires, ce qui rend la régression des cohérences mesurées sur le modèle plus complexe. Dans l'étude présentée, nous nous écartons de tels cas. À l'isotropie du champ acoustique nous ajoutons les hypothèses

suivantes : les microphones sont éloignés des parois, et omnidirectionnels.

Dans [139], Taghizadeh *et al.* étudient en détail l'applicabilité du modèle de champ diffus dans le cadre de la calibration géométrique d'un réseau de microphones. Ils montrent que les conditions expérimentales idéales de mesure sont un milieu fortement réverbérant, contenant des sources larges bandes. L'énergie acoustique due à la propagation directe des sources vers les microphones doit être négligeable par rapport à celle due aux réflexions sur les murs [46]. Par la suite, le principal cadre expérimental de l'étude sera un grand hall réverbérant.

### 3.3.2 Estimation des distances par paire entre microphones

La cohérence issue des mesures microphoniques est calculée par l'estimateur proposé dans [139]. Notons  $\hat{p}_m^b(f_l)$  la TFCT de pression associée à la trame temporelle  $b$ , mesurée par le microphone  $m$ . Nous considérons l'estimateur suivant :

$$\tilde{\Gamma}_{nm}(f_l) = \frac{1}{B} \sum_{b=1}^B \frac{\hat{p}_n^b(f_l) \hat{p}_m^{b*}(f_l)}{|\hat{p}_n^b(f_l)| |\hat{p}_m^b(f_l)|}. \quad (3.7)$$

En effet, le moyennage des cohérences instantanées mesurées sur plusieurs trames améliore la régression pour estimer les distances par paire  $d_{nm}$  [139]. Les TFCT sont calculées avec un fenêtrage de Blackman, efficace pour limiter la fuite spectrale [1, 31]. Les trames temporelles ont un recouvrement de 50 % pour augmenter le moyennage. Notons que l'estimateur  $\tilde{\Gamma}_{nm}(f)$  est à valeurs complexes de part la nature de la TFCT. La partie imaginaire est généralement non-nulle (cf sec. 3.5.2), puisque le champ diffus réel n'est pas parfaitement isotrope. La régression du modèle est faite sur la partie réelle de  $\tilde{\Gamma}_{nm}(f)$ .

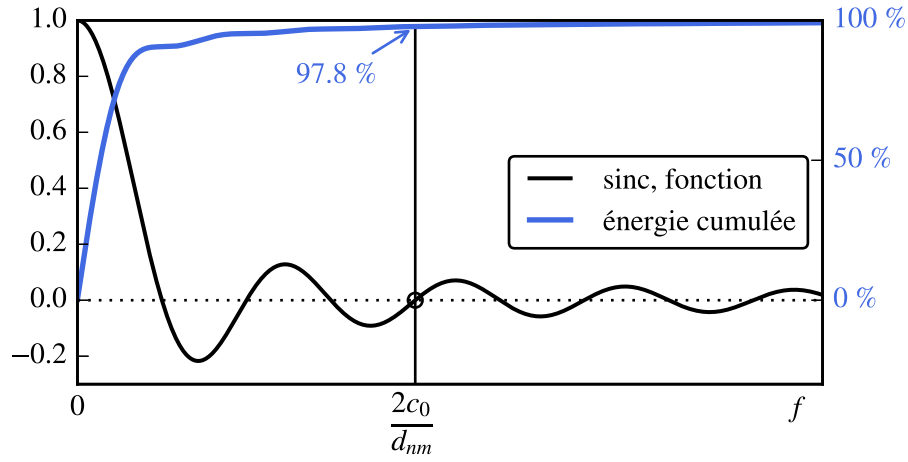
Notons  $\delta_{nm}$  la distance par paire estimée par régression ; celle-ci s'écrit comme une minimisation au sens des moindres carrés :

$$\delta_{nm} = \operatorname{argmin}_d \sum_{f_l} \left| \Re \left[ \tilde{\Gamma}_{nm}(f_l) \right] - \operatorname{sinc} \left( \frac{2\pi f_l d}{c_0} \right) \right|^2. \quad (3.8)$$

Comme la fonction sinus cardinal est non linéaire, le problème (3.8) est résoluble par une approche telle que la méthode de Nelder-Mead [165]. La résolution de (3.8) pour tous les couples  $(n, m)$  est effectuée  $M(M-1)/2$  fois, pour construire la matrice des distances estimées  $\mathbf{\Delta}$  dont les éléments sont  $\delta_{nm}$ .

### 3.3.3 Limitations de l'estimation des distances par régression

Parmi les différents aspects expérimentaux, la taille du réseau, la diffusivité acoustique (c'est à dire le caractère du champ acoustique à être plus ou moins proche du champ diffus) et le calcul



**FIGURE 3.2** – Fonction sinc : l'énergie cumulée de la fonction contient 97% de l'énergie totale (c'est à dire pour  $f \in [0, +\infty[$ ) entre  $f = 0$  et  $f = \frac{2c_0}{d_{nm}}$ .

de  $\tilde{\Gamma}_{nm}$  ont un impact sur l'étape de régression. Dans cette partie, nous définissons les limites expérimentales acceptables dans lesquelles  $\delta_{nm}$  peut être estimée.

Théoriquement, 97.8% de l'énergie du sinus cardinal est contenue dans les deux premières oscillations, c'est à dire entre l'origine et le 4-ème zéro de la fonction (voir figure 3.2). Cette portion influence particulièrement le résultat de la régression au sens des moindres carrés. Le 4-ème zéro de la fonction sinc correspond à l'égalité :

$$\frac{2\pi f d_{nm}}{c_0} = 4\pi. \quad (3.9)$$

L'allure de l'estimation  $\tilde{\Gamma}_{nm}$  est donc déterminante sur la régression au sens des moindres carrés, pour les fréquences se situant dans l'intervalle  $0 < f < \frac{2c_0}{d_{nm}}$ .

### 3.3.3.1 Limitations numériques

D'abord on définit les limites dues au calcul de  $\Gamma_{nm}$ . Les TFCTs, calculées sur des trames temporelles de durée  $T = (N - 1)/f_s$ , sont définies sur les fréquences entre 0 et  $f_s$ , au pas  $\Delta f = f_s/N$ . Le cas limite où les deux premières oscillations du sinc s'étendent entre 0 et  $f_s/2$  définit la plus petite distance accessible  $d_{nm}$ . C'est la distance satisfaisant l'égalité (3.9), pour  $f = f_s/2$  :

$$d_{\min} = \frac{4c_0}{f_s}. \quad (3.10)$$

Le cas limite opposé est fixé par la résolution fréquentielle : il faut un nombre minimum de valeurs de fréquence pour définir la forme de la courbe. De manière analogue à la limite de Shannon, on impose arbitrairement au moins huit valeurs de fréquence pour représenter les deux premières oscillations, ce qui donne :

$$d_{\max} = \frac{c_0 N}{4f_s}. \quad (3.11)$$

La taille globale du réseau pose une autre limite physique : le calcul de la cohérence suppose que toute paire de microphones mesure, en deux points de l'espace, le même ensemble d'ondes planes isotrope. Cependant pour garantir cette condition, la vitesse finie de propagation du son entre ces deux points doit être la moins apparente possible. En effet, rappelons que les TFCTs sont mesurées sur des longueurs de trames finies, de durée  $T$ . Donc, à la vitesse du son  $c_0$ , le temps de propagation d'une onde entre deux microphones doit être négligeable par rapport à la durée  $T$  de la trame sur laquelle est calculée chaque TFCT :

$$d_{nm}/c_0 \ll T. \quad (3.12)$$

Nous pouvons également formuler cette inégalité en fonction des paramètres du problème :

$$d_{nm} \ll \frac{Nc_0}{f_s}. \quad (3.13)$$

Cette inégalité est compatible avec la limite  $d_{\max}$  donnée dans (3.11). Elle sera expérimentalement validée en section 3.5.2, où nous verrons qu'augmenter les longueurs de trame améliore sensiblement l'estimation des cohérences.

### 3.3.3.2 Limitation physique : la diffusivité acoustique

Si le milieu de mesure est une salle réverbérante, la fréquence de Schroeder est un bon critère pour définir la diffusivité. Elle définit une fréquence de coupure, délimitant le caractère diffus du champ acoustique en basse fréquence. En effet, les modes propres excités par les sources acoustiques sont rares en basse fréquence ; cependant ils sont de plus en plus denses au fur et à mesure que la fréquence augmente. Au dessus de la fréquence de Schroeder, le nombre de modes devient suffisamment élevé pour créer un comportement chaotique par les réflexions. C'est en cela que le modèle de champ diffus suppose une distribution quasi uniforme des ondes planes. La fréquence de Schroeder est évaluée expérimentalement, en fonction du volume de la salle  $V$ , et du temps de réverbération  $T_{60}$  :

$$f_c = 2000 \sqrt{\frac{T_{60}}{V}} \quad (3.14)$$

De nouveau, cela définit une limite basse fréquence, imposant que la fréquence de Schroeder soit petite devant  $f = \frac{2c_0}{d_{nm}}$  donné par l'équation 3.9 (nous pouvons la choisir arbitrairement 8 fois plus petite). Cela donne la limite :

$$d_{\max} = \frac{c_0}{4f_c}. \quad (3.15)$$

Comme nous avons défini deux distances maximales (3.11) et (3.15), nous choisissons la valeur la plus contraignante. Au final,  $d_{nm}$  peut être *raisonnablement* retrouvé par régression dans l'intervalle :

$$d_{nm} \in \left[ \frac{4c_0}{f_s}, \min \left( \frac{c_0 N}{4f_s}; \frac{c_0}{4f_c} \right) \right]. \quad (3.16)$$

Notons qu'il faudrait davantage affiner les limites si la réponse en fréquence des microphones est restreinte. Nous écartons cette éventualité en considérant la bande passante des microphones MEMS suffisamment large pour ne pas influencer la mesure des cohérences pour  $d_{nm} \in [d_{\min}, d_{\max}]$ .

Avec une fréquence d'échantillonnage  $f_s = 50$  kHz, la borne inférieure vaut  $d_{\min} = 2.7$  cm, ce qui reste une contrainte négligeable. En pratique, la borne supérieure est avant tout donnée par la diffusivité acoustique, car la limitation numérique (éq. 3.11) est déterminé par le choix de la longueur de trame  $T$ . La borne  $d_{\max}$  dépend donc essentiellement de la fréquence de Schroeder de la salle. À titre d'exemple une étude a mesuré, pour 32 salles de classe différentes, des fréquences de Schroeder allant de 67 à 141 Hz [130]. Ces valeurs correspondent respectivement à un  $d_{\max}$  de 1.27 et 0.6 m. De tels ordres de grandeur montrent que la diffusivité est la contrainte majeure pour appliquer les régressions de cohérences avec des antennes de grande taille.

Les bornes (3.16) donnent un ordre de grandeur approximatif par rapport au dimensionnement de l'expérience. Dans l'étude expérimentale en section 3.5.2, nous verrons que la régression est particulièrement critique pour les grandes distances : elle peut être sujette à des fortes erreurs d'estimation. C'est la raison pour laquelle nous ne pouvons pas appliquer directement les solutions de l'état de l'art : la méthode de MDS choisie doit être expérimentalement robuste aux éventuels échecs de régression brisant la matrice  $\mathbf{\Delta}$ .

## 3.4 Estimation robuste de la géométrie du réseau

### 3.4.1 Le MDS pour la localisation de capteurs

#### 3.4.1.1 Approche par minimisation de la fonction *raw stress*

Nous souhaitons déterminer l'inconnue  $\mathbf{X}$  à partir de l'observable  $\mathbf{\Delta}$ . Nous pouvons construire ceci comme un problème inverse entre les distances  $d_{nm}$  issues de  $\mathbf{X}$ , et les distances mesurées  $\delta_{nm}$ . Pour cela nous prenons en compte conjointement :

- la connaissance *analytique* du problème, liant la donnée observable à la solution recherchée  $\mathbf{X}$  :  $d_{nm}$  est la distance euclidienne séparant les deux points  $n$  et  $m$  :

$$d_{nm}(\mathbf{X}) = \|\mathbf{x}_n - \mathbf{x}_m\| ; \quad (3.17)$$

- la partie *observée* du problème : l'ensemble des distances par paire "mesurées" notées  $\delta_{nm}$  (ou alors la matrice  $\Delta$ ). Pour représenter l'observation bruitée, le modèle le plus simple considère un bruit de type additif. En définissant  $\epsilon_{nm}$  comme une variable aléatoire suivant une loi de distribution centrée, le modèle de mesure écrit :

$$\delta_{nm} = d_{nm} + \epsilon_{nm} \quad (3.18)$$

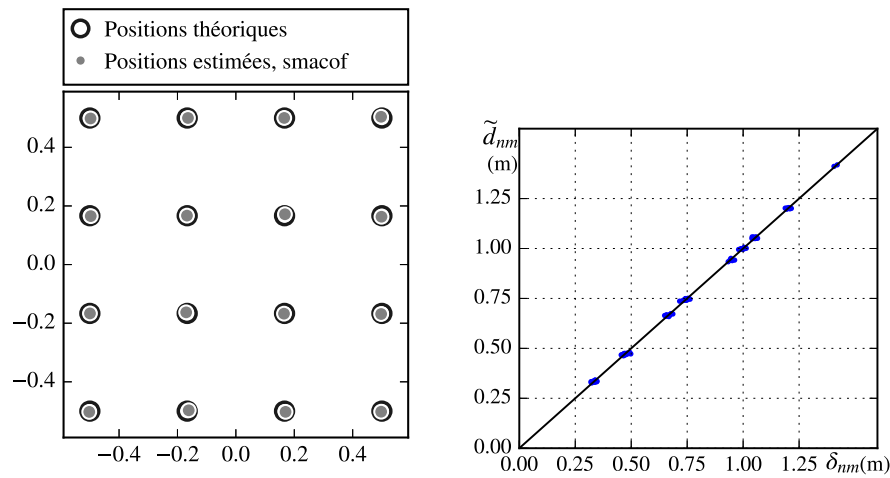
Dans le cas d'une approche MDS conventionnelle [42], ce problème inverse se traduit par la minimisation au sens des moindres carrés entre l'information analytique et l'information observée :

$$\begin{aligned} \tilde{\mathbf{X}} = \operatorname{argmin}_{\mathbf{X} \in \mathbb{R}^{D \times M}} \quad & f(\mathbf{X}) \\ \text{tel que} \quad & f(\mathbf{X}) = \sum_{n < m} [\delta_{nm} - d_{nm}(\mathbf{X})]^2 \end{aligned} \quad (3.19)$$

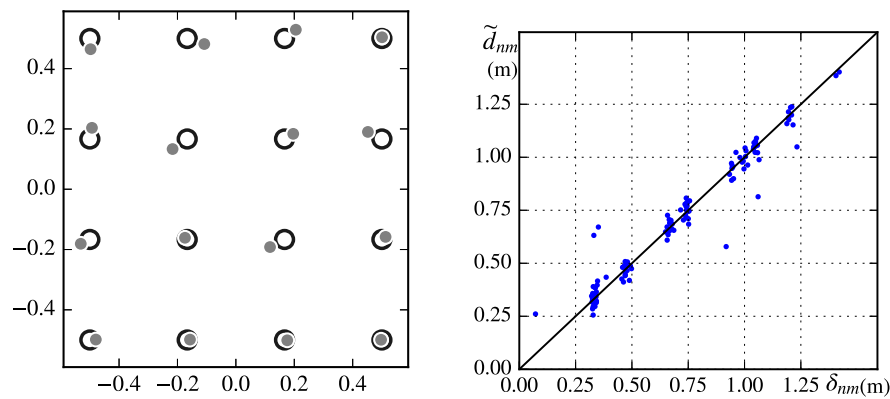
La fonction objective (également appelée fonction coût)  $f$  est appelée *raw stress* dans la littérature. Plusieurs méthodes de résolutions existent pour résoudre (3.19), comme par exemple l'algorithme itératif Smacof [45], fondé sur la transformée de Guttman (cf. section 3.4.4.2).

L'exemple, donné en figure 3.3, illustre le positionnement de 16 points, par la résolution du problème 3.19 avec l'algorithme Smacof. Dans le cas où les distances par paire contiennent un bruit de mesure  $\epsilon_{nm}$  relativement faible, cette approche est efficace pour reconstruire une géométrie (cf. partie gauche de la figure 3.3a). En effet, l'erreur de l'estimation géométrique, relativement à la géométrie théorique, vaut 0.76 %. Pour analyser la pertinence de la résolution par MDS, le diagramme de Shepard est la représentation la plus couramment utilisée. Tracé en partie droite de la figure 3.3a, il représente les points  $(\delta_{nm}, \tilde{d}_{nm})$ , c'est à dire les distances mesurées  $\delta_{nm}$  par rapport aux distances de la géométrie reconstruite  $\tilde{d}_{nm} = d_{nm}(\tilde{\mathbf{X}})$ . Une majorité de points proches de la droite  $\tilde{d}_{nm} = \delta_{nm}$  reflète une reconstruction géométrique pertinente par rapport aux distances mesurées (et vice-versa). Dans ce premier exemple, le diagramme de Shepard trace des points quasi superposés à la ligne, ce qui concorde bien avec l'estimation géométrique de très faible erreur.

La figure 3.3b représente le même exemple, mais 5 % de valeurs aberrantes sont ajoutés dans les distances mesurées  $\delta_{nm}$  (soit 6 valeurs aberrantes dans le cas de 16 positions). Elles contiennent en moyenne une erreur relative de 50 % par rapport aux valeurs théoriques. L'erreur géométrique d'estimation est alors de 7.2 %, ce qui est 10 fois supérieur au cas précédent sans valeur aberrante. La différence est également notable dans le diagramme de Shepard, où les points ne sont plus



(a) Distances  $\delta_{nm}$  avec un bruit tel que  $\text{std}(\epsilon_{nm}) = 0.01$  m, sans valeur aberrante.



(b) Distances  $\delta_{nm}$  avec un bruit tel que  $\text{std}(\epsilon_{nm}) = 0.01$  m, puis avec 6 valeurs aberrantes (5 %) contenant en moyenne 50 % d'erreur relative.

**FIGURE 3.3** – Un exemple de positionnement de 16 points par Smacof, avec ou sans valeurs aberrantes. Positions à gauche, diagramme de Shepard à droite.

superposés sur la droite  $\tilde{d}_{nm} = \delta_{nm}$ . Notons que les 6 points légèrement plus éloignés de cette droite correspondent aux 6 valeurs aberrantes rajoutées.

La minimisation au sens des moindres carrés est donc très sensible aux valeurs aberrantes. Le *raw stress*, composé de termes quadratiques, explique cette caractéristique. Si il y a une erreur de grande amplitude contenue dans les observables  $\delta_{nm}$ , elle aura une grande influence dans la valeur de la fonction coût.

### 3.4.1.2 Approches locales de MDS pour la localisation de capteurs

Les méthodes de MDS sont déjà souvent utilisées pour la localisation de capteurs dans un réseau sans fil, ou de transducteurs pour les systèmes de tomographie médicale [113]. L'envoi et la



réception de signaux par chacun des capteurs permet de mesurer les temps de vol, c'est à dire les distances entre les éléments. À cause de la portée limitée des signaux émis, il devient difficile de mesurer les distances entre deux capteurs éloignés. D'une part, la variance des distances mesurées augmente en fonction de la distance (ce qui augmente l'erreur), d'autre part si la portée du signal est trop limitée, alors la distance est inconnue.

L'approche adaptée à cette situation consiste à discriminer les distances *a priori* mauvaises, par rapport à celles des proches voisins. Une première façon consiste à pondérer la fonction stress (3.19) en fonction de chaque couple  $(n, m)$  :

$$\begin{aligned} \tilde{\mathbf{X}} &= \underset{\mathbf{X} \in \mathbb{R}^{D \times M}}{\operatorname{argmin}} f_{\mathbf{W}}(\mathbf{X}) \\ \text{tel que } f_{\mathbf{W}}(\mathbf{X}) &= \sum_{n < m} w_{nm} [\delta_{nm} - d_{nm}(\mathbf{X})]^2 \end{aligned} \quad (3.20)$$

où  $w_{nm}$  sont des poids positifs. La résolution par algorithme Smacof pondéré donne de bons résultats si le bruit de mesure est quantifiable par un modèle, ce qui permet un choix optimal des coefficients  $w_{nm}$ . Dans ce cas l'erreur de localisation tend vers la borne de Cramer-Rao [41]. Il est également possible d'appliquer la pondération par 1 ou 0, selon si la distance est connue ou inconnue.

Il existe un ensemble de MDS locaux fondés uniquement sur l'apprentissage des connectivités voisines [155]. Deux exemples connus sont Isomap [145] et Locally Linear Embedding (LLE) [122], typiquement utilisés pour positionner des *objets* dont les similarités (nom donné, dans la littérature, aux distances par paire entre les objets) sont globalement non-linéaires, mais qui approximativement sont localement linéaires. La clé de ces méthodes d'apprentissage est de s'appuyer uniquement sur les distances locales, en faisant une hypothèse d'approximation linéaire sur les *petites* similarités, ou *petites* distances. Ces méthodes ont montré de très bons résultats pour la localisation d'éléments [89, 114], et dans le cas de réseaux massif et étendus [157, 170].

Les approches *locales* semblent *a priori* adaptées à notre problème de calibration géométrique, puisque l'estimation des distances est bornée au-delà d'une limite  $d_{\max}$  (eq. (3.16)). Cependant, nous verrons dans la partie expérimentale qu'il n'est pas suffisant d'uniquement discriminer les grandes distances pour obtenir une estimation robuste de la géométrie  $\tilde{\mathbf{X}}$ . De plus, nous ne pouvons pas quantifier de modèle précis sur le bruit de mesure des  $\delta_{nm}$  entre microphones. En effet, nous verrons en partie 3.5.2 qu'il n'est pas possible d'établir un modèle d'erreur pertinent : les valeurs erronées sont des valeurs aberrantes issues d'une mauvaise régression (eq. (3.8)). À titre de comparaison, nous présenterons les résultats de calibration géométrique par deux approches de MDS locales dans la partie expérimentale 3.6.2.2.

### 3.4.2 Le RMDS garantit la robustesse de l'estimation

Forero *et al.* [59] proposent une méthode de MDS robuste aux valeurs aberrantes : le Robust MDS (RMDS). Il s'appuie sur un nouveau modèle de mesure, contenant explicitement un terme pour décrire la présence de valeurs aberrantes :

$$\delta_{nm} = d_{nm} + \epsilon_{nm} + o_{nm} \quad (3.21)$$

la mesure de  $d_{nm}$  contient deux bruits de natures différentes.  $\epsilon_{nm}$  est une variable aléatoire suivant une loi de distribution centrée, son ordre de grandeur est petit devant la valeur de  $d_{nm}$ . Le terme  $o_{nm}$  décrit l'erreur engendrée par une valeur aberrante, d'ordre de grandeur équivalent à  $d_{nm}$ . Notons  $\mathbf{O}$  la forme matricielle des termes  $o_{nm}$ .

Dans ce cas, la présence du terme  $o_{nm}$  dans le modèle de mesure permet d'écarter explicitement l'erreur d'ordre de grandeur élevé de l'estimation  $\tilde{d}_{nm}$ , donc de  $\tilde{\mathbf{X}}$ . Ceci pose donc un problème à deux inconnues matricielles  $(\mathbf{O}, \mathbf{X})$ . Forero *et al.* montrent que ce problème est résoluble s'il existe une condition de parcimonie sur le nombre de valeurs aberrantes, c'est à dire s'il y a peu d'éléments non-nuls dans la matrice  $\mathbf{O}$ . Le problème global est alors traduit par la minimisation de la fonction coût suivante :

$$\begin{aligned} (\tilde{\mathbf{O}}, \tilde{\mathbf{X}}) = & \underset{\mathbf{O} \in \mathbb{S}_h^M, \mathbf{X} \in \mathbb{R}^{D \times M}}{\operatorname{argmin}} f(\mathbf{O}, \mathbf{X}) \\ \text{tel que} & f(\mathbf{O}, \mathbf{X}) = \sum_{n < m} [\delta_{nm} - d_{nm}(\mathbf{X}) - o_{nm}]^2 + \nu \sum_{n < m} |o_{nm}| \end{aligned} \quad (3.22)$$

Il s'agit d'une formulation classique de problème parcimonieux sur la matrice  $\mathbf{O}$ , appelé LASSO [28].  $f$  est l'association de la norme  $\ell_2$ , puis de la norme  $\ell_1$  régularisée par le terme de pondération  $\nu$ , sur l'inconnue  $\mathbf{O}$  :

- le premier terme de somme est l'erreur quadratique entre les mesures  $\delta_{nm}$ , et la distance  $d_{nm}(\mathbf{X})$  associée à l'inconnue  $\mathbf{X}$ . Ce terme est semblable au *raw stress* (3.19), mais en soustrayant la valeur  $o_{nm}$ , l'erreur quadratique devient exempte de valeur aberrante. Ce point constitue la clé de la méthode RMDS;
- le deuxième terme de somme équivaut à la norme  $\ell_1$  sur la matrice  $\mathbf{O}$ . Le terme de régularisation  $\nu$  conditionne la parcimonie de  $\mathbf{O}$ . Si on augmente la valeur de  $\nu$ , on augmente la contrainte de parcimonie de l'estimée  $\tilde{\mathbf{O}}$ , c'est à dire  $\tilde{\mathbf{O}}$  contient plus d'éléments nuls. Le choix de  $\nu$  a donc une influence sur la solution finale  $(\tilde{\mathbf{O}}, \tilde{\mathbf{X}})$ .

L'approche proposée par Forero *et al.* est donc intéressante pour retrouver de manière optimale la géométrie du réseau, en supprimant les grandes erreurs dans  $\Delta$ .

### 3.4.3 Le LRMS, une approche locale du RMS adapté aux grands réseaux

Nous avons vu dans la partie précédente que les distances estimées par régression sont *a priori* bornées dans un intervalle  $d \in [d_{\min}, d_{\max}]$ . Ces limites seront analysées expérimentalement en section 3.5.2 : nous verrons empiriquement que la matrice  $\Delta$  obtenue est caractérisée par :

- des valeurs aberrantes en quantité relativement faible en dessous d'un certain seuil, que l'on note  $\delta_{\max}$  ;
- des valeurs aberrantes nombreuses, sans transition, au delà de  $\delta_{\max}$ .

Or pour garantir la condition de parcimonie sur  $\mathbf{O}$ , il est essentiel que le nombre de valeurs aberrantes dans  $\Delta$  soit faible. Ceci peut être assuré en rejetant les  $\delta_{nm}$  supérieures au seuil  $\delta_{\max}$ .

Nous proposons de pondérer la fonction coût originale du RMS, ce qui correspond au problème :

$$\begin{aligned}
 (\tilde{\mathbf{O}}, \tilde{\mathbf{X}}) &= \operatorname{argmin}_{\mathbf{O} \in \mathbb{S}_h^M, \mathbf{X} \in \mathbb{R}^{D \times M}} f_{\mathbf{W}}(\mathbf{O}, \mathbf{X}) \\
 \text{tel que} \quad f_{\mathbf{W}}(\mathbf{O}, \mathbf{X}) &= \sum_{n < m} w_{nm} [\delta_{nm} - d_{nm}(\mathbf{X}) - o_{nm}]^2 + \nu \sum_{n < m} w_{nm} |o_{nm}|
 \end{aligned} \tag{3.23}$$

Nous notons  $\mathbf{W}$  la matrice  $M \times M$  définie par les termes de pondération  $w_{nm}$ .

Il est possible de montrer que le problème pondéré du RMS suit la même méthode de résolution que le RMS classique, si les pondérations sont positives. Les détails sont donnés en annexe B. Par conséquent, pour rejeter les valeurs de  $\delta_{nm} \geq \delta_{\max}$ , la stratégie de pondération la plus simple est binaire :

$$\begin{aligned}
 w_{nm} &= 1 \quad \text{si } \delta_{nm} < \delta_{\max} \\
 &= 0 \quad \text{sinon}
 \end{aligned} \tag{3.24}$$

Par la suite, nous appellerons Local RMS (LRMS) le problème (3.23) suivant la condition (3.24).

### 3.4.4 Algorithme de résolution

La fonction objective  $f : (\mathbb{S}_h^M, \mathbb{R}^{M \times D}) \rightarrow \mathbb{R}$  ne permet pas d'exprimer la solution sous forme analytique. Il faut donc mettre en œuvre un algorithme pour obtenir une estimée, minimisant au mieux la fonction coût  $f$ . Deux aspects sont à prendre en compte pour l'approche à choisir :

- pour un réseau à un nombre d'éléments  $M$  grand, la taille du problème  $(\mathbb{S}_h^M, \mathbb{R}^{M \times D})$  augmente rapidement ;
- la fonction coût  $f$  est non-convexe et non-différentiable ; ce qui écarte toutes les méthodes par descente de gradient.

Forero *et al.* proposent d'utiliser l'approche de Majoration-Minimisation (MM). Elle s'appuie sur la caractérisation d'une approximation majorante de la fonction objective  $f$ , avec de bonnes

propriétés de convergence. L'algorithme MM est itératif. À chaque itération  $k$ , une approximation majorante est définie en fonction de l'état actuel du couple de solutions  $(\mathbf{O}^{(k)}, \mathbf{X}^{(k)})$ ; elle est minimisée pour donner un nouveau couple  $(\mathbf{O}^{(k+1)}, \mathbf{X}^{(k+1)})$ . Un critère de convergence détermine l'arrêt de l'algorithme itératif.

Une itération de l'algorithme contient deux étapes successives : la mise à jour de  $\mathbf{O}$ , puis de  $\mathbf{X}$ .

#### 3.4.4.1 Mise à jour de $\mathbf{O}$

La première étape est la résolution du problème parcimonieux : la fonction  $f$  correspond à la somme de  $M(M + 1)/2$  sous-problèmes de LASSO séparables sur les variables  $o_{nm}$ , qui sont :

$$\operatorname{argmin}_{o_{nm} \in \mathbb{R}} w_{nm} [\delta_{nm} - d_{nm}(\mathbf{X}) - o_{nm}]^2 + \nu w_{nm} |o_{nm}| \quad (3.25)$$

pour  $w_{nm} \neq 0$ . Leur solution s'exprime en fonction de l'opérateur de seuillage doux (*soft-thresholding* en anglais) :

$$o_{nm}^{(k+1)} = S_\nu (\delta_{nm} - d_{nm}(\mathbf{X}^{(k)})) \quad (3.26)$$

avec  $S_\nu$  dépendant du paramètre de régularisation :

$$S_\nu(u) = \begin{cases} u - \nu/2 & \text{si } u > \nu/2 \\ 0 & \text{si } u \in [-\nu/2; \nu/2] \\ -u + \nu/2 & \text{si } u < -\nu/2 \end{cases} \quad (3.27)$$

Le seuillage doux consiste à forcer les valeurs de  $\mathbf{O}^{(k+1)}$  à zéro si leur valeur absolue est inférieure à  $\nu/2$ . Nous retrouvons donc l'effet du paramètre de régularisation : plus  $\nu$  est grand, plus il contraint un grand nombre de valeurs de  $\mathbf{O}^{(k+1)}$  à zéro.

#### 3.4.4.2 Mise à jour de $\mathbf{X}$

Indépendante de la mise à jour  $\mathbf{O}^{(k+1)}$ , la solution de  $\mathbf{X}^{(k+1)}$  revient à la minimisation quadratique de l'approximation majorante utilisée par la méthode MM. Démontrée dans [58], elle s'écrit :

$$\mathbf{X}^{(k+1)} = \mathbf{X}^{(k)} \mathbf{L}_1(\mathbf{O}^{(k+1)}), \mathbf{X}^{(k)} \mathbf{L}^\dagger \quad (3.28)$$

Cette relation est très proche de la transformée de Guttman utilisée pour la mise à jour de  $\mathbf{X}$  dans l'algorithme Smacof. Il est montré en annexe B que l'ajout de termes de pondération  $w_{nm}$  est possible, à condition qu'ils soient positifs. Cela modifie l'écriture de la matrice  $\mathbf{L}_1(\mathbf{O}, \mathbf{X})$  par :

$$\mathbf{L}_1(\mathbf{O}, \mathbf{X}) = \text{diag}(\mathbf{A}_1(\mathbf{O}, \mathbf{X})\mathbf{1}_M) - \mathbf{A}_1(\mathbf{O}, \mathbf{X}) \quad (3.29)$$

avec  $\mathbf{1}_M$  le vecteur colonnes à  $M$  éléments égaux à 1, et  $\text{diag}(v)$  l'opérateur donnant la matrice carrée dont la diagonale contient les éléments du vecteur  $v$ . La matrice  $\mathbf{A}_1(\mathbf{O}, \mathbf{X})$ , de dimension  $M \times M$ , contient les éléments :

$$\begin{aligned} [\mathbf{A}_1(\mathbf{O}, \mathbf{X})]_{nm} &= w_{nm} \frac{\delta_{nm} - o_{nm}}{d_{nm}(\mathbf{X})} \quad \text{si } \delta_{nm} > o_{nm} \text{ et } d_{nm}(\mathbf{X}) \neq 0 \\ &= 0 \quad \text{sinon} \end{aligned} \quad (3.30)$$

Enfin  $\mathbf{L}^\dagger$  est la matrice pseudo-inverse de Moore-Penrose de la matrice  $\mathbf{L}$  telle que :

$$\mathbf{L} = \sum_{n < m} w_{nm} (\mathbf{e}_n - \mathbf{e}_m)^\top (\mathbf{e}_n - \mathbf{e}_m) \quad (3.31)$$

avec  $\mathbf{e}_n$  le vecteur ligne à  $M$  éléments, défini de telle manière que  $[\mathbf{e}_n]_m = 1$  si  $m = n$ , et 0 ailleurs.

### 3.4.4.3 Initialisation et convergence

Le LRMS est une approche itérative qui garantit la convergence vers un minimum local. Le choix d'une initialisation aléatoire peut suffire pour atteindre le minimum global, mais avec un taux de succès trop faible pour assurer la robustesse. Il est préférable de choisir une initialisation  $\mathbf{X}^{(0)}$  la plus proche possible du résultat final souhaité. Empiriquement, j'ai obtenu de bons résultats en initialisant avec une estimation approchée par MDS local (par exemple Isomap ou Smacof avec pondération locale).

Pour stopper l'algorithme, nous choisissons un critère qui quantifie la convergence de la mise à jour  $\mathbf{X}^{(k)}$ . L'algorithme itère la mise à jour tant que la convergence n'a pas atteint un certain seuil. Le critère de convergence adimensionné  $\|\mathbf{X}^{(k+1)} - \mathbf{X}^{(k)}\|_F / \|\mathbf{X}^{(k)}\|_F$ <sup>1</sup> stoppe l'algorithme si sa valeur est en dessous d'un seuil  $\kappa$  arbitrairement choisi;  $10^{-5}$  est un choix empiriquement satisfaisant pour la suite. Le dernier couple mis à jour  $(\mathbf{O}^{(k+1)}, \mathbf{X}^{(k+1)})$  donne l'estimée finale  $(\tilde{\mathbf{O}}, \tilde{\mathbf{X}})$ .

L'implémentation du LRMS est résumée dans l'algorithme 1. Les trois points clés de la méthode sont :

1. dans la matrice de distances estimée  $\mathbf{\Delta}$ , on garde uniquement les valeurs inférieures à une portée limite  $\delta_{\max}$  par l'intermédiaire de la matrice de pondération  $\mathbf{W}$ ;
2. sur ces données, le LRMS résout le problème de calibration géométrique, et conjointement détermine les valeurs aberrantes;

---

1.  $\|\cdot\|_F$  est la norme de Frobenius. Pour une matrice  $\mathbf{A}$ , elle est calculée par la relation  $\|\mathbf{A}\|_F = (\sum_{n,m} [\mathbf{A}]_{nm}^2)^{1/2}$ .

3.  $\tilde{\mathbf{X}}$  est donc la géométrie qui correspond *au mieux*<sup>2</sup> aux distances  $\delta_{nm} < \delta_{\max}$ , sauf celles issues d'une mauvaise régression sur les cohérences.

---

**ALGORITHME 1** RMDS proposé par Forero *et al.* [59].

---

**Entrées :** Initialiser  $\mathbf{O}^{(0)}$  à zéro, puis  $\mathbf{X}^{(0)}$

**for**  $k = 0, 1, 2, \dots$  **do**

Calculer  $\mathbf{O}^{(k+1)}$  via (3.26)

Calculer  $\mathbf{X}^{(k+1)}$  à jour via (3.28)-(3.31)

**if**  $\frac{\|\mathbf{X}^{(k+1)} - \mathbf{X}^{(k)}\|_F}{\|\mathbf{X}^{(k)}\|_F} < \kappa$  **then**

Arrêt de l'algorithme

**end if**

**end for**

**Sorties :**  $(\tilde{\mathbf{O}}, \tilde{\mathbf{X}}) = (\mathbf{O}^{(k+1)}, \mathbf{X}^{(k+1)})$

---

## 3.5 La réalisation expérimentale

Les motivations des expériences présentées dans cette partie sont doubles. Premièrement, nous ferons une analyse des cohérences mesurées  $\tilde{\Gamma}_{nm}$  et des distances obtenues par régression  $\mathbf{\Delta}$ ; cela nous permettra de voir les limitations énoncées dans la partie 3.4.3, et justifiera l'intérêt du LRMDS dans notre problème. Deuxièmement, nous illustrerons par une étude de cas le choix du LRMDS dans la méthode proposée. Nous discuterons de la démarche pratique à prendre, notamment sur le choix de paramètres les plus appropriés.

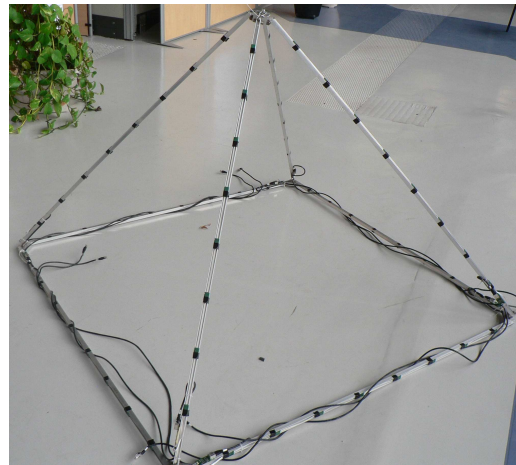
### 3.5.1 Scène et dispositif expérimentaux

Dans cette partie, les scénarios expérimentaux de validation se sont déroulés dans un grand hall de dimensions 10 m  $\times$  8 m  $\times$  20 m. C'est une salle où se déroulent différentes expériences de recherche ou de travaux pratiques. Ces activités sont la principale source du champ acoustique mesuré par le réseau de microphones. À partir de la mesure du temps de réverbération, nous estimons la fréquence de Schroeder à  $f_c = 77$  Hz. Le champ acoustique est issu de sources *naturelles* : dialogues, bancs d'appareils expérimentaux, ventilation *etc.* En effet, contrairement aux calibrations géométriques par temps de vol, il n'est pas nécessaire d'utiliser de sources contrôlées s'il existe déjà un environnement sonore large bande.

Le système d'acquisition *Megamicros*, décrit dans le chapitre 1, a été choisi pour toutes les expériences suivantes. Les signaux issus du réseau de microphones sont échantillonnés à  $f_s = 50$  kHz. Deux géométries d'antenne sont mises en œuvre :

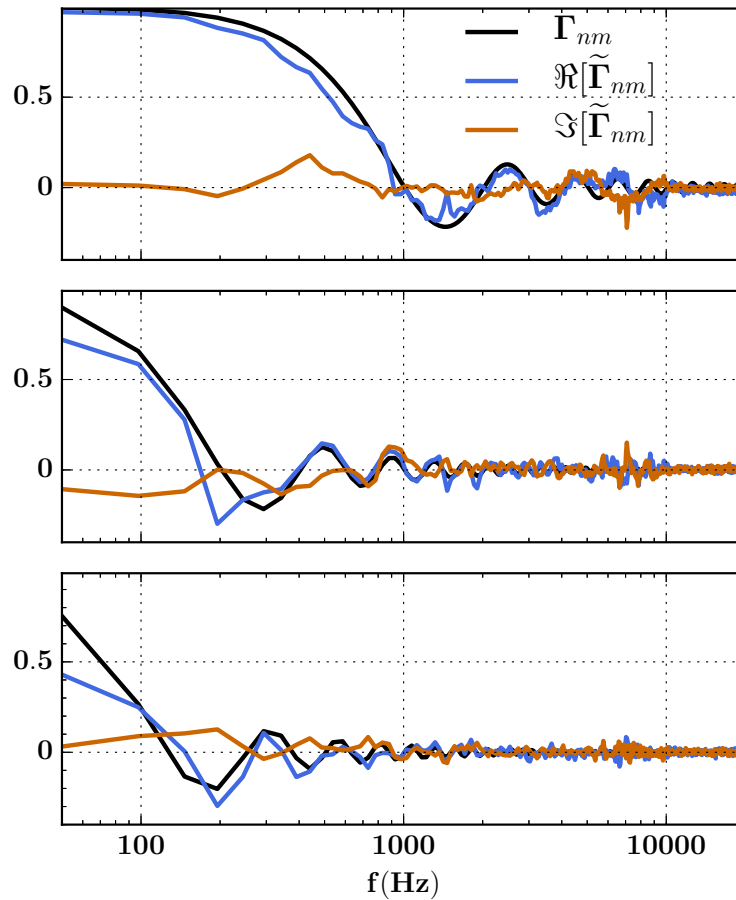
---

2. dans le sens des moindres carrés

(a) *Soleil* : réseau circulaire.(b) *Pyramide* : réseau pyramidal.(c) *Pyramide* : l'assemblage est constitué de pièces fraisées en atelier. La précision du montage est supposé millimétrique.**FIGURE 3.4** – Réseaux de microphones de grande taille *in situ* dans le hall réverbérant.

*Soleil* : une antenne plane à  $M = 128$  microphones. Elle est constituée de 16 barres de 8 éléments régulièrement espacés de 17 cm. Les barres constituent les rayons d'un disque d'environ 2.5 m de diamètre. À la fois large en dimension et en nombre d'éléments, ce réseau constitue quantitativement un cas de validation intéressant : 8128 distances par paire, dont les valeurs s'étendent sur l'intervalle  $d_{nm} \in [2.9 \text{ cm}; 2.78 \text{ m}]$ . Cette géométrie plane ( $\rightarrow D = 2$ ) est couramment utilisée en imagerie acoustique. La position réelle des éléments a été mesurée avec une précision d'un demi-millimètre ;

*Pyramide* : une antenne pyramidale à base carrée à 64 microphones. Les arêtes sont des barres longues de 1.5 m, constituées de 8 microphones régulièrement espacés de 17 cm. Comme il est difficile de mesurer des positions 3D à la main, la géométrie choisie est régulière ; nous estimons une précision millimétrique du montage. Cette antenne ne présente *a priori* pas un intérêt particulier pour l'imagerie acoustique. Elle est fabriquée spécifiquement pour la validation expérimentale de la méthode LRMS. Le choix de la forme pyramidale sert avant tout à obtenir la *vérité-terrain* de la géométrie au millimètre.



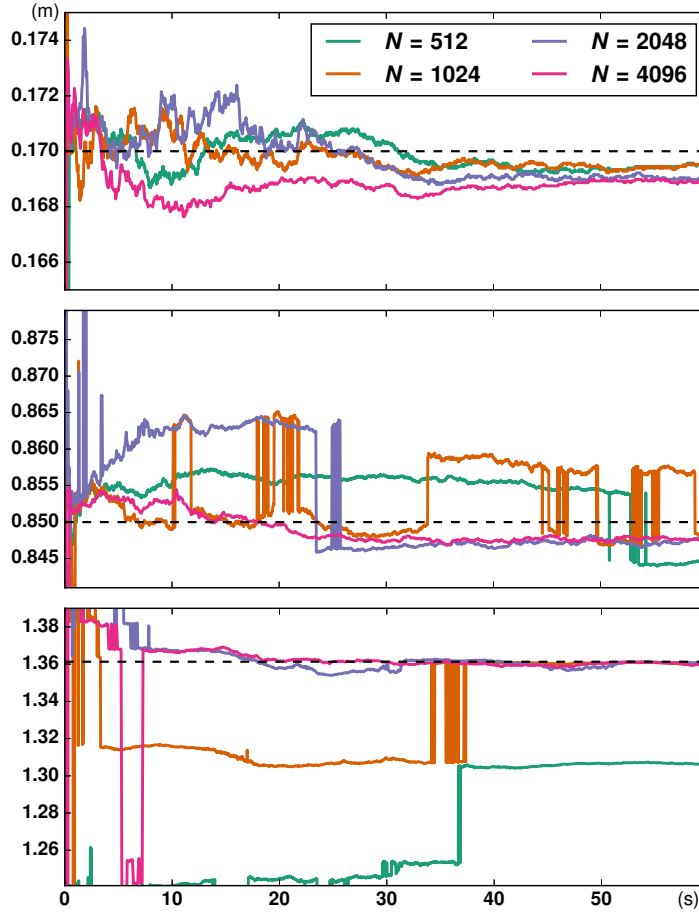
**FIGURE 3.5** – Estimation des cohérences : trois cas d'exemple, pour  $d_{nm} = 17, 85$  et  $136$  cm. La cohérence théorique est comparée avec les cohérences mesurées sur une minute. La partie imaginaire de la cohérence reste égale à  $0 \pm 0.2$ .

### 3.5.2 Estimation des cohérences et des distances

Des exemples de cohérences estimées sont tracés en figure 3.5, pour trois paires de microphones différents. La partie réelle est semblable au sinus cardinal, et plus particulièrement pour une petite distance. Le nombre de points définissant la forme de la courbe diminue rapidement quand la distance augmente, comme énoncé en partie 3.4.3. La partie imaginaire fluctue dans un intervalle de  $\pm 0.1$  ; c'est une amplitude non-négligeable compte tenu de la décroissance rapide du sinus cardinal. Il y a donc un intérêt à conserver uniquement la partie réelle dans la régression (3.8).

L'estimation des cohérences dépend de deux paramètres : la durée totale de l'acquisition, proportionnelle au nombre de trames moyennées  $B$ , puis la longueur de trame définie sur  $N$  échantillons. La valeur de  $\delta_{nm}$  dépend de ces paramètres comme le prouve la figure 3.6. La tendance énoncée en partie 3.4.3 sur le choix de  $N$  est illustrée par les courbes : si  $N$  est trop petit pour  $d_{nm}$  grand, la régression se dégrade car elle est sensible aux minimums locaux.





**FIGURE 3.6** – L’erreur de régression sur les cohérences dépend du calcul de l’estimateur  $\tilde{\Gamma}_{nm}$ . Ces courbes analysent l’effet issu (1) de la durée totale d’acquisition pour construire  $\tilde{\Gamma}_{nm}$ , (2) de la taille de la trame temporelle  $N$ . À  $f_s = 50$  kHz,  $N = (512, 1024, 2048, 4096)$  correspond aux durées de trame (10.2, 20.5, 41, 82) ms. De haut en bas,  $d_{nm} = 17, 85$  et  $136$  cm. Les courbes donnent la valeur de l’estimation  $\delta_{nm}$  en fonction de la durée d’acquisition. Valeur théorique  $d_{nm}$  en noir pointillé.

Le résultat des régressions se stabilise suite à environ 40 s d’acquisition. En augmentant la taille de trame, on réduit le nombre de trames moyennées  $B$  sur une même durée. Pourtant les figures ne mettent pas en évidence une convergence plus rapide de la régression en augmentant le moyennage. Ce point écarte donc un compromis sur le choix de  $N$ . Par la suite nous choisirons  $N = 2048$ , et nous gardons une durée totale d’acquisition égale à 1 min.

Pour une analyse globale de l’ensemble des régressions, les figures 3.7a et 3.8a présentent la répartition des couples  $(\delta_{nm}, d_{nm})$  sous forme d’un histogramme 2D. Les points sont regroupés et cumulés dans des rectangles élémentaires de  $3 \text{ cm} \times 3 \text{ cm}$ . Cette représentation traduit l’erreur  $|\delta_{nm} - d_{nm}|$  par la distance verticale ou horizontale à la droite  $d_{nm} = \delta_{nm}$ . Sur les deux figures, les résultats révèlent des comportements analogues :

- Les rectangles collectant le plus grand nombre se situent dans l’intervalle  $|\delta_{nm} - d_{nm}| < 2 \text{ cm}$ ,

- représentant 77.9 % de l'ensemble des distances. La plupart des distances sont donc estimées avec une erreur relativement faible. La variabilité de précision à  $\pm 2$  cm peut être associée au bruit  $\epsilon_{nm}$  ;
- le nombre de valeurs aberrantes dépend de l'ordre de grandeur de  $\delta_{nm}$ . Il augmente quand  $\delta_{nm}$  augmente, et ce de manière significative pour  $\delta_{nm} > 1.5$  m ;
  - le nombre de valeurs aberrantes dépend également de  $d_{nm}$  : il augmente aussi quand  $d_{nm}$  augmente. En effet, sur les 22.1 % de tous les couples à  $|\delta_{nm} - d_{nm}| > 2$  cm, seuls 8.4 % se situent à  $d_{nm} < 1$  m.

Ces observations coïncident avec les affirmations de la partie 3.4.3. La régression sur la cohérence entre les signaux des microphones éloignés est critique, car elle engendre des valeurs fortement erronées. Dans la matrice  $\Delta$ , ces cas critiques restent globalement situés au-delà d'un seuil  $\delta_{\max}$ , mais ne sont pas identifiables de manière précise. Ces résultats valident l'intérêt d'une recherche non supervisée des valeurs aberrantes, par le LRMDs. L'approche *locale* est indispensable pour garantir la parcimonie des valeurs aberrantes, en écartant des valeurs supérieures à une portée  $\delta_{\max}$  dans le processus de minimisation (3.23).

### 3.5.3 Le LRMDs en pratique : la configuration paramétrique

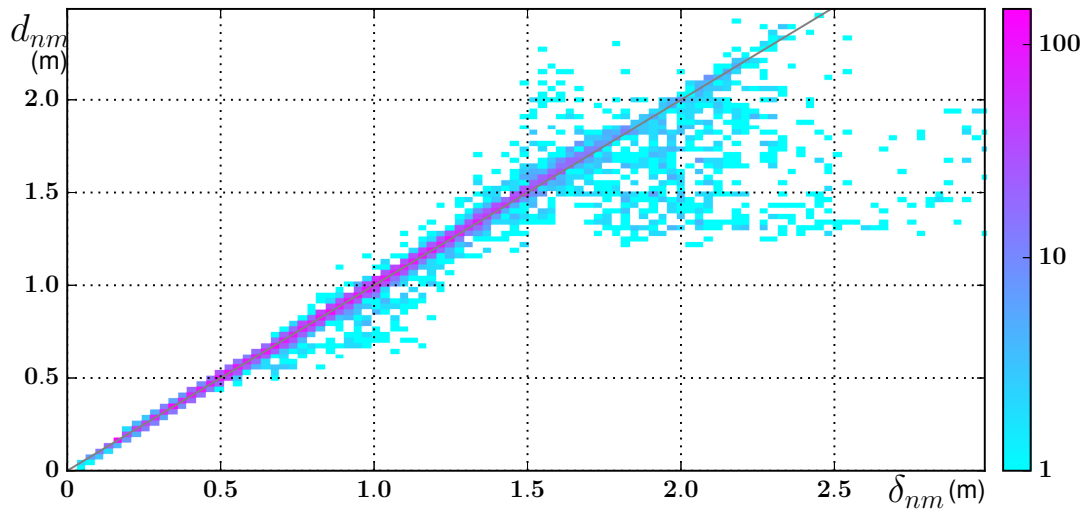
L'aspect fonctionnel du LRMDs peut être décrit par : les données d'entrée  $\Delta$ , la donnée de sortie  $\tilde{\mathbf{X}}$ , puis trois paramètres configurables  $D$ ,  $\nu$  et  $\delta_{\max}$ . La dimension spatiale du réseau  $D$  est ici supposée connue *a priori*. Si la calibration géométrique est faite à l'aveugle, il faut choisir  $D = 3$ .

Pour évaluer l'effet des deux autres paramètres sur le résultat final, nous proposons une représentation ne nécessitant pas la connaissance de la vérité-terrain. Pour cela nous utilisons un diagramme de Shepard [20]. Il trace le nuage de points  $(\delta_{nm}, \tilde{d}_{nm})$  issu d'un algorithme MDS. Si la matrice de distances  $\Delta$  n'est pas bruitée ( $\epsilon_{nm} = o_{nm} = 0 \forall (n, m)$ ), une reconstruction idéale par le MDS aboutit à un nuage de point aligné sur la droite  $\tilde{d}_{nm} = \delta_{nm}$ . Un point éloigné de cette droite traduit une distorsion entre les distances mesurées et les celles issues de la reconstruction par MDS. En réalité, les deux raisons possibles sont :

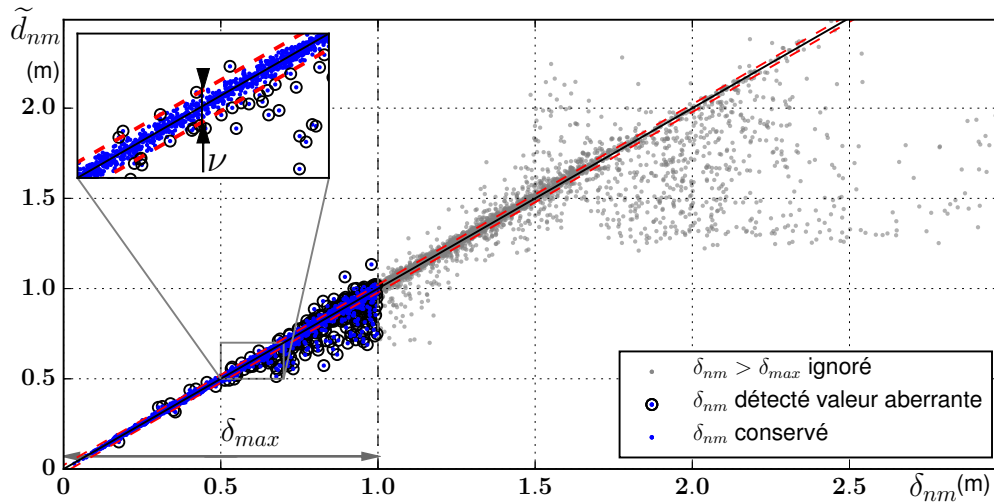
- une mauvaise reconstruction par MDS ; cela se produit quand l'algorithme choisi n'est pas entièrement adapté au problème ;
- la donnée d'entrée  $\Delta$  est bruitée.

Le deuxième point est la clé de la détection de valeurs aberrantes : à chaque itération, le LRMDs identifie les  $\delta_{nm}$  situés à une distance supérieure à  $\nu/2$  de la droite  $\tilde{d}_{nm} = \delta_{nm}$  (points cerclés dans les figures 3.7b et 3.8b). La donnée conservée pour trouver  $\tilde{\mathbf{X}}$  se situe entre les deux droites en rouge pointillé :  $\tilde{d}_{nm} - \nu/2 < \delta_{nm} < \tilde{d}_{nm} + \nu/2$ . On observe ici l'effet de l'opérateur de seuillage doux présenté en éq. (3.27) :  $\nu$  est un paramètre de seuil, en mètres, distinguant l'amplitude des deux types de bruits du modèle (3.21) :  $o_{mn}$  et  $\epsilon_{mn}$ .

Donc, pour faire un choix optimal de  $\nu$ , il faudrait avoir une connaissance *a priori* :

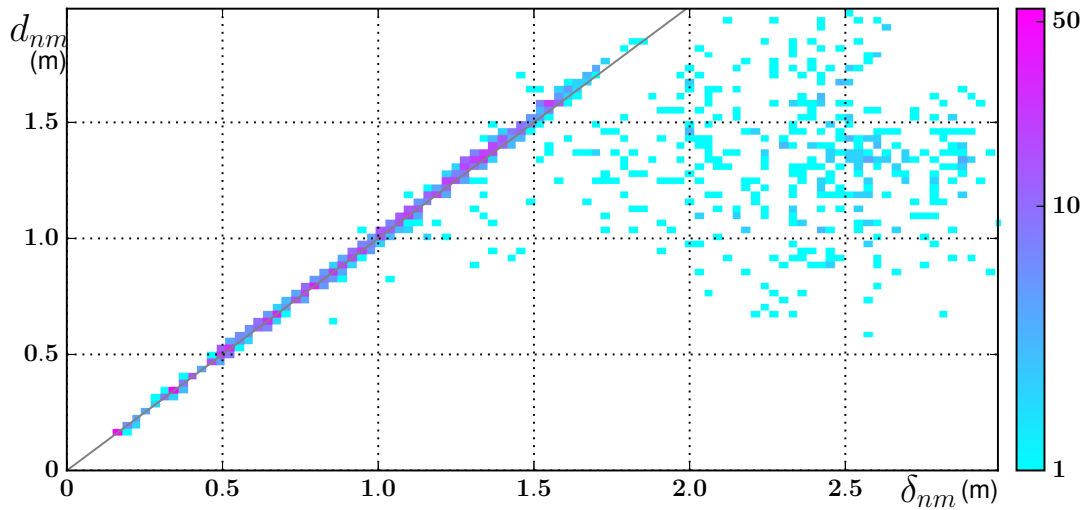


(a) Histogramme 2D des couples  $(\delta_{nm}, d_{nm})$ . L'erreur d'estimation  $|\delta_{nm} - d_{nm}|$  est la distance horizontale séparant les rectangles à la ligne  $d_{nm} = \delta_{nm}$ . Taille d'un carré :  $(3 \text{ cm} \times 3 \text{ cm})$ .

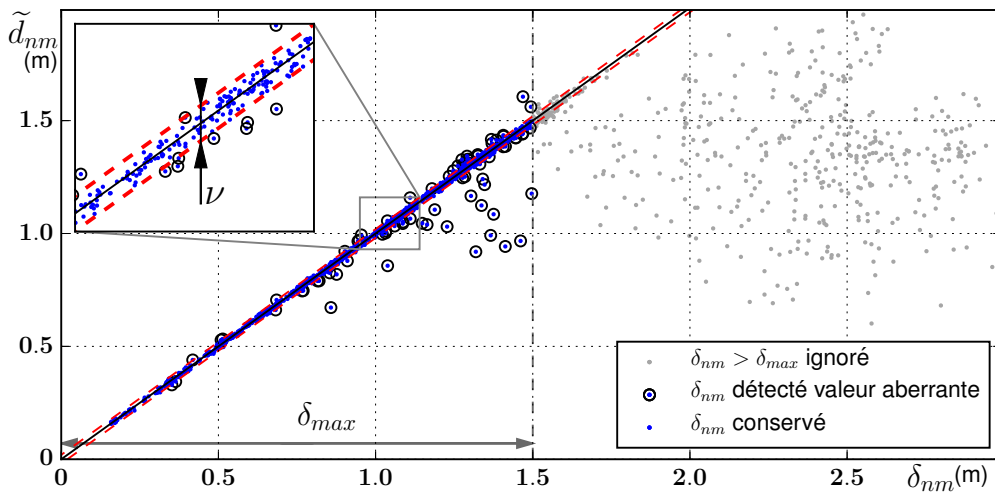


(b) Diagramme de Shepard du LRMDS : représentation des points  $(\delta_{nm}, \tilde{d}_{nm})$ .

**FIGURE 3.7 – Soleil** : évaluation de la méthode de calibration. En haut les couples  $(\delta_{nm}, d_{nm})$  sous forme d'histogramme 2D (fonction de la vérité-terrain). En bas les couples  $(\delta_{nm}, \tilde{d}_{nm})$ , ou diagramme de Shepard (obtenu sans connaissance de la vérité-terrain). 8128 distances pour  $M = 128$ .

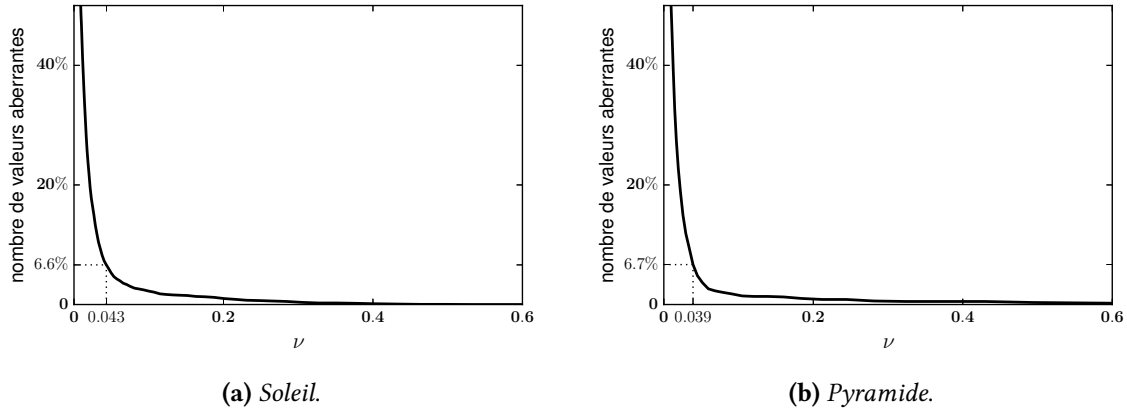


(a) Histogramme 2D des couples  $(\delta_{nm}, d_{nm})$ . L'erreur d'estimation  $|\delta_{nm} - d_{nm}|$  est la distance horizontale séparant les rectangles à la ligne  $d_{nm} = \delta_{nm}$ . Taille d'un carré :  $(3 \text{ cm} \times 3 \text{ cm})$ .



(b) Diagramme de Shepard du LRMDS : représentation des points  $(\delta_{nm}, \tilde{d}_{nm})$ .

**FIGURE 3.8** – *Pyramide* : évaluation de la méthode de calibration. En haut les couples  $(\delta_{nm}, d_{nm})$  (sous forme d'histogramme 2D), en bas les couples  $(\delta_{nm}, \tilde{d}_{nm})$  (diagramme de Shepard). 2016 distances pour  $M = 64$ .



**FIGURE 3.9** – Paramètre de régularisation : variation du nombre de valeurs aberrantes en fonction de  $\nu$ . Donné en pourcentage du nombre total de distances inférieures à  $\delta_{\max}$ .

- soit sur  $o_{mn}$  : le nombre total de valeurs aberrantes
- soit sur la loi de distribution caractérisant  $\epsilon_{mn}$  : les moments.

Ne connaissant aucun des deux, il n'existe pas de valeur objective optimale. Nous proposons de passer par une étude paramétrique, en exécutant le LRMDs pour différentes valeurs du paramètre de régularisation, comme proposé dans l'algorithme 2.

---

**ALGORITHME 2** RMDS : scan du paramètre de régularisation  $\nu$ .

---

```

for  $\nu_g$  de  $\nu_{\max} = \nu_1 > \nu_2 > \dots > \nu_G = \nu_{\min}$  espacés logarithmiquement do
  if  $\nu_g = \nu_1$  (Première fois) then
    Appliquer le RMDS 1 tel qu'il est décrit, avec  $\nu = \nu_g$ 
  else
    Appliquer l'algorithme RMDS avec  $\nu = \nu_g$ , en changeant l'initialisation :
    initialiser  $\mathbf{O}^{(0)}$  par l'estimée finale  $\tilde{\mathbf{O}}$  à  $\nu_{g-1}$ 
    initialiser  $\mathbf{X}^{(0)}$  par l'estimée finale  $\tilde{\mathbf{X}}$  à  $\nu_{g-1}$ 
  end if
end for

```

---

Nous pouvons obtenir la figure 3.9, qui représente le nombre de valeurs aberrantes identifiées en fonction du paramètre de régularisation :  $\|\tilde{\mathbf{O}}\|_0(\nu)$ .

Nous pouvons voir que réduire  $\nu$  relâche la contrainte de parcimonie, donc augmente la sensibilité de détection des valeurs aberrantes. Plus  $\nu$  est petit, plus  $\frac{\|\tilde{\mathbf{O}}\|_0}{2}$  augmente. La forme de la courbe en L distingue deux comportements : à gauche de la cassure,  $\frac{\|\tilde{\mathbf{O}}\|_0}{2}(\nu)$  augmente rapidement : le bruit  $\epsilon_{mn}$  est détecté comme valeur aberrante. Dans ce cas, la sensibilité est trop forte. Au contraire, prendre un  $\nu$  trop grand contraint une parcimonie tellement forte que aucune valeur aberrante est détectée. Dans ce cas  $\frac{\|\tilde{\mathbf{O}}\|_0}{2}$  est proche de 0. Le choix optimal de la valeur de  $\nu$  se

situé à la frontière située entre ces deux comportements : au niveau de la cassure de la courbe.

Enfin, les valeurs exclues supérieures à  $\delta_{\max}$  apparaissent en gris dans les deux diagrammes de Shepard. Il n'y a pas de critère objectif pour faire un choix optimal de  $\delta_{\max}$ . En pratique, deux points sont à prendre en compte :

- si l'algorithme 1 ne converge pas, c'est qu'il ne trouve pas de solution pour laquelle les valeurs aberrantes sont parcimonieuses. C'est que la valeur choisie de  $\delta_{\max}$  est trop grande. Nous avons observé, expérimentalement, une croissance rapide du nombre de valeurs aberrantes au delà de  $\delta_{nm} \approx 1.5$  m ;
- cependant plus on diminue  $\delta_{\max}$ , plus grand sera le nombre de pondérations  $w_{nm}$  à 0. D'une part, la diminution de la portée des distances  $\delta_{\max}$  dégrade la connectivité entre les éléments du réseau ; d'autre part, cela réduit la convexité de la fonction  $f_{\mathbf{W}}$ , ce qui rend la minimisation plus sensible aux minimums locaux. La solution  $\mathbf{X}$  en est fortement affectée.

Finalement, pour satisfaire ces deux conditions, la valeur de  $\delta_{\max}$  à prendre est la plus grande pour laquelle l'algorithme converge. Empiriquement, j'ai choisi  $1 \text{ m} \lesssim \delta_{\max} \lesssim 1.5 \text{ m}$ .

Le tableau 3.1 montre les résultats des calibrations géométriques, pour les antennes *Soleil* et *Pyramide*. Il quantifie le nombre d'éléments conservés par l'approche locale ( $< \delta_{\max}$ ), puis le nombre de valeurs aberrantes identifiées (avec une erreur d'amplitude  $> \nu/2$ ).

Antenne	$M$	$D$	$\delta_{\max}$ (m)	$\nu$ (m)	$\frac{\ \mathbf{W} \circ \Delta\ _0}{M^2}$	$\frac{\ \tilde{\mathbf{O}}\ _0}{\ \mathbf{W} \circ \Delta\ _0}$
<i>Soleil</i>	128	2	1	0.043	45.1%	6.6%
<i>Pyramide</i>	64	3	1.5	0.039	54.8%	6.7%

**TABLE 3.1** – Calibrations géométriques : paramètres choisis. Le nombre d'éléments conservés par l'approche locale est donnée par la norme  $\frac{\|\mathbf{W} \circ \Delta\|_0}{2}$  (ou  $\frac{\|\mathbf{W} \circ \Delta\|_0}{M^2}$  en ratio du nombre total). Le nombre de valeurs aberrantes identifiées est donnée par la norme  $\frac{\|\tilde{\mathbf{O}}\|_0}{2}$  (ou  $\frac{\|\tilde{\mathbf{O}}\|_0}{\|\mathbf{W} \circ \Delta\|_0}$  en ratio du nombre total).

## 3.6 Résultats expérimentaux

### 3.6.1 Performances de la méthode proposée

Nous calculons l'erreur géométrique de positionnement des microphones pour quantifier la performance de la méthode. Rappelons que la géométrie retrouvée dans l'espace n'est pas absolue : son orientation est arbitraire et son centroïde est l'origine. Dans la littérature, deux approches existent pour analyser l'erreur :

- à partir de  $\mathbf{X}$  et  $\tilde{\mathbf{X}}$ , l'erreur doit être indépendante de toute rotation, translation et réflexion. Représentant la *distance* entre les deux géométries [20], l'erreur est formulée de la façon

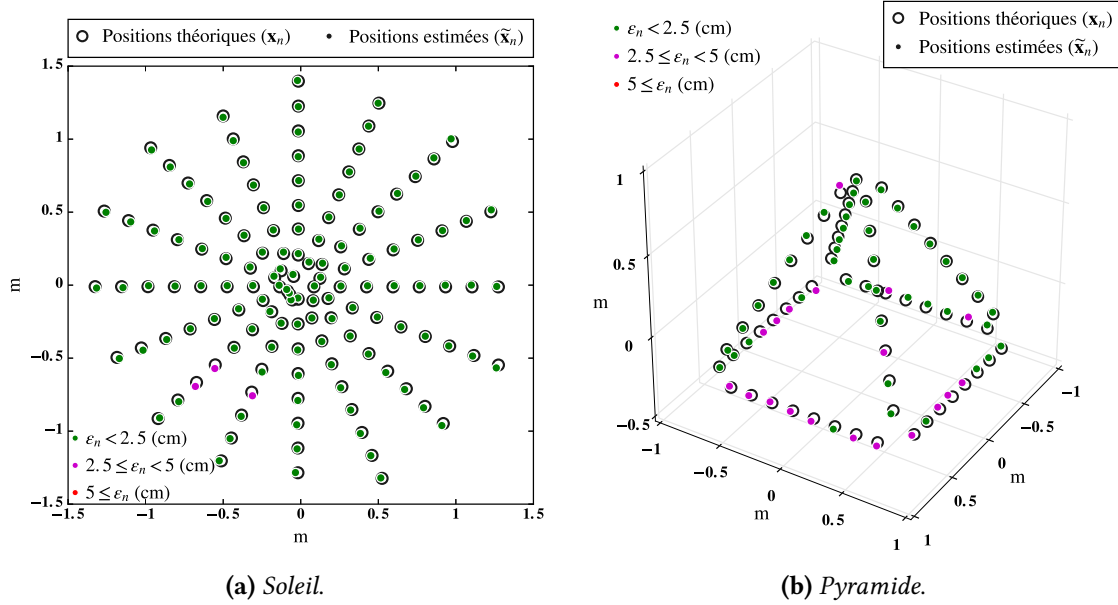


FIGURE 3.10 – Résultat de la calibration géométrique par la méthode proposée, après transformation de Procrustes. Configuration paramétrique en tableau 3.1, erreur en tableau 3.2.

suivante :

$$\text{dist}(\mathbf{X}, \tilde{\mathbf{X}}) = \frac{1}{M} \|\mathbf{J}\mathbf{X}^T\mathbf{X}\mathbf{J} - \mathbf{J}\tilde{\mathbf{X}}^T\tilde{\mathbf{X}}\mathbf{J}\|_F \quad (3.32)$$

avec  $\mathbf{J} = \mathbf{I}_M - (1/M)\mathbf{1}_M\mathbf{1}_M^T$ .  $\mathbf{J}$  applique une transformation de centrage à l'origine (*geometric centering matrix*).  $\mathbf{X}^T\mathbf{X}$  est le déterminant de Gram, qui a comme propriété d'être invariant en rotation et/ou réflexion. Ce critère quantifie une erreur globale de positionnement.

- à partir des positions individuelles  $\tilde{\mathbf{x}}_n$  et  $\mathbf{x}_n$ . Dans ce cas, il faut appliquer à l'une des deux géométries une transformation de translation et de rotation, pour ainsi les superposer *au mieux*. Cette technique est appelée transformation de Procrustes [61]. Nous appliquerons la méthode de Kabsch [75], résumée par l'algorithme 3. Notons qu'une erreur sur la célérité engendre une homothétie globale de la géométrie estimée ; nous supposons cette variabilité négligeable.

Ici nous préférons la deuxième approche. En effet, comme le montre la figure 3.11, elle représente plus fidèlement une erreur géométrique globale sur le réseau. Les résultats comparatifs des calibrations géométriques, après transformation de Procrustes, sont donnés en figure 3.10a et 3.10b. Notons l'erreur de positionnement par microphone  $\epsilon_n = \|\tilde{\mathbf{x}}_n - \mathbf{x}_n\|$ . Les erreurs des calibrations géométriques des réseaux *Soleil* et *Pyramide*, dans le tableau 3.2, sont données statistiquement par rapport aux erreurs individuelles.

Situés dans un lieu d'écoute identique (le grand hall), les résultats conduisent à une précision à l'échelle centimétrique sur le réseau pyramidal, millimétrique sur l'antenne circulaire. Pour le réseau *Soleil*, seuls 3 microphones sur 128 sont positionnés avec une erreur supérieure à 2.5 cm, ce qui est faible vis à vis de l'étendue spatiale totale. *Pyramide* a une erreur de reconstruction plus

---

**ALGORITHME 3** Transformation de Procrustes : algorithme de Kabsch.

---

**Entrées :**  $\tilde{\mathbf{X}}$  et  $\mathbf{X}$

1: Placer les centroïdes de  $\tilde{\mathbf{X}}$  et  $\mathbf{X}$  à l'origine :

$$\mathbf{c} = \frac{\sum_{n < M} \mathbf{x}_n}{M}, \quad \tilde{\mathbf{c}} = \frac{\sum_{n < M} \tilde{\mathbf{x}}_n}{M} \quad (3.33)$$

$$\mathbf{x}_n \leftarrow \mathbf{x}_n - \mathbf{c}, \quad \tilde{\mathbf{x}}_n \leftarrow \tilde{\mathbf{x}}_n - \tilde{\mathbf{c}} \quad (3.34)$$

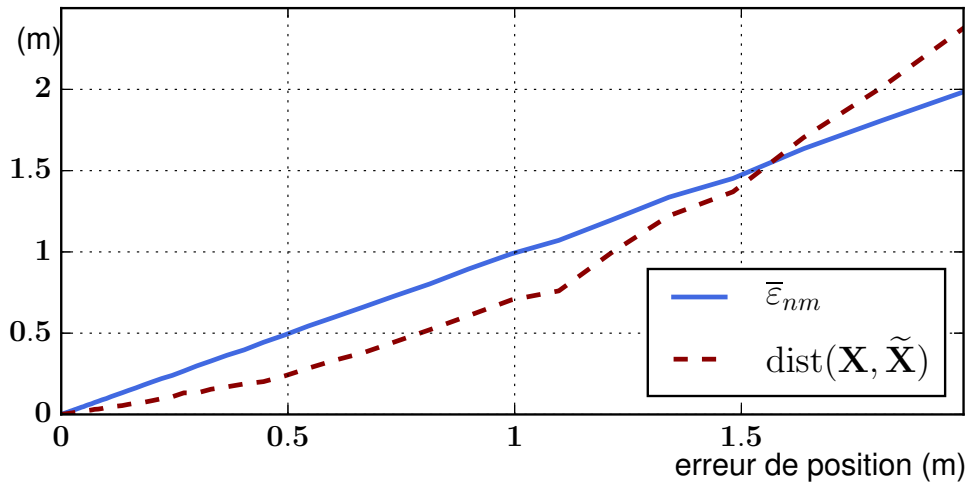
2: Sur la matrice de covariance  $\tilde{\mathbf{X}}\mathbf{X}^\top$ , appliquer une décomposition SVD :  $\tilde{\mathbf{X}}\mathbf{X}^\top = \mathbf{U}\mathbf{\Sigma}\mathbf{V}^\top$

3: En fonction du signe  $sgn = \text{sign}(\det(\tilde{\mathbf{X}}\mathbf{X}^\top))$ , calculer la matrice de rotation optimale :

$$\mathbf{\Omega} = \mathbf{V} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & sgn \end{pmatrix} \mathbf{U}^\top \quad (3.35)$$

4: **Sorties :**  $\mathbf{\Omega}\tilde{\mathbf{X}}$  et  $\mathbf{X}$

---



**FIGURE 3.11** – Comparaison des deux critères pour quantifier l'erreur globale :  $\text{dist}(\mathbf{X}, \tilde{\mathbf{X}})$  et  $\bar{\epsilon}_n$ . À un réseau tridimensionnel de 100 microphones disposés aléatoirement, une erreur de position est ajoutée à chaque élément, avec une amplitude homogène mais dans une direction aléatoire.  $\bar{\epsilon}_n$  traduit *a priori* plus précisément cette erreur de position que le critère  $\text{dist}(\mathbf{X}, \tilde{\mathbf{X}})$ .



Antenne	Méthode	$\bar{\varepsilon}_n$	$\min(\varepsilon_n)$	$\max(\varepsilon_n)$	$\text{std}(\varepsilon_n)$
<i>Soleil</i>					
	LRMDS	0.94 cm	0.08 cm	2.8 cm	0.50 cm
	Isomap	1.49 cm	0.19 cm	7.0 cm	1.21 cm
	Smacof local	1.50 cm	0.17 cm	6.0 cm	0.98 cm
<i>Pyramide</i>					
	LRMDS	2.2 cm	0.85 cm	4.6 cm	0.87 cm
	Isomap	10.0 cm	2.20 cm	18.8 cm	4.43 cm
	Smacof local	7.0 cm	0.73 cm	13.9 cm	3.11 cm

**TABLE 3.2** – Erreurs individuelles de positionnement pour des géométries *Soleil* et *Pyramide*, données en terme de moyenne, minimum, maximum, d'écart type (cm).

élevée, avec 19 microphones sur 64 ayant  $\varepsilon_n \geq 2.5$  cm.

L'antenne *Pyramide* a une taille inférieure, en revanche elle est plus complexe : sa géométrie est tridimensionnelle. La structure de la pyramide est plus creuse : 64 microphones sont situés uniquement sur les arêtes, alors que le réseau circulaire est constitué de 128 éléments répartis de façon relativement uniforme (la connectivité est donc plus grande). Par ailleurs, les microphones du réseau circulaire sont tous orientés dans la même direction, contrairement à la pyramide. Rappelons que les microphones sont supposés omnidirectionnels (cf. section 3.3.1). Selon la caractérisation des microphones MEMS, cette hypothèse semble pertinente. Cependant, le montage des microphones sur les barres en aluminium peut potentiellement affecter la directivité, et donc la mesure des cohérences [82].

### 3.6.2 Comparaison de résultats avec d'autres approches de la littérature

#### 3.6.2.1 Autre méthode de positionnement de microphones en champ diffus

Nous souhaitons étudier l'applicabilité des méthodes de la littérature [94, 139] sur un réseau de grande taille. Elles diffèrent de notre méthode par la manière d'écarter les valeurs aberrantes. Elles s'appuient sur le calcul, pour chaque paire de microphones ( $n, m$ ), de plusieurs estimations de cohérences calculées sur une durée courte, suivies d'une régression pour chacune d'elles. Comme le montre la figure 3.12, un ensemble de distances  $\{\delta_{nm}^1, \delta_{nm}^2, \delta_{nm}^3 \dots\}$  (environ 400 valeurs) par couple ( $n, m$ ) est construit. Une méthode de *clustering* extrait une valeur retenue traduisant la tendance globale des valeurs :

McCowan2008 [94] : la première approche exploite un *clustering* de type k-mean. À partir de points dont les coordonnées sont définies par une série de données, cette méthode consiste à diviser l'ensemble de ces points en  $k$  partitions (appelé *clusters*). Si l'ensemble de points contient des regroupements en amas, le *clustering* k-mean identifie ces regroupements. Ici, cette approche est effectuée sur les points dont les coordonnées sont les distances  $\{\delta_{nm}^1, \delta_{nm}^2, \delta_{nm}^3 \dots\}$ , en fonction des erreurs résiduelles issues de chaque minimisation qua-

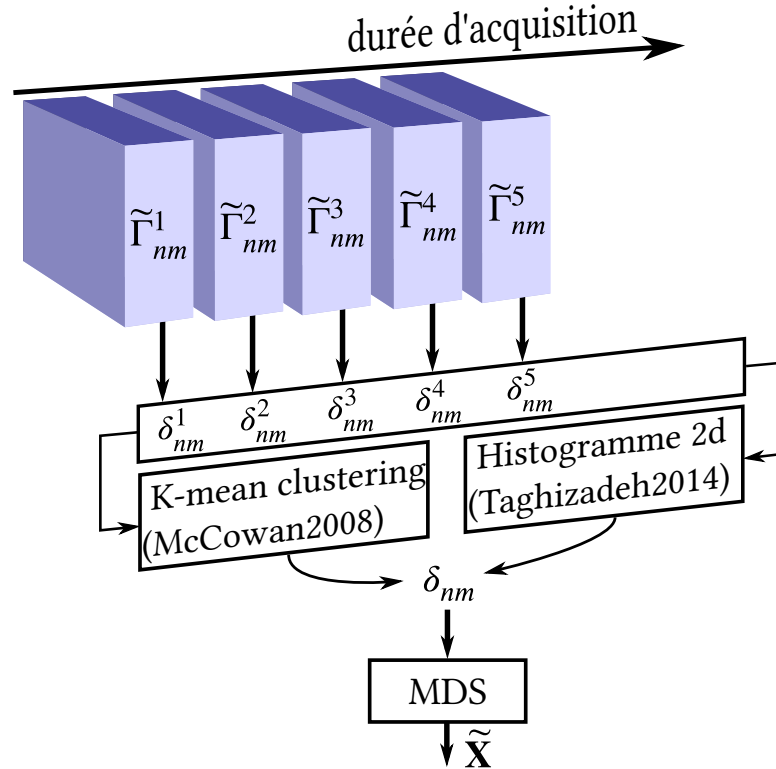


FIGURE 3.12 – Réjection des valeurs aberrantes : approche des méthodes de la littérature [94, 139].

dratique 3.8<sup>3</sup>. Par l’algorithme k-mean, deux clusters candidats sont déterminés ( $k = 2$ ). Parmi les centroïdes des deux clusters, la valeur associée à l’erreur résiduelle la plus faible constitue la valeur finale  $\delta_{nm}$ . Il s’ensuit une extraction de la géométrie globale par MDS classique.

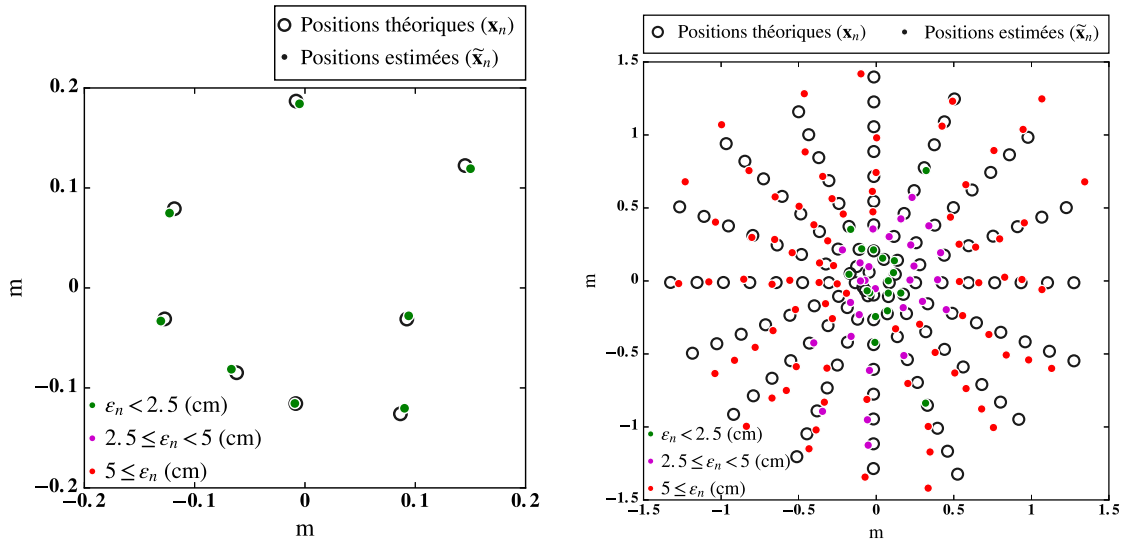
Taghizadeh2014 [139] : la deuxième approche exploite un *clustering* issu des mêmes points que précédemment, mais plutôt par histogramme 2D. Le rectangle de l’histogramme contenant la plus grande fréquence de valeurs traduit le comportement de régression le plus homogène, et détermine donc la valeur  $\delta_{nm}$  finale. L’extraction finale de la géométrie est issue d’une minimisation du stress (3.19). Nous choisirons l’algorithme Smacof pour cela.

Contrairement à notre méthode, ces approches agissent donc antérieurement au MDS. Nous nous focalisons sur la deuxième approche, dans laquelle les expériences exposent de bons résultats pour un réseau à 9 éléments, de 20 cm de largeur [139]. Pour la comparaison, nous utilisons deux réseaux :

- l’antenne circulaire *Soleil* utilisée précédemment.
- une antenne de petite taille, à  $M = 8$  éléments, contenue dans un disque de 15 cm de diamètre. Cette antenne est en fait un sous-réseau au centre de l’antenne *Soleil*.

Pour les deux antennes, 452 calculs de cohérence et régressions sont effectués pour chaque couple  $(n, m)$ , sur une durée de 1 min. Notons que cette étape, rapide pour la petite antenne

3. À ne pas confondre donc avec l’erreur de régression  $|\delta_{nm} - d_{nm}|$



(a) Antenne de 8 microphones, 15 cm de large. Erreur moyenne de reconstruction  $\bar{\varepsilon}_n = 1.1$  mm. (b) Soleil. Erreur moyenne de reconstruction  $\bar{\varepsilon}_n = 18.4$  cm.

**FIGURE 3.13** – Résultat de la calibration géométrique par Taghizadeh2014 [139], après transformation de Procrustes.

( $\approx 1$  min), devient relativement longue pour l'antenne complète ( $\approx 4$  h). En effet, la complexité de calcul est proportionnelle à  $M^2$ .

Les géométries estimées utilisant le *clustering* en histogramme 2D sont présentées en figure 3.13. Les 8 microphones du premier réseau sont positionnés avec une erreur de  $\bar{\varepsilon}_n = 0.11$  cm, ce qui valide l'efficacité de cette approche pour les réseaux de petite taille. En revanche, ce n'est pas le cas pour l'antenne circulaire (cf. figure 3.13b), reconstituée avec une erreur beaucoup plus grande :  $\bar{\varepsilon}_n = 18.4$  cm. C'est 20 fois moins précis que la méthode par LRMSD proposée dans ce chapitre. Le *clustering* par histogramme 2D devient inefficace pour les grandes distances, donc toutes les valeurs aberrantes ne sont pas écartées.

Cette comparaison permet de dégager deux conclusions. Premièrement, les approches par *clustering* semblent, en pratique, peu adaptées aux réseaux de grande taille. Deuxièmement, notre méthode est beaucoup moins coûteuse en temps de calcul que les méthodes de la littérature, pour un nombre  $M$  grand de microphones. En effet, elle ne nécessite pas d'étape de *clustering*, puisque les valeurs aberrantes sont explicitement identifiées, conjointement à  $\tilde{\mathbf{X}}$ .

### 3.6.2.2 Deux MDS locaux utilisés pour le positionnement de capteurs

Comme ce sont essentiellement les grandes distances qui sont erronées, ma première idée d'amélioration pour la calibration avait été d'utiliser les approches locales de MDS, puisque leur utilisation est déjà courante en localisation de capteurs (voir section 3.4.1.2). Les reconstructions géométriques sont nettement améliorées vis à vis d'un MDS global (cf fig. 3.13b). J'ai testé, à

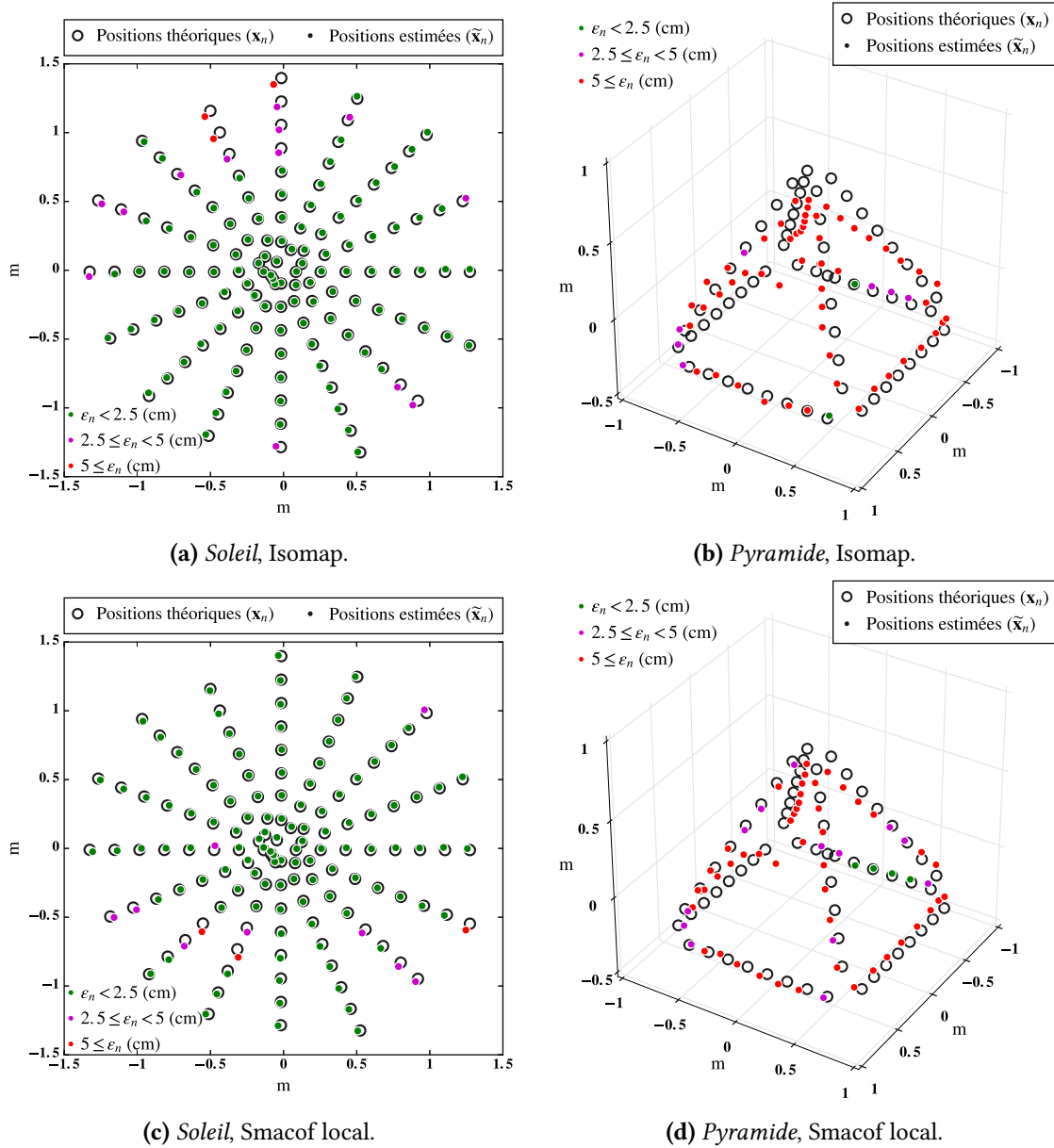
partir des même jeux de données  $\Delta$ , deux méthodes locales de MDS énoncées en section 3.4.1.2, appliquées sur les deux antennes *Soleil* et *Pyramide* :

- la minimisation de stress pondérée localement (3.20), avec Smacof. Comme avec le LRMDS, la pondération est de type 0/1 en fonction du seuil  $\delta_{\max}$  (cf eq. 3.24). Puisque cette méthode nécessite une condition initiale, il est judicieux de lancer plusieurs minimisations avec des initialisations aléatoires différentes, et de conserver l'estimée  $\tilde{\mathbf{X}}$  conduisant à la plus petite valeur de la fonction coût (3.20);
- l'Isomap, qui est un apprentissage de la géométrie en deux étapes. D'abord un graphe des proches voisins est construit : il représente, pour chaque microphone  $n$  du réseau, les microphones à proximité dans une portée inférieure à  $\delta_{\max}$ . À partir de ce graphe, toutes les distances par paire sont approximées par la distance géodésique. Elles sont construites par le plus court trajet menant du point  $n$  au point  $m$  dans le graphe (cf. algorithme de Dijkstra [145]). Le MDS classique est ensuite appliqué sur la matrice de distances géodésiques pour reconstruire la géométrie globale. Un point fort de cette méthode est qu'elle ne nécessite pas de condition initiale.

Les valeurs de seuil  $\delta_{\max}$  restent identique à celles choisis précédemment, à savoir 1 m pour *Soleil* et 1.5 m pour *Pyramide*. Les géométries obtenues sont en figure 3.14, et les erreurs sont comparées avec le LRMDS en tableau 3.2.

Les résultats montrent globalement une bonne fidélité de reconstruction pour le réseau de 128 microphones *Soleil*. Dans les deux cas, seuls 3 microphones sont reconstruits avec une erreur supérieure à 5 cm, et l'erreur de reconstruction est de 1.5 cm, donc 0.5 cm plus élevé que par LRMDS. Ces bons résultats s'expliquent par la répartition homogène et dense des microphones sur la surface plane, et s'opposent aux résultats obtenues avec l'antenne *Pyramide*. En effet, la pyramide est une structure creuse, qui convient mal à la construction de distances géodésiques dans Isomap. On peut d'ailleurs identifier une dégradation notable au niveau du sommet de la pyramide en figure 3.14b. L'erreur moyenne atteint 10 cm, supérieure à la reconstruction par Smacof (7 cm).

Ces deux approches sont intéressantes car elles permettent de retrouver au moins la structure géométrique, même si il n'y a pas de garantie sur l'erreur. Ces estimations géométriques sont exploitables en condition initiale pour le LRMDS, pour éviter le risque de convergence vers un minimum local.



**FIGURE 3.14** – Résultat de la calibration géométrique par deux approches locales utilisées en localisation de capteurs : Isomap et Smacof + pondération locale après transformation de Procrustes. Erreurs en tableau 3.2.

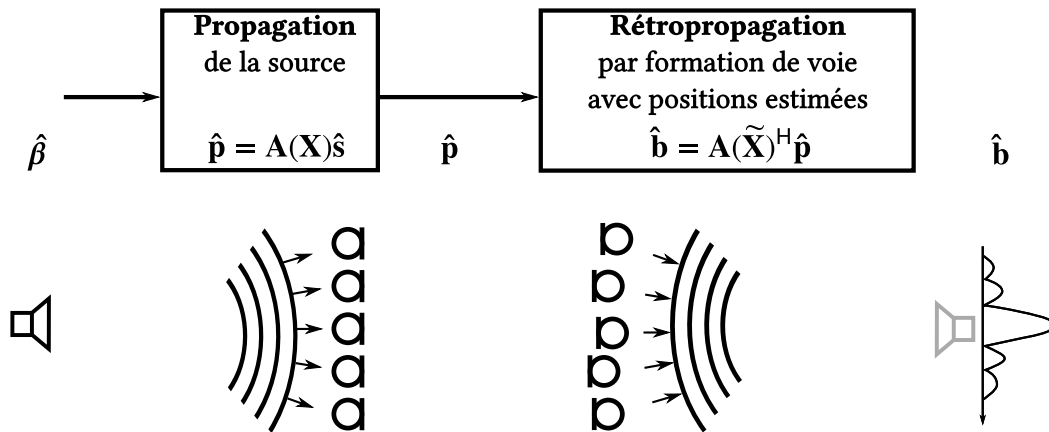


FIGURE 3.15 – Applicabilité à l'imagerie : principe de l'expérience numérique.

## 3.7 Discussions

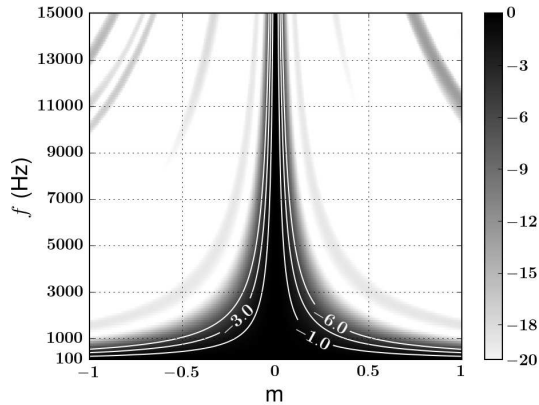
### 3.7.1 Applicabilité à l'imagerie acoustique

En soi, quantifier une erreur de positionnement est utile pour comparer différentes méthodes de calibration géométrique. Mais comment juger si une erreur de positionnement est acceptable ? L'estimation géométrique est-elle exploitable en vue d'une application à la localisation de sources ?

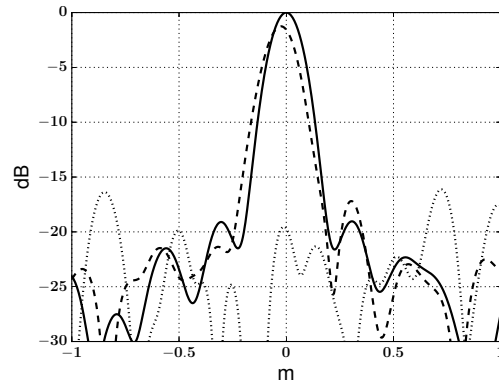
Nous proposons d'évaluer l'applicabilité de la méthode de calibration géométrique en imagerie acoustique, en particulier pour la formation de voies classique (type Delay&Sum), dans le domaine fréquentiel. Cette technique a été présentée dans le chapitre 2, dans lequel nous avons énoncé que l'erreur sur la connaissance *a priori* de la position des microphones dégrade la formation de voies (cf. section 2.5 du chapitre sur l'imagerie). En effet, comme le montrent Sachar *et al.* [125], le gain du lobe principal diminue en ajoutant une erreur de position par microphone suivant une loi de distribution gaussienne. Ce gain diminue également quand la fréquence augmente.

Ici, nous proposons par simulation d'analyser la formation de voies en champ proche<sup>4</sup>. Une source ponctuelle est placée à 5 m du centroïde des réseaux. Elle est caractérisée par un spectre unitaire large bande. Dans le domaine de fréquence, le niveau de la source est choisi de façon à ce que le gain par formation de voies soit théoriquement égal à 1 (ou 0 dB) dans la direction de la source. Schématisée en figure 3.15, l'expérience numérique consiste à appliquer la propagation de la source vers les positions réelles des microphones  $\mathbf{x}_n$ , puis d'appliquer la rétropropagation par formation de voies avec les positions de microphones connus *a priori*  $\tilde{\mathbf{x}}_n$ , donnés par l'expérience de calibration géométrique. À titre de comparaison, nous utiliserons les estimations par LRMS des antennes *Soleil* et *Pyramide* (cf fig. 3.10), puis également par la méthode de l'état de l'art de Taghizadeh *et al.* [139] de l'antenne *Soleil* (cf. fig 3.13b).

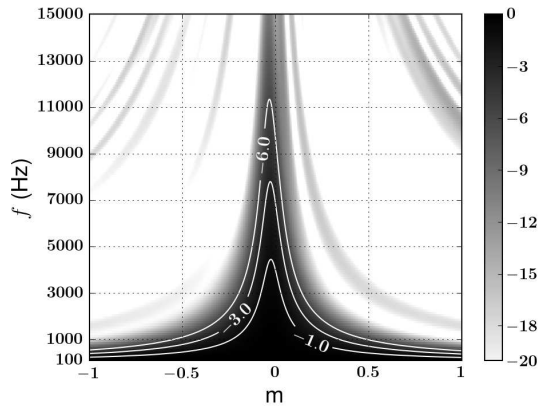
4. La rétro-propagation en champ proche suppose une propagation en onde sphériques.



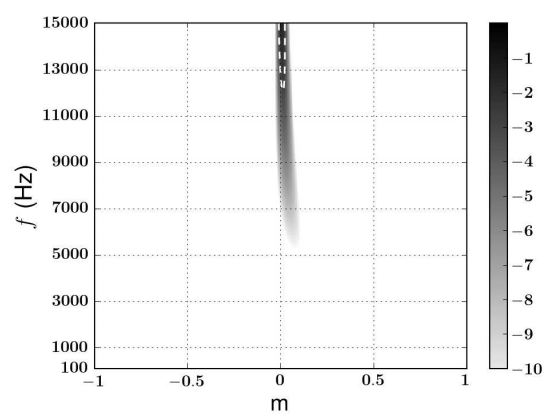
(a) Formation de voie avec la géométrie théorique  $\mathbf{X}$ , représentation en dB.



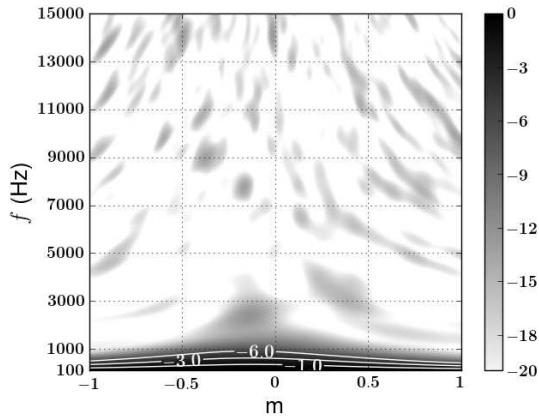
(d) Diagramme de directivité de 3.16a (trait continu), 3.16b (discontinu) et 3.16c (pointillé) à  $f = 5$  kHz.



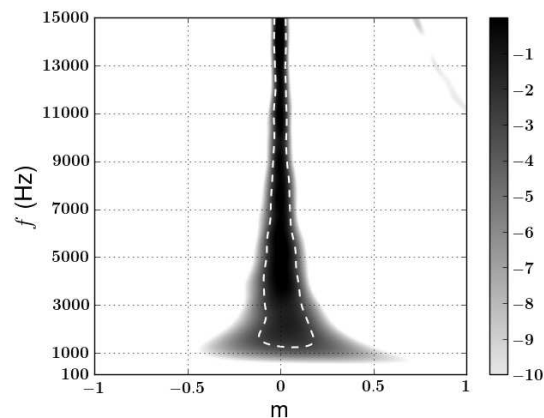
(b) Formation de voie avec la géométrie estimée  $\tilde{\mathbf{X}}$  par LRMDS ( $\bar{\epsilon}_n = 0.94$  cm, cf fig. 3.10a), représentation en dB.



(e) Formation de voie : erreur entre 3.16a et 3.16b, représentation en dB. Ligne de niveau à  $-3$  dB en blanc pointillé.



(c) Formation de voie avec la géométrie estimée  $\tilde{\mathbf{X}}$  par Taghizadeh *et al.* ( $\bar{\epsilon}_n = 18.4$  cm, cf fig. 3.13b), représentation en dB.



(f) Formation de voie : erreur entre 3.16a et 3.16c, représentation en dB. Ligne de niveau à  $-3$  dB en blanc pointillé.

**FIGURE 3.16** – Formation de voies en champ proche, avec l'antenne *Soleil*. La source ponctuelle est large bande, située à 5 m du centroïde du réseau. Comparaisons de reconstructions avec une connaissance géométrique *a priori* sans erreur (3.16a) et avec erreur (3.16b-3.16c).

Les résultats sont donnés en figure 3.16 pour *Soleil* et 3.17 pour *Pyramide*. Ils comparent le gain de formation de voies qui serait obtenu sans erreur de position (cf. fig. 3.16a - 3.17a), avec celui obtenu en utilisant les positions estimées (fig. 3.16b - 3.16c - 3.17b). Les erreurs relatives de gain de formation de voies sont représentées en fig. 3.16f - 3.16e - 3.17c. Nous considérons acceptable, arbitrairement, une baisse de gain due à l'erreur de positionnement si elle reste inférieure à 3 dB. Dans cette condition, la formation de voies est limitée en haute fréquence, à  $f_{\text{lim}} = 12$  kHz pour l'antenne circulaire avec l'estimation par LRMDS,  $f_{\text{lim}} = 1.2$  kHz pour l'antenne circulaire avec l'estimation par Taghizadeh *et al.*, et  $f_{\text{lim}} = 6$  kHz pour la pyramide par LRMDS.

L'évolution du gain limite en haute fréquence évolue avec l'erreur de positionnement donnée en tableau 3.2 :

- pour l'antenne *Soleil*, avec l'estimation par LRMDS, l'erreur moyenne est  $\bar{\epsilon}_n = 0.94$  cm. À  $f_{\text{lim}} = 12$  kHz, la longueur d'onde associée est de 2.8 cm ;
- pour l'antenne *Soleil*, avec l'estimation par Taghizadeh *et al.*, l'erreur moyenne est  $\bar{\epsilon}_n = 18.4$  cm. À  $f_{\text{lim}} = 1.2$  kHz, la longueur d'onde associée à cette fréquence est de 28 cm ;
- pour l'antenne *Pyramide*, l'erreur moyenne est  $\bar{\epsilon}_n = 2.05$  cm. À  $f_{\text{lim}} = 6$  kHz, la longueur d'onde associée est de 5.6 cm.

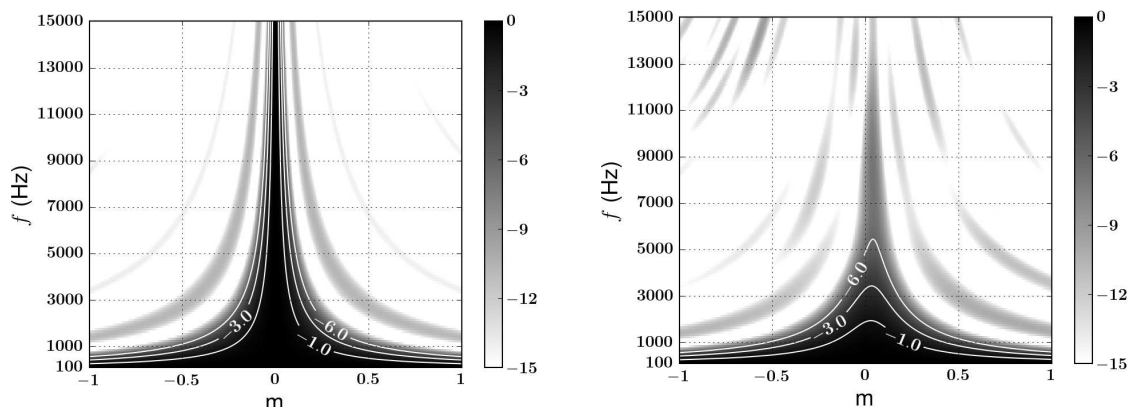
La longueur d'onde limite à la formation de voies semble donc évoluer dans le même sens que l'erreur de calibration. On retrouve donc la même tendance que celle montrée par Sachar *et al.* (voir la figure 2.6 dans le chapitre 2).

Enfin, il faut noter que l'erreur de positionnement a un second impact sur la directivité apparente de l'antenne. En plus de réduire le gain du lobe principal, elle augmente le niveau des lobes secondaires. La figure 3.16d montre la formation de voies à 5 kHz, qui est une vue en tranche des figures 3.16a, 3.16b et 3.16c. Avec une source large bande à spectre unitaire, ces courbes sont l'équivalent du diagramme de directivité. L'erreur de positionnement entraîne :

- dans le cas de la géométrie estimée par LRMDS :
  - une chute de 2 dB du gain du lobe principal ; de plus, la position apparente de la source (maximum) est décalée de 2.6 cm ;
  - une hausse du niveau maximal des lobes secondaires (premier lobe secondaire à gauche du lobe principal), passant de  $-19$  dB à  $-17$  dB ;
- dans le cas de la géométrie estimée par Taghizadeh *et al.* :
  - une chute de 20 dB du gain du lobe principal ; le niveau étant plus faible que celui des lobes secondaires, la formation de voies est inexploitable pour localiser la source ;
  - le niveau maximal des lobes secondaires passe de  $-19$  dB à  $-16$  dB.

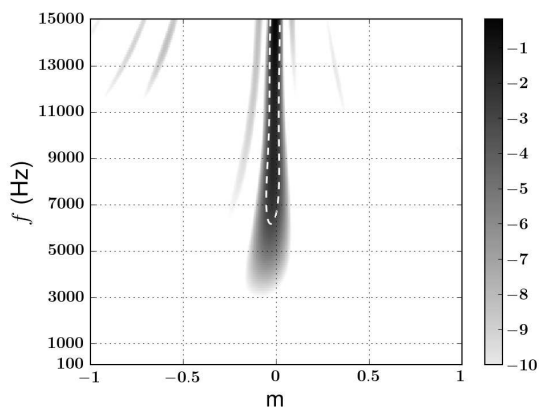
Globalement, l'erreur sur la géométrie a pour effet de dégrader la dynamique globale, définie par la différence de niveau entre le lobe principal et le niveau maximal des lobes secondaires. À une fréquence plus grande, les diagrammes de directivités apparents révéleraient une dynamique qui diminue. Fondamentalement, la localisation devient inexploitable quand la dynamique devient nulle, c'est à dire quand un lobe secondaire a la même amplitude que le lobe principal.





(a) Formation de voie avec la géométrie théorique  $\mathbf{X}$ , représentation en dB.

(b) Formation de voie avec la géométrie estimée  $\tilde{\mathbf{X}}$  par LRMS ( $\bar{\epsilon}_n = 2.05$  cm, cf fig. 3.10b), représentation en dB.



(c) Erreur sur la formation de voies, entre 3.17a et 3.17b, représentation en dB. Ligne de niveau à  $-3$  dB en blanc pointillé.

**FIGURE 3.17** – Formation de voies en champ proche, avec l’antenne *Pyramide*. La source ponctuelle est large bande, située à 5 m du centroïde du réseau. Comparaison de la reconstruction avec une connaissance géométrique *a priori* sans erreur (3.17a) et avec erreur (3.17b).

En conclusion, le gain en précision géométrique apporté par le LRMS se montre indispensable pour l’application à l’imagerie acoustique avec des grandes antennes. En effet, l’erreur géométrique induite par les méthodes de l’état de l’art est trop élevée pour pouvoir localiser des sources. Parmi les exemples concrets d’application, ces deux antennes estimées par LRMS fonctionneraient pour la localisation d’un locuteur, le spectre de la parole intelligible se situant en dessous de 4 kHz. En revanche la pyramide serait inappropriée pour des sources à spectre large bande au-delà de 6 kHz.

### 3.7.2 Au delà du milieu réverbérant

Les expériences précédentes ont permis de valider la méthodologie, notamment grâce à la bonne connaissance des géométries. Nous nous sommes également placés dans de bonnes conditions expérimentales pour mesurer un champ diffus. Cependant, est-il possible d’utiliser cette méthode de calibration géométrique quand les champs acoustiques ne sont pas parfaitement diffus ? Cette

section présente d'autres essais qui se sont déroulés dans des scénarios hors des conditions de laboratoire. Nous nous plaçons dans deux scénarios différents.

### 3.7.2.1 Cocktail party en environnement intérieur

La première expérience se déroule dans une salle de taille moyenne, moyennement réverbérante, dans laquelle sont réparties une vingtaine de personnes en train de discuter. C'est un environnement de type *Cocktail Party*. 24 microphones sont collés sur les murs, et 8 sur le plafond. La géométrie du réseau tridimensionnel était trop complexe pour être mesurée : ne connaissant pas la géométrie théorique, le résultat est comparé à une photographie.

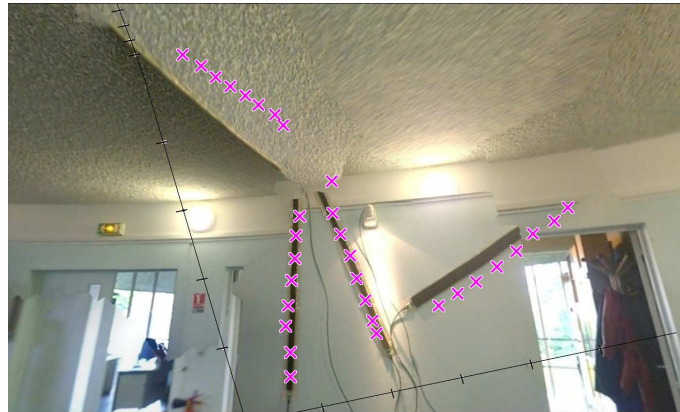
Le résultat de cette première expérience est montré en figure 3.18a. La géométrie estimée (croix violettes) est superposée sur une photo de la scène, avec une approximation sur la perspective de la photo. Il est impossible de quantifier la précision du positionnement, toutefois la géométrie estimée coïncide bien avec la positions des 4 faisceaux noirs de microphones sur les murs.

Dans un milieu faiblement réverbérant, la présence d'un grand nombre de sources décorréelées avec les premières réflexions semble suffisante pour la mesure de la cohérence. Nous tirons parti ici du modèle isotrope d'ondes planes décorréelées pour décrire le champ acoustique. Visiblement, il est possible aller au delà de la condition de diffusivité classique du milieu aérien associée au milieu réverbérant. La prochaine expérience se place dans le cas extrême où le milieu est non réverbérant, et dans lequel l'isotropie du champ est créé directement par la distribution des sources dans le milieu.

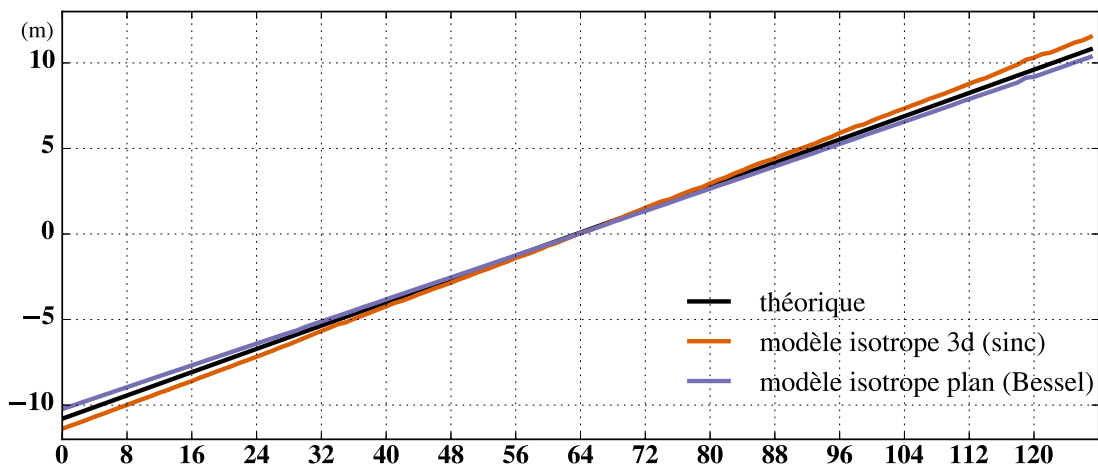
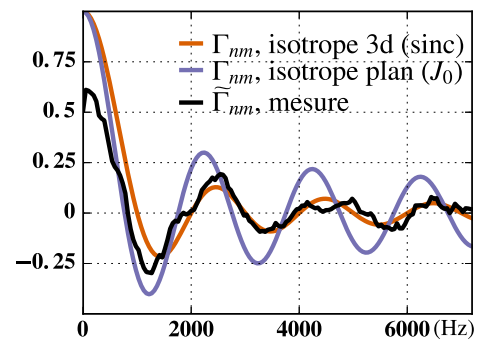
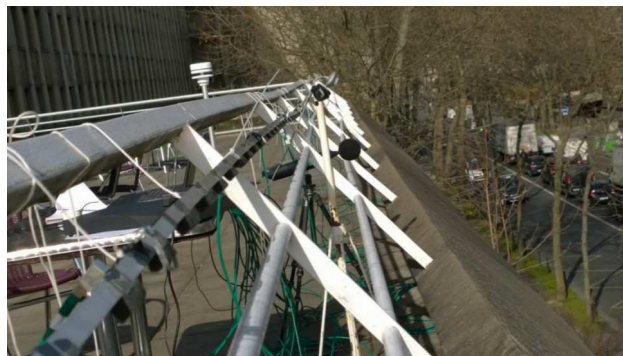
### 3.7.2.2 Bruit urbain en environnement extérieur

L'expérience présente se place en environnement extérieur, le long des Quais de Seine. Un réseau de 128 microphones a été mis en place pour faire de l'imagerie acoustique, en vue d'une étude sur le bruit urbain. C'est une antenne linéaire régulière, avec un pas de 17 cm ; nous connaissons les positions de manière approximative (précision centimétrique).

Dans ce deuxième scénario en plein air, le sol est pratiquement la seule cause de réflexion acoustique. L'antenne est positionnée le long d'une voie de circulation avec un trafic régulier de véhicules. La propagation directe des sources vers les microphones est donc l'origine principale du bruit. Ici, le champ isotrope d'ondes planes existe grâce au bruit global issu de tous les événements acoustiques urbains environnants. Ce phénomène est bien connu est exploité en acoustique sous-marine ou sismique [121] (cf. fin de l'annexe A). Dans l'expérience, les régressions des cohérences mesurées dans ce milieu sont en grande partie pertinentes. Cependant, il a été constaté que le nombre de valeurs aberrantes est élevé. Dans ce cas, nous ne sommes plus en adéquation avec la condition de parcimonie nécessaire à la résolution LRMDs. En pratique, cela s'est traduit par une non convergence de l'algorithme LRMDs. En revanche, la reconstruction avec le MDS local par



(a) Environnement cocktail party : une vingtaine de personnes sont en conversation. 4 faisceaux de 8 microphones sont placés sur les murs d'une salle moyennement réverbérante et à fréquence de Schroeder élevée. Les croix magenta représentent le positionnement par la méthode proposée, et les axes sont en noir. Les perspectives photographique et du rendu graphique des points sont différents, ce qui explique la superposition approximative.



(b) Environnement Urbain : 128 microphones alignés le long de la route du Quai Saint-Bernard (Jussieu, Paris). En haut à droite : la cohérence mesurée met en évidence une situation intermédiaire entre un champ isotrope plan et 3D. En bas : positionnement obtenu par Isomap [145] pour  $D = 1$  ; estimation avec modèle isotrope plan (mauve) et 3D (orange), en fonction de l'indice de microphone.

FIGURE 3.18 – Deux expériences *in situ* dans des conditions ne satisfaisant pas la condition de champ diffus : peu ou pas de réverbération, mais de nombreuses sources environnantes.

Modèle de champ	Homothétie	$\bar{\varepsilon}_n$	$\min(\varepsilon_n)$	$\max(\varepsilon_n)$	$\text{std}(\varepsilon_n)$
isotrope 3D	sans correction, $\tilde{\mathbf{X}} \approx 1.064\mathbf{X}$	35.2 cm	0.2 cm	74.5 cm	19.5 cm
	corrigée	3.2 cm	0.08 cm	13.4 cm	2.5 cm
isotrope plan	sans correction, $\tilde{\mathbf{X}} \approx 0.953\mathbf{X}$	25.5 cm	0.4 cm	56.1 cm	14.8 cm
	corrigée	2.7 cm	0.16 cm	13.8 cm	1.8 cm

**TABLE 3.3** – Calibration géométrique en environnement urbain. Erreurs individuelles de positionnement pour la géométrie (cf fig. 3.18b), données en terme de moyenne, minimum, maximum, d'écart type (cm). L'homothétie globale est issue de modèles de champs isotropes, approximant la distribution de sources. Cependant l'erreur sur la forme reste faible dans les deux cas.

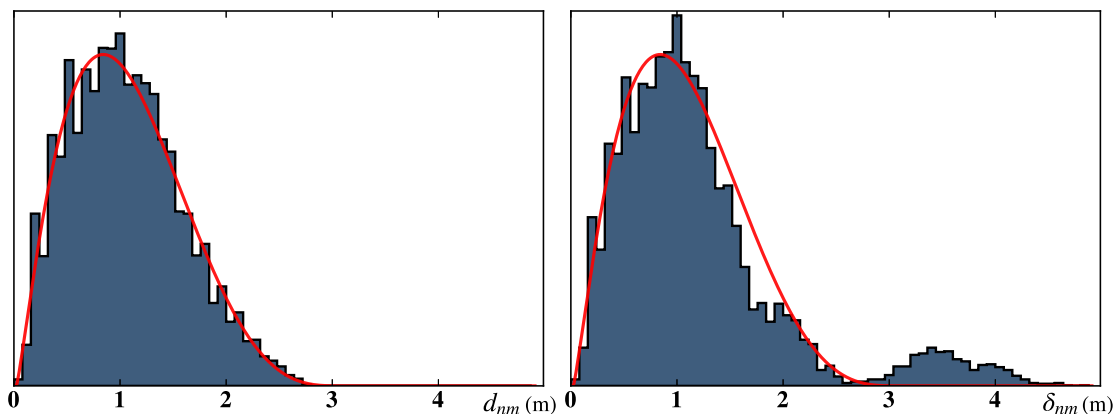
Isomap donne un résultat pertinent. Le positionnement résultant, pour  $D = 1$ , est comparé avec la théorie en figure 3.18b.

Le résultat est très encourageant : la forme de l'estimation correspond à la théorie, cependant il y a un facteur d'homothétie de 1.064. Il est faible mais notable en terme d'erreur à cause de la grande taille du réseau. L'erreur entre les deux géométries, de 35.2 cm, chute à 3.2 cm en corrigeant cette homothétie. La dilatation de l'estimation géométrique est trop grande pour être liée à une erreur sur la vitesse du son  $c_0$ . En revanche, nous pouvons remettre en question le choix du modèle de cohérence. En effet, le bruit venant de la route est inclus dans un angle solide restreint, ce qui modifie l'expression analytique [19]. Nous pouvons même conduire cette idée à l'extrême : faire l'hypothèse que les voitures sont contenues dans le même plan que le réseau de microphones. Dans ce cas, le modèle de cohérence peut être calculé à partir d'un champ isotrope 2D ; comme démontré en annexe A, ce modèle conduit à l'expression analytique :

$$\Gamma_{nm}(f) = J_0\left(\frac{2\pi f d_{nm}}{c_0}\right) \quad (3.36)$$

où  $J_0$  est la fonction de Bessel à l'ordre 0. En haut à droite de la figure 3.18b, une cohérence mesurée semble être intermédiaire entre les deux modèles de cohérence. En bas de la même figure, la géométrie théorique est comparée avec les reconstructions par Isomap, pour les deux choix de modèle. Le choix du modèle 2D montre une tendance inverse par rapport au modèle 3D : cette fois l'homothétie contracte la forme géométrique (0.953). Le tableau 3.3 montre que l'erreur a baissé de 10.3 cm par rapport au modèle isotrope 3D. Donc les sources issues des voitures sont plus proches d'une distribution plane que volumique, mais la réalité se positionne entre ces deux cas extrêmes.

Ces calibrations géométriques *in situ* sont prometteuses : il est possible d'aller au-delà du milieu fortement réverbérant, bien que la stratégie de calibration soit fondée sur un modèle de champ diffus. Plus encore, la méthode pourrait être applicable aux réseaux de capteurs en acoustique sous-marine.



**FIGURE 3.19** – Antenne *Soleil* : histogramme des distances théoriques (à gauche), et issues des cohérences (à droite). La régression de la loi de distribution beta sur l’histogramme des distances théoriques est tracée en rouge.

### 3.7.3 Matrice de distances et loi de distribution

Comme nous l’avons énoncé dans l’introduction, il existe des alternatives de calibration géométrique exploitant directement les propriétés mathématiques sur la matrice de distances. Le problème impliquant la méconnaissance des grande distances est connu, en téléphonie par exemple [41]. Des méthodes de complétion de matrice de distances arrivent à pallier ce problème.

Je n’ai trouvé qu’une méthode qui identifie la présence de valeurs aberrantes directement dans la matrice de distances observée  $\Delta$ . Cela dit, l’exploration de la matrice de distances, compte tenu des propriétés liées [52], serait une piste intéressante. Cette étude va au-delà des travaux de ma thèse, mais j’ai trouvé un point intéressant lié à la distribution statistique des distances. Un histogramme des valeurs  $d_{nm}$  pour le réseau circulaire est présenté à gauche de la figure 3.19. Sa forme en cloche asymétrique rappelle celle de la densité de probabilité de distributions continues positives.

Bien que l’antenne circulaire soit structurée en barres, sa géométrie plane reste considérée pseudo-aléatoire dans le cadre de l’imagerie (voir chapitre 2). Or, un modèle récurrent de distribution uniforme indépendante de points suit un processus de Poisson dans  $\mathbb{R}^D$  pour un domaine à  $D$  dimensions [98]. Selon des hypothèses supplémentaires (domaine fini ou infini par exemple) et selon  $D$ , il est démontré que les distances entre les points suivent certaines loi de distribution positives :

- le domaine plan  $\mathbb{R}^2$  a été étudié expérimentalement à plusieurs reprises. Par exemple, dans le domaine des télécommunications [159], la distance séparant un téléphone mobile des différentes stations de bases cellulaires environnantes suit la loi de Rayleigh ;
- Srinivasa *et al.* ont généralisé l’expression de la distribution des distances par paire [137]. Pour un ensemble de points finis, situé dans une boule de dimension arbitraire, la distribution

des distances par paire euclidiennes suit une loi beta ;

Dans les deux cas, il y a hypothèse d'une répartition de points dans un espace borné, en adéquation avec la forme finie d'un réseau de microphones. Donc, il est possible d'exploiter la distribution des distances par paire d'un réseau de microphones, à condition que les microphones soient en grand nombre, de densité spatiale homogène, et positionnés aléatoirement.

À partir de la géométrie de l'antenne *Soleil*, la figure 3.19 trace les histogrammes des distances théoriques ( $d_{nm}$ ) à gauche, et issues des régressions de cohérences ( $\delta_{nm}$ ) à droite. La régression de la loi beta, appliquée sur les distances théoriques, est tracée en rouge. Elle se révèle pertinente, et traduit donc la répartition quasi aléatoire uniforme des microphones dans le plan de la géométrie circulaire. Cependant, la même distribution se révèle moins pertinente sur l'histogramme des  $\delta_{nm}$  (présence d'une bosse entre 3 et 4 m, puis d'un creux à 1.8 m).

Si les microphones à positionner sont disposés aléatoirement et de façon homogène, je pense que cette démarche permettrait d'étiqueter les valeurs qui ont peu de probabilité d'être bonnes. Cette approche serait alors une autre possibilité d'écarter les valeurs aberrantes dans l'ensemble des distances  $\Delta$ .

### 3.8 Conclusion

Ce chapitre est consacré à la proposition d'une méthode de calibration géométrique en champ diffus. Par rapport aux méthodes de la littérature, l'enjeu particulier est d'avoir une méthode expérimentalement robuste, adaptée aux réseaux de grande taille, et à grand nombre de microphones. La méthode se fonde sur l'estimation des distances par paire entre les microphones, issues de la mesure des cohérences. De ces données géométriques, il est possible d'extraire une géométrie spatiale du réseau de microphones, par Multidimensional Scaling.

La régression d'une cohérence pouvant potentiellement aboutir à une valeur aberrante, l'ensemble des distances obtenues contient des valeurs fortement bruitées. Cet ensemble étant proportionnel au carré du nombre de microphones, il est inadapté de valider visuellement chaque régression. Il faut identifier les données aberrantes de façon non supervisée.

Ce chapitre propose l'algorithme LRMSD : il est *robuste* via l'identification des données aberrantes, conjointement à l'extraction de la géométrie globale. Pour garantir l'hypothèse nécessaire de parcimonie des données aberrantes, la méthode est *locale* : seules les distances entre les éléments voisins, situées dans une certaine portée, sont exploitées.

Les expériences montrent plusieurs points. D'une part, elles valident l'applicabilité de la méthode pour des réseaux de grande taille, à géométrie plane et tridimensionnelle. D'autre part, elles montrent que le processus de calibration géométrique est dans l'ensemble rapide, d'environ quelques minutes. Le champ acoustique mesuré est constitué de sources non contrôlées : il n'y a pas besoin de mettre en place un protocole de mesure expérimental particulier. La méthode est

donc passive. Enfin l'applicabilité pour l'imagerie acoustique en champ proche a été quantifiée ; l'erreur de positionnement des microphones dégrade la localisation en haute fréquence, de façon proportionnelle.

Globalement, la méthode ouvre des possibilités intéressantes pour la mise en œuvre de réseaux de grande taille. Le nombre d'éléments, ainsi que la taille du réseau ne sont *a priori* pas limités. La seule contrainte géométrique est de conserver une bonne connectivité, autrement dit de garder une densité volumique de capteurs suffisante. Notons également que la méthode proposée *étend* l'applicabilité aux grands réseaux, ce qui signifie qu'elle fonctionne également pour les petits réseaux.

Le caractère isotrope du champ est la clé. En effet, deux expériences de terrain ont montré des résultats prometteurs, dans un milieu peu ou pas réverbérant, mais contenant des sources décorrélées, réparties dans l'espace. Jusqu'à présent les conditions ont donc concerné soit (a) une salle réverbérante, dans lequel il existe peu ou de nombreuses sources, soit (b) un champ libre, mais exigeant de nombreuses sources distribuées dans l'espace. Un dernier cas n'ayant pas été traité est le champ libre contenant un petit nombre de sources, pour lequel le modèle de champ diffus est inapproprié. La calibration géométrique dans un tel milieu s'appuie sur la connaissance des temps de vol entre les sources et les microphones. Le chapitre suivant et final a pour but de proposer une approche robuste de calibration par temps de vol ou différence de temps de vol, ce qui permettra de compléter toutes les conditions acoustiques possibles.

## Chapitre 4

# Calibration géométrique robuste de grands réseaux par temps ou différence de temps de vol

### Sommaire

---

<b>4.1</b>	<b>Introduction</b>	<b>92</b>
<b>4.2</b>	<b>Calibration géométrique par temps de vol</b>	<b>94</b>
4.2.1	Le protocole classique	94
4.2.2	Une approche par RMDS robuste aux phénomènes de réflexion	99
4.2.3	Expérience numérique et comparaison	100
4.2.3.1	La simulation de Monte-Carlo	100
4.2.3.2	Résultats	100
<b>4.3</b>	<b>Calibration géométrique par différence de temps de vol</b>	<b>103</b>
4.3.1	Le protocole classique	103
4.3.2	Première approche robuste : débruitage préalable de l'observable	106
4.3.3	Deuxième approche robuste : calibration conjointe à l'identification des observables bruités	109
4.3.4	Expérience numérique et comparaison	112
4.3.4.1	La simulation de Monte-Carlo	112
4.3.4.2	Résultats	113
<b>4.4</b>	<b>Analyse globale des approches dans une expérience <i>in situ</i></b>	<b>116</b>
<b>4.5</b>	<b>Conclusion</b>	<b>118</b>

---



## 4.1 Introduction

La mesure de temps de vol (TOA : Time Of Arrival) ou de différences de temps de vol (TDOA : Time Difference Of Arrival) est employée dans différentes applications. Le premier type est la localisation de sources par une antenne [50] à partir des signaux acoustiques temporels. Dans ce cas, il s'agit de retrouver la position de sources à partir d'une géométrie d'antenne connue. La deuxième application concerne le positionnement d'appareils connectés à un réseau sans fil. Ce principe a été étendu à la calibration géométrique d'antennes acoustiques en champ libre. Généralement, le principe consiste à émettre des signaux de sources fixes en différents points de l'espace. Dans ce deuxième cas, il s'agit de retrouver à la fois les positions de sources et la géométrie de l'antenne. Les méthodes de calibration géométrique nécessitent la présence de sources fixes contrôlées ou non. On distingue alors deux cas de figure :

- le cas TOA : les sources sont contrôlées et leur signal est enregistré, synchronisé avec celui des récepteurs. On peut alors mesurer les temps de vol entre sources et récepteurs, par le calcul de l'intercorrélacion entre les signaux des sources et des récepteurs ;
- le cas TDOA : les sources sont inconnues et leur signal n'est pas enregistré. On ne peut avoir accès qu'aux différences de temps de vol entre les récepteurs, calculées par intercorrélacion entre les signaux des récepteurs.

Dans les deux cas, la littérature expose des méthodes de résolution, pour retrouver à la fois les positions de sources et la géométrie de l'antenne. Il en existe de deux types. Le premier concerne les méthodes *itératives* (par TOA [60] ou TDOA [65, 110]). Elles font intervenir une minimisation au sens des moindres carrés non-linéaire. Leur inconvénient principal est de converger vers un minimum local si les positions initiales de sources et de microphones sont trop différentes de la géométrie réelle. La deuxième famille de méthodes surmonte cet inconvénient en exprimant une solution *explicite* du problème de calibration géométrique. Il en existe pour le cas de mesures de TOAs [43, 44], et plus récemment avec des mesures de TDOAs [84, 85]. Ce chapitre traite de l'adaptation des méthodes par TOA et TDOA aux grands réseaux de microphones en environnements réels non réverbérants.

Ces méthodes ont fait l'objet de validations expérimentales, dans le cas TOA (128 [76] et 448 [125] microphones) et TDOA (121 microphones [86]). Elles mettent toutes en évidence une amélioration de l'estimation, quand le nombre de microphones et de sources augmente. Elles sont donc *a priori* bien adaptées aux applications des grands réseaux tels que *Megamicros*.

Généralement, l'environnement réel est considéré comme un champ libre, ce qui constitue une approximation forte. En effet, de nombreux éléments réfléchissants, tels que le sol ou des murs, bornent le milieu où se propagent les signaux acoustiques. Ceci implique que, pour une source et un microphone donnés, il existe potentiellement plusieurs chemins de propagation. Or la mesure

des TOAs et TDOAs par intercorrélation est perturbée par les réflexions. La figure 4.1 illustre le problème en présence d'un sol. Le TOA ou TDOA, supposé traduire la propagation directe entre les sources et microphones, peut en réalité être la signature d'un chemin indirect issu d'une réflexion par le sol. Dans ce cas, l'erreur engendrée peut être grande et donner lieu à une mesure aberrante du temps de vol ou de la différence de temps de vol. Pour une calibration efficace de la géométrie de l'antenne, ces erreurs doivent être éliminées de l'observable [116].

Dans ce chapitre, l'objectif est d'établir des méthodes de calibration géométrique par TOA/TDOA non-supervisées, robustes aux réflexions. Pour apporter de la robustesse vis-à-vis du bruit, Hennecke *et al.* proposent une estimation par maximum de vraisemblance de la géométrie, dans le cas où le bruit de mesure est gaussien [65]. Mais du fait du nombre relativement faible de valeurs aberrantes par rapport à l'ensemble de l'observable, il faut aller au-delà de l'hypothèse d'un bruit suivant une loi normale. Pour une application à un réseau Wi-Fi en intérieur, Batstone *et al.* [10] proposent de coupler une solution explicite de positionnement par TOA avec la méthode RANdom SAMple Consensus (RANSAC). Pour supprimer les valeurs aberrantes, RANSAC est un processus itéré comprenant deux étapes [55] :

1. la première détermine la géométrie de réseau à partir d'un sous-ensemble des temps de vol. Le problème de calibration géométrique par TOA étant un problème surdéterminé, la résolution reste possible avec une partie restreinte de l'observable. Le tirage du sous-ensemble des TOAs est aléatoire. L'hypothèse clé de RANSAC est que le sous-ensemble ne contient pas (ou très peu) de valeurs aberrantes ;
2. la deuxième étape analyse la pertinence de chaque temps de vol, non utilisé, avec la solution géométrique déterminée dans l'étape 1. Si la pertinence est mauvaise, le temps de vol est considéré valeur aberrante, et est rejeté. Dans le cas contraire le temps de vol est retenu, puis ajouté dans l'*ensemble de consensus*.

L'étape 1 et 2 est répétée plusieurs fois. L'estimation géométrique retenue correspond à celle pour laquelle l'*ensemble de consensus* a le cardinal le plus élevé. Notons que l'approche RANSAC a aussi été exploitée pour la calibration géométrique par TDOA [169].

Plutôt que d'exploiter un nouveau paradigme avec RANSAC, ce chapitre propose une approche unifiée avec la méthode de calibration robuste en champ diffus, proposée au chapitre précédent. Nous rappelons qu'elle s'appuie sur un modèle de bruit étendu, dans lequel les erreurs de grande amplitude sont définies explicitement, mais de valeur inconnue. De cette manière, la minimisation originale au sens des moindres carrés peut s'exprimer par une erreur quadratique excluant les grandes erreurs. Par une hypothèse de parcimonie des valeurs aberrantes, les grandes erreurs sont estimées à l'aide d'une méthode de Lasso [146]. La résolution globale est obtenue en alternant (a) la minimisation au sens des moindres carrés de l'erreur quadratique exempte des grandes erreurs, avec (b) la méthode de Lasso.

Pour rappel, dans le cas du MDS robuste, Forero *et al.* montrent [59] que l'erreur quadratique fait intervenir la régression des distances par paire, par une fonction coût non-linéaire. Dans

d'autres travaux, She *et al.* montrent que l'erreur quadratique peut représenter un problème linéaire quelconque [131]. L'idée exploitée dans ce chapitre est d'étendre le même principe, en utilisant une erreur quadratique exprimant les problèmes de calibration géométrique, par TOA et TDOA.

La première partie du chapitre 4.2 se focalise sur la calibration géométrique par TOA. Elle expose d'abord la méthodologie classique ; deuxièmement nous proposons une solution robuste fondée sur le RMDS exposé dans le chapitre précédent. Troisièmement, nous validerons l'intérêt de l'approche robuste par rapport à l'approche classique, à partir d'une simulation de Monte-Carlo. Sous une structure identique, la deuxième partie 4.3 se concentre sur la calibration par TDOA. Suite à la présentation de la méthodologie classique, nous proposons cette fois deux approches robustes. La première est fondée sur un débruitage préalable des TDOAs, tandis que la deuxième effectue la calibration géométrique conjointement à l'identification des valeurs aberrantes. Une simulation de Monte-Carlo permettra de comparer les deux méthodes robustes à la méthode classique. Enfin, en partie 4.4, dans une expérience *in situ* nous tenterons de reconstituer la géométrie d'une antenne de 256 microphones et de 20 m de longueur. À partir de données réelles, nous comparerons l'ensemble des méthodes exposées dans ce chapitre.

## 4.2 Calibration géométrique par temps de vol

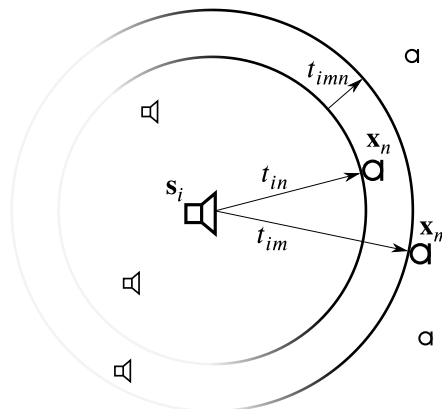
Les notations utilisées sont les mêmes qu'au chapitre précédent : nous considérons un problème de dimension spatiale  $D \leq 3$  dans lequel sont situés les sources et microphones. Un réseau comprend  $M$  microphones dont les positions spatiales sont données par les coordonnées  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M$ , ou la matrice globale  $\mathbf{X} = [\mathbf{x}_1, \dots, \mathbf{x}_M]$  dans  $\mathbb{R}^{D \times M}$ . En revanche, l'environnement est dorénavant un champ libre. Successivement,  $I$  sources fixes émettent un signal en différents points. Notons  $\mathbf{s}_1, \mathbf{s}_2, \dots, \mathbf{s}_I$  les positions des sources, ainsi que la forme matricielle  $\mathbf{S} = [\mathbf{s}_1, \dots, \mathbf{s}_I]$  dans  $\mathbb{R}^{D \times I}$ . Nous notons  $\beta_i(t)$  le signal temporel émis par la source  $i$ , et  $p_m(t)$  la pression mesurée par le microphone  $m$ . Enfin, nous verrons que l'ensemble des observables géométriques énoncés par la suite s'exprime en fonction de la distance séparant la source  $i$  du microphone  $m$  :  $d_{im} = \|\mathbf{s}_i - \mathbf{x}_m\|$ .

### 4.2.1 Le protocole classique

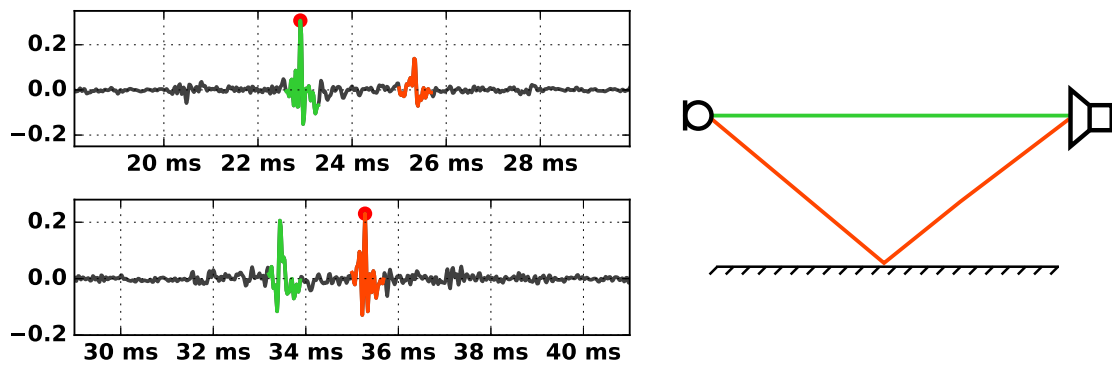
L'approche par temps de vol (TOA) suppose que les signaux sources  $\beta_i(t)$  émis sont connus, avec la référence temporelle identique à celle des mesures microphoniques. Dans ce cas, la mesure  $p_m(t)$  du microphone  $m$  contient le signal source, avec un retard de propagation  $t_{im}$  tel que :

$$t_{im}(\mathbf{X}, \mathbf{S}) = \frac{d_{im}}{c_0} = \frac{\|\mathbf{s}_i - \mathbf{x}_m\|}{c_0} \quad (4.1)$$

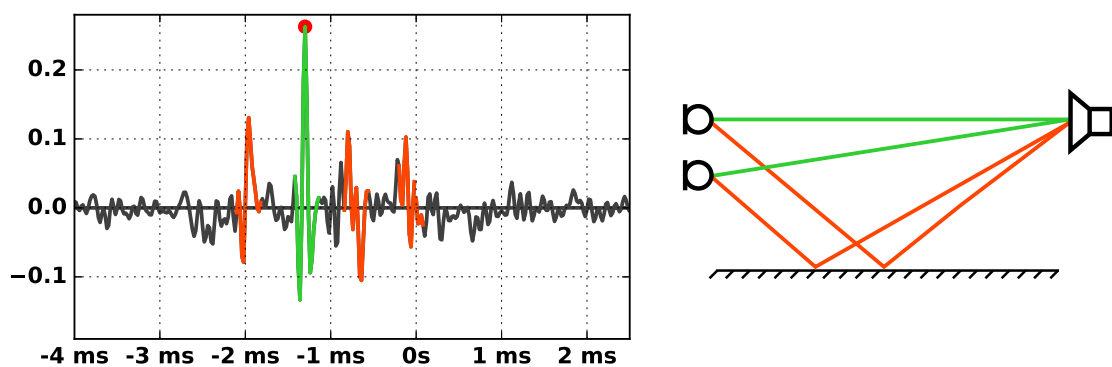
$c_0$  étant la vitesse du son. L'étape initiale nécessite de déterminer l'ensemble des  $t_{im}$  (voir le



(a) Notations utilisées pour les calibrations géométriques dans ce chapitre.



(b) Cas du TOA. Deux calculs d'intercorrélation, faisant apparaître le chemin direct (vert) et la réflexion (orange). Valeur maximale de l'intercorrélation en rouge. En bas, la fausse détection engendre une erreur de 2 ms, équivalente à 68 cm en distance.



(c) Cas du TDOA. Un calcul d'intercorrélation, faisant apparaître le chemin direct (vert) et les réflexions (orange). Valeur maximale de l'intercorrélation en rouge. Un pic détecté issu de réflexions peut engendrer une erreur de 1.3 ms, équivalente à 44 cm en distance.

FIGURE 4.1 – Les temps ou différences de temps de vol sont déterminés par la valeur maximale de l'intercorrélation. Les pics issus des réflexions sont sources de fausses détections, engendrant des valeurs aberrantes.

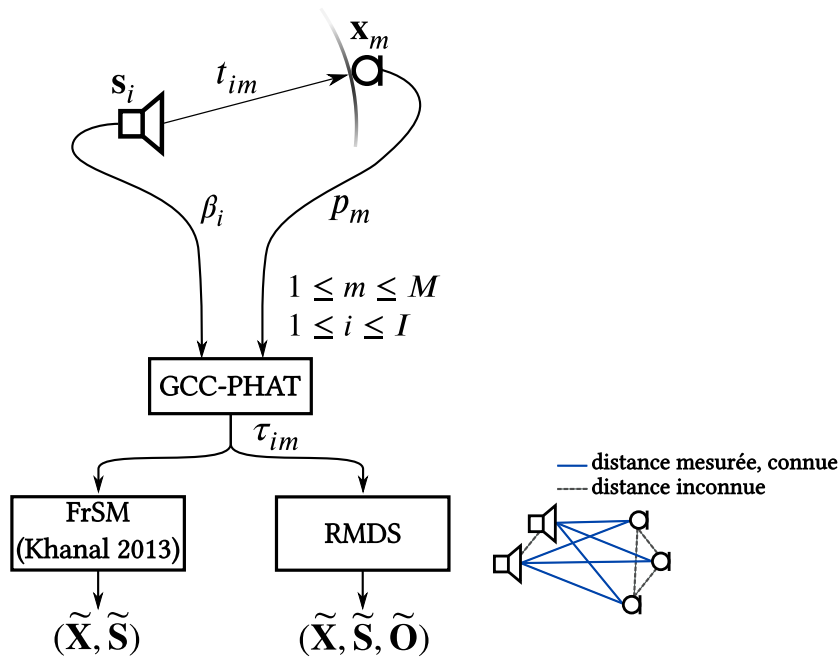


FIGURE 4.2 – Schéma de la méthode par TOA, par approche classique (gauche) ou robuste (droite).

schéma en figure 4.2). En principe, l'intercorrélacion temporelle dans sa définition initiale est rarement employée. Elle est sensible aux effets de la réverbération, se traduisant par des pics d'amplitude issus des réflexions élevés.

Pour réduire l'impact de la réverbération, l'intercorrélacion généralisée (en anglais *Generalized Cross Correlation*, ou GCC) s'appuie sur la transformée de Fourier inverse de l'interspectre entre le signal de la source  $\beta_i(t)$  et le signal du microphone  $p_m(t) : S_{im}(f)$ . Elle s'écrit :

$$r_{\text{PHAT}}(\tau) = \int_{\mathbb{R}} \Psi(f) S_{im}(f) e^{j2\pi f \tau} df \quad (4.2)$$

où  $\Psi$  est une fonction de pondération fréquentielle de l'interspectre  $S_{im}$ . Aujourd'hui, la fonction de pondération PHASE Transform (PHAT) est la plus répandue [79]. Son effet est de *blanchir* l'interspectre, autrement dit de donner un poids identique à chaque fréquence indépendamment du spectre du signal :

$$\Psi(f) = \frac{1}{|S_{im}(f)|} \quad (4.3)$$

Si l'étendue spectrale du signal source est connue et contenue dans l'intervalle  $f \in [f_{\min}; f_{\max}]$ , il est préférable de restreindre la transformée de Fourier inverse sur cette plage pour réduire le

bruit. L'estimateur par GCC-PHAT s'écrit alors :

$$r_{\text{PHAT}}(\tau) = 2 \int_{f_{\min}}^{f_{\max}} \frac{S_{im}(f)}{|S_{im}(f)|} e^{j2\pi f \tau} df \quad (4.4)$$

Finalement, l'estimée du temps de vol  $t_{im}$ , notée  $\tau_{im}$ , est donnée par le niveau maximal de la GCC-PHAT  $r_{\text{PHAT}}(\tau)$ , au niveau du pic :

$$\tau_{im} = \underset{\tau}{\operatorname{argmax}} r_{\text{PHAT}}(\tau) \quad (4.5)$$

L'expérience montre que les sinus glissants, ou *chirps*, sont des signaux efficaces pour mesurer  $r_{\text{PHAT}}$  [76, 125].

La résolution du problème de calibration géométrique par TOA, sous forme itérative, se traduit par la minimisation au sens des moindres carrés entre la mesure  $\tau_{im}$  et le modèle des temps de vol  $t_{im}(\mathbf{X}, \mathbf{S})$  :

$$\begin{aligned} (\tilde{\mathbf{X}}, \tilde{\mathbf{S}}) &= \underset{\mathbf{X}, \mathbf{S}}{\operatorname{argmin}} f_{\text{TOA}}(\mathbf{X}, \mathbf{S}) \\ \text{tel que } f_{\text{TOA}}(\mathbf{X}, \mathbf{S}) &= \sum_{i=1}^I \sum_{m=1}^M [\tau_{im} - t_{im}(\mathbf{X}, \mathbf{S})]^2 \end{aligned} \quad (4.6)$$

Il peut être montré que le problème 4.6 est équivalent à une estimation par maximum de vraisemblance du couple de matrices  $(\tilde{\mathbf{X}}, \tilde{\mathbf{S}})$  [18]. Parmi les différentes résolutions proposées dans l'état de l'art, nous choisissons celle proposée par Khanal *et al.* [76], nommée Free-Source Method (FrSM) et dédiée aux grands réseaux. À partir d'une condition initiale donnée par une solution explicite, ce dernier propose d'employer une descente de gradient à pas fixe.

Pour obtenir les gradients de  $f_{\text{TOA}}(\mathbf{X}, \mathbf{S})$ , Khanal *et al.* construisent une fonction  $g$  convexe sur les deux matrices  $\mathbf{X}$  et  $\mathbf{S}$ . De cette manière l'expression des dérivées se retrouvera simplifiée. Premièrement notons que  $f_{\text{TOA}}$  peut se re-écrire sous la forme :

$$f_{\text{TOA}}(\mathbf{X}, \mathbf{S}) = \sum_{i=1}^I \sum_{m=1}^M t_{im}(\mathbf{X}, \mathbf{S})^2 \left( 1 - \frac{\tau_{im}}{t_{im}(\mathbf{X}, \mathbf{S})} \right)^2 \quad (4.7)$$

$$= \sum_{i=1}^I \sum_{m=1}^M \frac{1}{c_0^2} \|\mathbf{s}_i - \mathbf{x}_m\|^2 \left( 1 - \frac{c_0 \tau_{im}}{\|\mathbf{s}_i - \mathbf{x}_m\|} \right)^2 \quad (4.8)$$

$$= \sum_{i=1}^I \sum_{m=1}^M \frac{1}{c_0^2} \left\| (\mathbf{s}_i - \mathbf{x}_m) \left( 1 - \frac{c_0 \tau_{im}}{\|\mathbf{s}_i - \mathbf{x}_m\|} \right) \right\|^2 \quad (4.9)$$

$$= \sum_{i=1}^I \sum_{m=1}^M \frac{1}{c_0^2} \left\| \mathbf{s}_i - \mathbf{x}_m - c_0 \tau_{im} \frac{\mathbf{s}_i - \mathbf{x}_m}{\|\mathbf{s}_i - \mathbf{x}_m\|} \right\|^2 \quad (4.10)$$

En notant  $\mathbf{e}$  l'ensemble des vecteurs unitaires, tel que  $\mathbf{e}_{im}$  est orienté du capteur  $m$  vers la source  $i$

$$\mathbf{e}_{im} = \frac{\mathbf{s}_i - \mathbf{x}_m}{\|\mathbf{s}_i - \mathbf{x}_m\|} \quad (4.11)$$

nous obtenons la fonction  $g_{\text{TOA}}$ , dépendant de la variable auxiliaire  $\mathbf{e}$  :

$$g_{\text{TOA}}(\mathbf{X}, \mathbf{S}, \mathbf{e}) = \sum_{i=1}^I \sum_{m=1}^M \frac{1}{c_0^2} \|\mathbf{s}_i - \mathbf{x}_m - c_0 \tau_{im} \mathbf{e}_{im}\|^2 \quad (4.12)$$

Dans une itération, la première étape est le calcul de la variable auxiliaire  $\mathbf{e}$ . La descente de gradient est appliquée indépendamment sur chaque position de source  $\mathbf{s}_i$  et de capteur  $\mathbf{x}_m$ . Les gradients sont déterminés par les dérivées partielles de  $g_{\text{TOA}}$  sur les variables à mettre à jour à chaque itération :

$$\frac{\partial g_{\text{TOA}}(\mathbf{X}, \mathbf{S}, \mathbf{e})}{\partial \mathbf{s}_i} = \sum_{m=1}^M \frac{2}{c_0^2} (\mathbf{s}_i - \mathbf{x}_m - c_0 \tau_{im} \mathbf{e}_{im}) \quad (4.13)$$

$$\frac{\partial g_{\text{TOA}}(\mathbf{X}, \mathbf{S}, \mathbf{e})}{\partial \mathbf{x}_m} = \sum_{i=1}^I \frac{-2}{c_0^2} (\mathbf{s}_i - \mathbf{x}_m - c_0 \tau_{im} \mathbf{e}_{im}) \quad (4.14)$$

La mise à jour est basée sur une descente à pas constant, noté  $\alpha > 0$  et choisi arbitrairement. À l'itération  $k$  nous avons alors :

$$\mathbf{s}_i^{(k+1)} = \mathbf{s}_i^{(k)} - \frac{\alpha}{M} \sum_{m=1}^M (\mathbf{s}_i^{(k)} - \mathbf{x}_m^{(k)} - c_0 \tau_{im} \mathbf{e}_{im}) \quad (4.15)$$

$$\mathbf{x}_m^{(k+1)} = \mathbf{x}_m^{(k)} + \frac{\alpha}{I} \sum_{i=1}^I (\mathbf{s}_i^{(k)} - \mathbf{x}_m^{(k)} - c_0 \tau_{im} \mathbf{e}_{im}) \quad (4.16)$$

où  $\mathbf{e}_{im}$  est calculé comme en équation 4.11 en fonction de  $\mathbf{x}_m^{(k)}$  et  $\mathbf{s}_i^{(k)}$ . Notons que les termes  $\frac{2}{c_0^2}$  constants apparaissant dans l'expression des dérivées 4.13, sont contenus dans le paramètre  $\alpha$ . Enfin les sommes sont normalisées pour garder une vitesse de descente identique dans les deux expressions des mises à jour, avec le paramètre de pas commun  $\alpha$ .

Les règles de mises à jour 4.15-4.16 donnent lieu à l'algorithme FrSM. Le choix de  $\alpha$  a un impact sur la rapidité de convergence. Plus  $\alpha$  est grand, plus la mise à jour des variables converge rapidement vers le minimum local de la fonction coût. Mais il faut également éviter de prendre

une valeur trop élevée, qui augmente le risque d'instabilité de l'algorithme. Finalement, le critère d'arrêt traduit la convergence de l'estimation recherchée. Dans le cas présent, il existe les deux matrices de positions  $\tilde{\mathbf{S}}$  et  $\tilde{\mathbf{X}}$ , mais nous choisissons  $\tilde{\mathbf{X}}$  qui correspond à la géométrie de l'antenne (ce que nous recherchons initialement). Le critère calculé  $\|\mathbf{X}^{(k+1)} - \mathbf{X}^{(k)}\|_F / \|\mathbf{X}^{(k)}\|_F$  définit la dernière itération  $k$  quand sa valeur est inférieure à un seuil positif noté  $\kappa$ .

#### 4.2.2 Une approche par RMDS robuste aux phénomènes de réflexion

À cause de la nature quadratique de la fonction  $f_{\text{TOA}}$  minimisée, nous verrons par l'expérience que la résolution du problème 4.6 est sensible aux valeurs aberrantes des temps de vol mesurés. Pour apporter de la robustesse, nous allons utiliser la même stratégie que dans le cas du champ diffus. Nous déclarons explicitement une variable, inconnue, qui va représenter le bruit issu des valeurs aberrantes. L'observable  $\tau_{im}$  suit le modèle de bruit suivant :

$$\tau_{im} = t_{im} + \epsilon_{im} + o_{im} \quad (4.17)$$

où  $\epsilon_{im}$  est un bruit d'amplitude faible suivant une loi de distribution centrée, et  $o_{im}$  l'erreur contenue dans la valeur aberrante. La matrice  $\mathbf{O}$  contenant les éléments  $o_{im}$ , de taille  $I \times M$ , devient une inconnue supplémentaire à déterminer. Nous supposons que le nombre de valeurs aberrantes engendrées par les réflexions est petit devant le nombre total  $I \times M$  de temps de vol. Il s'agit d'une hypothèse de parcimonie sur la matrice  $\mathbf{O}$ . Dans ce cas, le problème peut se traduire par :

$$(\tilde{\mathbf{X}}, \tilde{\mathbf{S}}, \tilde{\mathbf{O}}) = \underset{\mathbf{X}, \mathbf{S}, \mathbf{O}}{\operatorname{argmin}} \sum_{i=1}^I \sum_{m=1}^M [\tau_{im} - t_{im}(\mathbf{X}, \mathbf{S}) - o_{im}]^2 + \nu \sum_{i=1}^I \sum_{m=1}^M |o_{im}| \quad (4.18)$$

avec  $\nu$  le paramètre de régularisation de la norme  $\ell_1$  sur la matrice de bruit  $\mathbf{O}$ . Le problème 4.18 est une extension du problème de TOA posé en équation 4.6. La minimisation au sens des moindres carrés exclut les grandes erreurs  $o_{im}$  de l'erreur quadratique, pour rendre l'estimation géométrique optimale. Pour retrouver la matrice inconnue  $\mathbf{O}$  *a priori* parcimonieuse, le terme de droite minimise la norme  $\ell_1$  de  $\mathbf{O}$ . Il s'agit donc d'un problème de type Lasso.

Notons que, au facteur  $c_0$  près, les temps de vol traduisent les distances. Donc le problème 4.18 est identique au problème de RMDS 3.23 posé dans le chapitre 3. En effet, si on considère l'ensemble des points (sources + microphones), connaître les temps de vol revient à connaître une partie de la matrice des distances par paire de l'ensemble des points. Il s'agit du cas particulier de MDS dans lequel nous connaissons uniquement les distances entre deux groupes de points. Ce type de problème est généralement appelé MultiDimensional Unfolding [52].

Pour l'approche robuste par temps de vol, nous proposons donc de reprendre l'algorithme RMDS 1 du chapitre 3, en adaptant les termes de pondérations, égaux à 1 pour les couples (sources, microphones), et 0 ailleurs (cf. figure 4.2). Dans le cas du RMDS, le critère d'arrêt évalue la



convergence de l'ensemble des coordonnées  $\mathbf{X}$  et  $\mathbf{S}$ .

### 4.2.3 Expérience numérique et comparaison

#### 4.2.3.1 La simulation de Monte-Carlo

Nous comparons la solution des deux problèmes énoncés 4.6 (par FrSM) et 4.18 (par RMDS), suivant une étude paramétrique selon deux critères susceptibles de varier dans des conditions expérimentales :

1. la part de temps de vol contaminés par des valeurs aberrantes ;
2. l'écart entre la géométrie initiale donnée  $(\mathbf{X}^{(0)}, \mathbf{S}^{(0)})$  et la vérité-terrain.

Selon ces deux paramètres, une simulation de Monte-Carlo est établie en calculant 100 réalisations différentes. La configuration géométrique globale, représentée en figure 4.8b, est fixée sur toute la simulation. Une antenne plane constituée de 256 microphones, est étendue sur une longueur de 19.6 m et une hauteur de 2.25 m. Elle est entourée de 25 sources.

L'ensemble des temps de vol contient 0, 1, 3, 5, 8 ou 11% de valeurs aberrantes. Ces dernières sont calculées en considérant le temps de vol issu de la réflexion sur le sol au lieu du chemin de propagation direct. Le choix des couples  $(i, m)$ , pour lesquels le temps de vol  $\tau_{im}$  est contaminé, est aléatoire. La géométrie initiale  $(\mathbf{X}^{(0)}, \mathbf{S}^{(0)})$  fournie aux deux algorithmes est construite à partir des positions théoriques perturbées par un bruit. Ce bruit suit une loi normale sur les trois coordonnées, avec un écart-type allant de 0 à 1 m, avec un pas de 10 cm.

Le but de la simulation de Monte-Carlo est avant tout d'évaluer la robustesse vis à vis des réflexions sur le sol. Pour focaliser l'étude sur ce point, dans une série de 100 réalisations :

- les couples  $(i, m)$  contenant une valeur aberrante par réflexion **sont tirés aléatoirement à chaque réalisation** ;
- l'initialisation  $(\mathbf{X}^{(0)}, \mathbf{S}^{(0)})$  est **identique sur les 100 réalisations**.

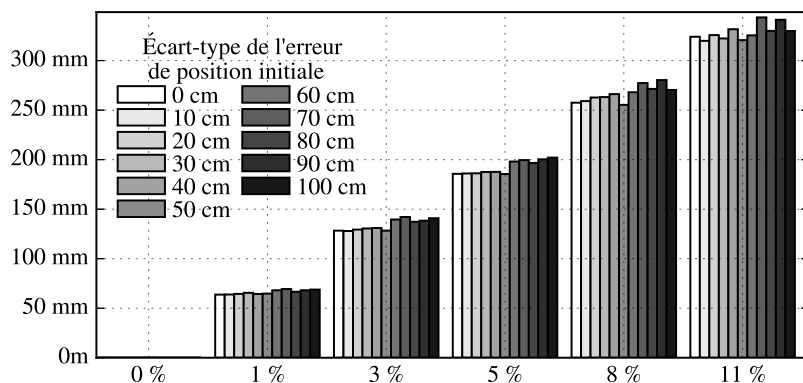
Les différents algorithmes sont finalement exécutés à partir du même jeu de données pour établir la comparaison.

#### 4.2.3.2 Résultats

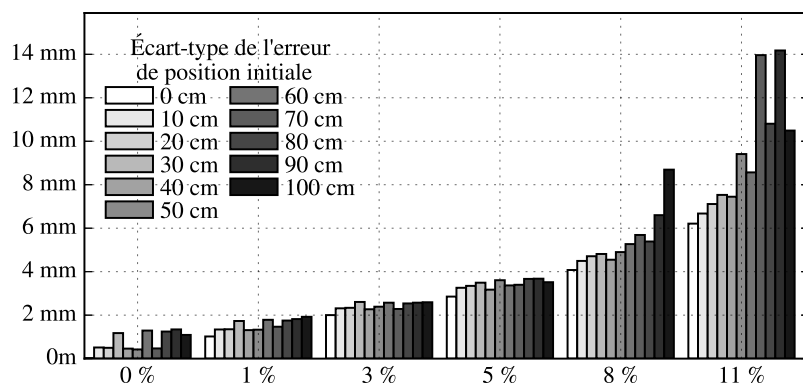
Les choix de paramètres des algorithmes sont fixés sur toute la simulation. Ils sont les suivants :

- l'algorithme FrSM résout l'ensemble des calibrations de façon stable avec un pas  $\alpha = 1.0$ . Le seuil de convergence est fixé à  $\kappa = 10^{-6}$  ;
- pour l'algorithme RMDS, le paramètre de régularisation  $\nu$  est fixé à  $\nu = 6 \times 10^{-5}$  s (équivalent à 0.2 m en distance).

La figure 4.3 représente les moyennes des 100 simulations de Monte-Carlo pour chaque choix de paramètres. Premièrement, on remarque que l'erreur sur la position initiale n'a pratiquement



(a) Résolution avec FrSM, par descente de gradient.

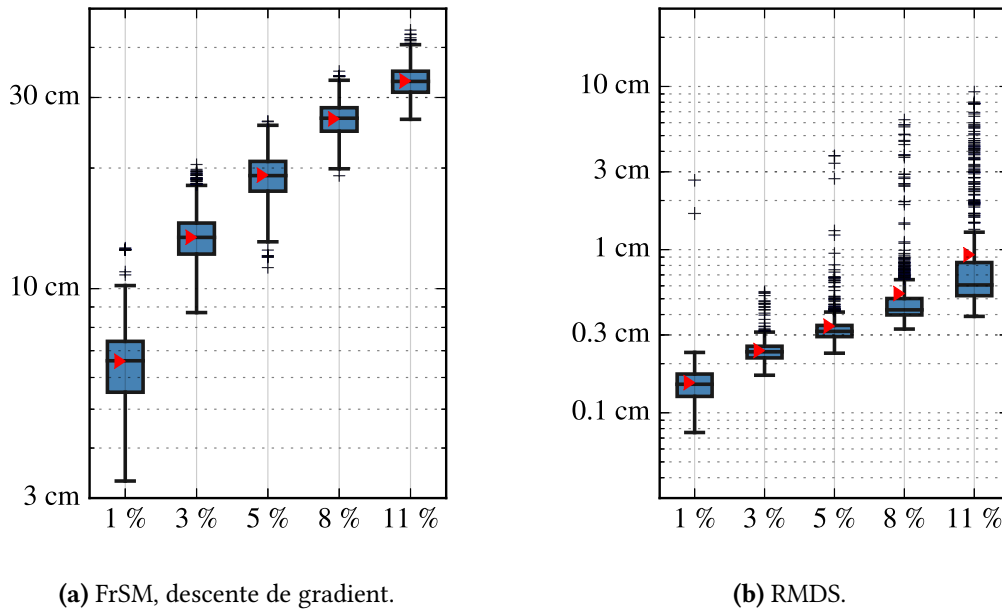


(b) Résolution par RMDS.

**FIGURE 4.3** – Simulation de Monte-Carlo des calibrations géométriques par TOA : erreurs d'estimation dans  $\tilde{\mathbf{X}}$ . Moyennes sur 100 réalisations. De 0 à 11% de valeurs aberrantes sont issues de la réflexion sur le sol. L'écart type de l'erreur dans les positions initiales ( $\mathbf{X}^{(0)}, \mathbf{S}^{(0)}$ ) vont de 0 à 1 m.

pas d'impact sur l'erreur finale d'estimation  $\tilde{\mathbf{X}}$  par FrSM. En effet, pour un pourcentage de valeur aberrantes donné, les valeurs de moyennes sont peu variables. Pour le RMDS, l'influence ne commence à être visible qu'à partir de 8% de valeurs aberrantes. L'inconvénient majeur des algorithmes itératifs employés est qu'ils sont sensibles au minimum locaux, c'est à dire sur la position initiale fournie. L'erreur relative entre la géométrie initiale et la vérité-terrain est un bon indicateur du risque de tomber dans minimum local. Or, la géométrie d'antenne utilisée est grande (20 m  $\times$  2 m). Une erreur d'écart-type de 1 m sur les positions correspond à 19.8% d'erreur relative, ce qui reste relativement faible. Ici, la grande taille de l'antenne devient un atout, car l'erreur relative devient faible devant l'erreur absolue apparente.

En revanche, nous validons la différence nette de performance entre les deux approches en fonction du pourcentage de valeurs aberrantes. Quand 1% des intercorrélations détectent le pic issu de la réflexion sur le sol, l'erreur d'estimation par FrSM est en moyenne de 6.5 cm. De même que dans le chapitre précédent, l'approche classique de minimisation au sens des moindres



**FIGURE 4.4** – Simulation de Monte-Carlo des calibrations géométriques par TOA : erreurs d'estimation dans  $\tilde{\mathbf{X}}$ . Boîtes à moustaches contenant 1000 réalisations, pour toutes les erreurs d'initialisation entre 10 et 100 cm. De 0 à 11 % de valeurs aberrantes sont issues de la réflexion sur le sol.

carrés montre une forte sensibilité aux valeurs aberrantes. Cette tendance est confirmée par l'augmentation de l'erreur en fonction du pourcentage, allant jusqu'à 33 cm pour 11%. À l'inverse, l'erreur d'estimation par RMDS reste relativement faible, inférieure à 1.5 cm, toutes initialisations ( $\mathbf{X}^{(0)}, \mathbf{S}^{(0)}$ ) confondues. L'augmentation légère de l'erreur en fonction du pourcentage est due au paramètre de régularisation, qui est fixe. Comme  $\nu = 6 \times 10^{-5}$  s, les TOAs issus d'une réflexion contenant une erreur inférieure à  $3 \times 10^{-5}$  s ne sont pas identifiés. Donc la régression de la géométrie globale est dégradée. Si le choix de  $\nu$  est optimal (par exemple par la technique de courbe en L, cf. algorithme 2 chapitre 3), alors l'erreur ne doit *a priori* pas augmenter.

Pour analyser l'homogénéité des résultats donnés sur l'ensemble du test de Monte-Carlo, la figure 4.4 trace les boîtes à moustaches des erreurs de l'estimation  $\tilde{\mathbf{X}}$  à chaque réalisation. Une boîte à moustaches correspond à 1000 réalisations pour un pourcentage donné de valeurs aberrantes. Dans une boîte à moustaches :

- la flèche rouge donne la moyenne des 1000 réalisations ;
- la boîte bleue contient 50 % des 1000 valeurs (deuxième et troisième quartiles). Le trait noir inclus dans la boîte définit la valeur médiane de l'ensemble ;
- une moustache (trait vertical) correspond aux premier ou dernier quartile. La taille maximale d'une moustache est de 1.5 fois la taille de la boîte. Les réalisations dépassant ce seuil sont représentées par les croix noires. En limitant la taille d'une moustache, il est possible d'analyser si l'ensemble est dispersé ou homogène, par la présence ou non de croix.

La comparaison des figures 4.4a et 4.4b met en évidence une dispersion relativement plus élevée

par RMDS. En effet, le nombre de points en dehors de la boîte à moustaches est globalement supérieur dans la figure 4.4b. De plus ce phénomène s'accroît fortement quand le pourcentage de valeurs aberrantes augmente. À 11%, la plus grande erreur sur 1000 réalisations est de 9.1 cm. Au delà de 5%, le RMDS présente une dispersion significative dans les erreurs d'estimations. Cependant dans tous les cas, la plus mauvaise estimation par RMDS reste plus fidèle que le meilleur cas par FrSM.

Dans l'ensemble, le RMDS reste bien plus performant que le FrSM quand les réflexions sur le sol brulent les TOAs. Mais au delà de 8% de TOAs bruités, le RMDS est plus sensible aux positions initiales choisies. En pratique, pour réduire ce problème, il est possible de lancer l'algorithme avec différentes conditions initiales, puis de choisir l'estimation associée à la valeur la plus faible de la fonction coût du problème 4.18.

En conclusion, proposer une approche robuste de calibration par TOA s'est avéré être dans le prolongement naturel des travaux présentés au chapitre 3. En effet, elle peut se traduire par un problème de MultiDimensional Scaling, dans lequel les seules distances connues sont celles entre les sources et micros. Il suffit alors de reprendre l'algorithme de RMDS, en adaptant la bonne pondération de distances. Dans le cas de calibration par TDOA, le problème est plus compliqué car nous ne connaissons que des différences de distance. Il faut donc implémenter une autre forme d'algorithme. La deuxième partie de ce chapitre se focalise sur ce problème.

### 4.3 Calibration géométrique par différence de temps de vol

Cette partie explore la dernière stratégie de localisation de microphones, dans laquelle les conditions expérimentales sont quasiment identiques à la stratégie évoquée précédemment. La seule différence est que seuls les signaux mesurés par l'antenne sont utilisés. En pratique, cette technique est intéressante quand il n'est pas possible d'acquérir simultanément les signaux des sources et des microphones.

#### 4.3.1 Le protocole classique

La mesure d'une différence de temps de vol (TDOA) est illustrée dans la figure 4.5. Nous supposons la source  $i$  fixe et deux microphones  $n$  et  $m$ . Notons  $t_{imn}$  la différence de temps de vol entre les microphones  $m$  et  $n$ , issus de la source  $i$  :

$$t_{imn}(\mathbf{X}, \mathbf{S}) = t_{im}(\mathbf{X}, \mathbf{S}) - t_{in}(\mathbf{X}, \mathbf{S}) \quad (4.19)$$

$$= \frac{\|\mathbf{s}_i - \mathbf{x}_m\| - \|\mathbf{s}_i - \mathbf{x}_n\|}{c_0} \quad (4.20)$$

Cette expression est valable dans le cas où le réseau est synchrone. Si le réseau est asynchrone, le décalage temporel de chacune des voies devient une inconnue supplémentaire à déterminer. Pour un réseau de  $M$  microphones, ceci se traduit par  $M$  inconnues supplémentaires. La littérature considère généralement le cas d'un réseau asynchrone [60]. Cependant ce manuscrit se focalise sur l'étude de réseaux synchrones, comme c'est le cas pour le système *Megamicros*, donc les décalages temporels entre les canaux sont nuls.

La variable  $t_{imn}$  peut être estimée par l'intercorrélacion entre les signaux mesurés du couple de microphones  $(m, n)$ . De la même manière que pour estimer les temps de vol, le GCC-PHAT est utilisé pour réduire la dégradation causée par le bruit de mesure et les échos. Dans le cas présent, le GCC-PHAT est calculé à partir de l'interspectre  $S_{mn}$ , et s'écrit :

$$r_{\text{PHAT}}(\tau) = 2 \int_{f_{\min}}^{f_{\max}} \frac{S_{nm}(f)}{|S_{mn}(f)|} e^{j2\pi f \tau} df \quad (4.21)$$

En supposant que les sources sont activées à tour de rôle, l'estimée de la différence de temps de vol  $\tau_{imn}$  associée à la source  $i$  est donnée par le pic d'intercorrélacion :

$$\tau_{imn} = \underset{\tau}{\operatorname{argmax}} r_{\text{PHAT}}(\tau) \quad (4.22)$$

Contrairement à l'approche par TOA, les étapes de calcul de GCC-PHAT et de régression nécessitent plus de calcul, notamment :

- dans le cas TOA, le calcul global des interspectres  $S_{mn}$  et des  $\tau_{imn}$  a une complexité  $\mathcal{O}(IM)$ ;
- dans le cas TDOA, le calcul des interspectres  $S_{mn}$  et des  $\tau_{imn}$  a une complexité  $\mathcal{O}(IM^2)$ .

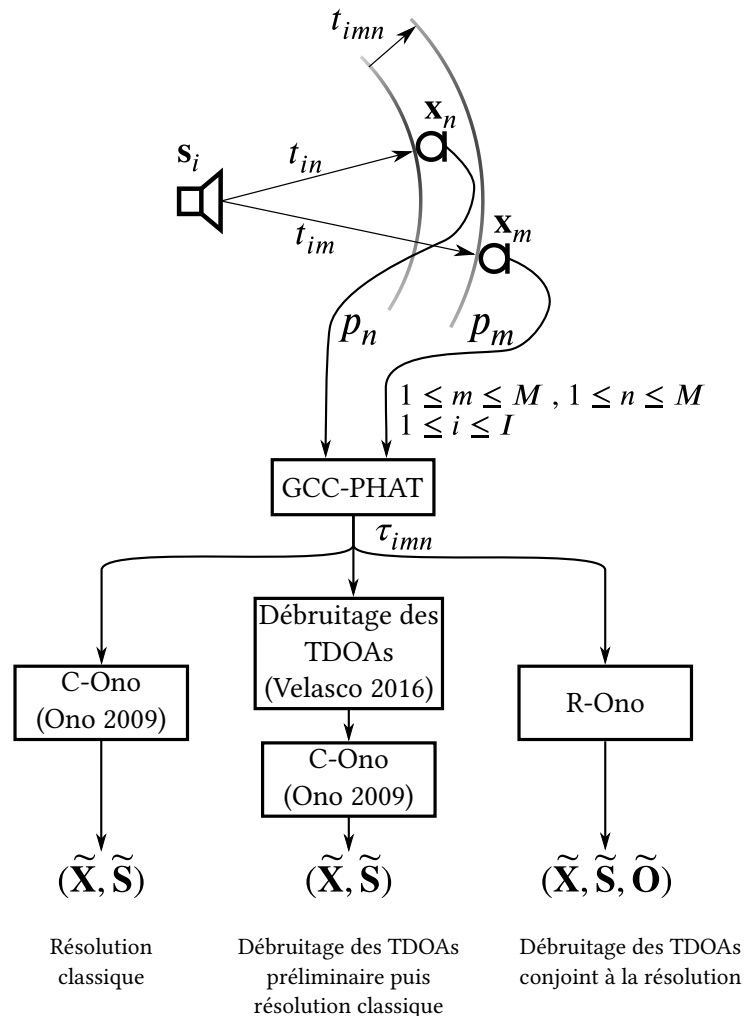
Appliqué à un réseau à grand nombre de microphones, le deuxième cas devient beaucoup plus long. Pour réduire le coût de calcul, notons qu'il existe une asymétrie des différences de temps de vol :  $\tau_{imn} = -\tau_{inm}$ . Donc, il est possible de diviser le temps de calcul par deux.

L'estimation géométrique au sens des moindres carrés correspond à la régression du modèle (en équation 4.19) sur l'observable donné par la relation 4.22. Comme  $t_{imn}$  est fonction à la fois des positions de source et de micros, la calibration géométrique nécessite une estimation couplée des deux inconnues  $\mathbf{X}$  et  $\mathbf{S}$ , ce qui donne la forme du problème classique :

$$(\tilde{\mathbf{X}}, \tilde{\mathbf{S}}) = \underset{\mathbf{X}, \mathbf{S}}{\operatorname{argmin}} \sum_{i=1}^I \sum_{m,n=1}^M [\tau_{imn} - t_{imn}(\mathbf{X}, \mathbf{S})]^2 \quad (4.23)$$

La minimisation de cette fonction non linéaire pourrait être résolue par une descente de gradient classique, mais souffre d'une sensibilité forte aux minima locaux. Face à cette contrainte, Ono *et al.* exposent une résolution par Minimisation Maximisation (MM) [110]. Le principe et l'application à ce problème seront détaillés en section 4.3.3. Par la suite, nous appellerons C-Ono l'algorithme proposé dans [110], qui sera utilisé comme référence de l'état de l'art dans les simulations.

Pour permettre la résolution de la relation 4.23, l'unique condition concerne le nombre minimum



**FIGURE 4.5** – Schéma de la méthode par différence de temps de vol, par approche classique (gauche) ou robuste (milieu et droite).

de microphones et de sources ( $(I - 4)(M - 4) \geq 9$  [110]). En pratique, plus grands sont  $I$  et  $M$ , plus l'erreur d'estimation diminue. Il a été démontré expérimentalement que l'erreur d'estimation réduit jusqu'à ce que  $I + M \approx 100$  avec C-Ono [86]. Au-delà, la diminution de l'erreur devient négligeable.

Globalement, l'approche par TDOA semble proche de l'approche TOA. En outre, dans une application avec un grand nombre de microphones, il s'agit d'un problème surdéterminé. En effet, l'observable connu augmente proportionnellement à  $IM^2$  alors que le nombre d'inconnues géométriques est  $3I + 3M$ . Une redondance importante existe dans l'observable connu  $\{\tau_{imn}\}$ . Nous verrons que cette redondance est exploitée dans la première approche robuste de calibration par TDOA, proposée ci-dessous.

### 4.3.2 Première approche robuste : débruitage préalable de l'observable

Nous proposons une première approche de calibration par TDOA robuste aux réflexions. L'objectif de cette première approche est d'utiliser uniquement des algorithmes de l'état de l'art, dont les performances sont déjà connues et validées. Illustrée en figure 4.5, elle fonctionne en deux étapes :

1. la première étape consiste à débruiter l'ensemble de l'observable  $\{\tau_{imn}\}$ . En pratique, elle implique le débruitage disjoint des  $I$  sous-ensembles  $\{\tau_{1mn}\}, \dots, \{\tau_{imn}\}, \dots, \{\tau_{Imn}\}$  ;
2. la deuxième étape résout la calibration géométrique par une méthode de résolution classique, à partir de l'observable débruité. Dans un souci d'établir une comparaison pertinente dans les prochaines validations expérimentales, nous choisissons C-Ono, énoncé dans la section précédente.

Pour apporter de la robustesse, nous conservons toujours le même paradigme consistant à ajouter un terme de bruit, inconnu et à déterminer. Pour cela la différence de temps de vol observée  $\tau_{imn}$  se décrit par :

$$\tau_{imn} = t_{imn}(\mathbf{X}, \mathbf{S}) + \epsilon_{imn} + o_{imn} \quad (4.24)$$

dans lequel  $\epsilon_{imn}$  est un bruit de faible amplitude suivant une loi de distribution centrée, et  $o_{imn}$  le terme d'erreur contenu dans une valeur aberrante. Pour une source  $i$  donnée, cette égalité peut être exprimée de manière équivalente sous forme matricielle :

$$\mathbf{T}_i = \mathbf{T}_i + \mathbf{E}_i + \mathbf{O}_i \quad (4.25)$$

les quatre matrices appartenant à l'ensemble des matrices  $M \times M$  antisymétriques à valeurs diagonales nulles, noté  $\mathbb{A}_h^M$ . Elles sont construites telles que  $[\mathbf{T}_i]_{mn} = \tau_{imn}$ ,  $[\mathbf{T}_i]_{mn} = t_{imn}$ ,  $[\mathbf{E}_i]_{mn} = \epsilon_{imn}$  et  $[\mathbf{O}_i]_{mn} = o_{imn}$ .

Velasco *et al.* [154] démontre que la somme 4.25 est une décomposition de  $\mathbf{T}_i$  en la somme :

- d'une **matrice  $\mathbf{T}_i$**  de rang égal à 2, donc **de rang faible** ( $2 \ll M$ ),
- d'une **matrice  $\mathbf{O}_i$  parcimonieuse**,
- et d'une matrice de bruit résiduel  $\mathbf{E}_i$ .

L'analyse en composantes principales robuste (RPCA) est un ensemble de techniques permettant de décomposer  $\mathbf{T}_i$ , à partir de l'unique hypothèse de rang faible de  $\mathbf{T}_i$  et de parcimonie sur  $\mathbf{O}_i$ . Elle est souvent étudiée pour les applications au traitement vidéo [22]. La décomposition découle de la résolution du problème suivant :

$$\begin{aligned} (\tilde{\mathbf{T}}_i, \tilde{\mathbf{O}}_i) = & \underset{\mathbf{T}_i \in \mathbb{A}_h^M, \mathbf{O}_i \in \mathbb{A}_h^M}{\operatorname{argmin}} \quad \|\mathbf{T}_i + \mathbf{O}_i - \mathbf{T}_i\|_F^2 \\ & \text{tel que} \quad \|\mathbf{O}_i\|_0 < 2K \text{ et } \operatorname{rang}(\mathbf{T}_i) = 2 \end{aligned} \quad (4.26)$$

Généralement, les techniques de RPCA sont conçues pour trouver *une* approximation fidèle de  $\mathbf{T}_i$ ,

ayant le rang le plus faible. Or ici, le rang de  $\mathbf{T}_i$  est *a priori* connu. Dans un tel cas, Zhou *et al.* [171] montrent que la formulation 4.26 est équivalente au problème alterné :

$$\begin{cases} \mathbf{T}_i^{(k+1)} = \underset{\mathbf{T}_i \in \mathbb{A}_h^M}{\operatorname{argmin}} \left\| \mathbf{T}_i + \mathbf{O}_i^{(k)} - \mathbf{T}_i \right\|_F^2 & \text{tel que } \operatorname{rang}(\mathbf{T}_i) = 2 \\ \mathbf{O}_i^{(k+1)} = \underset{\mathbf{O}_i \in \mathbb{A}_h^M}{\operatorname{argmin}} \left\| \mathbf{T}_i^{(k+1)} + \mathbf{O}_i - \mathbf{T}_i \right\|_F^2 & \text{tel que } \|\mathbf{O}_i\|_0 < 2K \end{cases} \quad (4.27)$$

où  $K$  est le nombre d'éléments non-nuls de  $\mathbf{O}_i$  tel que  $2K \ll M^2$  (condition de parcimonie). Si il existe une forme explicite pour exprimer  $\mathbf{T}_i^{(k+1)}$ , Zhou *et al.* démontrent que le problème alterné converge (soit vers un minimum local soit vers le minimum global).

Velasco *et al.* propose et démontre [154] une forme explicite de  $\mathbf{T}_i^{(k+1)}$ , telle que :

$$\mathbf{T}_i^{(k+1)} = \frac{(\mathbf{T}_i - \mathbf{O}_i^{(k)}) \mathbf{1}_M \mathbf{1}_M^\top + \mathbf{1}_M \mathbf{1}_M^\top (\mathbf{T}_i - \mathbf{O}_i^{(k)})}{M} \quad (4.28)$$

où  $\mathbf{1}_M$  est un vecteur colonne à  $M$  éléments égaux à 1.

Ensuite,  $\mathbf{O}_i^{(k+1)}$  est calculé tel que seuls les  $2K$  éléments de plus grande amplitude soient conservés, tous les autres éléments étant forcés à 0. Cette opération, appelée seuillage dur (*hard thresholding*), donne l'expression de  $\mathbf{O}_i^{(k+1)}$  :

$$\mathbf{O}_i^{(k+1)} = \mathcal{H}_{2K}(\mathbf{T}_i - \mathbf{T}_i^{(k+1)}) \quad (4.29)$$

$\mathcal{H}_{2K}(\mathbf{A})$  étant défini tel que :

$$[\mathcal{H}_{2K}(\mathbf{A})]_{nm} = \begin{cases} [\mathbf{A}]_{nm} & \text{pour les } 2K \text{ éléments de } \mathbf{A} \text{ de plus grande amplitude} \\ 0 & \text{sinon} \end{cases} \quad (4.30)$$

Cependant le choix d'un seuillage dur, énoncé par Zhou et Velasco, nécessite une connaissance au moins approximative du degré de parcimonie de  $\mathbf{O}_i$ . Ce n'est pas le cas d'une matrice  $\mathbf{T}_i$  obtenue expérimentalement. En effet le choix de  $K$  dépend de la présence plus ou moins forte d'échos, donc du cadre expérimental. Un modèle prédisant le choix de  $K$  en fonction de l'environnement semble compliqué. De plus, en présence d'un bruit résiduel  $\mathbf{E}_i$  sur tous les éléments de la matrice de TDOAs  $\mathbf{T}_i$ , la matrice de valeurs aberrantes  $\mathbf{O}_i^{(k+1)}$  contiendra toujours  $2K$  éléments non nuls, quelle que soit la valeur de  $K$  choisie. Une étude paramétrique sur  $K$  n'est donc pas envisageable.

En s'inspirant de [172], nous proposons de changer la contrainte sur le degré de parcimonie par une régularisation sur la norme  $\ell_1$  de  $\mathbf{O}_i$  par le paramètre  $\nu$  :



$$\begin{aligned}
(\tilde{\mathbf{T}}_i, \tilde{\mathbf{O}}_i) = & \underset{\mathbf{T}_i \in \mathbb{A}_h^M, \mathbf{O}_i \in \mathbb{A}_h^M}{\operatorname{argmin}} \quad \|\mathbf{T}_i + \mathbf{O}_i - \mathbf{T}_i\|_F^2 + \nu \|\mathbf{O}_i\|_1 \\
& \text{tel que} \quad \operatorname{rang}(\mathbf{T}_i) = 2
\end{aligned} \tag{4.31}$$

Ce simple changement modifie uniquement la mise à jour  $\mathbf{O}_i^{(k+1)}$ , en remplaçant l'opérateur de seuillage dur par l'opérateur de seuillage doux  $\mathcal{S}_\nu$ . Ce dernier a déjà été utilisé pour le RMDS dans le chapitre précédent. La mise à jour de chaque élément de  $\mathbf{O}_i$  est faite par l'opération :

$$o_{imn}^{(k+1)} = \mathcal{S}_\nu \left( \tau_{imn} - t_{imn}^{(k+1)} \right) \tag{4.32}$$

l'opérateur  $\mathcal{S}_\nu$  étant défini en équation 3.27 chapitre 3. Dans ce cas,  $\nu$  peut être défini de manière optimale par une étude paramétrique<sup>1</sup>. Le débruitage robuste est finalement résumé en algorithme 4. Appliqué sur les  $I$  matrices  $\mathbf{T}_i$ , nous obtenons la totalité des différences de temps de vol débruitées  $\tilde{\mathbf{T}}_1, \dots, \tilde{\mathbf{T}}_I, \dots, \tilde{\mathbf{T}}_I$ .

---

**ALGORITHME 4** Débruitage de différences de temps de vol par décomposition de matrice, fondé sur Velasco *et al.* [154] et Zhou *et al.* [172].

---

**Entrées :**  $\nu, \kappa$ , Initialiser  $\mathbf{O}^{(0)}$  à zéro puis  $\mathbf{T}_i^{(0)} = \mathbf{T}_i$

**for**  $k = 0, 1, 2, \dots$  **do**

    Calculer  $\mathbf{T}_i^{(k+1)}$  via (4.28)

    Calculer  $\mathbf{O}_i^{(k+1)}$  par seuillage doux via (4.32)

**if**  $\frac{\|\mathbf{T}_i^{(k+1)} - \mathbf{T}_i^{(k)}\|_F^2}{\|\mathbf{T}_i^{(k)}\|_F^2} < \kappa$  **then**

        Arrêt de l'algorithme

**end if**

**end for**

**Sorties :**  $(\tilde{\mathbf{T}}_i, \tilde{\mathbf{O}}_i) = (\mathbf{T}_i^{(k+1)}, \mathbf{O}_i^{(k+1)})$

---

En résumé, l'approche robuste proposée suit deux étapes successives. La première étape implique le débruitage disjoint et indépendant de  $I$  matrices de différences de temps de vol. Les  $IM^2$  différences de temps de vol exemptes de valeurs aberrantes sont ensuite utilisées pour une résolution classique énoncée en section 4.3.1, via C-Ono. Bien que cette stratégie semble intuitive, elle représente en totalité la résolution du problème par l'assemblage de  $I + 1$  algorithmes indépendants. La partie suivante tente de simplifier la structure globale, en couplant la calibration géométrique et le débruitage sous un unique problème.

---

1. Cette technique est expliquée à la fin de l'annexe B, et a été employée dans le chapitre 3 section 3.5.3

### 4.3.3 Deuxième approche robuste : calibration conjointe à l'identification des observables bruités

Dans cette section, nous proposons une approche unifiée, ne consistant *in fine* qu'en un seul algorithme de résolution, que nous appellerons R-Ono. Pour cela, il faut identifier à la fois la géométrie du réseau et les valeurs aberrantes de l'ensemble des TDOAs. Pour parvenir à ce résultat, nous nous inspirons du processus de résolution du RMDS [59], ou encore de She *et al.* résolvant un problème linéaire à partir de mesures contenant des valeurs aberrantes [131]. Il est possible de construire une fonction coût pour résoudre la calibration géométrique par TDOA, semblable à l'équation 4.23, et enlevant la contribution des termes de bruit  $o_{inn}$  de l'erreur quadratique :

$$\begin{aligned} (\tilde{\mathbf{X}}, \tilde{\mathbf{S}}, \tilde{\mathbf{O}}) &= \underset{\mathbf{X}, \mathbf{S}, \mathbf{O}}{\operatorname{argmin}} f_{\text{TDOA}}(\mathbf{X}, \mathbf{S}, \mathbf{O}) \\ \text{tel que } f_{\text{TDOA}}(\mathbf{X}, \mathbf{S}, \mathbf{O}) &= \sum_{i=1}^I \sum_{m,n=1}^M [\tau_{inn} - t_{inn}(\mathbf{X}, \mathbf{S}) - o_{inn}]^2 + \nu \sum_{i=1}^I \sum_{m,n=1}^M |o_{inn}| \end{aligned} \quad (4.33)$$

Dans le cas présent la norme  $\ell_1$  régularisée par le paramètre  $\nu$ , contrôle la parcimonie de la matrice à trois dimensions  $\mathbf{O} = [\mathbf{O}_1, \dots, \mathbf{O}_I, \dots, \mathbf{O}_I]$  de taille  $I \times M \times M$ . À titre de comparaison, dans la section précédente la régularisation intervenant dans le débruitage contrôlait la parcimonie des matrices  $\mathbf{O}_i$  de manière disjointe. Le problème 4.33 associe dans la même opération la régularisation, sur l'ensemble des termes  $o_{inn}$ .

La fonction  $f_{\text{TDOA}}(\mathbf{X}, \mathbf{S}, \mathbf{O})$  couple un problème au sens des moindres carrés (terme de gauche) et de Lasso (norme  $\ell_1$  régularisée). Pour résoudre le problème 4.33, nous proposons de nous inspirer des approches qui ont été exposées jusque là dans le manuscrit [59, 131, 172], par une minimisation alternée de deux sous-problèmes :

$$(\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)}) = \underset{\mathbf{X}, \mathbf{S}}{\operatorname{argmin}} f_{\text{TDOA}}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}, \mathbf{O}^{(k)}) \quad (4.34)$$

$$\mathbf{O}^{(k+1)} = \underset{\mathbf{O}}{\operatorname{argmin}} f_{\text{TDOA}}(\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)}, \mathbf{O}^{(k)}) \quad (4.35)$$

Cette approche a l'avantage majeur de conduire à un algorithme itératif simple, comprenant deux étapes successives. La première fait intervenir la minimisation au sens des moindres carrés classique. Il est *a priori* possible d'utiliser toute résolution déjà proposée dans l'état de l'art. Nous choisissons l'approche MM utilisé par Ono *et al.* [110], en ajoutant les termes d'erreur inconnus  $o_{inn}$ . La deuxième étape met à jour la matrice des grandes erreurs  $\mathbf{O}^{(k+1)}$ , de dimension 3, en minimisant sa norme  $\ell_1$  régularisée.

### Première étape : mise à jour de $\mathbf{X}$ et $\mathbf{S}$

La Majoration Minimisation est une technique qui permet l'optimisation d'une fonction coût non-convexe. Généralement, elle s'appuie sur la construction d'une approximation quadratique (que nous notons  $g_{\text{TDOA}}$ ), majorant la fonction  $f_{\text{TDOA}}$ . La Majoration Minimisation est exploitée de manière itérative. À chaque itération  $k$ ,  $g_{\text{TDOA}}$  est construite telle que :

- (a)  $g_{\text{TDOA}}$  est une majorante :  $g_{\text{TDOA}} \geq f_{\text{TDOA}} \forall (\mathbf{X}, \mathbf{S})$
- (b)  $g_{\text{TDOA}}$  est tangente à la fonction coût  $f_{\text{TDOA}}$  au point  $f_{\text{TDOA}}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}, \mathbf{O}^{(k)})$
- (c)  $g_{\text{TDOA}}$  est quadratique sur chaque position de micro  $\mathbf{x}_i$  et de source  $\mathbf{s}_i$ .

Ono *et al.* démontrent que la fonction

$$g_{\text{TDOA}}(\mathbf{X}, \mathbf{S}, \boldsymbol{\mu}, \mathbf{e}) = \sum_{i=1}^I \sum_{m,n=1}^M \|\mathbf{s}_i - \mathbf{x}_m - \mathbf{e}_{im} \mu_{mimn}\|^2 + \|\mathbf{s}_i - \mathbf{x}_n - \mathbf{e}_{in} \mu_{nimn}\|^2 \quad (4.36)$$

suit les trois conditions (a), (b) et (c).  $\boldsymbol{\mu}$  et  $\mathbf{e}$  sont deux variables auxiliaires constantes, contenant respectivement les vecteurs  $\mathbf{e}_{im}$ ,  $\mathbf{e}_{in}$  et les éléments  $\mu_{mimn}$ ,  $\mu_{nimn}$ . Elles sont calculées au début de chaque itération  $k$ , tel que :

$$\mathbf{e}_{im} = \frac{\mathbf{s}_i^{(k)} - \mathbf{x}_m^{(k)}}{\|\mathbf{s}_i^{(k)} - \mathbf{x}_m^{(k)}\|} \quad (4.37)$$

$$\mathbf{e}_{in} = \frac{\mathbf{s}_i^{(k)} - \mathbf{x}_n^{(k)}}{\|\mathbf{s}_i^{(k)} - \mathbf{x}_n^{(k)}\|} \quad (4.38)$$

$$\mu_{mimn} = \|\mathbf{s}_i^{(k)} - \mathbf{x}_m^{(k)}\| - \frac{1}{2} \varepsilon_{imn} \quad (4.39)$$

$$\mu_{nimn} = \|\mathbf{s}_i^{(k)} - \mathbf{x}_n^{(k)}\| + \frac{1}{2} \varepsilon_{imn} \quad (4.40)$$

$\varepsilon_{imn}$  étant le terme de l'erreur quadratique dans la fonction coût  $f_{\text{TDOA}}$ , associé aux indices de somme  $(i, m, n)$ , et à l'itération  $k$ . Pour que  $g_{\text{TDOA}}$  soit une majorante de  $f_{\text{TDOA}}(\mathbf{X}, \mathbf{S}, \mathbf{O})$ , il faut inclure la présence du terme supplémentaire  $o_{imn}$ , ce qui donne :

$$\varepsilon_{imn} = [\tau_{imn} - t_{imn}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}) - o_{imn}^{(k)}]^2 \quad (4.41)$$

À partir de l'expression de  $g_{\text{TDOA}}$ , il est facile de trouver son minimum car elle est quadratique sur les termes  $\mathbf{x}_m$  et  $\mathbf{s}_i$ . Les solutions doivent vérifier que les dérivées  $\frac{\partial g_{\text{TDOA}}}{\partial \mathbf{x}_m}$  et  $\frac{\partial g_{\text{TDOA}}}{\partial \mathbf{s}_i}$  sont nulles. Cette condition mène finalement aux règles de mises à jour de l'approche MM :

$$\mathbf{x}_m^{(k+1)} = \frac{1}{I} \left( \sum_{i=1}^I \mathbf{s}_i^{(k)} - \frac{1}{M} \mathbf{e}_{im} \sum_{m'=1}^M \mu_{nim'n} \right) \quad (4.42)$$

$$\mathbf{s}_i^{(k+1)} = \frac{1}{M} \left( \sum_{m=1}^M \mathbf{x}_m^{(k)} + \frac{1}{M} \mathbf{e}_{im} \sum_{n=1}^M \mu_{mimn} \right) \quad (4.43)$$

Enfin, notons que les variables  $\mathbf{x}_m$  et  $\mathbf{x}_n$ , présentes dans la fonction  $g_{\text{TDOA}}$ , représentent  $\mathbf{X}$  de façon identique. Il est d'ailleurs possible de montrer que la solution  $\mathbf{x}_m$  vérifiant  $\frac{\partial g_{\text{TDOA}}}{\partial \mathbf{x}_m} = 0$  mène à la même expression que celle de  $\mathbf{x}_n$  vérifiant  $\frac{\partial g_{\text{TDOA}}}{\partial \mathbf{x}_n} = 0$ .

### Deuxième étape : mise à jour de $\mathbf{O}$

Le choix du problème de type Lasso sur  $\mathbf{O}$  permet de conserver une formulation homogène avec celle du premier problème robuste par TDOA. En effet, la mise à jour implique de nouveau l'opérateur de seuillage doux, fonction du paramètre de régularisation  $\nu$  de la norme  $\ell_1$  de  $\mathbf{O}$ . Le sous-problème de Lasso 4.35 faisant suite au sous-problème MM 4.34, les nouvelles mises à jour  $\mathbf{X}^{(k+1)}$  et  $\mathbf{S}^{(k+1)}$  sont utilisées pour obtenir  $\mathbf{O}^{(k+1)}$ , tel que :

$$o_{imn}^{(k+1)} = S_\nu (\tau_{imn} - t_{imn}(\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)})) \quad (4.44)$$

On notera que la séquence des erreurs résiduelles  $\{f_{\text{TDOA}}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}, \mathbf{O}^{(k)})\}$  est non-croissante en fonction de l'indice d'itération  $k$ . Ceci est vérifié car la minimisation alternée du problème MM et du problème de Lasso donne successivement :

$$f_{\text{TDOA}}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}, \mathbf{O}^{(k)}) = g_{\text{TDOA}}(\mathbf{X}^{(k)}, \mathbf{S}^{(k)}, \boldsymbol{\mu}, \mathbf{e}) \quad (4.45)$$

$$\geq g_{\text{TDOA}}(\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)}, \boldsymbol{\mu}, \mathbf{e}) \quad (4.46)$$

$$\geq f_{\text{TDOA}}(\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)}, \mathbf{O}^{(k)}) \quad (4.47)$$

$$\geq f_{\text{TDOA}}(\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)}, \mathbf{O}^{(k+1)}) \quad (4.48)$$

Ceci garantit la convergence, vers le minimum global ou un minimum local. Finalement on utilise le même critère de convergence que précédemment, qui porte sur la géométrie  $\tilde{\mathbf{X}}$  de l'antenne. Ce deuxième processus itératif robuste est synthétisé par l'algorithme 5 intitulé R-Ono.

En conclusion, l'algorithme R-Ono que je propose est une version *enrichie* de C-Ono, faisant intervenir le modèle de bruit étendu 4.24. En effet, en supprimant l'élément  $o_{imn}$  dans le terme de bruit résiduel 4.41, ainsi que la mise à jour de  $\mathbf{O}$  en équation 4.44, nous retrouvons l'algorithme original C-Ono. De la même manière, on peut constater que le RMDS est une version enrichie de Smacof, qui est une méthode déjà existante du MDS [45, 59].

---

**ALGORITHME 5** R-Ono : algorithme proposé de calibration géométrique robuste par TDOA.

---

**Entrées :**  $\{\tau_{imm}\}$ ,  $\mathbf{X}^{(0)}$ ,  $\mathbf{S}^{(0)}$ ,  $\mathbf{O}^{(0)}$  initialisé par des zéros,  $\nu$ ,  $\kappa$

**for**  $k = 0, 1, 2, \dots$  **do**

Calculer la variable auxiliaire  $\boldsymbol{\mu}$  via eqs. 4.39, 4.40 et 4.41

Calculer la variable auxiliaire  $\mathbf{e}$  via eq. 4.37

Mise à jour  $\mathbf{X}^{(k+1)}$  via eq. 4.42

Mise à jour  $\mathbf{S}^{(k+1)}$  via eq. 4.43

Mise à jour  $\mathbf{O}^{(k+1)}$  via eq. 4.44

**if**  $\frac{\|\mathbf{X}^{(k+1)} - \mathbf{X}^{(k)}\|_F^2}{\|\mathbf{X}^{(k+1)}\|_F^2} < \kappa$  **then**

Arrêt de l'algorithme

**end if**

**end for**

**Sorties :**  $(\tilde{\mathbf{X}}, \tilde{\mathbf{S}}, \tilde{\mathbf{O}}) = (\mathbf{X}^{(k+1)}, \mathbf{S}^{(k+1)}, \mathbf{O}^{(k+1)})$

---

### 4.3.4 Expérience numérique et comparaison

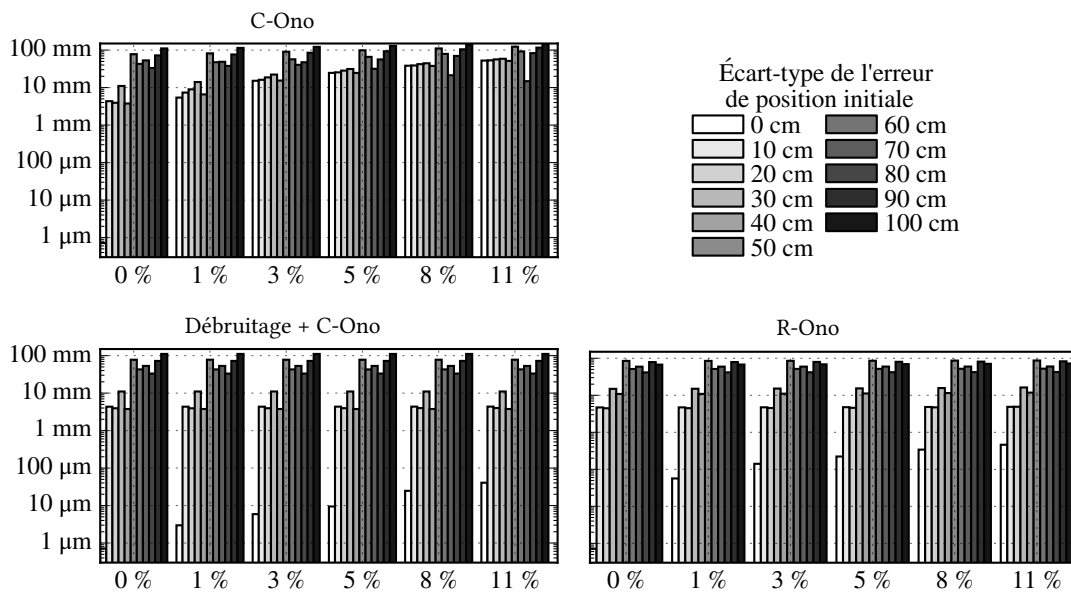
#### 4.3.4.1 La simulation de Monte-Carlo

Nous cherchons à valider l'intérêt d'une approche robuste de calibration géométrique par TDOA, en présence d'échos parasitant la mesure des  $\tau_{imm}$ . Pour cela, nous comparons les performances des trois méthodes énoncées précédemment, résumées en figure 4.5. Elle sont :

- la méthode "classique" énoncée section 4.3.1, par **C-Ono**,
- la première méthode robuste, proposée dans la section 4.3.2, par débruitage *préalable* de l'ensemble  $\{\tau_{imm}\}$  suivit de la méthode classique : **Débruitage+C-Ono**,
- la deuxième méthode robuste, proposée en section 4.3.3 par débruitage *conjoint* à la calibration géométrique : **R-Ono**.

Une simulation de Monte-Carlo est mise en oeuvre. On suit la même démarche que la simulation numérique dans le cas des méthodes par TOA, en section 4.2.3. L'objectif est de réaliser une étude paramétrique, (a) sur différents pourcentages de valeurs aberrantes, puis (b) pour des positions initiales  $(\mathbf{X}^{(0)}, \mathbf{S}^{(0)})$  plus ou moins erronées par rapport à la vérité-terrain. Les conditions restent identiques, à cela près que :

1. l'observable contaminé contient des TDOAs. Dans ce cas, il existe différents cas possibles de réflexions parasites comme le montre la figure 4.1c. Pour une valeur de  $\tau_{imm}$  contaminée, le choix de réflexion parasite est aléatoire ;
2. comme la complexité de calcul est plus élevée ( $\mathcal{O}(IM^2)$ ), nous réduisons le nombre de microphones à  $M = 64$ . À partir de l'antenne schématisée en figure 4.8b, nous sélectionnons 1 microphone sur 4 pour conserver l'étendue globale de 19.6 m.



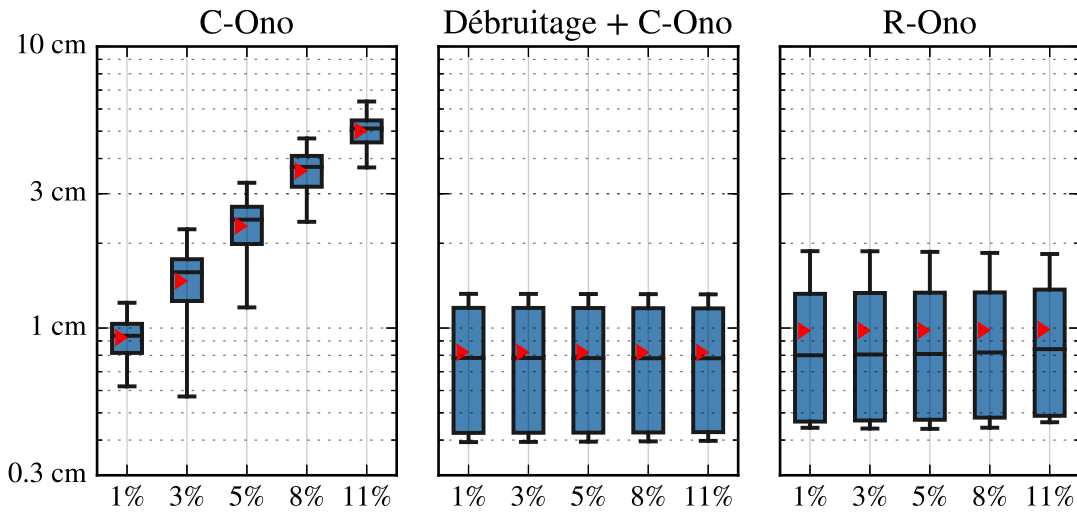
**FIGURE 4.6** – Simulation de Monte-Carlo des calibrations géométriques par TDOA : erreurs d’estimation de  $\tilde{\mathbf{X}}$ . Moyennes sur 100 réalisations. De 0 à 11% de valeurs aberrantes sont issues de la réflexion sur le sol. L’écart type de l’erreur dans les positions initiales ( $\mathbf{X}^{(0)}$ ,  $\mathbf{S}^{(0)}$ ) vont de de 0 à 1 m.

#### 4.3.4.2 Résultats

Pour établir une comparaison homogène des deux méthodes robustes Débruitage+C-Ono et R-Ono, le paramètre de régularisation choisi est le même, et vaut  $\nu = 1.17 \times 10^{-5}$  s (équivalent à 4 mm de distance pour une célérité  $c_0 = 340 \text{ m s}^{-1}$ ). De même, le critère d’arrêt vaut  $\kappa = 10^{-6}$  pour l’ensemble des algorithmes employés.

La figure 4.6 donne les moyennes des séries de 100 réalisations pour chaque jeu de paramètre. Tout d’abord, on remarque que l’estimation est très sensible à la condition initiale. Au delà de 40 cm sur l’erreur de position initiale, l’impact est visible sur les trois méthodes. Comparé aux méthodes par TOA, la résolution par TDOA est relativement sensible aux minimums locaux. En revanche pour une erreur d’initialisation inférieure ou égale à 40 cm, les erreurs d’estimation sont notablement réduites avec les deux méthodes robustes. Avec 11% de valeurs aberrantes, C-Ono estime en moyenne  $\mathbf{X}$  avec 5.7 cm d’erreur, tandis que Débruitage+C-Ono et R-Ono réduisent respectivement l’erreur à 1.3 cm et 1.9 cm.

Pour analyser l’homogénéité de la simulation de Monte-Carlo, des boites à moustaches sont tracées en figure 4.7 en fonction du pourcentage de valeurs aberrantes. Pour réduire l’impact du à la position initiale, chaque boite à moustache réduit les 400 réalisations pour des erreurs d’initialisation entre 10 et 40 cm. Nous voyons clairement l’amélioration apportée par les méthodes robustes. En augmentant le pourcentage de valeurs aberrantes, l’erreur d’estimation par C-Ono est croissante, allant jusqu’à 5.5 cm pour 11%. En revanche, les estimations géométriques données par deux algorithmes robustes ont la même erreur quelle que soit le pourcentage. Par Débruitage+C-



**FIGURE 4.7** – Simulation de Monte-Carlo des calibrations géométriques par TDOA : erreurs d'estimation dans  $\tilde{\mathbf{X}}$ . Boîtes à moustaches contenant 400 réalisations, pour des erreurs d'initialisation entre 10 et 40 cm. De 0 à 11 % de valeurs aberrantes sont issues de la réflexion sur le sol.

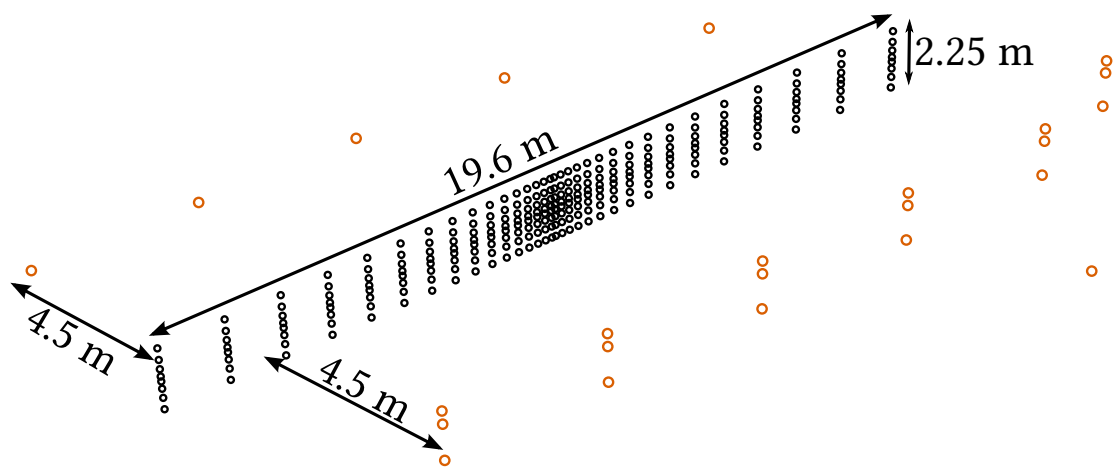
Ono, la moyenne sur les 400 réalisations est de 8.1 mm, légèrement meilleure qu'avec R-Ono (9.8 mm).

Enfin, la hauteur des boîtes à moustache met en évidence une bonne homogénéité par les trois méthodes. En effet, la dispersion n'est pas critique puisque aucune réalisation ne se situe en dehors des moustaches. Notons toutefois qu'elle est légèrement plus élevée en faisant appel à une méthode robuste.

En conclusion, l'expérience numérique a montré que les deux méthodes robustes réduisent significativement l'erreur d'estimation géométrique par rapport à C-Ono. Nous obtenons des résultats légèrement meilleurs par débruitage préalable (Débruitage+C-Ono) que par débruitage conjoint à la calibration (R-Ono). Cependant, l'erreur initiale de position est critique dans un problème de TDOA. Au delà d'une erreur de 50 cm à l'initialisation, l'estimation géométrique se dégrade, avec les trois méthodes. Sur ce point, l'approche par TOA est meilleure l'approche par TDOA.



(a) L'antenne de 256 microphones sur le circuit.



(b) Configuration géométrique de l'expérience de calibration géométrique. Microphones en noir, sources en orange.

FIGURE 4.8 – Le système *Megamicros* installé pour une série de mesure de bruit au passage de véhicules.



#### 4.4 Analyse globale des approches dans une expérience *in situ*

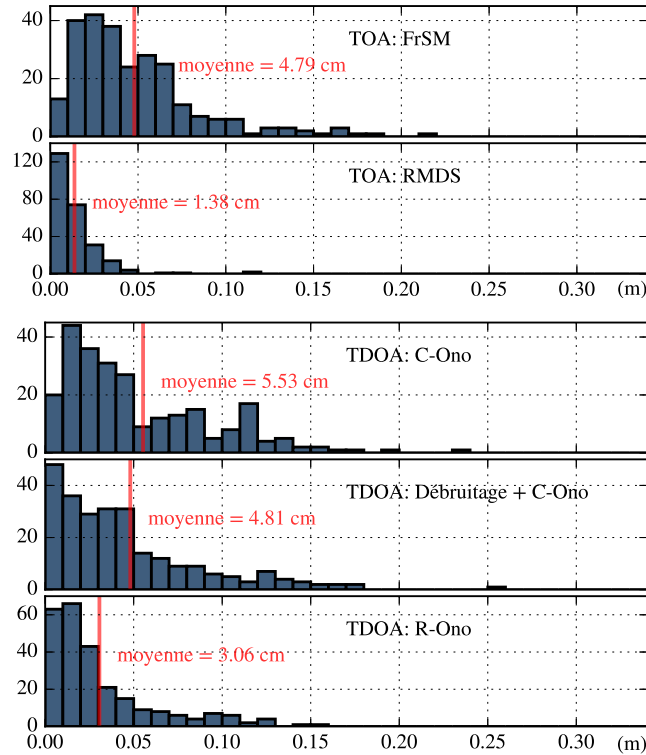
Cette dernière expérience nous permet de tester les 5 méthodes de calibration géométrique énoncées dans ce chapitre, à partir d'un jeu de données réel. Le but est de retrouver la géométrie du réseau de 256 microphones dans la photo 4.8a, à partir de signaux acoustiques mesurés. Notons que l'objectif initial de l'expérience est de faire une série de mesure de bruit au passage de différents véhicules. L'antenne est donc installée sur un circuit d'essais bitumé. Nous considérons le sol comme la seule source de réflexions dans l'environnement.

25 sources de calibration sont disposées autour des microphones. Elle sont placées à différentes hauteurs (allant de 20 cm à 2.3 m par rapport au sol) pour garantir de retrouver la structure tridimensionnelle de l'antenne. Les signaux émis sont mesurés par les entrées analogiques du système *Megamicros* (cf. schéma du système 1.3a au chapitre 1). De cette façon, ils sont synchronisés avec les mesures des microphones MEMS. Conformément aux indications tirées de l'état de l'art [76], les signaux sources choisis sont des sinus dont la fréquence évolue suivant une progression logarithmique (ou *chirps* logarithmiques). Ils balayent le spectre sur l'intervalle  $f \in [150 \text{ Hz}, 12 \text{ kHz}]$ , sur une durée de 10 secondes. Pour mesurer les GCC-PHATs, les interpectres sont estimés par la méthode de Welch, en moyennant des trames de 10240 échantillons. À la fréquence d'échantillonnage  $f_s = 50 \text{ kHz}$ , la moyenne est effectuée sur 48 trames de 0.2 secondes.

Les GCC-PHATs permettent d'obtenir les temps de vol  $\tau_{im}$  et différences de temps de vol  $\tau_{imn}$ . Les figures 4.1b-4.1c illustrant l'impact des réflexions sur le sol sont construites avec des données de cette expérience. Bien que le sol soit l'unique source de réflexion, son impact est nettement visible dans les tracés des GCC-PHATs sur les figures. Donc nous savons *a priori* que des pics d'intercorrélation issus de la réflexion sur le sol engendrent des valeurs aberrantes dans les données réelles  $\tau_{im}$  et  $\tau_{imn}$ .

Pour employer les trois méthodes de calibration robuste de manière optimale, le paramètre de régularisation  $\nu$  est déterminé par une étude paramétrique, avec la technique de courbe en L. Le principe a déjà été employé pour la calibration en champ diffus dans le chapitre précédent, section 3.5.3. La courbe analysée est le nombre de valeurs aberrantes identifiées en fonction de  $\nu$ , c'est à dire la norme  $\ell_0$  de l'estimation  $\tilde{\mathbf{O}}$ . Elle met en évidence deux comportements selon la valeur du paramètre de régularisation. Si  $\nu$  est trop petit, la détection des valeurs aberrantes est trop sensible. Dans ce cas  $\|\tilde{\mathbf{O}}\|_0$  augmente rapidement. Au contraire, si  $\nu$  est trop élevé, la contrainte de parcimonie sur  $\mathbf{O}$  est trop forte, menant à une norme  $\|\tilde{\mathbf{O}}\|_0$  quasi constante et proche de 0. Donc l'allure de  $\|\tilde{\mathbf{O}}\|_0$  en fonction de  $\nu$  est une courbe en L. Le choix supposé optimal de  $\nu$  est à la frontière entre ces deux régimes. Il en résulte d'une valeur  $\nu = 8.85 \times 10^{-5} \text{ s}$  (équivalent à 3.01 cm en distance) pour le RMDS, et  $\nu = 3.7 \times 10^{-4} \text{ s}$  (équivalent à 12.6 cm) pour R-Ono et Débruitage+C-Ono.

Étant donné la taille relativement grande de l'antenne, la vérité-terrain est compliquée à obtenir



**FIGURE 4.9** – Comparaison des cinq méthodes de calibration géométrique par TOA/TDOA, sur l’antenne de 256 microphones (figure 4.8a). Histogramme des 256 erreurs par microphone, entre les estimations et la vérité-terrain.

précisément. L’antenne a été assemblée sur un cadre, sur lequel sont positionnées 32 barres de 8 microphones. Il est impossible en pratique de garantir un positionnement au millimètre sur la totalité de l’antenne. En revanche, la position relative des micros sur une même barre est d’une précision millimétrique. Donc pour évaluer les estimations géométriques  $\tilde{\mathbf{X}}$ , nous faisons une transformation de Procrustes par barre de 8 micros. Autrement dit, la géométrie estimée de chaque barre est alignée sur la géométrie théorique, et le processus est répété indépendamment sur les 32 barres. L’alignement est réalisé par l’algorithme 3 de Kabsch [75], décrit au chapitre 3. Enfin, les 256 erreurs de positionnement par microphone sont calculées :  $\varepsilon_m = \|\mathbf{x}_m - \tilde{\mathbf{x}}_m\|$ .

Les estimations géométriques des cinq algorithmes par TOA (FrSM, RMDS) et par TDOA (C-Ono, Débruitage+C-Ono, R-Ono) sont comparées dans la figure 4.9. Les 256 erreurs par microphone  $\varepsilon_m$  sont représentées sous forme d’histogramme, avec la moyenne en rouge. Une première observation globale montre que :

- dans le cas par TOA, la robustesse réduit l’erreur d’estimation moyenne de 3.41 cm, en passant de 4.79 cm par FrSM à 1.38 cm par RMDS ;
- dans le cas par TDOA, la robustesse apporte une amélioration dans les deux méthodes proposées. Contrairement aux test de Monte-Carlo précédent, R-Ono est plus performant (3.06 cm d’erreur moyenne) que le Débruitage+C-Ono (4.81 cm d’erreur).

Au delà de la tendance moyenne, les barres d'histogrammes reflètent une dispersion des erreurs plus élevée par les méthodes classiques. Sur ce point, les algorithmes apportant la meilleure homogénéité en précision sont le RMDS, suivi de R-Ono.

Notons qu'en terme d'erreur moyenne, les approches par TOA apportent de meilleurs résultats que par TDOA. Selon les analyses par Monte-Carlo, le point différenciant les deux approches est la sensibilité à l'erreur sur la position initiale, supérieure dans le cas TDOA. Il est probable que les positions initiales fournies ( $\mathbf{X}^{(0)}$ ,  $\mathbf{S}^{(0)}$ ) soient la raison expliquant les meilleures performances des les approches TOA. En conséquence, le meilleur résultat sur les cinq méthodes est donné par le RMDS en TOA dont l'erreur moyenne vaut 1.38 cm.

En conclusion, les résultats de l'expérience réelle montrent l'importance d'utiliser une approche robuste de calibration géométrique par TOA ou TDOA. En effet, l'approximation de champ libre devient fausse dès qu'il existe des obstacles réfléchissants dans l'espace de propagation. Nous validons l'amélioration apportée par les trois méthodes robustes de calibration géométrique, dans le cas d'une application *in situ*.

## 4.5 Conclusion

Ce chapitre est dédié à la calibration géométrique par temps de vol et différences de temps de vol. Les deux méthodes utilisent des sources fixes dont les signaux sont contrôlés ou non. Comparés aux méthodes proposées dans la littérature, les développements présentés ici permettent d'identifier la position des microphones avec robustesse, en particulier dans les conditions d'un environnement réel. Elles permettent en effet de s'affranchir des mesures aberrantes des temps ou différences de temps de vol induites par des réflexions parasites sur des obstacles (sol, parois, bâti, ...), présents dans le milieu supposé libre.

Pour parvenir à ces résultats, nous proposons plusieurs solutions qui exploitent le principe de parcimonie des valeurs aberrantes introduit au chapitre précédent. Ces solutions parviennent conjointement à éliminer les grandes erreurs de l'observable (TOA ou TDOA) et à identifier la géométrie du réseau. La méthode exploite une minimisation alternée. D'une part, la méthode Lasso résout le problème parcimonieux pour éliminer les mesures aberrantes; d'autre part, la régression de l'observable par une minimisation aux moindres carrés permet d'obtenir la géométrie recherchée du réseau. Les expériences numériques valident le gain apporté par les méthodes robustes par rapport aux méthodes classiques. Pour finir, une expérience *in situ* illustre l'apport de l'algorithme robuste, en présence du sol qui constitue l'unique réflecteur. Il reste à établir que la robustesse subsiste en présence de réflecteurs multiples, comme par exemple dans une salle.

D'un point de vue formel, le travail cherche à généraliser le principe de calibration robuste afin de couvrir toutes les stratégies existantes sous une approche unifiée. Au delà de l'application aux réseaux à grands nombre de microphones, les méthodes présentées peuvent traiter le cas

des petites antennes. Il s'agit d'un chapitre d'ouverture, et plusieurs questions n'ont pas encore trouvé de réponse. Premièrement, il serait intéressant de comparer les méthodes proposées avec les approches exploitant RANSAC [10, 169] énoncées dans l'introduction. Notamment, il faudrait comparer la limite en robustesse, ainsi que la complexité en temps de calcul en fonction du nombre  $M$  de micros. Deuxièmement, la validation expérimentale s'est limitée à la reconstruction d'une grande antenne par sous réseau. Pour finaliser cette validation et établir les limites de la méthode robuste, il faudrait disposer d'un réseau de grande taille dont la géométrie globale est connue précisément. En outre, les expériences pourraient être effectuées dans un environnement plus complexe, dans lequel il existe un plus grand nombre de réflecteurs tels que des murs. Ainsi, il serait possible d'identifier plus précisément la limite des méthodes proposées.



# Conclusions

Les travaux présentés traitent des différents verrous scientifiques et technologiques liés à l'emploi de réseaux à grand nombre de microphones. Jusqu'à récemment leur utilisation était bridée par la technologie, nécessitant un matériel trop lourd et complexe. Pour repousser cette limite, le projet *Megamicros* s'appuie sur l'utilisation de microphones MEMS. J'ai commencé par une description matérielle du système. J'ai ensuite décrit la caractérisation des microphones MEMS ainsi qu'une application temps réel d'imagerie acoustique, qui montrent que ces microphones sont performants pour des applications liées à l'antennerie.

Les travaux se focalisent essentiellement sur une problématique commune : comment appliquer les techniques d'antennerie acoustique avec un réseau fortement étendu et à très grand nombre de microphones ? Mon travail a donc consisté à proposer des méthodologies adaptées aux grands réseaux, notamment en terme (a) de traitement massif de données, et (b) de déploiement du système *in situ*. Je me suis concentré sur deux problèmes inverses : la localisation spatiale de sources acoustiques et la calibration géométrique d'un réseau de microphones. Bien que ces deux problèmes soient l'objet de nombreuses recherches dans la littérature, les cas d'application avec un réseau massivement multi-canal restent relativement rares.

## Localisation de sources par les grands réseaux

Premièrement, le problème de localisation de sources a relevé un défi essentiellement technique. En effet, j'ai cherché avant tout à construire un outil de diagnostic dynamique par imagerie acoustique. Fonctionnant en temps réel, il permet l'observation de phénomènes transitoires. Ce dernier point a posé le problème de la vitesse de traitement de données. Il s'est posé non seulement à cause du flux important des données issues des microphones, mais également parce que le processus d'imagerie acoustique couvre le spectre audible. Ainsi il est possible de cartographier les sources large bande. Grâce au calcul parallèle sur GPU, il a été possible de réaliser ce traitement en temps réel. L'algorithme de formation de voies a été choisi naturellement pour sa robustesse et sa simplicité d'implémentation. On a pu montrer qu'il est, de plus, particulièrement bien adapté à l'architecture parallélisée du GPU. La perspective naturelle de ce travail est l'implémentation d'autres méthodes de localisation de sources à haute résolution, par exemple par déconvolution

ou par approximation parcimonieuse. De telles méthodes sont parallélisables, dans une moindre mesure que la formation de voies car elles font souvent appel à des algorithmes itératifs. Bien que la parallélisation de ces algorithmes soit restreinte à l'échelle de l'itération, elle devient intéressante si la complexité de calcul est fonction du nombre de microphones.

## La calibration géométrique des grands réseaux

Le déploiement *in situ* de grands réseaux a révélé le besoin impératif de méthodes de calibration géométrique pour retrouver automatiquement la position des microphones. De plus, ces méthodes doivent couvrir l'ensemble des champs acoustiques réels, du champ réverbérant au champ libre. Je me suis d'abord focalisé sur une stratégie impliquant l'écoute d'un champ acoustique diffus isotrope non contrôlé, qui permet d'extraire la matrice des distances par paire entre les microphones. Grâce au MDS, cet observable géométrique permet de connaître la géométrie du réseau, hormis son orientation et sa position absolues. Les approches de la littérature se sont montrées inadaptées au grand nombre de microphones et à l'extension spatiale. En effet, l'ensemble des  $M^2$  distances pour  $M$  microphones est issu de régressions sur les cohérences. Il contient inévitablement des valeurs aberrantes (1) quand  $M \gg 1$ , (2) quand le réseau est étendu et (3) quand le champ acoustique n'est pas parfaitement isotrope. Ma principale contribution a été de proposer une méthode robuste face à ces trois contraintes. Les expériences avec le système *MegaMicros* ont montré que la méthode de calibration géométrique proposée est plus précise que celles de l'état de l'art.

Enfin, le dernier chapitre tente d'exploiter l'approche algorithmique utilisée dans la calibration en champ diffus, mais cette fois-ci en champ *quasi* libre. Dans ce cas on se fonde soit sur l'intercorrélation des signaux des microphones et des sources pour obtenir les temps de vol, soit sur l'intercorrélation des signaux mesurés pour obtenir les différences de temps de vol. Or, les intercorrélations sont rapidement affectées par les réflexions sur le sol ou les murs, qui peuvent dégrader la mesure des temps de vol ou différences de temps de vol. Là encore, cela se traduit par la présence de valeurs aberrantes qui détériorent le positionnement des microphones. Dans le cas d'un grand réseau, il est primordial d'identifier et d'écarter ces données gênantes de façon automatique et non supervisée. Les expériences numériques, ainsi que celles menées avec le système *MegaMicros*, ont montré l'intérêt d'utiliser une approche robuste aux réflexions. En effet, les estimations géométriques sont sensiblement plus précises qu'avec les méthodes de l'état de l'art étudiées.

## Perspectives

Au cours de la thèse, l'utilisation *in situ* du système *Megamicros* a offert les perspectives les plus originales en terme d'application à l'antennerie. Parmi les projets en cours, je cite d'abord les

travaux de thèse effectués par Raphaël Leiba : le système est déployé dans des milieux urbains afin de localiser les sources de bruit des véhicules. Grâce à la grande taille de l'antenne, il est possible d'établir une cartographie du bruit d'une route sur une portion longue de 20 m, avec une résolution permettant d'identifier les différentes sources présentes sur un même véhicule (bruit de moteur et de roulement). Le second projet en cours consiste à identifier les sources aéroacoustiques dans une veine de soufflerie. Le grand nombre de microphones permet d'entourer entièrement la veine, et donc les sources. De nouvelles méthodes peuvent alors être exploitées, tels que le retournement temporel [92] ou l'identification de la directivité des sources aéroacoustiques [115].

Enfin, un projet à venir appelé les Murs Ont Des Oreilles (MODO) démarrera fin 2016. L'ambition est de construire une salle contenant 1000 microphones encastrés dans les murs et le plafond. Cela permettra de mesurer le champ acoustique tout autour de l'espace de la salle réverbérante. Parmi les applications potentielles, nous avons (a) l'exploitation de techniques de déréverbération du champ acoustique de la salle par séparation du champ réverbérant et du champ libre [33, 107], puis (b) l'étude du problème de *cocktail party* par un grand réseau. En effet, les algorithmes de séparation de source exploitent la covariance spatiale entre les canaux pour extraire une source [111]. Mais jusqu'à présent il était difficile d'utiliser plus de 6 canaux [9]. En effet, une étape algorithmique coûteuse en temps de calcul nécessite l'inversion de la matrice de covariance. FitzGerald *et al.* ont récemment proposé une approche évitant cette étape, ce qui permet d'exploiter un nombre de canaux beaucoup plus grand, en un temps de calcul raisonnable [56].



# Liste des publications

## Articles dans des revues à comité de lecture

- Charles Vanwynsberghe, Pascal Challande, Jacques Marchal, Régis Marchiano, and François Ollivier, *A robust and passive method for geometric calibration of large arrays*, The Journal of the Acoustical Society of America **139** (2016), no. 3, 1252–1263
- Charles Vanwynsberghe, Régis Marchiano, François Ollivier, Pascal Challande, Hélène Moingeon, and Jacques Marchal, *Design and implementation of a multi-octave-band audio camera for realtime diagnosis*, Applied Acoustics **89** (2015), 281–287

## Conférences avec acte

- Charles Vanwynsberghe, Pascal Challande, Jacques Marchal, Régis Marchiano, and François Ollivier, *Blind geometric calibration of large arrays in diffuse field*, 12th Western Pacific Acoustics Conference, 2015
- Charles Vanwynsberghe, Jacques Marchal, François Ollivier, Pascal Challande, and Régis Marchiano, *Développement d'un sonomètre imageur temps réel à représentation par bande de fréquences*, Congrès Français d'Acoustique, 2014, pp. 403–409

## Conférences sans acte

- Charles Vanwynsberghe, Pascal Challande, Raphael Leiba, Jacques Marchal, Régis Marchiano, François Ollivier, Gilles Puy, and Pierre Vanderghenst, *Localisation et identification spectrale conjointe de sources large bande par parcimonie groupée*, Congrès Français d'Acoustique, 2016
- Charles Vanwynsberghe, Pascal Challande, Jacques Marchal, François Ollivier, and Régis Marchiano, *Localisation des microphones d'un grand réseau en champ diffus*, Congrès Français d'Acoustique, 2016
- Thibaut Meurisse, François Ollivier, Jacques Marchal, and Charles Vanwynsberghe, *Analyse modale en temps réel par holographie de champ proche et calculs parallèles*, Congrès Français d'Acoustique, 2016
- Régis Marchiano, Philippe Druault, Raphael Leiba, Jacques Marchal, François Ollivier, Vincent Valeau, and Charles Vanwynsberghe, *Localisation de sources aéroacoustiques par une méthode de retournement temporel tri-dimensionnelle*, Congrès Français d'Acoustique, 2016

# Annexe A

## Le modèle de cohérence en champ diffus

### A.1 Cas du champ isotrope 3D

La démonstration proposée s'appuie sur [73], en incluant le formalisme complexe utilisée dans l'étude du champ diffus électromagnétique [66].

La cohérence est fonction des interspectre et autospectres du couple de pressions  $(p_n, p_m)$ . Nous considérons une configuration tridimensionnelle, dans laquelle se propage un champ acoustique d'ondes planes harmoniques. Ces ondes planes sont caractérisées par une direction de propagation  $\Theta$ , un vecteur d'onde  $\mathbf{k}$ , et une phase à l'origine  $\phi$ . L'angle  $\Theta$  est défini par les coordonnées polaires  $\Theta = (\psi, \theta)$ .

Pour simplifier le calcul, les deux microphones sont positionnés sur l'axe  $(0, z)$ , et sont espacés d'une distance  $d_{nm}$ , donc  $\mathbf{x}_m - \mathbf{x}_n = d_{nm} \vec{z}$ .

#### A.1.1 Cas d'un champ diffus harmonique, à $\omega = \omega_0$

On considère en premier temps un champ diffus harmonique à la pulsation  $\omega = \omega_0$ , donc les vecteurs d'ondes valent

$$\mathbf{k}_0(\Theta) = k_0 \Theta = \frac{\omega_0}{c_0} \Theta. \quad (\text{A.1})$$

Les phases des ondes planes sont décorréliées en fonction de l'angle de propagation : le modèle considère  $\phi$  comme une variable aléatoire suivant  $\mathcal{U}(0, 2\pi)$ , la distribution uniforme continue sur l'intervalle  $[0, 2\pi]$ . Le champ d'onde planes est isotrope, donc l'amplitude est invariante selon  $\Theta$ ; nous considérons des ondes planes d'amplitude 1.

Dans le formalisme complexe, et selon les hypothèses énoncées, les pressions mesurées  $p_n$  et  $p_m$  s'expriment de la façon suivante :

$$p_n(t) = \iint_{4\pi} e^{j(\mathbf{k}_0(\Theta_n) \cdot \mathbf{x}_n - \omega_0 t + \phi(\Theta_n))} d\Theta_n \quad (\text{A.2})$$

$$p_m(t) = \iint_{4\pi} e^{j(\mathbf{k}_0(\Theta_m) \cdot \mathbf{x}_m - \omega_0 t + \phi(\Theta_m))} d\Theta_m \quad (\text{A.3})$$

Si le rayonnement du champ diffus est un processus stationnaire au cours du temps, l'intercorrélation du couple de signaux  $(p_n, p_m)$  est calculée statistiquement par l'espérance suivante :

$$r_{nm}(\tau) = \mathbb{E} [p_n^*(t) p_m(t + \tau)] \quad (\text{A.4})$$

En injectant les équations du modèle de champ (A.2) et (A.3), on trouve que :

$$r_{nm}(\tau) = \mathbb{E} \left[ \iint_{4\pi} e^{-j(\mathbf{k}_0(\Theta_n) \cdot \mathbf{x}_n - \omega_0 t + \phi(\Theta_n))} d\Theta_n \iint_{4\pi} e^{j(\mathbf{k}_0(\Theta_m) \cdot \mathbf{x}_m - \omega_0(t+\tau) + \phi(\Theta_m))} d\Theta_m \right] \quad (\text{A.5})$$

$$= \mathbb{E} \left[ \iint_{4\pi} \iint_{4\pi} e^{j(\mathbf{k}_0(\Theta_n) \cdot \mathbf{x}_n - \mathbf{k}_0(\Theta_m) \cdot \mathbf{x}_m)} e^{j(\phi(\Theta_m) - \phi(\Theta_n))} e^{-j\omega_0 \tau} d\Theta_n d\Theta_m \right] \quad (\text{A.6})$$

La phase à l'origine  $\phi$  est une variable aléatoire, cependant toutes les autres variables sont déterministes, ce qui permet de simplifier l'intégrale :

$$r_{nm}(\tau) = e^{-j\omega_0 \tau} \iint_{4\pi} \iint_{4\pi} e^{j(\mathbf{k}_0(\Theta_n) \cdot \mathbf{x}_n - \mathbf{k}_0(\Theta_m) \cdot \mathbf{x}_m)} \mathbb{E} [e^{j(\phi(\Theta_m) - \phi(\Theta_n))}] d\Theta_n d\Theta_m \quad (\text{A.7})$$

Avec  $\phi \in \mathcal{U}(0, 2\pi)$ , pour deux processus différents  $n$  et  $m$ , la variable  $(\phi(\Theta_m) - \phi(\Theta_n))$  suit la loi  $\mathcal{U}(-2\pi, 2\pi)$ , qui est également uniforme et continue, mais sur l'intervalle  $[-2\pi, 2\pi]$ . Or l'espérance d'une exponentielle complexe à phase uniformément répartie sur cet intervalle est nulle si  $n \neq m$ . Par contre, pour deux processus identiques  $n = m$ , la phase est toujours nulle, donc  $\mathbb{E} [e^{j(\phi(\Theta_m) - \phi(\Theta_n))}] = \delta(\Theta_m - \Theta_n)$  où  $\delta$  est la fonction de Dirac. L'intégrale se simplifie donc par :

$$r_{nm}(\tau) = e^{-j\omega_0 \tau} \iint_{4\pi} \iint_{4\pi} e^{j(\mathbf{k}_0(\Theta_n) \cdot (\mathbf{x}_n - \mathbf{x}_m))} \delta(\Theta_m - \Theta_n) d\Theta_n d\Theta_m \quad (\text{A.8})$$

On voit apparaître une convolution avec la fonction Dirac de la forme  $f(x) = \int f(t) \delta(x - t) dt$ .

La double intégrale sur l'angle solide  $4\pi$  se transforme finalement en intégrale simple sur  $\Theta_m$  :

$$r_{nm}(\tau) = e^{-j\omega_0\tau} \iint_{4\pi} e^{j(\mathbf{k}_0(\Theta_m) \cdot (\mathbf{x}_n - \mathbf{x}_m))} d\Theta_m \quad (\text{A.9})$$

Pour calculer l'intégrale, nous faisons varier les coordonnées sphériques  $\Theta = (\psi, \theta)$  telles que  $\psi \in [-\pi, \pi]$ , et  $\theta \in [0, \pi]$ . Rappelons que l'angle élémentaire sphérique s'écrit  $d\Theta = \sin \theta d\theta d\psi$ . Dans la configuration géométrique donnée des microphones, l'intégrale donne :

$$r_{nm}(\tau) = e^{-j\omega_0\tau} \iint_{4\pi} e^{jk_0 d_{nm} \cos \theta} d\Theta = e^{-j\omega_0\tau} \int_{-\pi}^{\pi} d\psi \int_0^{\pi} e^{jk_0 d_{nm} \cos \theta} \sin \theta d\theta \quad (\text{A.10})$$

$$r_{nm}(\tau) = 4\pi e^{-j\omega_0\tau} \text{sinc}(k_0 d_{nm}) \quad (\text{A.11})$$

L'intercorrélacion temporelle de deux signaux est donc aussi fonction du sinus cardinal de la distance entre les microphones. De cela, il est également possible d'obtenir l'interspectre : le théorème de Wiener-Khintchine énonce que, dans le cas d'un processus stationnaire, l'interspectre vaut la transformée de Fourier de l'intercorrélacion :

$$S_{nm}(\omega) = 4\pi \int_{\tau \in \mathbb{R}} \text{sinc}(k_0 d_{nm}) e^{-j\omega_0\tau} e^{-j\omega\tau} d\tau \quad (\text{A.12})$$

$$= 4\pi \text{sinc}(k_0 d_{nm}) \delta(\omega + \omega_0) \quad (\text{A.13})$$

$$S_{nm}(\omega) = 4\pi \text{sinc}(k_0 d_{nm}) \delta(\omega - \omega_0) \quad (\text{A.14})$$

Nous avons pu voir que l'intercorrélacion est l'interspectre ont la même forme, avec une fonction du terme  $\text{sinc}(k_0 d_{nm})$ . Notons également que les dépendances spatiale et fréquentielle sont identiques, et les expériences dans la littérature ont été menées pour valider ces deux dépendances.

### A.1.2 Cas d'un champ diffus large bande

L'expression (A.14) donne l'interspectre pour un champ diffus harmonique à  $\omega_0$ . Pour un champ diffus étendu sur tout le spectre, on intègre cette expression précédente sur  $\omega_0 \in [-\infty, \infty]$  :

$$S_{nm}(\omega) = 4\pi \int_{\omega_0 \in \mathbb{R}} \text{sinc}\left(\frac{d_{nm}}{c_0} \omega_0\right) \delta(\omega - \omega_0) d\omega_0 \quad (\text{A.15})$$

On retrouve de nouveau une convolution avec la fonction de Dirac, ce qui donne directement :

$$S_{nm}(\omega) = 4\pi \operatorname{sinc}(kd_{nm}) \quad (\text{A.16})$$

On en déduit finalement la cohérence ; en effet ici l'autospectre est  $S_{nn}(\omega) = 4\pi \operatorname{sinc}(0) = 4\pi$ .  
Donc :

$$\Gamma_{nm}(f) = \frac{S_{nm}(f)}{\sqrt{S_{nn}(f)S_{mm}(f)}} = \frac{4\pi \operatorname{sinc}(kd_{nm})}{\sqrt{(4\pi)^2}} = \operatorname{sinc}(kd_{nm}) \quad (\text{A.17})$$

On peut facilement montrer que le résultat final est le même :

- si l'amplitude des ondes planes varie en fonction de la fréquence. Le champ diffus peut donc avoir une signature spectrale quelconque et ceci n'influe pas sur la mesure expérimentale de la cohérence ;
- si les sensibilités des microphones  $n$  et  $m$  sont différents : la mesure de la cohérence est invariante quel que soit les réponses en fréquence des microphones.

## A.2 Cas particulier du champ isotrope plan

Dans un problème situé dans un plan  $(O, x, y)$ , les étapes de calcul sont les mêmes que précédemment, seules les conditions géométriques changent :

- nous supposons que les microphones sont parallèles à l'axe  $(O, y)$  :  $\mathbf{x}_m - \mathbf{x}_n = d_{nm}\vec{y}$  ;
- le champ isotrope est plan, d'amplitude constante selon l'angle polaire  $\theta \in [-\pi; \pi]$ . En posant le nombre d'onde  $\mathbf{k}_0(\theta) = \omega_0/c_0(\cos \theta, \sin \theta)$ , les pressions mesurées s'expriment par l'intégrale :

$$p_n(t) = \int_{2\pi} e^{j(\mathbf{k}_0(\theta_n) \cdot \mathbf{x}_n - \omega_0 t + \phi(\theta_n))} d\theta_n \quad (\text{A.18})$$

$$p_m(t) = \int_{2\pi} e^{j(\mathbf{k}_0(\theta_m) \cdot \mathbf{x}_m - \omega_0 t + \phi(\theta_m))} d\theta_m \quad (\text{A.19})$$

Alors le calcul de l'intercorrélation équivalent à (A.10) en 2D aboutit à l'intégrale :

$$r_{nm}(\tau) = e^{-j\omega_0 \tau} \int_{2\pi} e^{jk_0 d_{nm} \sin \theta} d\theta \quad (\text{A.20})$$

C'est la forme de l'intégrale de Bessel de première espèce à l'ordre 0 :

$$r_{nm}(\tau) = 2\pi e^{-j\omega_0 \tau} J_0(k_0 d_{nm}) \quad (\text{A.21})$$

Si le champ diffus plan est large bande, la cohérence vaut alors :

$$\Gamma_{nm}(f) = J_0(kd_{nm}) \quad (\text{A.22})$$

### **A.3 Estimation de la fonction de Green du milieu par l'écoute du champ diffus**

En acoustique sous-marine ou sismique [121], il a déjà été constaté que le bruit global issu de tous les événements acoustiques environnants crée un champ isotrope. Dans ces circonstances, il a été démontré que la dérivée de l'intercorrélation par deux capteurs révèle la fonction de Green du milieu [121] :

$$\frac{d}{dt}r_{12}(\tau) = \frac{1}{4\pi d_{nm}/c_0} \left[ \delta\left(\tau + \frac{d_{nm}}{c_0}\right) - \delta\left(\tau - \frac{d_{nm}}{c_0}\right) \right] \quad (\text{A.23})$$

où  $\delta(t)$  est la fonction Dirac. Sans dérivée,  $r_{12}(\tau)$  est donc caractérisée par une fonction Porte. Or dans le cas d'un processus temporel stationnaire, l'interspectre est la transformée de Fourier de l'intercorrélation. Par conséquent cette expression fonction du temps est analogue à la formule de l'interspectre révélant le sinus cardinal. Des résultats retrouvant la fonction de Green en acoustique aérienne sont montrées dans [106].



## Annexe B

# LRMDS : majorante et conditions de pondération

L'objectif n'est pas de présenter la totalité de la démonstration [59] menant à l'algorithme final 1. Nous souhaitons avant tout montrer comment adapter et implémenter l'algorithme original pour une version pondérée. Nous reprenons l'énoncé des étapes principales permettant de retrouver la fonction de substitution  $g$  proposée pour le RMDS.

### B.1 Construction de la fonction majorante

Rappelons la fonction coût proposée dans [58] :

$$f(\mathbf{O}, \mathbf{X}) = \sum_{n < m} [\delta_{nm} - d_{nm}(\mathbf{X}) - o_{nm}]^2 + \nu \sum_{n < m} |o_{nm}| \quad (\text{B.1})$$

$f$  est une fonction *compliquée* à minimiser, dans le sens où elle est à la fois non-convexe et non-différentiable. L'approche MM consiste à minimiser, itérativement, une fonction de substitut semblable à  $f$ , convexe et différentiable. Elle s'appuie sur une fonction  $g$  appelée approximation majorante de  $f$ . À chaque itération, la fonction  $g$  est construite puis minimisée, jusqu'à satisfaire un critère de convergence.

Dans le cas de la fonction coût (B.1), à partir du couple de variable  $(\mathbf{O}^{(k)}, \mathbf{X}^{(k)})$  à l'itération  $k$ , la minimisation de  $g$  détermine un nouveau couple de variable  $(\mathbf{O}^{(k+1)}, \mathbf{X}^{(k+1)})$ .  $g$  est donc fonction de  $(\mathbf{O}, \mathbf{X})$ , puis de deux constantes  $\mathbf{O}^{(k)}$  et  $\mathbf{X}^{(k)}$ . Quatre variables sont nécessaires pour construire  $g$ , car pour toutes les itérations  $k$ , deux conditions doivent être respectées :

- la fonction  $g$  est tangente à la fonction  $f$  au point de contact  $(\mathbf{O}^{(k)}, \mathbf{X}^{(k)})$ . Notons les variables



$(\mathbf{Q}, \mathbf{Z})$  qui définissent le point de contact par lequel passe  $g$ ; on a l'égalité :

$$g(\mathbf{O} = \mathbf{O}^{(k)}, \mathbf{X} = \mathbf{X}^{(k)}, \mathbf{Q} = \mathbf{O}^{(k)}, \mathbf{Z} = \mathbf{X}^{(k)}) = f(\mathbf{O}^{(k)}, \mathbf{X}^{(k)}) \quad (\text{B.2})$$

–  $g$  est une fonction majorante de  $f$  :

$$g(\mathbf{O}, \mathbf{X}, \mathbf{Q}, \mathbf{Z}) \geq f(\mathbf{O}, \mathbf{X}) \quad \forall (\mathbf{O}, \mathbf{X}) \in (\mathbb{S}_h^M, \mathbb{R}^{M \times D}) \quad (\text{B.3})$$

À chaque itération, la mise à jour  $(\mathbf{O}^{(k+1)}, \mathbf{X}^{(k+1)})$  s'exprime par le problème de minimisation :

$$(\mathbf{O}^{(k+1)}, \mathbf{X}^{(k+1)}) = \min_{(\mathbf{O}, \mathbf{X})} g(\mathbf{O}, \mathbf{X}, \mathbf{Q} = \mathbf{O}^{(k)}, \mathbf{Z} = \mathbf{X}^{(k)}) \quad (\text{B.4})$$

### B.1.1 Construction de $g$ : cas du RMSD, sans pondération

Notons les matrices  $\mathbf{A}$  et  $\mathbf{D}$  contenant respectivement l'ensemble des éléments  $\delta_{nm}$  et  $d_{nm}$ . La fonction coût (B.1) s'écrit sous forme matricielle :

$$f(\mathbf{O}, \mathbf{X}) = \frac{1}{2} \|(\mathbf{A} - \mathbf{O})\|_F^2 + \text{tr } \mathbf{X} \mathbf{L} \mathbf{X}^\top - 2 \sum_{n < m} [\delta_{nm} - o_{nm}] d_{nm}(\mathbf{X}) + \frac{\nu}{2} \|\mathbf{O}\|_1 \quad (\text{B.5})$$

en sachant que  $\mathbf{L}$  se construit de la façon suivante :

$$\begin{aligned} \sum_{n < m} d_{nm}^2(\mathbf{X}) &= \text{tr } \mathbf{X} \left( \sum_{n < m} (\mathbf{e}_n - \mathbf{e}_m)^\top (\mathbf{e}_n - \mathbf{e}_m) \right) \mathbf{X}^\top \\ &= \text{tr } \mathbf{X} \mathbf{L} \mathbf{X}^\top \end{aligned} \quad (\text{B.6})$$

Le troisième terme dans B.5 est critique car non-différentiable et non-convexe. C'est ce dernier qui doit être approximé pour obtenir  $g$ , avec les deux conditions nécessaires pour utiliser l'approche type MM.

Notons ce troisième terme  $\phi(\mathbf{O}, \mathbf{X}) = -2 \sum_{n < m} [\delta_{nm} - o_{nm}] d_{nm}(\mathbf{X})$ . Il est décomposable en deux sous-sommes  $\phi(\mathbf{O}, \mathbf{X}) = \phi_+(\mathbf{O}, \mathbf{X}) + \phi_-(\mathbf{O}, \mathbf{X})$  telles que :

- $\phi_+$  soit la somme sur les couples  $(n, m) \in \mathcal{Q}_+$ , tels que  $\delta_{nm} - o_{nm} > 0$
- $\phi_-$  soit la somme sur les couples  $(n, m) \in \mathcal{Q}_-$ , tels que  $\delta_{nm} - o_{nm} \leq 0$

$\phi_+$  est donc non-convexe, mais peut être majoré; en passant par l'inégalité de Cauchy Schwartz  $(\mathbf{z}_n - \mathbf{z}_m)^\top (\mathbf{x}_n - \mathbf{x}_m) \leq d_{nm}(\mathbf{Z}) d_{nm}(\mathbf{X})$ , et appliqué sur tous les couples  $(n, m)$  dans  $\mathcal{Q}_+$ , l'inégalité par rapport à  $\phi_+$  est obtenue :

$$\phi_+(\mathbf{O}, \mathbf{X}) \leq -2 \sum_{(n, m) \in \mathcal{Q}_+} \frac{\delta_{nm} - o_{nm}}{d_{nm}(\mathbf{Z})} (\mathbf{z}_n - \mathbf{z}_m)^\top (\mathbf{x}_n - \mathbf{x}_m) \quad (\text{B.7})$$

en sachant qu'il y a bien égalité pour  $\mathbf{Z} = \mathbf{X}$  (première condition nécessaire sur  $g$ ).

$\phi_-$  doit également être approximé car il est non différentiable (mais convexe). L'idée est de partir sur l'inégalité suivante :  $\sqrt{x} \leq \sqrt{x_0} + \frac{1}{2}x_0^{-\frac{1}{2}}(x - x_0)$  toujours vraie pour  $x_0 > 0$ . En choisissant  $x = d_{nm}^2(\mathbf{X})$  et  $x_0 = d_{nm}^2(\mathbf{Z})$ , et en sommant sur tout les couples  $(n, m)$  dans  $\mathcal{Q}_-$ , l'inégalité par rapport à  $\phi_-$  est obtenue :

$$\phi_-(\mathbf{O}, \mathbf{X}) \leq - \sum_{(n,m) \in \mathcal{Q}_-} \frac{\delta_{nm} - o_{nm}}{d_{nm}(\mathbf{Z})} d_{nm}^2(\mathbf{X}) - \sum_{(n,m) \in \mathcal{Q}_-} [\delta_{nm} - o_{nm}] d_{nm}(\mathbf{Z}) \quad (\text{B.8})$$

Par ailleurs il y a également égalité si  $\mathbf{Z} = \mathbf{X}$ , qui est la première condition nécessaire sur  $g$ .

En posant  $\mathbf{A}_1$  et  $\mathbf{A}_2$  contenant les termes constants qui interviennent dans les sommes B.7 et B.8 :

$$\begin{aligned} [\mathbf{A}_1(\mathbf{O}, \mathbf{Z})]_{nm} &= \frac{\delta_{nm} - o_{nm}}{d_{nm}(\mathbf{Z})} \quad \text{si } (n, m) \in \mathcal{Q}_+ \\ &= 0 \quad \text{sinon} \end{aligned} \quad (\text{B.9})$$

$$\begin{aligned} [\mathbf{A}_2(\mathbf{O}, \mathbf{Z})]_{nm} &= -\frac{\delta_{nm} - o_{nm}}{d_{nm}(\mathbf{Z})} \quad \text{si } (n, m) \in \mathcal{Q}_- \\ &= 0 \quad \text{sinon} \end{aligned} \quad (\text{B.10})$$

Forero montre que ces deux inégalités peuvent s'écrire sous la forme matricielle :

$$\phi_+(\mathbf{O}, \mathbf{X}) \leq -2\text{tr}(\mathbf{X}\mathbf{L}_1(\mathbf{O}, \mathbf{Z})\mathbf{Z}^\top) \quad (\text{B.11})$$

$$\phi_-(\mathbf{O}, \mathbf{X}) \leq \text{tr}(\mathbf{X}\mathbf{L}_2(\mathbf{O}, \mathbf{Z})\mathbf{X}^\top) + C_0(\mathbf{O}, \mathbf{Z}) \quad (\text{B.12})$$

avec :

$$\mathbf{L}_1(\mathbf{O}, \mathbf{Z}) = \text{diag}(\mathbf{A}_1(\mathbf{O}, \mathbf{Z})\mathbf{1}_M) - \mathbf{A}_1(\mathbf{O}, \mathbf{Z}) \quad (\text{B.13})$$

$$\mathbf{L}_2(\mathbf{O}, \mathbf{Z}) = \text{diag}(\mathbf{A}_2(\mathbf{O}, \mathbf{Z})\mathbf{1}_M) - \mathbf{A}_2(\mathbf{O}, \mathbf{Z}) \quad (\text{B.14})$$

$$C_0(\mathbf{O}, \mathbf{Z}) = - \sum_{(n,m) \in \mathcal{Q}_-} (\delta_{nm} - o_{nm}) d_{nm}(\mathbf{Z}) \quad (\text{B.15})$$

Finalement en sommant B.11 et B.12, nous obtenons un majorant du terme  $\phi$ . En l'injectant

dans l'équation B.5, nous obtenons la fonction de substitution  $g$  :

$$g(\mathbf{O}, \mathbf{X}, \mathbf{Z}) = \frac{1}{2} \|\Delta - \mathbf{O}\|_F^2 + \text{tr}(\mathbf{X} (\mathbf{L} + \mathbf{L}_2(\mathbf{O}, \mathbf{Z})) \mathbf{X}^\top) - 2\text{tr}(\mathbf{X} \mathbf{L}_1(\mathbf{O}, \mathbf{Z}) \mathbf{X}^\top) + C_0(\mathbf{O}, \mathbf{Z}) + \frac{1}{2} \lambda \|\mathbf{O}\|_1 \quad (\text{B.16})$$

À partir de cette fonction, l'article fait la démonstration qui mène à l'algorithme 1 proposé.

### B.1.2 Construction de $g$ : cas du RMSD pondéré

Établissons l'expression de  $g$  dans le cas d'une fonction coût pondérée. Nous nous appuyerons sur la démonstration précédente car le cheminement des étapes de calcul est identique.

Rappelons l'expression de la fonction coût pondérée :

$$f_{\mathbf{W}}(\mathbf{O}, \mathbf{X}) = \sum_{n < m} w_{nm} [\delta_{nm} - d_{nm}(\mathbf{X}) - o_{nm}]^2 + \nu \sum_{n < m} w_{nm} |o_{nm}|$$

Notons  $\mathbf{W}$  la matrice contenant les éléments  $w_{nm}$ . Comme dans la partie précédente nous retrouvons la forme matricielle :

$$f_{\mathbf{W}}(\mathbf{O}, \mathbf{X}) = \frac{1}{2} \|\mathbf{W} \circ (\Delta - \mathbf{O})\|_F^2 + \text{tr} \mathbf{X} \mathbf{L} \mathbf{X}^\top - 2 \sum_{n < m} w_{nm} [\delta_{nm} - o_{nm}] d_{nm}(\mathbf{X}) + \frac{1}{2} \nu \|\mathbf{W} \circ \mathbf{O}\|_1 \quad (\text{B.17})$$

$\circ$  étant l'opérateur de Hadamard, le produit terme à terme de deux matrices. L'expression du deuxième terme avec la présence de la pondération est démontrée dans [20, p. 188]. Elle peut s'exprimer en fonction de  $\mathbf{X}$  par :

$$\sum_{n < m} w_{nm} d_{nm}^2(\mathbf{X}) = \text{tr} \mathbf{X} \left( \sum_{n < m} w_{nm} (\mathbf{e}_n - \mathbf{e}_m)^\top (\mathbf{e}_n - \mathbf{e}_m) \right) \mathbf{X}^\top \quad (\text{B.18})$$

$$= \text{tr} \mathbf{X} \mathbf{L} \mathbf{X}^\top \quad (\text{B.19})$$

Nous ajoutons l'hypothèse importante suivante :  $w_{nm} \geq 0$ . En effet, elle assure la convexité du premier et deuxième terme dans B.17. EN revanche, le troisième terme (que nous noterons de nouveau  $\phi$ ) reste toujours non-convexe et non-différentiable ; si il existe un majorant de ce terme, alors il existe un majorant  $g$  de la fonction  $f_{\mathbf{W}}$ .

Nous avons montré que  $\phi$  est décomposable en deux sous sommes  $\phi = \phi_+ + \phi_-$  selon le signe

des termes de la somme. Si  $w_{nm} \geq 0$ , alors les ensembles  $Q_+$  et  $Q_-$  sont les mêmes que s'il n'y a pas de pondération (c'est à dire quand  $w_{nm} = 1 \ \forall (n, m)$ ). De plus, la démonstration précédente des inégalités B.8 et B.7 reste valide, et nous avons :

$$\phi_+(\mathbf{O}, \mathbf{X}) \leq -2 \sum_{(n,m) \in Q_+} w_{nm} \frac{\delta_{nm} - o_{nm}}{d_{nm}(\mathbf{Z})} (\mathbf{z}_n - \mathbf{z}_m)^\top (\mathbf{x}_n - \mathbf{x}_m) \quad (\text{B.20})$$

$$\phi_-(\mathbf{O}, \mathbf{X}) \leq - \sum_{(n,m) \in Q_-} w_{nm} \frac{\delta_{nm} - o_{nm}}{d_{nm}(\mathbf{Z})} d_{nm}^2(\mathbf{X}) - \sum_{(n,m) \in Q_-} w_{nm} (\delta_{nm} - o_{nm}) d_{nm}(\mathbf{Z}) \quad (\text{B.21})$$

Le reste de la démonstration mène à une expression de  $g$  qui reste inchangée (cf B.16), en redéfinissant  $\mathbf{A}_1$ ,  $\mathbf{A}_2$  et  $C_0$  en fonction des  $w_{nm}$  :

$$\begin{aligned} [\mathbf{A}_1(\mathbf{O}, \mathbf{X})]_{nm} &= w_{nm} \frac{\delta_{nm} - o_{nm}}{d_{nm}(\mathbf{X})} \quad \text{si } (n, m) \in Q_+ \\ &= 0 \quad \text{sinon} \end{aligned} \quad (\text{B.22})$$

$$\begin{aligned} [\mathbf{A}_2(\mathbf{O}, \mathbf{X})]_{nm} &= -w_{nm} \frac{\delta_{nm} - o_{nm}}{d_{nm}(\mathbf{X})} \quad \text{si } (n, m) \in Q_- \\ &= 0 \quad \text{sinon} \end{aligned} \quad (\text{B.23})$$

$$C_0(\mathbf{O}, \mathbf{Z}) = - \sum_{(nm) \in Q_-} (\delta_{nm} - o_{nm}) d_{nm}(\mathbf{Z}) \quad (\text{B.24})$$

## B.2 Propriétés du RMDS

Le RMDS garantit la convergence vers un minimum local, mais pas vers le minimum global. Il est important d'optimiser le choix de la condition initiale  $\mathbf{X}^{(0)}$ . C'est pour cela que, dans les expériences, nous partons d'une géométrie initiale construite par une méthode de MDS locale, telle que Smacof pondéré ou ISOMAP.

Le paramètre de régularisation  $\nu$  peut être déterminé de manière optimale. Pour cela, à partir le modèle de mesure  $\delta_{nm} = d_{nm} + \epsilon_{nm} + o_{nm}$ , au moins une information sur l'un des deux termes du bruit doit être *a priori* connue :

- le nombre de valeurs aberrantes dans la mesure  $\mathbf{\Delta}$ , autrement dit  $\|\mathbf{O}\|_0$ . C'est une information sur le degré de parcimonie de l'ensemble associé au deuxième terme de bruit  $o_{nm}$  ;
- les moments du premier terme de bruit  $\epsilon_{nm}$ . En supposant une loi normale, la connaissance de la variance sur  $\epsilon_{nm}$  permet un choix quasi optimal [59].

Si aucun paramètre sur le bruit est quantifiable (ce qui est généralement le cas), il faut choisir  $\nu$  de manière supervisée, en analysant la courbe du nombre de valeurs non-nulles dans  $\mathbf{O}$  en fonction de  $\nu$  (utiliser l'algorithme 2). Cette courbe doit avoir une forme en L, le paramètre  $\nu$  optimal se situant au niveau du coude de la courbe. En effet :

- un paramètre  $\nu$  surévalué donne une contrainte de parcimonie sur  $\mathbf{O}$  trop forte. En pratique, peu de valeurs aberrantes sont détectées : la courbe reste proche de 0 ;
- au contraire, un paramètre  $\nu$  trop faible ne donne pas assez de contrainte sur la parcimonie de  $\mathbf{O}$ . Il se pose le problème inverse, dans lequel trop de mesures  $\delta_{nm}$  sont identifiées comme contenant une valeur aberrante. Dans ce cas, la résolution sur  $\mathbf{O}$  devient sensible au premier terme de bruit  $\epsilon_{nm}$ , qui n'a aucune condition de parcimonie. En dessous de la valeur optimale, plus  $\nu$  est petit plus le bruit associé à  $\epsilon_{nm}$  fuite dans le terme de valeur aberrante  $o_{nm}$ .

En cela,  $\nu$  distingue l'amplitude  $\epsilon_{nm}$  de  $o_{nm}$ , tel que :

$$|\epsilon_{nm}| \leq \frac{\nu}{2} \quad (\text{B.25})$$

$$|o_{nm}| > \frac{\nu}{2} \quad (\text{B.26})$$

Notons enfin que la présence du facteur  $\frac{1}{2}$  vient de la manière par laquelle l'opérateur de seuillage doux  $\mathcal{S}_\nu(u)$  est défini (cf. eq (3.27)).

# Bibliographie

- [1] *Understanding FFT windows*, Application note AN014, LDS Dactron, 2003. (cité page 54)
- [2] *Aircraft noise (excerpt from the Oakland international airport master plan update)*, 2006. (cité page 32)
- [3] *Laboratoire National de métrologie et d'Essai*, [http://www.lne.eu/en/tests/performance\\_tests/acoustics.asp](http://www.lne.eu/en/tests/performance_tests/acoustics.asp), accessed June 2016. (cité page 17)
- [4] Jon Petter Asen, Jo Inge Buskenes, Carl-Inge C. Nilsen, Andreas Austeng, and Sverre Holm, *Implementing Capon beamforming on the GPU for real time cardiac ultrasound imaging*, Ultrasonics Symposium (IUS), 2012 IEEE International, IEEE, 2012, pp. 2133–2136. (cité page 45)
- [5] Jon Petter Asen, Jo Inge Buskenes, Carl-Inge Colombo Nilsen, Andreas Austeng, and Sverre Holm, *Implementing Capon beamforming on a GPU for real-time cardiac ultrasound imaging*, Ultrasonics, Ferroelectrics, and Frequency Control, IEEE Transactions on **61** (2014), no. 1, 76–85. (cité pages 26 et 45)
- [6] Mingsian R Bai, Jeong-Guon Ih, and Jacob Benesty, *Acoustic array systems : Theory, implementation, and application*, John Wiley & Sons, 2013. (cité page 1)
- [7] José A Ballesteros, Ennes Sarradj, Marcos D Fernandez, Thomas Geyer, and Ma Jesus Ballesteros, *Noise source identification with beamforming in the pass-by of a car*, Applied Acoustics **93** (2015), 106–119. (cité page 26)
- [8] Yu Bao and Huawei Chen, *Design of passband shaping broadband beamformer robust against microphone position errors*, Signal Processing, Communications and Computing (ICSPCC), 2014 IEEE International Conference on, Aug 2014, pp. 521–526. (cité page 44)
- [9] Jon Barker, Ricard Marxer, Emmanuel Vincent, and Shinji Watanabe, *The third CHiME speech separation and recognition challenge : Dataset, task and baselines*, 2015 IEEE Workshop on Automatic Speech Recognition and Understanding (ASRU), IEEE, 2015, pp. 504–511. (cité page 123)

- [10] K. Batstone, M. Oskarsson, and K. Åström, *Robust time-of-arrival self calibration and indoor localization using Wi-Fi round-trip time measurements*, Proc. IEEE Int. Conf. Communications Workshops (ICC), May 2016, pp. 26–31. (cité pages 93 et 119)
- [11] Jose A. Belloch, Alberto Gonzalez, Antonio M. Vidal, and Maximo Cobos, *On the performance of multi-GPU-based expert systems for acoustic localization involving massive microphone arrays*, Expert Systems with Applications **42** (2015), no. 13, 5607 – 5620. (cité page 7)
- [12] Ilan Ben Hagai, Martin Pollow, Michael Vorländer, and Boaz Rafaely, *Acoustic centering of sources measured by surrounding spherical microphone arrays*, The Journal of the Acoustical Society of America **130** (2011), no. 4, 2003–2015. (cité page 3)
- [13] Julius S. Bendat and Allan G. Piersol, *Engineering applications of correlation and spectral analysis*, Wiley, 1980. (cité page 17)
- [14] Nancy Bertin, Srđan Kitić, and Rémi Gribonval, *Joint estimation of sound source location and boundary impedance with physics-driven cosparse regularization*, Proc. Speech and Signal Processing (ICASSP) 2016 IEEE Int. Conf. Acoustics, March 2016, pp. 6340–6344. (cité page 48)
- [15] Cagdas Bilen, Gilles Puy, Rémi Gribonval, and Laurent Daudet, *Convex optimization approaches for blind sensor calibration using sparsity*, Signal Processing, IEEE Transactions on **62** (2014), no. 18, 4847–4856. (cité page 44)
- [16] John Billingsley and R Kinns, *The acoustic telescope*, Journal of Sound and Vibration **48** (1976), no. 4, 485–510. (cité page 2)
- [17] Stanley T. Birchfield and Amarnag Subramanya, *Microphone array position calibration by basis-point classical multidimensional scaling*, Speech and Audio Processing, IEEE Transactions on **13** (2005), no. 5, 1025–1034. (cité page 50)
- [18] Rahul Biswas and Sebastian Thrun, *A passive approach to sensor network localization*, Proc. IEEE/RSJ Int. Conf. Intelligent Robots and Systems (IROS 2004), vol. 2, September 2004, pp. 1544–1549 vol.2. (cité page 97)
- [19] William K. Blake and Richard V. Waterhouse, *The use of cross-spectral density measurements in partially reverberant sound fields*, Journal of Sound and Vibration **54** (1977), no. 4, 589–599. (cité pages 53 et 87)
- [20] Ingwer Borg and Patrick J.F. Groenen, *Modern multidimensional scaling : Theory and applications*, Springer, 2005. (cité pages 69, 73 et 134)
- [21] Petros T. Boufounos, Paris Smaragdis, and Bhiksha Raj, *Joint sparsity models for wideband array processing*, 2011, pp. 81380K–81380K–10. (cité page 45)

- [22] Thierry Bouwmans and El Hadi Zahzah, *Robust PCA via principal component pursuit : A review for a comparative evaluation in video surveillance*, Computer Vision and Image Understanding **122** (2014), 22 – 34. (cité page 106)
- [23] Ann Bowles and Brigitte Schulte-Fortkamp, *Noise as an indicator of quality of life : advances in measurement of noise and noise effects on humans and animals in the environment*, Acoustics Today (2008), 35–39. (cité page 26)
- [24] Thomas F. Brooks and William M. Humphreys, *A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays*, Journal of Sound and Vibration **294** (2006), no. 4, 856–879. (cité pages 3 et 38)
- [25] Janusz Bryzek, Shad Roundy, Brian Bircumshaw, Charles Chung, Kenneth Castellino, Joseph R Stetter, and Michael Vestel, *Marvellous mems : Advanced IC sensors and microstructures for high volume applications*, IEEE Circuits and Devices Magazine **22** (2006), no. 2, 8–28. (cité page 6)
- [26] Jo Inge Buskenes, Jon Petter Asen, Carl-Inge Colombo Nilsen, and Andreas Austeng, *Adapting the minimum variance beamformer to a graphics processing unit for active sonar imaging systems*, vol. 133, Acoustical Society of America, 2013, pp. 3613–3613. (cité page 26)
- [27] ———, *An optimized gpu implementation of the MVDR beamformer for active sonar imaging*, Oceanic Engineering, IEEE Journal of **PP** (2014), no. 99, 1–13. (cité page 26)
- [28] Peter Bühlmann and Sara Van De Geer, *Statistics for high-dimensional data : methods, theory and applications*, Springer Science & Business Media, 2011. (cité page 61)
- [29] Jack Capon, *High-resolution frequency-wavenumber spectrum analysis*, Proceedings of the IEEE **57** (1969), no. 8, 1408–1418. (cité page 3)
- [30] Lawrence Cayton and Sanjoy Dasgupta, *Robust euclidean embedding*, Proceedings of the 23rd international conference on machine learning, ACM, 2006, pp. 169–176. (cité page 50)
- [31] Michael Cerna and Audrey F Harvey, *The fundamentals of FFT-based signal analysis and measurement*, National Instruments, Junho (2000). (cité page 54)
- [32] Gilles Chardon, Wolfgang Kreuzer, and Markus Noisternig, *Design of spatial microphone arrays for sound field interpolation*, IEEE Journal of Selected Topics in Signal Processing **9** (2015), no. 5, 780–790. (cité page 2)
- [33] Gilles Chardon, Thilbault Nowakowski, Julien de Rosny, and Laurent Daudet, *A blind dereverberation method for narrowband source localization*, Selected Topics in Signal Processing, IEEE Journal of **9** (2015), no. 5, 815–824. (cité pages 3, 48 et 123)



- [34] Huawei Chen, Yu Bao, and Wee Ser, *Effects of sensor position errors on farfield/nearfield wideband beamformers for microphone arrays*, Sensors Journal, IEEE **15** (2015), no. 9, 4812–4825. (cité page 43)
- [35] Huawei Chen, Wee Ser, and Jianjiang Zhou, *Robust nearfield wideband beamformer design using worst case mean performance optimization with passband response variance constraint*, IEEE Transactions on Audio, Speech, and Language Processing **20** (2012), no. 5, 1565–1572. (cité page 43)
- [36] John Cheng, Max Grossman, and Ty McKercher, *Professional Cuda C Programming*, John Wiley & Sons, 2014. (cité page 34)
- [37] CF Chien and WW Soroka, *Spatial cross-correlation of acoustic pressures in steady and decaying reverberant sound fields*, Journal of Sound and Vibration **48** (1976), no. 2, 235–242. (cité page 48)
- [38] Ning Chu, Ali Mohammad-Djafari, and José Picheral, *A bayesian sparse inference approach in near-field wideband aeroacoustic imaging*, Proc. 19th IEEE Int. Conf. Image Processing, September 2012, pp. 2529–2532. (cité page 45)
- [39] Ning Chu, José Picheral, and Ali Mohammad-Djafari, *A robust super-resolution approach with sparsity constraint for near-field wideband acoustic imaging*, Proc. IEEE Int. Symp. Signal Processing and Information Technology (ISSPIT), December 2011, pp. 310–315. (cité page 45)
- [40] Richard K Cook, RV Waterhouse, RD Berendt, Seymour Edelman, and MC Thompson Jr, *Measurement of correlation coefficients in reverberant sound fields*, The Journal of the Acoustical Society of America **27** (1955), no. 6, 1072–1077. (cité pages 48 et 53)
- [41] Jose A. Costa, Neal Patwari, and Alfred O. Hero, *Distributed weighted-multidimensional scaling for node localization in sensor networks*, ACM Transactions On Sensor Networks **2** (2006), no. 1, 39–64. (cité pages 60 et 88)
- [42] Trevor F. Cox and Michael AA Cox, *Multidimensional scaling*, CRC Press, 2000. (cité page 58)
- [43] Marco Crocco, Alessio Del Bue, Matteo Bustreo, and Vittorio Murino, *A closed form solution to the microphone position self-calibration problem*, Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on, IEEE, 2012, pp. 2597–2600. (cité page 92)
- [44] Marco Crocco, Alessio Del Bue, and Vittorio Murino, *A bilinear approach to the position self-calibration of multiple sensors*, IEEE Transactions On Signal Processing **60** (2012), no. 2, 660–673. (cité page 92)
- [45] Jan de Leeuw and Patrick Mair, *Multidimensional scaling using majorization : Smacof in R*, Journal of Statistical Software **31** (2009), no. 3, 1–30. (cité pages 58 et 111)

- [46] Giovanni Del Galdo, Maja Taseska, Oliver Thiergart, Jukka Ahonen, and Ville Pulkki, *The diffuse sound field in energetic analysis*, The Journal of the Acoustical Society of America **131** (2012), no. 3, 2141–2151. (cité page 54)
- [47] Analog Devices, *Omnidirectional microphone with bottom port and I2S digital output ADMP441*, 2011. (cité pages 6, 12 et 20)
- [48] Maurizio di Bisceglie, Michele Di Santo, Carmela Galdi, Riccardo Lanari, and Nadia Ranaldo, *Synthetic aperture radar processing with GPGPU*, IEEE Signal Processing Magazine **27** (2010), no. 2, 69–78. (cité page 26)
- [49] EU Directive, *Directive 2002/49/ec of the european parliament and the council of 25 june 2002 relating to the assessment and management of environmental noise*, 2002. (cité page 26)
- [50] Hoang Do, Harvey F. Silverman, and Ying Yu, *A real-time SRP-PHAT source location implementation using stochastic region contraction (SRC) on a large-aperture microphone array*, Proc. IEEE Int. Conf. Acoustics Speech and Signal Processing - ICASSP '07, vol. 1, April 2007, pp. I–121–I–124. (cité page 92)
- [51] Simon Doclo and Marc Moonen, *Design of broadband beamformers robust against gain and phase errors in the microphone array characteristics*, IEEE Transactions on Signal Processing **51** (2003), no. 10, 2511–2526. (cité page 43)
- [52] Ivan Dokmanić, Reza Parhizkar, Juri Ranieri, and Martin Vetterli, *Euclidean distance matrices : Essential theory, algorithms, and applications*, IEEE Signal Processing Magazine **32** (2015), no. 6, 12–30. (cité pages 50, 88 et 99)
- [53] Ivan Dokmanić, Reza Parhizkar, Andreas Walther, Yue M. Lu, and Martin Vetterli, *Acoustic echoes reveal room shape*, Proceedings of the National Academy of Sciences **110** (2013), no. 30, 12186–12191. (cité page 48)
- [54] Petros Drineas, Asif Javed, Malik Magdon-Ismail, Gopal Pandurangant, Reino Virrankoski, and Andreas Savvides, *Distance matrix reconstruction from incomplete distance information for sensor network localization*, Proc. 3rd Annual IEEE Communications Society Sensor and Ad Hoc Communications and Networks, vol. 2, September 2006, pp. 536–544. (cité page 50)
- [55] Martin A. Fischler and Robert C. Bolles, *Random sample consensus : A paradigm for model fitting with applications to image analysis and automated cartography*, Commun. ACM **24** (1981), no. 6, 381–395. (cité page 93)
- [56] Derry FitzGerald, Antoine Liutkus, and Roland Badeau, *PROJET - Spatial audio separation using projections*, 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), March 2016, pp. 36–40. (cité page 123)

- [57] Brian P. Flanagan and Kristine L. Bell, *Array self-calibration with large sensor position errors*, *Signal Processing* **81** (2001), no. 10, 2201–2214. (cité page 44)
- [58] Pedro A. Forero and Georgios B. Giannakis, *Robust multi-dimensional scaling via outlier-sparsity control*, *Signals, Systems and Computers (ASILOMAR)*, 2011 Conference Record of the Forty Fifth Asilomar Conference on, 2011, pp. 1183–1187. (cité pages 63 et 131)
- [59] ———, *Sparsity-exploiting robust multidimensional scaling*, *IEEE Transactions On Signal Processing* **60** (2012), no. 8, 4118–4134. (cité pages 50, 61, 65, 93, 109, 111, 131 et 135)
- [60] Nikolay D. Gaubitch, W. Bastiaan Kleijn, and Richard Heusdens, *Auto-localization in ad-hoc microphone arrays*, *Acoustics, Speech and Signal Processing (ICASSP)*, 2013 IEEE International Conference on, 2013, pp. 106–110. (cité pages 92 et 104)
- [61] John C. Gower, *Procrustes methods*, *WIREs Comp Stat* **2** (2010), no. 4, 503–508. (cité page 74)
- [62] Rémi Gribonval, Gilles Chardon, and Laurent Daudet, *Blind calibration for compressed sensing by convex optimization*, *Proc. Speech and Signal Processing (ICASSP) 2012 IEEE Int. Conf. Acoustics*, March 2012, pp. 2713–2716. (cité page 44)
- [63] Ines Hafizovic, Carl-Inge Colombo Nilsen, Morgan Kjolerbakken, and Vibeke Jahr, *Design and implementation of a MEMS microphone array system for real-time speech acquisition*, *Applied Acoustics* **73** (2012), no. 2, 132–143 (English). (cité pages 6, 15 et 26)
- [64] Marius Hennecke, Thomas Plotz, Gernot A Fink, Jorg Schmalenstroer, and Reinhold Hab-Umbach, *A hierarchical approach to unsupervised shape calibration of microphone array networks*, *Statistical Signal Processing, 2009. SSP '09. IEEE/SP 15th Workshop on*, 2009, pp. 257–260. (cité pages 50 et 51)
- [65] Marius H. Hennecke and Gernot A. Fink, *Towards acoustic self-localization of ad hoc smart-phone arrays*, *Hands-free Speech Communication and Microphone Arrays (HSCMA)*, 2011 Joint Workshop on, 2011, pp. 127–132. (cité pages 92 et 93)
- [66] David A. Hill, *Spatial correlation function for fields in a reverberation chamber*, *IEEE Transactions on Electromagnetic Compatibility* **37** (1995), no. 1, 138. (cité page 125)
- [67] ———, *Plane wave integral representation for fields in reverberation chambers*, *IEEE Transactions on Electromagnetic Compatibility* **40** (1998), no. 3, 209–217. (cité page 49)
- [68] Ivan Himawan, Sridha Sridharan, and Iain McCowan, *Dealing with uncertainty in microphone placement in a microphone array speech recognition system*, 2008 IEEE International Conference on Acoustics, Speech and Signal Processing, March 2008, pp. 1565–1568. (cité page 43)
- [69] Charlotte Hurltley, *Night noise guidelines for europe*, WHO Regional Office Europe, 2009. (cité page 26)

- [70] InvenSense, *Applicatios note 1003 : Recommandations for mounting and connecting invensense mems microphones*, dec 2013. (cité page 15)
- [71] InvenSense, *Omnidirectional microphone with bottom port and I2S digital output INMP441*, 2014. (cité page 13)
- [72] ———, *ICS-52000 product brief : Low-noise microphone with TDM digital output*, 2016. (cité page 13)
- [73] Finn Jacobsen, *The diffuse sound field*, Tech. Report Report No. 27, The Acoustics Laboratory, Technical University of Denmark, Lyngby, Denmark, 1979. (cité pages 48 et 125)
- [74] Finn Jacobsen and Thibaut Roisin, *The coherence of reverberant sound fields*, The Journal of the Acoustical Society of America **108** (2000), no. 1, 204–210. (cité page 53)
- [75] Wolfgang Kabsch, *A solution for the best rotation to relate two sets of vectors*, Acta Crystallographica Section A : Crystal Physics, Diffraction, Theoretical and General Crystallography **32** (1976), no. 5, 922–923. (cité pages 74 et 117)
- [76] Sarthak Khanal, Harvey F. Silverman, and Rahul R. Shakya, *A free-source method (FrSM) for calibrating a large-aperture microphone array*, IEEE Transactions On Audio Speech and Language Processing **21** (2013), no. 8, 1632–1639. (cité pages 8, 92, 97 et 116)
- [77] David Blair Kirk and Wen mei W. Hwu, *Programming massively parallel processors - a hands-on approach.*, Morgan Kaufmann, 2010. (cité page 34)
- [78] Andreas Klöckner, Nicolas Pinto, Yunsup Lee, B. Catanzaro, Paul Ivanov, and Ahmed Fasih, *PyCUDA and PyOpenCL : A scripting-based approach to gpu run-time code generation*, Parallel Computing **38** (2012), no. 3, 157–174. (cité page 35)
- [79] Charles H. Knapp and G. Clifford Carter, *The generalized correlation method for estimation of time delay*, IEEE Transactions on Acoustics, Speech, and Signal Processing **24** (1976), no. 4, 320–327. (cité page 96)
- [80] Hyung-Seok Kook, G. B. Moebs, Patricia Davies, and J. S. Bolton, *An efficient procedure for visualizing the sound field radiated by vehicles during standardized passby tests*, Journal of sound and vibration **233** (2000), no. 1, 137–156 (English). (cité page 26)
- [81] Yuji Koyano, Kohei Yatabe, Yusuke Ikeda, and Yasuhiro Oikawa, *Physical-model based efficient data representation for many-channel microphone array*, Proc. Speech and Signal Processing (ICASSP) 2016 IEEE Int. Conf. Acoustics, March 2016, pp. 370–374. (cité page 7)
- [82] Martin Kuster, *Spatial correlation and coherence in reverberant acoustic fields : Extension to microphones with arbitrary first-order directivity*, The Journal of the Acoustical Society of America **123** (2008), no. 1, 154–162. (cité pages 53 et 76)

- [83] Yenming M. Lai, Radu Balan, Heiko Claussen, and Justinian Rosea, *Broadband sensor location selection using convex optimization in very large scale arrays*, Applications of Signal Processing to Audio and Acoustics (WASPAA), 2013 IEEE Workshop on, Oct 2013, pp. 1–4. (cité page 7)
- [84] Trung-Kien Le and Nobutaka Ono, *Reference-distance estimation approach for TDOA-based source and sensor localization*, Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on, 2015, pp. 2549–2553. (cité page 92)
- [85] ———, *Closed-form and near closed-form solutions for TOA-based joint source and sensor localization*, IEEE Transactions on Signal Processing **64** (2016), no. 18, 4751–4766. (cité page 92)
- [86] Trung-Kien Le, Nobutaka Ono, Thibault Nowakowski, Laurent Daudet, and Julien De Rosny, *Experimental validation of TOA-based methods for microphones array positions calibration*, Proc. Speech and Signal Processing (ICASSP) 2016 IEEE Int. Conf. Acoustics, March 2016, pp. 3216–3220. (cité pages 8, 92 et 105)
- [87] Raphael Leiba, Francois Ollivier, Régis Marchiano, Nicolas Misdariis, and Jacques Marchal, *Imagerie acoustique en milieu urbain : de la mesure à la perception du paysage sonore*, Congrès Français d’Acoustique, 2016. (cité pages 21 et 22)
- [88] Jerad Lewis and Paul Schreier, *Application note : Low self noise : The first step to high-performance mems microphone applications*, InvenSense, dec 2013. (cité page 20)
- [89] Bing Li, Yigang He, Fengming Guo, and Lei Zuo, *A novel localization algorithm based on isomap and partial least squares for wireless sensor networks*, IEEE Transactions On Instrumentation and Measurement **62** (2013), no. 2, 304–314. (cité page 60)
- [90] Jorge Lorente, Gema Piñero, Antonio M. Vidal, Jose Antonio Belloch, and Alberto Gonzalez, *Parallel implementations of beamforming design and filtering for microphone array applications*, Signal Processing Conference, 2011 19th European, 2011, pp. 501–505. (cité page 7)
- [91] Robert G. Lorenz and Stephen P. Boyd, *Robust minimum variance beamforming*, IEEE Transactions on Signal Processing **53** (2005), no. 5, 1684–1696. (cité pages 3 et 38)
- [92] Régis Marchiano, Philippe Druault, Raphael Leiba, Jacques Marchal, François Ollivier, Vincent Valeau, and Charles Vanwynsberghe, *Localisation de sources aéroacoustiques par une méthode de retournement temporel tri-dimensionnelle*, Congrès Français d’Acoustique, 2016. (cité pages 21, 22 et 123)
- [93] Carlos J. Martin-Arguedas, David Romero-Laorden, Oscar Martinez-Graullera, Manuel Perez-Lopez, and Luis Gomez-Ullate, *An ultrasonic imaging system based on a new soft approach and a gpu beamformer*, IEEE transactions on ultrasonics ferroelectrics and frequency control **59** (2012), no. 7, SI, 1402–1412 (English). (cité page 26)

- [94] Iain McCowan, Mike Lincoln, and Ivan Himawan, *Microphone array shape calibration in diffuse noise fields*, IEEE Transactions On Audio Speech and Language Processing **16** (2008), no. 3, 666–670. (cité pages 49, 51, 76 et 77)
- [95] Cyril Mellet, Fabien Létourneaux, Frank Poisson, and Corinne Talotte, *High speed train noise emission : Latest investigation of the aerodynamic/rolling noise contribution*, Journal of sound and vibration **293** (2006), no. 3, 535–546. (cité page 26)
- [96] Thibaut Meurisse, François Ollivier, Jacques Marchal, and Charles Vanwynsberghe, *Analyse modale en temps réel par holographie de champ proche et calculs parallèles*, Congrès Français d’Acoustique, 2016. (cité pages 21 et 22)
- [97] Nicolas Misdariis, Régis Marchiano, Patrick Susini, François Ollivier, Raphael Leiba, and Jacques Marchal, *Mobility and life quality relationships—measurement and perception of noise in urban context*, Internoise-International Congress on Noise Control Engineering, 2014. (cité pages 21 et 22)
- [98] Dmitri Moltchanov, *Distance distributions in random networks*, Ad Hoc Networks **10** (2012), no. 6, 1146–1166. (cité page 88)
- [99] Aaftab Munshi, Benedict Gaster, Timothy G Mattson, and Dan Ginsburg, *Opencl programming guide*, Pearson Education, 2011. (cité page 34)
- [100] Arulkumaran Muthukumarasamy and Kevin D. Donohue, *Impact of microphone placement errors on speech intelligibility*, IEEE Southeastcon 2009, March 2009, pp. 323–328. (cité page 43)
- [101] Boon Poh Ng, Meng Hwa Er, and C. Kot, *Array gain/phase calibration techniques for adaptive beamforming and direction finding*, Sonar and Navigation IEE Proceedings-Radar **141** (1994), no. 1, 25–29. (cité page 44)
- [102] Boon Poh Ng, Joni Polili Lie, Meng Hwa Er, and Aigang Feng, *A practical simple geometry and gain/phase calibration technique for antenna array processing*, IEEE Transactions on Antennas and Propagation **57** (2009), no. 7, 1963–1972. (cité page 44)
- [103] C. I. C. Nilsen and I. Hafizovic, *Digital beamforming using a GPU*, Proc. Speech and Signal Processing 2009 IEEE Int. Conf. Acoustics, April 2009, pp. 609–612. (cité page 26)
- [104] R. C. Nongpiur, *Design of minimax broadband beamformers that are robust to microphone gain, phase, and position errors*, IEEE/ACM Transactions on Audio, Speech, and Language Processing **22** (2014), no. 6, 1013–1022. (cité pages 43 et 44)
- [105] Thibault Nowakowski, Laurent Daudet, and Julien De Rosny, *Méthode parcimonieuse pour la localisation de sources en milieux réverbérants hétérogènes*, Actes du Congrès Français d’Acoustique, Poitiers, 2014. (cité page 48)

- [106] Thibault Nowakowski, Laurent Daudet, and Julien de Rosny, *Localization of acoustic sensors from passive green's function estimation*, The Journal of the Acoustical Society of America **138** (2015), no. 5, 3010–3018. (cité pages 53 et 129)
- [107] Thibault Nowakowski, Julien De Rosny, Laurent Daudet, and François Ollivier, *Blind source localization in a room based on wavefield separation*, The Journal of the Acoustical Society of America **137** (2015), no. 4, 2388–2388. (cité page 123)
- [108] Hugues Néglise and Jean Nicolas, *Characterization of a diffuse field in a reverberant room*, The Journal of the Acoustical Society of America **101** (1997), no. 6, 3517–3524. (cité page 53)
- [109] Adam O'Donovan, Ramani Duraiswami, and Nail A. Gumerov, *Real time capture of audio images and their use with video*, Proc. IEEE Workshop Applications of Signal Processing to Audio and Acoustics, October 2007, pp. 10–13. (cité page 34)
- [110] Nobutaka Ono, Hitoshi Kohno, Nobutaka Ito, and Shigeki Sagayama, *Blind alignment of asynchronously recorded signals for distributed microphone array*, Applications of Signal Processing to Audio and Acoustics, 2009. WASPAA '09. IEEE Workshop on, Oct 2009, pp. 161–164. (cité pages 92, 104, 105 et 109)
- [111] Alexey Ozerov, Emmanuel Vincent, and Frédéric Bimbot, *A general flexible framework for the handling of prior information in audio source separation*, IEEE Transactions on Audio, Speech, and Language Processing **20** (2012), no. 4, 1118–1133. (cité page 123)
- [112] Thomas Padois, Philippe-Aubert Gauthier, and Alain Berry, *Inverse problem with beamforming regularization matrix applied to sound source localization in closed wind-tunnel using microphone array*, Journal of Sound and Vibration **333** (2014), no. 25, 6858–6868. (cité page 38)
- [113] Reza Parhizkar, Amin Karbasi, Sewong Oh, and Martin Vetterli, *Calibration using matrix completion with application to ultrasound tomography*, IEEE Transactions on Signal Processing **61** (2013), no. 20, 4923–4933. (cité page 59)
- [114] Neal Patwari and Alfred O. Hero, *Manifold learning algorithms for localization in wireless sensor networks*, Proc. IEEE Int. Conf. Acoustics, Speech, and Signal Processing (ICASSP '04), vol. 3, May 2004, pp. iii–857–60 vol.3. (cité page 60)
- [115] Antoine Peillot, *Imagerie acoustique par approximations parcimonieuses des sources*, Ph.D. thesis, Université Pierre et Marie Curie-Paris VI, 2012. (cité pages 3, 4, 8, 38, 45 et 123)
- [116] François Perrodin, Janosch Nikolic, Joël Busset, and Roland Siegwart, *Design and calibration of large microphone arrays for robotic applications*, Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on, Oct 2012, pp. 4596–4601. (cité page 93)

- [117] Axel Plinge and Gernot A. Fink, *Geometry calibration of multiple microphone arrays in highly reverberant environments*, Proc. 14th Int Acoustic Signal Enhancement (IWAENC) Workshop, September 2014, pp. 243–247. (cité page 8)
- [118] Axel Plinge, Florian Jacob, Reinhold Haeb-Umbach, and Gernot A. Fink, *Acoustic microphone geometry calibration : An overview and experimental evaluation of state-of-the-art algorithms*, IEEE Signal Processing Magazine **33** (2016), no. 4, 14–29. (cité page 7)
- [119] Guinness World Records, *Largest microphone array*, <http://www.guinnessworldrecords.com/world-records/largest-microphone-array/>, 2014 (accessed June 2016). (cité page 6)
- [120] David Romero-Laorden, Javier Villazon-Terrazas, Oscar Martinez-graullera, Alberto Ibanez, Montserrat Parrilla, and Matilde Santos, *Analysis of parallel computing strategies to accelerate ultrasound imaging processes*, IEEE Transactions on Parallel and Distributed Systems **PP** (2016), no. 99, 1. (cité page 26)
- [121] Philippe Roux, Karim G. Sabra, William A. Kuperman, and Andre Roux, *Ambient noise cross correlation in free space : Theoretical approach*, The Journal of the Acoustical Society of America **117** (2005), no. 1, 79–84. (cité pages 85 et 129)
- [122] Sam T. Roweis and Lawrence K. Saul, *Nonlinear dimensionality reduction by locally linear embedding*, Science **290** (2000), no. 5500, 2323–2326. (cité page 60)
- [123] Olivier Roy, Ivana Jovanović, Neb Durić, Louis Poulo, and Martin Vetterli, *Robust array calibration using time delays with application to ultrasound tomography*, SPIE Medical Imaging, International Society for Optics and Photonics, 2011, pp. 796806.1–796806.11. (cité page 50)
- [124] Joshua M. Sachar, Harvey F. Silverman, and William R. Patterson, *Large vs small aperture microphone arrays : performance over a large focal area*, Acoustics, Speech, and Signal Processing, 2001. Proceedings. (ICASSP '01). 2001 IEEE International Conference on, vol. 5, 2001, pp. 3049–3052 vol.5. (cité page 3)
- [125] ———, *Microphone position and gain calibration for a large-aperture microphone array*, IEEE Transactions On Speech and Audio Processing **13** (2005), no. 1, 42–52. (cité pages 8, 42, 43, 81, 92 et 97)
- [126] Shuichi Sakamoto, Jun'ichi Kodama, Satoshi Hongo, Takuma Okamoto, Yukio Iwaya, and Yôiti Suzuki, *A 3d sound-space recording system using spherical microphone array with 252ch microphones*, Proceedings of 20th International Congress on Acoustics, 2010. (cité page 4)
- [127] Ralph O. Schmidt, *Multiple emitter location and signal parameter estimation*, IEEE Transactions on Antennas and Propagation **34** (1986), no. 3, 276–280. (cité page 3)



- [128] Manfred R Schroeder, *New method of measuring reverberation time*, The Journal of the Acoustical Society of America **37** (1965), no. 3, 409–412. (cité page 48)
- [129] Manfred R. Schroeder, *The "schroeder frequency" revisited*, The Journal of the Acoustical Society of America **99** (1996), no. 5, 3240–3241. (cité page 48)
- [130] P Jafari Shalkouhi, *Room modes with regard to speech intelligibility in classrooms : Comments on knecht et al.(2002)*, American journal of audiology **19** (2010), no. 2, 139–141. (cité page 57)
- [131] Yiyuan She and Art B Owen, *Outlier detection using nonconvex penalized regression*, Journal of the American Statistical Association (2012). (cité pages 94 et 109)
- [132] Pieter Sijtsma and Robert Stoker, *Determination of absolute contributions of aircraft noise components using fly-over array measurements*, 10th AIAA/CEAS Aeroacoustics Conference, vol. 2958, 2004, p. 10. (cité page 26)
- [133] Harvey F. Silverman and William R. Patterson, *Visualizing the performance of large-aperture microphone arrays*, Proc. Conf. IEEE Int Acoustics, Speech, and Signal Processing, vol. 2, March 1999, pp. 969–972 vol.2. (cité page 4)
- [134] Harvey F. Silverman, William R. Patterson, and James L. Flanagan, *The huge microphone array (HMA)*, The Journal of the Acoustical Society of America **101** (1997), no. 5, 3119–3119. (cité page 4)
- [135] ———, *The huge microphone array*, Concurrency, IEEE **6** (1998), no. 4, 36–46. (cité page 4)
- [136] Harvey F. Silverman, William R. Patterson, and Joshua Sachar, *Factors affecting the performance of large-aperture microphone arrays*, The Journal of the Acoustical Society of America **111** (2002), no. 5, 2140–2157. (cité page 4)
- [137] Sunil Srinivasa and Martin Haenggi, *Distance distributions in finite uniformly random networks : Theory and applications*, Vehicular Technology, IEEE Transactions on **59** (2010), no. 2, 940–949. (cité page 88)
- [138] Mohammad J. Taghizadeh, *Theoretical analysis of euclidean distance matrix completion for ad hoc microphone array calibration*, Idiap-RR Idiap-RR-20-2014, Idiap, 11 2014. (cité pages 8 et 50)
- [139] Mohammad J. Taghizadeh, Philip N. Garner, and Herve Boursard, *Enhanced diffuse field model for ad hoc microphone array calibration*, Signal Processing **101** (2014), 242–255 (English). (cité pages 49, 51, 54, 76, 77, 78 et 81)
- [140] Mohammad J. Taghizadeh, Reza Parhizkar, Philip N. Garner, and Herve Boursard, *Euclidean distance matrix completion for ad-hoc microphone array calibration*, Digital Signal Processing (DSP), 2013 18th International Conference on, July 2013, pp. 1–7. (cité page 50)

- [141] Mohammad J. Taghizadeh, Reza Parhizkar, Philip N. Garner, Hervé Boursard, and Afsaneh Asaei, *Ad hoc microphone array calibration : Euclidean distance matrix completion algorithm and theoretical guarantees*, Signal Processing **107** (2015), 123–140, Special Issue on ad hoc microphone arrays and wireless acoustic sensor networks Special Issue on Fractional Signal Processing and Applications. (cité page 50)
- [142] Tomoya Takagi, Hiroki Noguchi, Koji Kugata, Masahiko Yoshimoto, and Hiroshi Kawaguchi, *Microphone array network for ubiquitous sound acquisition*, Proc. Speech and Signal Processing 2010 IEEE Int. Conf. Acoustics, March 2010, pp. 1474–1477. (cité page 4)
- [143] Yuki Tamai, Satoshi Kagami, Hiroshi Mizoguchi, Yutaka Amemiya, Koichi Nagashima, and Tachio Takano, *Real-time 2 dimensional sound source localization by 128-channel huge microphone array*, Proc. 13th IEEE Int. Workshop Robot and Human Interactive Communication ROMAN 2004, September 2004, pp. 65–70. (cité page 4)
- [144] Yuki Tamai, Satoshi Kagami, Hiroshi Mizoguchi, Kota Sakaya, Koichi Nagashima, and Tachio Takano, *Circular microphone array for meeting system*, Sensors, 2003. Proceedings of IEEE, vol. 2, 2003, pp. 1100–1105 Vol.2. (cité page 4)
- [145] Joshua B. Tenenbaum, Vin De Silva, and John C. Langford, *A global geometric framework for nonlinear dimensionality reduction*, Science **290** (2000), no. 5500, 2319–2323. (cité pages 60, 79 et 86)
- [146] Joel A. Tropp and Stephen J. Wright, *Computational methods for sparse solution of linear inverse problems*, Proceedings of the IEEE **98** (2010), no. 6, 948–958. (cité page 93)
- [147] S. D. Valente, M. Tagliasacchi, F. Antonacci, P. Bestagini, A. Sarti, and S. Tubaro, *Geometric calibration of distributed microphone arrays from acoustic source correspondences*, Proc. IEEE Int Multimedia Signal Processing (MMSP) Workshop, October 2010, pp. 13–18. (cité page 8)
- [148] Charles Vanwysberghe, Pascal Challande, Raphael Leiba, Jacques Marchal, Régis Marchiano, François Ollivier, Gilles Puy, and Pierre Vandergheynst, *Localisation et identification spectrale conjointe de sources large bande par parcimonie groupée*, Congrès Français d’Acoustique, 2016. (cité page 21)
- [149] Charles Vanwysberghe, Pascal Challande, Jacques Marchal, Régis Marchiano, and François Ollivier, *A robust and passive method for geometric calibration of large arrays*, The Journal of the Acoustical Society of America **139** (2016), no. 3, 1252–1263. pas de citations
- [150] Charles Vanwysberghe, Pascal Challande, Jacques Marchal, Régis Marchiano, and François Ollivier, *Blind geometric calibration of large arrays in diffuse field*, 12th Western Pacific Acoustics Conference, 2015. pas de citations

- [151] Charles Vanwysberghe, Pascal Challande, Jacques Marchal, François Ollivier, and Régis Marchiano, *Localisation des microphones d'un grand réseau en champ diffus*, Congrès Français d'Acoustique, 2016. *pas de citations*
- [152] Charles Vanwysberghe, Jacques Marchal, François Ollivier, Pascal Challande, and Régis Marchiano, *Développement d'un sonomètre imageur temps réel à représentation par bande de fréquences*, Congrès Français d'Acoustique, 2014, pp. 403–409. *pas de citations*
- [153] Charles Vanwysberghe, Régis Marchiano, François Ollivier, Pascal Challande, Hélène Moingeon, and Jacques Marchal, *Design and implementation of a multi-octave-band audio camera for realtime diagnosis*, Applied Acoustics **89** (2015), 281–287. (*cité page 26*)
- [154] Jose Velasco, Daniel Pizarro, Javier Macias-Guarasa, and Afsaneh Asaei, *TDOA matrices : Algebraic properties and their application to robust denoising with missing data*, arXiv preprint arXiv :1601.04485 (2016). (*cité pages 106, 107 et 108*)
- [155] Jarkko Venna and Samuel Kaski, *Local multidimensional scaling*, Neural Networks **19** (2006), no. 6, 889–899, Advances in Self Organising Maps. (*cité page 60*)
- [156] Sergiy A. Vorobyov, Alex B. Gershman, and Zhi-Quan Luo, *Robust adaptive beamforming using worst-case performance optimization : a solution to the signal mismatch problem*, IEEE Transactions on Signal Processing **51** (2003), no. 2, 313–324. (*cité page 43*)
- [157] Chengqun Q. Wang, Jiming M. Chen, Youxian X. Sun, and Xuemin S. Shen, *Wireless sensor networks localization with isomap*, Communications, 2009. ICC'09. IEEE International Conference on, IEEE, 2009, pp. 28–32. (*cité page 60*)
- [158] Lin Wang, Tsz-Kin Hon, Joshua D. Reiss, and Andrea Cavallaro, *Self-localization of ad-hoc arrays using time difference of arrivals*, IEEE Transactions on Signal Processing **64** (2016), no. 4, 1018–1033. (*cité page 7*)
- [159] Siyi Wang, Weisi Guo, and Mark D McDonnell, *Distance distributions for real cellular networks*, Computer Communications Workshops (INFOCOM WKSHPS), 2014 IEEE Conference on, IEEE, 2014, pp. 181–182. (*cité page 88*)
- [160] Yanwei Wang, Jian Li, Petre Stoica, Mark Sheplak, and Toshikazu Nishida, *Wideband relax and wideband clean for aeroacoustic imaging*, The Journal of the Acoustical Society of America **115** (2004), no. 2, 757–767. (*cité page 45*)
- [161] Darren B. Ward, Eric A. Lehmann, and Robert C. Williamson, *Particle filtering algorithms for tracking an acoustic source in a reverberant environment*, Speech and Audio Processing, IEEE Transactions on **11** (2003), no. 6, 826–836. (*cité page 48*)

- [162] Eugene Weinstein, Kenneth Steele, Anant Agarwal, and James Glass, *A 1020-node modular microphone array and beamformer for intelligent computing spaces*, Tech. report, 2004. (cité pages 4 et 5)
- [163] ———, *A 1020-node microphone array and acoustic beamformer*, International Congress on Sound and Vibration (ICSV), July 2007. (cité page 4)
- [164] Earl G. Williams, *Fourier acoustics : sound radiation and nearfield acoustical holography*, Academic press, 1999. (cité pages 2 et 27)
- [165] Margaret H Wright, *Direct search methods : Once scorned, now respectable*, Pitman Research Notes in Mathematics Series (1996), 191–208. (cité page 54)
- [166] Angeliki Xenaki and Peter Gerstoft, *Grid-free compressive beamforming*, The Journal of the Acoustical Society of America **137** (2015), no. 4, 1923–1935. (cité page 3)
- [167] Angeliki Xenaki, Peter Gerstoft, and Klaus Mosegaard, *Compressive beamforming*, The Journal of the Acoustical Society of America **136** (2014), no. 1, 260–271. (cité page 3)
- [168] Billy Y. S. Yiu and Alfred C. H. Yu, *Gpu-based minimum variance beamformer for synthetic aperture imaging of the eye*, Ultrasound in medicine & biology **41** (2015), no. 3, 871–883. (cité page 26)
- [169] Simayijiang Zhayida, Frederik Andersson, Yubin Kuang, and Kalle Åström, *An automatic system for microphone self-localization using ambient sound*, 2014 22nd European Signal Processing Conference (EUSIPCO), Sept 2014, pp. 954–958. (cité pages 93 et 119)
- [170] Ziguu Zhong and Tian He, *Achieving range-free localization beyond connectivity*, Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems, ACM, 2009, pp. 281–294. (cité page 60)
- [171] Tianyi Zhou and Dacheng Tao, *Godec : Randomized low-rank & sparse matrix decomposition in noisy case*, Proceedings of the 28th International Conference on Machine Learning (ICML-11) (New York, NY, USA) (Lise Getoor and Tobias Scheffer, eds.), ICML '11, ACM, June 2011, pp. 33–40. (cité page 107)
- [172] Tianyi Zhou and Dacheng Tao, *Shifted subspaces tracking on sparse outlier for motion segmentation.*, IJCAI, Citeseer, 2013. (cité pages 107, 108 et 109)

