

Acquisition comprimée en acoustique audible

L. Daudet^a, N. Bertin^b, V. Emiya^c et R. Gribonval^b

^aInstitut Langevin, 1 rue Jussieu, 75005 Paris, France

^bCNRS, IRISA, Campus de Beaulieu, 35042 Rennes Cedex, France

^cLaboratoire d'informatique Fondamentale de Marseille (LIF), CMI Technopôle de Château-Gombert, 39,
rue F. Joliot Curie, 13453 Marseille Cedex 13, France
laurent.daudet@espci.fr

L'acoustique est un domaine se prêtant particulièrement bien aux nouvelles théories d'échantillonnage appelées acquisition comprimée, car les équations sont linéaires, et le nombre d'échantillons à acquérir, ou le nombre de capteurs, est souvent très important. Cette acquisition comprimée suppose d'une part une forme de parcimonie du signal à acquérir, et d'autre part des mesures qui sont incohérentes avec cette base de parcimonie. Plusieurs scénarios sont envisagés suivant le type de parcimonie, qui peut être dans le nombre de sources ponctuelles, ou dans une paramétrisation du champ acoustique en un faible nombre de fonctions élémentaires (par exemple des ondes planes ou des fonctions de Fourier-Bessel). Nous présentons ici plusieurs études en acoustique audible, produites dans le cadre du projet ANR ECHANGE : holographie acoustique en champ proche, réponses impulsionnelles de salles, localisation de sources en milieu réverbérant. Des questions intéressantes sont soulevées par la mise en oeuvre pratique de tels systèmes, comme la discrétisation du dictionnaire choisi, ou la sensibilité à l'étalonnage du réseau de capteurs.

1 Introduction

L'acquisition comprimée, ou échantillonnage compressé (*compressed sensing / compressive sampling*, CS), et plus généralement les techniques reposant sur la parcimonie, ont permis de radicalement repenser un grand nombre de problèmes en sciences de l'ingénieur - typiquement, de type "problèmes inverses" -, par exemple pour des applications à l'optique, l'astrophysique, ou l'imagerie par résonance magnétique. Au-delà d'une nouvelle méthode d'analyse des données et de l'introduction de nouveaux *a priori*, ce changement de paradigme cherche à optimiser conjointement l'acquisition des données (jusqu'à faire des recommandations sur la géométrie des réseaux de capteurs) et leur traitement (partie algorithmique).

L'imagerie acoustique possède un certain nombre de caractéristiques qui rendent ce champ d'application particulièrement adapté à l'échantillonnage compressé, ou plus généralement aux modèles parcimonieux :

- des données de grandes dimensions : l'acquisition par exemple d'un champ de pression acoustique génère une quantité de données très élevée, ce qui nécessite des capacités adaptées pour le stockage, la transmission et le traitement ;
- des données hautement structurées : les réseaux de capteurs comme les antennes acoustiques acquièrent des signaux fortement corrélés ;
- des capteurs linéaires : le comportement de la plupart des capteurs acoustiques tels que les microphones ou les hydrophones est approximativement linéaire dans les régimes d'utilisation courants ;
- des équations linéaires : de même, en régime standard, l'équation d'onde a un comportement linéaire, ainsi que la plupart des EDPs connexes qui déterminent la physique des phénomènes acoustiques, de telle sorte que les phénomènes observés dépendent linéairement des sources de production, que ce soit dans les scénarios actifs ou passifs.

Ces caractéristiques nous poussent naturellement à formaliser l'imagerie acoustique comme une série de problèmes inverses linéaires, où la réduction de dimension peut être mise à profit pour acquérir ces objets de grande dimension avec un faible nombre de mesures linéaires. Cependant, la mise en oeuvre pratique d'une régularisation parcimonieuse et des outils d'acquisition comprimée soulèvent un certain nombre de questions :

- quelles sont les sources de parcimonie ou de basse dimensionnalité ?

- quels sont les gains à espérer en pratique : gain de performance, économie en nombre de capteurs, limitation de la complexité algorithmique ou réduction du nombre de données ?
- quelles marges de manœuvre a-t-on pour effectuer vraiment une acquisition comprimée, c'est-à-dire ici pour optimiser le système de mesure de façon à maximiser les bénéfices de la régularisation parcimonieuse ?

Dans cet article, nous présentons un certain nombre de scénarios en acoustique audible, pour lesquels des résultats récents ont montré l'intérêt de l'acquisition comprimée. La plupart de ces travaux ont été effectués dans le cadre du projet ANR ECHANGE¹, réunissant des compétences en acoustique, traitement du signal et mathématiques appliquées. Cette présentation n'a bien entendu aucune prétention à l'exhaustivité, mais cherche avant tout à dégager un certain nombre de problématiques liées à la modélisation et la mise en oeuvre pratique de ces techniques.

Cet article présente tout d'abord brièvement le principe de l'acquisition comprimée, puis décrit son application à trois contextes d'imagerie acoustique : l'holographie acoustique en champ proche, l'échantillonnage de la "fonction plénacoustique", et enfin la localisation de sources (dans une première étude, pour des sources à bande étroite en milieu réverbérant ; dans une seconde étude, pour une localisation et une caractérisation jointes de sources directionnelles). La conclusion aborde certaines questions ouvertes par la mise en oeuvre pratique - par exemple la sensibilité à la connaissance du système d'acquisition -, ainsi que les pistes de travaux actuels.

2 Qu'est ce que l'acquisition comprimée ?

L'acquisition comprimée nous montre comment acquérir des signaux parcimonieux avec un nombre de mesures plus restreint que ce que nous dit le théorème d'échantillonnage classique de Shannon-Nyquist (qui ne fait des hypothèses que sur la bande passante), tout en apportant certaines garanties théoriques sur la qualité de la reconstruction. Il est à noter que l'hypothèse de parcimonie (ou plus généralement de régularité et / ou de structure liées à la parcimonie) est en fait vérifiée pour de nombreux signaux naturels, comme par exemple les signaux de parole ou les images : c'est d'ailleurs cette propriété qui est exploitée pour leur codage.

¹ le site <http://echange.inria.fr> indique l'ensemble des références bibliographiques sur ce projet

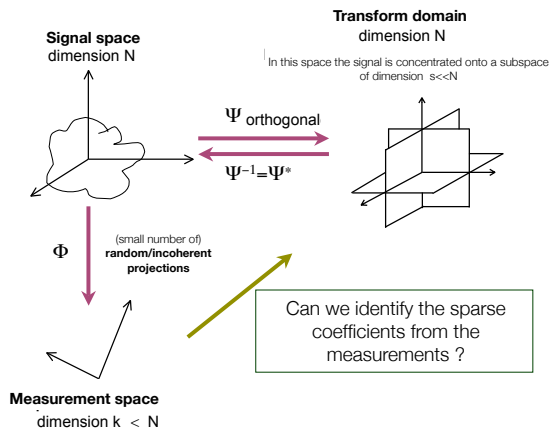


FIGURE 1 – Principe de l’acquisition comprimée.

La figure 1 représente le principe de l’acquisition comprimée. On suppose que le signal à acquérir, de dimension N , est parcimonieux, c’est-à-dire qu’il peut être représenté, exactement ou approximativement, par un faible nombre s de coefficients dans une base Ψ (ou plus généralement un dictionnaire qui peut être redondant) :

$$x = \psi\alpha \quad (1)$$

où α est un vecteur de coefficients dont seulement s composantes au maximum sont non-nulles (on utilise souvent la pseudo-norme ℓ_0 qui compte le nombre de coefficients non-nuls d’un vecteur, avec cette notation on a donc $\|\alpha\|_0 \leq s$). Pour acquérir ce signal, on effectue un certain nombre k de mesures linéaires par produit scalaire avec les k colonnes de la matrice de mesure Φ . Retrouver les s coefficients non nuls à partir des k mesures s’exprime alors comme un problème inverse linéaire sous-déterminé ($k < N$), sous contrainte de parcimonie. Les articles fondateurs de l’acquisition comprimée [11, 5] montrent que, sous réserve d’une condition d’incohérence entre Φ et Ψ , la minimisation de la norme ℓ_1 des coefficients permet de retrouver la solution parcimonieuse de façon unique, avec un nombre de mesures k proportionnel à $s \log N$, c’est-à-dire essentiellement proportionnel au degré de parcimonie du signal, qui est une forme de mesure de complexité. Si $s \ll N$, le gain en termes de nombre d’échantillons peut être très significatif. Notons que le fait de minimiser la norme ℓ_1 des coefficients renforce la parcimonie mais d’une façon différente que celle définie ci-dessus par la pseudo-norme ℓ_0 . En fait, chercher à minimiser exactement ℓ_0 se heurte à un problème combinatoire, on peut donc soit chercher à minimiser de la norme ℓ_1 , ce qui est un problème d’optimisation convexe, soit chercher des solutions approchées au minimum de la norme ℓ_0 , par exemple avec des algorithmes gloutons (*greedy*) comme Orthogonal Matching Pursuit (OMP).

Maintenant mûr d’une dizaine d’années de recherches intense, le CS est un domaine extrêmement actif, d’un point de vue algorithmique et théorique². De très nombreuses applications ont également vu le jour, dans tous les domaines

2. Le blog Nuit Blanche <http://nuit-blanche.blogspot.com>, mis à jour quotidiennement, fournit de très nombreuses références sur le CS et les techniques voisines.

de l’ingénierie. Pour résumer, le CS est une technique d’acquisition des signaux qui repose sur deux principes :

- **la parcimonie** des signaux à acquérir, dans une base connue à l’avance ; il est à noter que cette contrainte de parcimonie peut être relâchée par celle de “compressibilité”, qui exprime le fait que si l’on classe les coefficients de α par amplitude décroissante, leur décroissance rapide permet de contrôler la qualité de l’approximation à s termes.
- **une incohérence** des mesures relativement à la base de parcimonie, afin que chaque mesure apporte une information globale sur l’ensemble des coefficients. Si l’on dispose d’une entière liberté de choisir sa matrice de mesure, les matrices aléatoires offrent des garanties probabilistes sur le degré d’incohérence, et ce quelle que soit la base Φ , elles sont donc très souvent utilisées. En pratique, il existe souvent de fortes contraintes, physiques et matérielles, sur le choix de la matrice de mesure.

Lorsque le projet ANR ECHANGE a commencé, en 2009, très peu d’applications pratiques du CS en acoustique avait été développées. Le but de ce projet a été d’estimer quels sont les gains possibles et contraintes, dans différents scénarios d’imagerie acoustique, audible ou ultrasonore.

3 Holographie Acoustique en Champ Proche

L’holographie acoustique en champ proche, en anglais *nearfield acoustic holography* (NAH) est une technique d’imagerie visant à estimer la vibration d’une structure à partir de mesures de la pression acoustique rayonnée en champ proche (mesure d’ondes évanescentes). L’approche traditionnelle de la NAH, qui repose sur une rétropropagation du champ de pression, régularisée par la méthode de Tikhonov, implique une durée d’acquisition et un nombre de microphones élevés. L’échantillonnage compressif, appliqué à ce cas par le choix d’un modèle parcimonieux adapté et par la conception d’une nouvelle antenne, permet de réduire ces contraintes en préservant la qualité des images reconstruites.

3.1 Méthodes standard d’Holographie Acoustique (NAH)

Après discrétisation, et à une fréquence modale donnée ω , le champ de pression acoustique $p(\vec{r}, t)$ mesuré à une distance z_0 d’une structure vibrante (en l’occurrence, une plaque incluse dans le plan (Oxy)) s’exprime en fonction du champ de vitesse normale à la plaque $u(\vec{r}, t)$ sous la forme :

$$p = F^{-1}GFu = Au \quad (2)$$

où :

- u est le vecteur des vitesses normales, discrétisé sur une grille rectangulaire, dans le plan (Oxy) ,
- p est le vecteur des pressions acoustiques mesurées, discrétisées dans le plan de l’hologramme,
- $A = F^{-1}GF$ est la “matrice de mesure” obtenue par la composition de la transformée de Fourier spatiale discrète (DFT) F , l’opérateur de propagation G , et la transformée inverse F^{-1} .

Même lorsque les matrices sont carrées, l'inversion naïve du problème ($u = A^{-1}p$) est instable, pour des raisons de mauvais conditionnement dues à la présence d'ondes évanescentes dans le propagateur G . La NAH classique contourne ce problème par l'ajout d'une régularisation, le plus souvent "de Tikhonov" qui consiste à résoudre le problème de minimisation suivant [14] :

$$\hat{u} = \min_u \|p - Au\|_2^2 + \lambda \|Lu\|_2^2 \quad (3)$$

où L est un opérateur de régularisation et λ le paramètre correspondant. Cette méthode présente notamment deux inconvénients : d'une part, la régularisation de Tikhonov tend à se traduire par un filtrage passe-bas dans le domaine de Fourier spatial, qui crée des artefacts dans l'image reconstruite ; d'autre part, elle suppose un échantillonnage du plan hologramme à la fréquence de Nyquist, ce qui implique jusqu'à plusieurs centaines voire milliers de points de mesure.

3.2 Modélisation parcimonieuse

L'approche standard repose seulement sur une hypothèse de largeur de bande finie pour les fréquences spatiales du champ acoustique étudié. Afin de réduire le nombre de mesures nécessaire, un modèle plus précis est nécessaire, et nous allons voir ici que la propriété de parcimonie du champ dans un dictionnaire bien choisi Ψ permet de décrire et d'exploiter un tel modèle.

La théorie nous enseigne que des combinaisons linéaires d'ondes planes constituent de bonnes approximations des solutions de l'équation de Helmholtz ou de Kirchhoff-Love sur des plaques de forme étoilée [17, 6]. Ainsi, on peut exprimer les vitesses normales u sur la plaque comme la somme d'un petit nombre d'ondes planes :

$$u(\vec{r}) \approx \left(\sum_j c_j e^{i\vec{k}_j \cdot \vec{r}} \right) \mathbf{1}_S(\vec{r}) \quad (4)$$

$\mathbf{1}_S(\vec{r})$ est la fonction indicatrice qui restreint les ondes planes au domaine S délimité par la plaque, les vecteurs \vec{k}_j sont les vecteurs d'onde, $\vec{r} = (x, y)$ le vecteur position et c_j les coefficients de la combinaison linéaire. Le dictionnaire Ψ est donc construit en générant des ondes planes de vecteurs d'ondes \vec{k}_j régulièrement espacés dans le plan de Fourier, puis en les restreignant au domaine S . Ce processus est illustré sur la figure 2.

3.3 Problème inverse régularisé par la parcimonie

Dans ce cadre, l'hypothèse de parcimonie des coefficients c est utilisée pour régulariser le problème de la NAH, que l'on reformule comme suit :

$$\operatorname{argmin}_c \|c\|_0 \quad \text{subject to} \quad p = A\Psi c. \quad (5)$$

Autrement dit, il s'agit de trouver les coefficients c qui soient à la fois les plus parcimonieux possibles et compatibles avec les observations. Ce problème d'optimisation peut être approximativement résolu par divers algorithmes tels que les algorithmes gloutons (Matching Pursuit et ses variantes) ou la relaxation ℓ_1 (Basis Pursuit Denoising [10]).

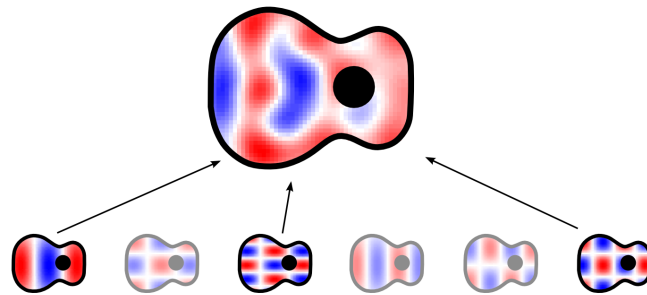


FIGURE 2 – A une fréquence donnée, la vibration complexe d'une structure (ici, une table d'harmonie de guitare) peut être approximativement décomposée en une somme parcimonieuse d'ondes planes. Avec l'aimable autorisation de F. Ollivier, UPMC.

Ce deuxième cas revient essentiellement à remplacer le terme de régularisation en norme ℓ_2 de l'approche classique (par Tikhonov, équation (3)) par un terme favorisant la parcimonie :

$$\operatorname{argmin}_c \|p - A\Psi c\|_2^2 + \lambda \|c\|_1 \quad (6)$$

3.4 Antenne optimisée pour l'acquisition compressive

La modélisation parcimonieuse du problème permet en outre de se débarrasser de la contrainte d'échantillonnage complet du plan hologramme, à condition que les mesures de la pression aient une diversité suffisante pour capturer les différents degrés de liberté du problème. Combien de mesures sont nécessaires ? Peut-on réduire encore ce nombre par une conception adaptée de l'antenne d'acquisition ? Dans le cas d'un signal parcimonieux dans le plan de Fourier, on sait qu'il suffit d'un petit nombre de mesures ponctuelles pour reconstruire parfaitement le signal, même en présence de bruit [19]. La théorie suggère l'utilisation d'une antenne acoustique dont les microphones sont positionnés aléatoirement, plutôt que régulièrement, augmentant l'incohérence entre la mesure et le domaine dans lequel le signal est parcimonieux.

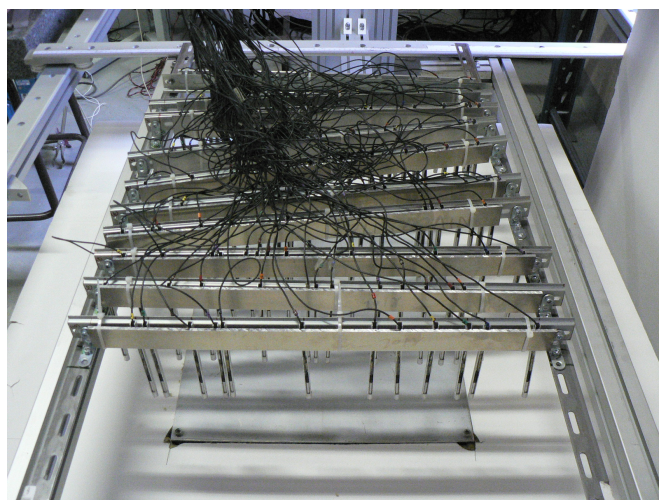


FIGURE 3 – Antenne aléatoire pour l'holographie compressée. Avec l'aimable autorisation de F. Ollivier, UPMC.

En pratique, il est difficile de construire une antenne purement aléatoire tout en préservant une connaissance précise de la position de chaque microphone. Une solution expérimentale proposée, après validation en simulation numérique, est l'emploi de plusieurs barres métalliques fixées sur un châssis avec lequel elles forment un angle déviant de 90° d'une petite valeur aléatoire, et de placer aléatoirement les micros sur chacune de ces barres (Fig. 3).

3.5 Résultats et sensibilité à la connaissance du système d'acquisition

La figure 4 présente quelques résultats de reconstruction des cartes de vitesse de deux plaques de forme différente, obtenues d'une part par holographie classique et d'autre part par holographie compressée. On montre qu'avec l'antenne aléatoire précédemment décrite (120 microphones) et en une seule acquisition, on obtient des performances au moins similaires à celles obtenues par holographie classique et régularisation de Tikhonov à partir de 1920 mesures (antenne régulière de 120 microphones déplacée 16 fois). A certaines fréquences et pour certaines formes de plaques, les artefacts basse fréquence sont supprimés et le nombre de mesures peut même être encore abaissé (40 microphones pour une plaque rectangulaire).

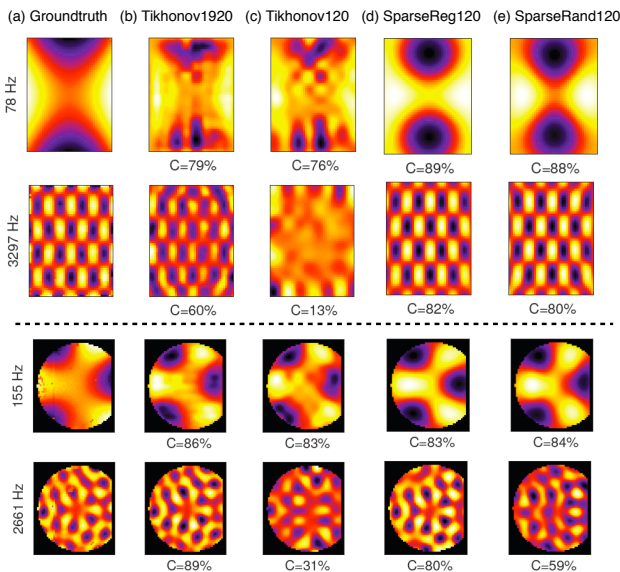


FIGURE 4 – Résultats d'imagerie par NAH. Les colonnes (b) et (c) sont obtenues par NAH classique avec 1920 et 120 microphones respectivement. Les colonnes (d) et (e) exploitent le modèle parcimonieux avec 120 mesures, respectivement régulières et aléatoires. C représente l'intercorrélacion normalisée de l'image reconstruite avec des mesures de référence (a) obtenues par vibrométrie laser. *Figure adaptée de [9].*

Il faut cependant noter que la mise en œuvre de l'holographie compressée implique une connaissance très précise de la position de chaque microphone, tâche moins aisée sur une antenne aléatoire [7]. Par ailleurs, l'holographie compressée se révèle beaucoup plus sensible aux éventuelles décalibrations des microphones que la NAH classique, même à partir de quelques dB de gain [13].

Pour résumer les principaux enseignements de cette étude sur la NAH, appliquer l'échantillonnage comprimé à un problème inverse expérimental met en jeu trois composantes essentielles :

- *la parcimonie*, qui est ici déduite de la physique du problème. Prendre en compte la parcimonie dans la formulation du problème inverse peut déjà produire des gains de performance significatifs, au prix d'une augmentation des temps de calcul ;
- *des mesures* adaptées, qui peuvent être conçues pour tirer parti de façon optimale de la parcimonie, tout en gardant à l'esprit qu'il existe souvent de fortes contraintes physiques : les mesures complètement aléatoires telles qu'utilisées dans la plupart des études théoriques ne sont presque jamais accessibles en pratique ;
- *des difficultés supplémentaires*, non seulement en termes de complexité de calcul, mais aussi une plus grande sensibilité par rapport à la connaissance de l'ensemble du système de mesure.

4 Echantillonnage de la fonction plénacoustique

L'acquisition comprimée peut de même être utilisée pour échantillonner la *fonction plénacoustique* \mathcal{P} , qui rassemble l'ensemble des réponses impulsionnelles pour toutes les positions de sources et récepteurs (x_s et x_p , respectivement) dans une salle donnée [1]. On la note $\mathcal{P}(x_s, x_p, t)$, et elle dépend à la fois de la géométrie de la pièce et des propriétés acoustiques de matériaux pour les conditions aux limites. Dans un cadre linéaire, cette fonction \mathcal{P} caractérise complètement l'acoustique de la salle : "Combien de microphones devons-nous placer dans la salle afin de reconstruire complètement le champ sonore à n'importe quelle position?" [1]. Échantillonner régulièrement \mathcal{P} , au sens de Shannon, dans l'ensemble de la gamme audible (avec des fréquences allant jusqu'à 20 kHz) semble sans espoir, car il faudrait être en mesure de déplacer la source et le capteur (microphone) sur une grille 3D avec un taille de pas de moins de 1 cm (moitié de la plus petite longueur d'onde), conduisant à plus de 1 million de positions de capteurs par m^3 . Cependant, la nature des ondes acoustiques introduit des contraintes fortes sur \mathcal{P} , que l'on retrouve par exemple dans le théorème intégral de Kirchoff. Pour profiter de ces contraintes dans un cadre d'acquisition comprimée, il faut trouver des modèles parcimonieux pour le champ acoustique p . Supposons que la source soit à une position fixée dans la salle, l'objectif est alors d'estimer la réponse impulsionnelle $p(x, t)$ dans un domaine spatial \mathcal{D} , où $x \in \mathcal{D}$ est la position du récepteur. La parcimonie est ici issue de deux points de vue différents, complémentaires : **Un point de vue temporel** : pour tout $x = x_0$ fixé, $p(x_0, t)$ est parcimonieux dans le début des réponses impulsionnelles, soit à $t < t_{mix}$, avant le temps de mélange t_{mix} (uniquement son direct et réflexions précoces) :

$$p(x, t) = \sum_{k=0}^K \beta_k \frac{\delta(t - \|s_k - x\|)}{4\pi \|s_k - x\|} \quad (7)$$

pour $x \in \mathcal{D}$, $t < t_{mix}$, où K est le nombre de sources réelles et virtuelles dans la boule de rayon ct_{mix} autour de la source

réelle, s_k est la position de la source virtuelle k et β_k est l'intensité correspondante - en tenant compte de certaines atténuations possibles aux réflexions. Ainsi, pour $t < t_{mix}$, la fonction plénacoustique (pour une source fixée en x_0) est entièrement déterminée par une combinaison linéaire d'un faible nombre d'impulsions correspondant à un ensemble de sources virtuelles dans une sphère de taille finie.

A partir de mesures aux points d'échantillonnage, le modèle est estimé en recherchant la position d'un faible nombre de sources (la source réelle et les sources virtuelles dans la sphère d'intérêt), dont la combinaison explique de manière optimale les données observées. Compte tenu de la taille du domaine d'intérêt, cette recherche de sources virtuelles se fait de manière gloutonne, les sources étant identifiées l'une après l'autre.

Une hypothèse similaire de parcimonie dans le domaine temporel des des réponses impulsionnelles de salle peut également être exploitée pour la mesure simultanée des réponses impulsionnelles à partir de différentes positions de source[2]. On peut montrer que ce problème est équivalent à l'estimation des filtres de mélange dans le contexte de la séparation de sources d'un mélange convolutif.

Un point de vue fréquentiel : aux basses fréquences (sous la fréquence de Schroeder f_{Sch} , le spectre des réponses impulsionnelles montre des pics isolés qui correspondent au comportement modal de la salle. On postule donc un modèle parcimonieux en fréquence en dessous de f_{Sch} . Les modes eux-mêmes ont une répartition spatiale très spécifique : pour une fréquence modale f_0 donnée, suffisamment loin des murs ce qui permet de négliger les ondes évanescentes, le mode ne contient que la longueur d'onde $\lambda = c/f_0$. En d'autres termes, les modes sont entièrement décrits par les ondes planes $e^{i(\vec{k}\vec{x}-2\pi f_0 t)}$, avec le vecteur d'onde \vec{k} de module fixe $\|\vec{k}\| = 2\pi f_0/c$. Le modèle discrétisé spatio-temporel complet s'écrit

$$p(x, t) = \sum_{r=1}^R \sum_{p=1}^P \alpha_{r,p} e^{i(\vec{k}_{r,p}\vec{x}-2\pi f_r t)} \quad (8)$$

avec $\|\vec{k}_{r,p}\| = 2\pi f_r/c$, R est le nombre de modes (parcimonie en fréquence), P est le nombre d'ondes planes utilisées pour discrétiser les directions la sphère 3D, et les $\alpha_{r,p}$ sont les coefficients correspondants. Compte tenu des mesures aux points d'échantillonnage, ce modèle est estimé en deux étapes : d'abord, l'ensemble parcimonieux de fréquences modales est estimé, conjointement à travers les signaux des microphones. Puis, à une fréquence modale f_r , les coefficients $\alpha_{r,p}$ sont estimés par projections aux moindres carrés sur l'ensemble discret des ondes planes $e^{i(\vec{k}\vec{x}-2\pi f_0 t)}$, discrétisées aux positions des capteurs.

Les deux modèles décrits ci-dessus ont été testés dans des conditions expérimentales réelles [15, 16], avec 120 microphones répartis dans un volume de $2 \times 2 \times 2$ m (avec une répartition des microphones à peu près uniforme), dans une grande salle possédant une forte réverbération. La fig. 6 compare la qualité de l'interpolation pour le modèle temporel. Il est important de rappeler que les méthodes présentées ici effectuent l'acquisition comprimée uniquement sur la partie de la fonction plénacoustique où certains modèles parcimonieux ont pu être identifiés : au début de la réponse, et dans les basses fréquences. Notons aussi que les deux points de vue temporel et fréquentiel sont

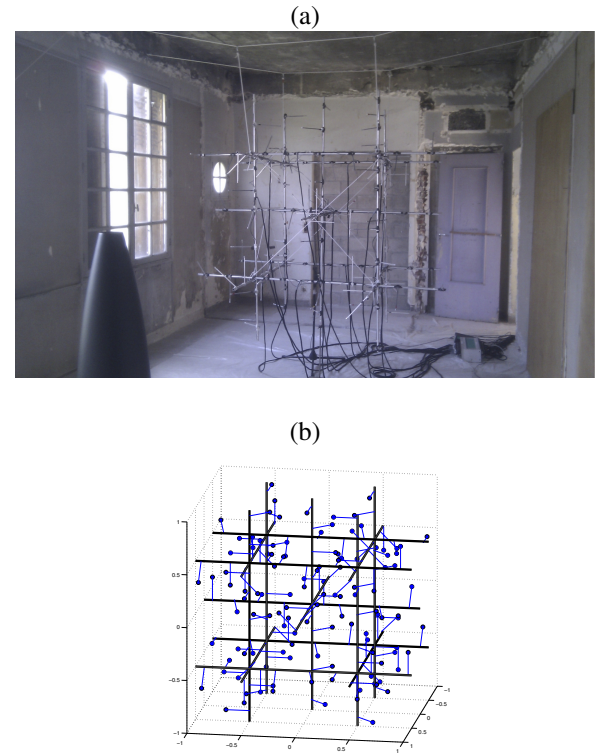


FIGURE 5 – Antenne de microphones utilisée pour l'acquisition comprimée de la fonction plénacoustique. (a) Photographie du réseau de 120 microphones dans la pièce. La source omnidirectionnelle se trouve au premier plan, à gauche. (b) Géométrie de l'antenne, le points bleus indiquant la position des capsules des microphones. D'après[16].

ici traités indépendamment, un traitement conjoint serait certainement bénéfique, et fera l'objet de travaux futurs.

5 Localisation de sources

5.1 Localisation de sources à bande étroite, en milieu réverbérant

Le modèle fréquentiel développé ci-dessus pour modéliser la fonction plénacoustique par combinaison linéaire d'ondes planes peut également être utilisé pour séparer le champ réverbérant du champ direct lorsque des sources sont présentes dans le domaine étudié. Ainsi, on obtient une estimée du champ déréverbéré ce qui permet de localiser les sources. Si le nombre de sources est faible, il est naturel d'utiliser également à cet effet un modèle basé sur la parcimonie. Il est ainsi possible de localiser des sources dans une configuration a priori extrêmement défavorable : bande étroite (une seule fréquence), et milieu fortement réverbérant inconnu (ni la géométrie du domaine, ni les conditions aux limites ne sont connus).

Ainsi, en 2 dimensions, le champ acoustique produit par S sources émettant à la fréquence ω en milieu réverbérant peut se modéliser comme [8] :

$$p(\vec{x}) \approx \sum_{l=1}^L \alpha_l e^{i\vec{k}_l \vec{x}} + \sum_{j=1}^S s_j(\omega) y_0(k|\vec{x} - \vec{x}_j|) \quad (9)$$

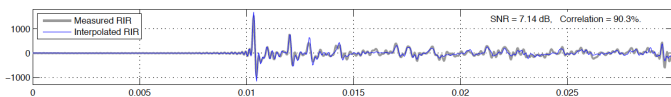


FIGURE 6 – Début de la réponse impulsionnelle enregistrée sur un microphone (trait gris), et réponse impulsionnelle interpolée (trait bleu) au même point à partir des mesures aux autres microphones, en utilisant le modèle temporel. D'après [15].

où le premier terme représente le champ réverbérant avec les L ondes planes $e^{i\vec{k}_l \cdot \vec{x}}$ de vecteur d'onde \vec{k}_l tel que $\|\vec{k}_l\| = k$, et α_l les coefficients correspondants ; et où le second terme représente les termes sources, avec les y_0 des fonctions de Bessel de 2ème espèce, les x_j les positions des sources, et les $s_j(\omega)$ les amplitudes correspondantes.

La généralisation de ces travaux à l'acoustique audible 3D fait l'objet de travaux actuels, en milieu homogène et hétérogènes, présentés dans une communication spécifique de ce congrès [18]. Notons également que le problème de la déréverbération présente une formulation spécifique pour des réseaux sphériques de microphones [21].

5.2 Localisation et caractérisation jointes de sources

Les modèles parcimonieux décrits ci-dessus peuvent être adaptés pour effectuer de façon simultanée la localisation et la caractérisation de la directivité des sources (on suppose maintenant que le champ réverbérant est négligeable). La mesure des diagrammes de directivité en 3D nécessite généralement un protocole de mesure très long, avec un grand nombre de microphones positionnés à une distance fixe de la source (généralement 1 m ou 2 m). Le rayonnement est décrit par son expansion en harmoniques sphériques d'ordre croissant. En se limitant à un ordre fini L ($L = 0$ pour monopolaire seulement, $L = 1$ pour monopolaire et dipolaire, etc), on peut écrire le champ sonore en coordonnées polaires (r, θ, φ) au nombre d'onde k pour une source située à l'origine :

$$p(kr, \theta, \varphi) = \sum_{l=0}^L \sum_{q=-l}^l \gamma_l^q(k) h_l(kr) Y_l^q(\theta, \varphi) \quad (10)$$

où les Y_l^q sont les harmoniques sphériques de degré l et d'ordre q , h_l sont les fonctions propagatives de Hankel d'ordre l , et γ_l^q les coefficients correspondants. Cela peut être utilisé pour construire un modèle *parcimonieux structuré* pour le champ produit par un faible nombre de sources avec un rayonnement non uniforme : la parcimonie dans l'espace restreint le nombre d'emplacements actifs, et pour un emplacement actif donné tous les coefficients de la décomposition en harmoniques sphériques sont non nuls. Ce type de modèle peut être identifié à partir de mesures en utilisant une pénalité mixte ℓ_1/ℓ_2 sur l'ensemble des coefficients d'activité γ_l^q (réorganisés en colonne pour chaque emplacement), ou par des algorithmes de type groupe-OMP.

Les résultats expérimentaux soulèvent ici la question du choix de la grille d'échantillonnage spatial. Si les sources sont exactement sur des points d'échantillonnage, ou très

proches, le modèle identifie correctement le diagramme de rayonnement au moins jusqu'à l'ordre $L = 2$ (rayonnement quadripolaire). Toutefois, une source située entre des points d'échantillonnage apparaît comme une combinaison linéaire de deux sources ou plus, avec des coefficients complexes de rayonnement : par exemple, une source dipolaire peut apparaître comme une combinaison de deux monopôles voisins en opposition de phase, mais des combinaisons plus complexes apparaissent également, sans interprétation physique simple. Les travaux futurs devront considérer une optimisation parcimonieuse sur un espace continu de paramètres spatiaux [12].

Si on ne se préoccupe pas de la localisation des sources mais uniquement de la représentation du champ sonore à directivité complexe, il est également possible d'estimer des ordres élevés d'un modèle spatial de type *ambisonics* à partir d'un faible nombre de mesures [20].

6 Discussion et conclusion

À travers un panorama non-exhaustif de scénarios en acoustique audible, cet article a montré comment l'acquisition compressée pouvait être un outil utile à l'imagerie acoustique en général. Nous avons en particulier vu comment l'ingrédient principal pour utiliser cet outil, l'hypothèse de parcimonie, était naturellement présente dans la physique des problèmes rencontrés, et ce selon plusieurs dimensions : distribution parcimonieuse des objets dans l'espace, des réflexions précoces, des modes, etc. Par ailleurs, l'acquisition des données par les capteurs traditionnels présente naturellement des caractéristiques d'incohérence, autre ingrédient permettant l'application de techniques d'acquisition compressée. Les conditions sont ainsi réunies pour bénéficier des promesses de cette nouvelle technologie : réduction du coût des équipements par diminution du nombre de capteur ; réduction de la quantité de données acquises, transmises et/ou stockées ; super-résolution, c'est-à-dire amélioration de la qualité de l'imagerie au-delà des limites de Shannon-Nyquist.

Ces constats permettent de conclure d'une façon encourageante, et de dresser une liste de défis à relever à la fois par les mathématiciens et les acousticiens. Premièrement, la théorie de l'acquisition compressée suppose une connaissance parfaite du dispositif d'acquisition alors qu'en pratique, elle peut être imparfaite ou coûteuse à déterminer (calibration, positionnement des capteurs, etc). L'enjeu est ici d'autant plus important que l'on souhaite manipuler de grands réseaux de capteurs et des capteurs bon marché. Certains travaux préliminaires [4, 3] ont montré que dans le cas de gains et de déphasages inconnus, l'effet de régularisation de la parcimonie pouvait aider à estimer automatiquement ces paramètres du dispositif d'acquisition. Deuxièmement, les interactions entre mathématiques et acoustique peuvent permettre de développer de nouveaux modèles parcimonieux, par exemple pour intégrer la directivité à l'émission, à la réflexion ou à la réception dans le problème, ou pour modéliser le speckle. De même, l'imagerie 3D de grandes régions représente un véritable défi algorithmique. Finalement, de nombreuses questions restent ouvertes quant à la conception de nouveaux dispositifs plus adaptés à l'acquisition compressée, en s'appuyant davantage sur l'aléatoire dans l'échantillonnage pour augmenter le

caractère incohérent des mesures.

Remerciements

Les auteurs tiennent à remercier chaleureusement François Ollivier pour les figures, ainsi que Gilles Chardon, Rémi Mignot, Antoine Peillot, et les collègues des projets ECHANGE et PLEASE dont les contributions ont été essentielles pour le travail décrit dans cet article. Ce travail a été partiellement financé par l'ANR, projet ECHANGE (ANR-08-EMER-006) ainsi que par l'ERC, projet PLEASE (ERC-StG-2011-277906).

Références

- [1] Thibaut Ajdler, Luciano Sbaiz, and Martin Vetterli. The plenacoustic function and its sampling. *Signal Processing, IEEE Transactions on*, 54(10) :3790–3804, 2006.
- [2] Alexis Benichoux, Emmanuel Vincent, and Rémi Gribonval. A compressed sensing approach to the simultaneous recording of multiple room impulse responses. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics 2011*, page 1, New Paltz, NY, États-Unis, October 2011.
- [3] Cagdas Bilen, Gilles Puy, Rémi Gribonval, and Laurent Daudet. Blind Phase Calibration in Sparse Recovery. In *EUSIPCO - 21st European Signal Processing Conference - 2013*, Marrakech, Maroc, September 2013.
- [4] Cagdas Bilen, Gilles Puy, Rémi Gribonval, and Laurent Daudet. Blind Sensor Calibration in Sparse Recovery Using Convex Optimization. In *SAMPTA - 10th International Conference on Sampling Theory and Applications - 2013*, Bremen, Allemagne, 2013.
- [5] E. J. Candès, J. Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Comm. Pure Appl. Math*, 59 :1207–1223, 2006.
- [6] G. Chardon, A. Leblanc, and L. Daudet. Plate impulse response spatial interpolation with subnyquist sampling. *Journal of Sound and Vibration*, 330(23) :5678 – 5689, 2011.
- [7] Gilles Chardon, Nancy Bertin, and Laurent Daudet. Multiplexage spatial aléatoire pour l'échantillonnage compressif - application à l'holographie acoustique. In *XXIIIe Colloque GRETSI*, Bordeaux, France, September 2011.
- [8] Gilles Chardon and Laurent Daudet. Narrowband source localization in an unknown reverberant environment using wavefield sparse decomposition. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 9 –12, Kyoto, Japon, 2012.
- [9] Gilles Chardon, Laurent Daudet, Antoine Peillot, François Ollivier, Nancy Bertin, and Rémi Gribonval. Nearfield Acoustic Holography using sparsity and compressive sampling principles. *Journal of the Acoustical Society of America*, 2012. Code & data for reproducing the main figures of this paper are available at <http://echange.inria.fr/nah>.
- [10] Scott S. Chen, David L. Donoho, and Michael A. Saunders. Atomic Decomposition by Basis Pursuit. *SIAM J. Scientific Comp.*, 20(1) :33–61, 1999.
- [11] David L. Donoho. Compressed sensing. *IEEE Trans. Inform. Theory*, 52(4) :1289–1306, 2006.
- [12] Chaitanya Ekanadham, Daniel Tranchina, and Eero P Simoncelli. Recovery of sparse translation-invariant signals with continuous basis pursuit. *Signal Processing, IEEE Transactions on*, 59(10) :4735–4744, 2011.
- [13] Rémi Gribonval, Gilles Chardon, and Laurent Daudet. Blind Calibration For Compressed Sensing By Convex Optimization. In *IEEE International Conference on Acoustics, Speech, and Signal Processing*, Kyoto, Japon, 2012.
- [14] J.D. Maynard, E.G. Williams, and Y. Lee. Nearfield acoustic holography : I, theory of generalized holography and the development of NAH. *J. Acoust. Soc. Am.*, 78(4) :1395–1413, 1985.
- [15] R. Mignot, G. Chardon, and L. Daudet. Low frequency interpolation of room impulse responses using compressed sensing. *IEEE/ACM Transactions on Audio, Speech, and Language Processing.*, 22(1) :205–216, Jan 2014.
- [16] R. Mignot, L. Daudet, and F. Ollivier. Room reverberation reconstruction : Interpolation of the early part using compressed sensing. *IEEE Transactions on Audio, Speech, and Language Processing.*, 21(11) :2301–2312, Nov. 2013.
- [17] A. Moiola, R. Hiptmair, and I. Perugia. Plane wave approximation of homogeneous helmholtz solutions. *Zeitschrift fur Angewandte Mathematik und Physik (ZAMP)*, 62 :809–837, 2011. 10.1007/s00033-011-0147-y.
- [18] T. Nowakowski, L. Daudet, and J. De Rosny. Méthode parcimonieuse pour la localisation de sources en milieux réverbérants hétérogènes. In *Actes du Congrès Français d'Acoustique, Poitiers*, 2014.
- [19] Holger Rauhut. Random sampling of sparse trigonometric polynomials. *Appl. Comput. Harmon. Anal.*, 22(1) :16–42, 2007.
- [20] Andrew Wabnitz, Nicolas Epain, Alistair McEwan, and Craig Jin. Upscaling ambisonic sound scenes using compressed sensing techniques. In *Applications of Signal Processing to Audio and Acoustics (WASPAA), 2011 IEEE Workshop on*, pages 1–4. IEEE, 2011.
- [21] Ping Kun Tony Wu, Nicolas Epain, and Craig Jin. A dereverberation algorithm for spherical microphone arrays using compressed sensing techniques. In *Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on*, pages 4053–4056. IEEE, 2012.