

Localisation de sources en milieu réverbérant par décomposition parcimonieuse du champ acoustique sur un modèle hybride pression-vitesse

Thibault NOWAKOWSKI¹, Laurent DAUDET¹, Julien DE ROSNY¹, Gilles CHARDON²

¹Institut Langevin
1 rue Jussieu, 75005 Paris, France

²Acoustics Research Institute
Wohllebengasse 12-14, 1040 Wien, Austria

thibault.nowakowski@espci.fr, laurent.daudet@espci.fr
julien.derosny@espci.fr, gchardon@kfs.oeaw.ac.at

Résumé – Une méthode de localisation de sources en milieu réverbérant, fondée sur la parcimonie du signal source est ici adaptée au cas expérimental où de l’information géométrique est réintroduite sous forme d’hypothèses sur les conditions limites. Ces hypothèses permettent de compléter le champ acoustique échantillonné par les microphones en y ajoutant des contraintes de vitesse normale nulle sur les bords du domaine de mesure. Finalement, l’influence de la réverbération est amoindrie et la source est ainsi mieux localisée. Du fait de ces nouvelles données homogènes à des vitesses, un nouveau modèle est proposé pour résoudre le problème dans le cas de données mixtes : des mesures en pression et des contraintes additionnelles en vitesses.

Abstract – In this paper we adapt a source localization method for reverberant spaces which is based on the sparsity of the source signal. For strong reverberant rooms, some geometric information can be reintroduced as assumptions on the boundary conditions. These assumptions allow us to complete the information about the acoustic field sampled by the microphones by adding null normal velocities constraints on the field boundaries. Finally, the influence of the reverberation is reduced and the source is better located. Because of these new data homogeneous to velocities, a new model is proposed to solve the problem in the case of mixed data : measured pressures and added velocity constraints.

1 Introduction

Localiser des sources acoustiques est un enjeu important dans nombre de domaines (télécommunications, débruitage en industrie, surveillance, etc.). Cependant, si de multiples méthodes ont été développées pour résoudre ce problème dans des milieux ouverts où les fonctions de transfert entre émetteurs et récepteurs sont connues, le cas de milieux réverbérants reste encore mal résolu, d’autant plus si aucune information sur la géométrie ni sur les conditions limites du domaine n’est accessible.

Une méthode originale de séparation du champ acoustique, échantillonné sur des positions aléatoires du domaine, en deux contributions directe et diffuse et mettant à profit des a priori de parcimonie spatiale du signal direct, a été proposée dans [1]. Si l’application des méthodes de *Compressed Sensing* (CS) [4], [5] permet de reconstruire le signal par résolution du problème inverse, la non-connaissance du milieu nécessite un nombre important de mesures.

Reprenant les principes de cette méthode, nous proposons de limiter le nombre de mesures en introduisant de l’information

partielle sur la géométrie du milieu. A cette fin, le modèle est adapté au cas de mesures mixtes en pression et en vitesse.

2 Modèle de séparation de champs pour la localisation de sources

Dans la suite, nous traitons le cas à deux dimensions et mono-fréquentiel. Les quantités sont exprimées dans l’espace de Fourier.

Soit \mathcal{D} un domaine quelconque et $\Omega \in \mathcal{D}$ une de ses sous-parties pouvant être considérée comme un espace libre et contenant les sources $s(\mathbf{x}, \omega)$. Le champ acoustique est échantillonné aux points $\mathbf{p} \in \Omega$. Le modèle établi dans [1] propose de découper \mathbf{p} en deux composantes : une solution particulière \mathbf{p}_s obéissant à l’équation de Helmholtz inhomogène (1) dans l’espace libre Ω

$$\Delta p_s + k^2 p_s = s \quad (1)$$

et une solution \mathbf{p}_0 de l’équation de Helmholtz homogène sur Ω (2) modélisant tous les phénomènes acoustiques ayant lieu sur $\mathcal{D} \setminus \{\Omega\}$ et mesurés dans Ω .

$$\Delta p_0 + k^2 p_0 = 0 \quad (2)$$

Dans le modèle [1], la solution particulière \mathbf{p}_s est une somme de termes sources décrits par des fonctions de Bessel de 2^{nde}

This work is supported by DGA (Direction Générale de l’Armement) and by LABEX WIFI (Laboratory of Excellence within the French Program "Investments for the Future") under references ANR-10-LABX-24 and ANR-10-IDEX-0001-02 PSL.

espèce $Y_0(kr_{in})$ où k est le nombre d'onde et r_{in} la distance entre la $i^{ème}$ source d'amplitude s_i et de coordonnées \vec{x}_i avec les points de mesure \vec{x}_n . A trois dimensions, la seule différence est l'emploi des fonctions de Bessel sphériques de 2^{nde} espèce. La solution homogène \mathbf{p}_0 peut, quant à elle, s'écrire comme une somme de L ondes planes provenant de diverses directions définies par leur vecteur d'onde \vec{k}_l , tel que $\|\vec{k}_l\| = k$. Dans la pratique, ces vecteur d'ondes planes sont pris régulièrement sur le cercle (en 2D) ou la sphère (en 3D) de rayon k . Ainsi :

$$\mathbf{p} = \mathbf{p}_s + \mathbf{p}_0 \approx \sum_{i=1}^S s_i Y_0(kr_{in}) + \sum_{l=1}^L \varepsilon_l e^{j\vec{k}_l \vec{x}_n} \quad (3)$$

Dans le cas où peu de sources sont présentes sur Ω , \mathbf{p}_s est parcimonieux. Cet a priori est essentiel dans la suite du problème, car il permet de le résoudre par régularisation parcimonieuse. En définissant les dictionnaires de sources \mathbf{S} (tel que $S_{ni} = Y_0(kr_{ij})$ avec $\|S_{\cdot,i}\| = 1$ pour ne pas favoriser les zones plus densément échantillonnées) et d'ondes planes \mathbf{W} (tel que $W_{nl} = e^{j\vec{k}_l \vec{x}_n}$), l'équation (3) se réécrit sous la forme d'une équation matricielle :

$$\mathbf{p} = \mathbf{S}\beta + \mathbf{W}\alpha \quad (4)$$

où β est parcimonieux, les indices de ses coefficients non nuls indiquant la position des sources. Aucun a priori sur la structure de α n'est établi.

Grâce à la parcimonie, le système peut être résolu à l'aide de plusieurs types d'algorithmes, en ajoutant une étape de suppression du champ réverbéré. Ainsi, la fonction de coût dans le cas d'une résolution par minimisation de la norme ℓ_1 sur β et ℓ_2 sur α s'écrit :

$$(\hat{\alpha}, \hat{\beta}) = \min_{\alpha, \beta} \|\alpha\|_2 + \|\beta\|_1 \quad t.q. \quad \|\mathbf{S}\hat{\beta} + \mathbf{W}\hat{\alpha} - \mathbf{p}\| < \sigma \quad (5)$$

où σ est un paramètre de bruit lié aux mesures.

La résolution par algorithmes itératifs (*Orthogonal Matching Pursuit*) est précédée d'une étape durant laquelle la contribution de \mathbf{p}_0 est approchée par la méthode des moindres carrés (projections préalables de mesures dans un espace orthogonal à \mathbf{W}) puis retranchée aux mesures \mathbf{p} :

$$\mathbf{p}_s = (\mathbb{I} - \mathbf{W}\mathbf{W}^\dagger)\mathbf{p} \quad (6)$$

où \mathbf{W}^\dagger est la pseudo-inverse de \mathbf{W} . Le vecteur source \mathbf{p}_s est ensuite corrélé avec chaque colonne de \mathbf{S} , le maximum de corrélation indiquant la position de la source. Dans le cas multi-sources, la colonne de \mathbf{S} correspondant à la source isolée est ôtée au dictionnaire et ajoutée à \mathbf{W} puis l'algorithme est réitéré à partir de (6).

3 Mesures « virtuelles » aux parois

Dans les applications de la méthode [1], aucune information n'est connue sur la géométrie du milieu de propagation. Ceci se traduit en la nécessité d'un nombre important de mesures afin de séparer au mieux la solution homogène de la solution particulière.

D'autres techniques de localisation se placent au contraire dans le cadre d'une connaissance parfaite du milieu, modélisé sur ordinateur par exemple par éléments finis, pour déterminer toutes les fonctions de transfert entre récepteurs et points du maillage [3]. Le nombre de mesures peut ainsi être considérablement réduit par modélisation de la physique du domaine, au détriment d'un coût calculatoire important et d'une connaissance préalable parfaite du milieu et de sa géométrie.

Dans le cas de salles réverbérantes, il est possible de se placer entre les méthodes décrites dans [1] et [2] afin d'allier leurs avantages respectifs. En effet, en acoustique des salles, les parois peuvent être considérées comme rigides. L'espace \mathcal{D} est alors un domaine clos réverbérant, tel que :

$$\pm \frac{\partial p}{\partial e_n} \Big|_{\partial \mathcal{D}} = 0 \quad (7)$$

où e_n est la normale aux parois. La zone d'échantillonnage lui est confondue ($\Omega = \mathcal{D}$) ou bien contient une partie de ses parois.

Ainsi, il est possible d'introduire des capteurs de vitesse normale « virtuels » qui, du fait des conditions limites, mesurent une vitesse nulle. Si les autres composantes de la vitesse ne sont pas connues, l'information supplémentaire apportée suffit pour entraîner une meilleure séparation des champs \mathbf{p}_s et \mathbf{p}_0 et donc compenser le faible nombre de microphones (figure 1).

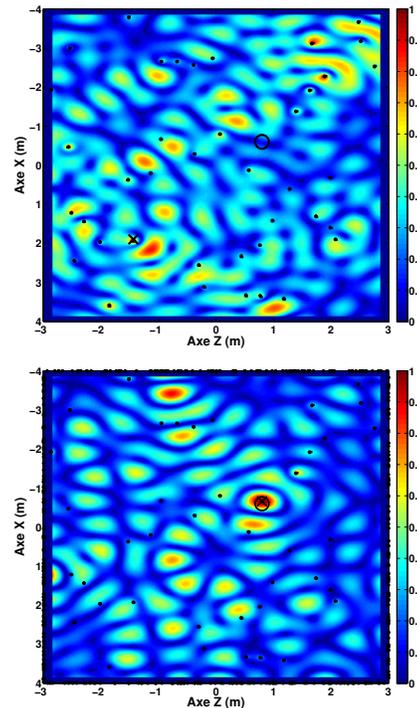


FIGURE 1 – Corrélations normalisées du champ direct estimé et du dictionnaire de sources. (Haut) Antenne de 40 microphones, sans mesures virtuelles. (Bas) Même antenne avec 100 mesures virtuelles par paroi. Cercle : position de la source ; Croix : position estimées ; Points : points d'échantillonnage.

Ici, le dictionnaire d’ondes planes \mathbf{W} est construit en sur-échantillonnant le cercle de rayon k . Le nombre de vecteurs de \mathbf{W} utilisables lors de la reconstruction étant limité par la dimension du vecteur de mesures \mathbf{p} , le rang de \mathbf{W} est réduit par décomposition en valeurs singulières (svd). Seuls les vecteurs associés aux valeurs singulières les plus fortes sont conservés. Les mesures virtuelles présentent alors l’avantage de permettre un rang de \mathbf{W} plus élevé, qui peut être nécessaire à la reconstruction.

Il est intéressant de remarquer (figure 2), que si l’information apportée permet une réduction de l’erreur moyenne sur la localisation de la source, il semble exister un premier seuil au-delà duquel l’ajout de mesures virtuelles n’est plus efficace et pour lequel l’erreur est minimale. La droite en pointillés est le seuil pour lequel deux mesures virtuelles voisines seront distantes de moins de $\lambda/2$. L’information devient alors redondante.

Lorsqu’un trop grand nombre de mesures virtuelles est employé, l’erreur semble remonter. Ce phénomène pourrait être dû à l’émergence de lobes causés par le sur-échantillonnage aux parois, mais reste à étudier davantage.

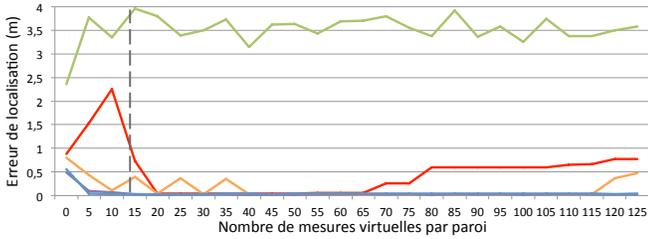


FIGURE 2 – Erreur de localisation moyennée sur 10 reconstructions en fonction du nombre de mesures virtuelles imposées sur chaque paroi. Antenne de (Vert) 20, (Rouge) 50, (Orange) 80, (Violet) 90, (Bleu) 100 microphones. Courbe en pointillés : seuil à $\lambda/2$.

Il faut toutefois prendre en compte dans ces simulations le caractère aléatoire de l’antenne qui, pour certaines reconstructions peut ne pas englober la source, causant alors l’échec de la localisation, alors qu’une répartition optimisée du même nombre de microphones pourrait permettre une localisation efficace. De récentes études se sont penchées sur ce problème dans le cas particulier où Ω est un disque [2], mais une extension directe au cas général serait de placer davantage de microphones aux bords de Ω et une petite partie à l’intérieur.

4 Modèle mixte en pression-vitesse

Pour obtenir les résultats précédents, la méthode présentée dans [1] a dû être adaptée pour permettre la reconstruction du champ à partir de mesures réelles en pression et de mesures virtuelles homogènes à des vitesses. Un modèle hybride pour la reconstruction algorithmique est proposé. Le vecteur mixte de

mesure $\mathbf{p}_{\text{mixte}}$ se décompose maintenant de la façon suivante :

$$\mathbf{p}_{\text{mixte}} = [\vec{p}_n \quad \vec{v}_{x,N_x} \quad \vec{v}_{z,N_z}]^T \quad (8)$$

avec \vec{p}_i les mesures en pression (pour $n = 1, \dots, N$ microphones), \vec{v}_{x,N_x} les N_x mesures en vitesse imposées nulles sur les parois de normales $\vec{e}_n = \pm \vec{e}_x$ et \vec{v}_{z,N_z} les N_z mesures placées sur les parois de normales $\vec{e}_n = \pm \vec{e}_z$.

Dans un premier temps, les composantes des dictionnaires de sources et d’ondes planes correspondant aux mesures virtuelles imposées en vitesse sont dérivées de leurs expressions en pression à partir de l’équation d’Euler (9),

$$v_n = \frac{1}{jZk} \cdot \frac{\partial p}{\partial \vec{e}_n} \quad (9)$$

où $Z = \rho c$ est l’impédance du milieu de propagation (avec ρ la densité et c la célérité du milieu).

Un problème subsiste pour pouvoir appliquer les algorithmes de reconstruction parcimonieuse et espérer pouvoir localiser la source. En effet, le déroulé d’OMP implique la projection de composantes en pression sur des composantes en vitesses, notamment pour le calcul du produit scalaire entre mesures et atomes du dictionnaire. Ces deux grandeurs n’étant pas directement comparables, il est nécessaire d’homogénéiser le jeu de données. La définition de la densité d’énergie acoustique (10) permet ainsi de relier les données en pression et en vitesse.

$$E = \frac{1}{2} \rho v^2 + \frac{p^2}{2\rho c^2} \quad (10)$$

L’équation (9) entraîne $v = p/Z$. Toutes les composantes en vitesses de $\mathbf{p}_{\text{mixte}}$, de \mathbf{S} et de \mathbf{W} sont multipliées par Z afin d’être rendues homogènes à des pressions. Ainsi, on obtient finalement pour \mathbf{S} et \mathbf{W} les expressions suivantes :

$$\mathbf{S} = [\mathbf{S}_{n,i}(\vec{p}_n) \quad Z \cdot \mathbf{S}_{N_x,i}(\vec{v}_{x,N_x}) \quad Z \cdot \mathbf{S}_{N_z,i}(\vec{v}_{z,N_z})]^T \quad (11)$$

$$\mathbf{W} = [\mathbf{W}_{n,l}(\vec{p}_n) \quad Z \cdot \mathbf{W}_{N_x,l}(\vec{v}_{x,N_x}) \quad Z \cdot \mathbf{W}_{N_z,l}(\vec{v}_{z,N_z})]^T \quad (12)$$

concaténations des sous-matrices exprimées en pression et vitesses normales puis homogénéisées, pour les i points sources et les l ondes planes échantillonnés sur les $n + N_x + N_z$ positions de capteurs réels et de mesures virtuelles.

Le produit scalaire renvoie une somme de termes homogènes à p^2 . La source est alors localisée au point où la corrélation est maximale (figure 3).

5 Résultats expérimentaux

Afin de valider les modèles présentés, des mesures ont été effectuées en chambre réverbérante à l’aide d’une antenne de 42 microphones omnidirectionnels. La zone d’échantillonnage Ω est un cube d’environ 1 m de côté positionné près d’une des parois de la salle, à la surface duquel ont été placés les capteurs. Quelques microphones ont également été placés à l’intérieur de Ω . La salle de mesures présente un temps de réverbération proche de 4 s.

Les positions des microphones ont été calibrées automatiquement à l’aide d’une méthode décrite dans [6]. La position

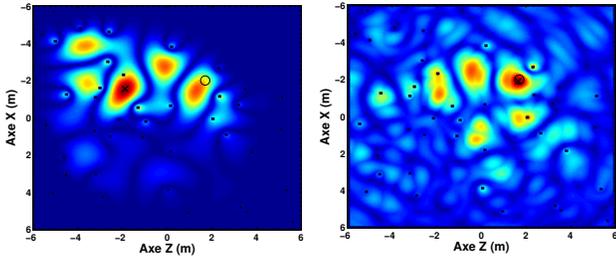


FIGURE 3 – Comparaison d’une reconstruction sans étape d’homogénéisation (haut) et avec données homogénéisées (bas). Champs normalisés. Cercle : position de la source ; Croix : position estimée ; Points : points d’échantillonnages.

de la source présente à l’intérieur de Ω a été mesurée à l’aide de cette même méthode et est ensuite vérifiée par la méthode de localisation parcimonieuse.

La reconstruction est effectuée à 617 Hz, le dictionnaire \mathbf{W} est constitué de 2000 ondes planes, mais est tronqué par svd au rang 40. Une grille de 100 mesures virtuelles est ajouté au niveau de la paroi la plus proche de l’antenne (figure 4). Les ré-

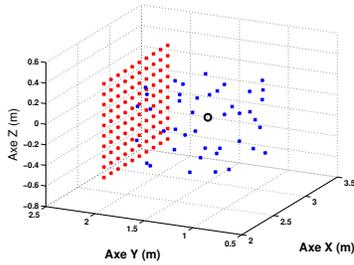


FIGURE 4 – Schéma du dispositif expérimental. Croix bleues : microphones ($\sim \vec{p}_n$). Croix rouges : mesures virtuelles ($\sim \vec{v}_{y,Ny}$). Rond : Position de la source.

sultats obtenus (figure 5) montrent une nette amélioration de la localisation de la source grâce à l’ajout des mesures virtuelles.

6 Conclusion

La méthode de localisation en milieu réverbérant a été étudiée dans le cas où la réverbération est suffisamment importante dans le domaine pour permettre de formuler une hypothèse de rigidité des parois. L’approximation faite permet d’introduire de l’information sur le champ acoustique et de mieux séparer les contributions de la solution homogène et de la solution particulière, diminuant ainsi l’erreur de localisation de la source et compensant en partie un faible nombre de capteurs. L’adaptation en un modèle mixte permet de prendre en compte ces mesures imposées en vitesse. Enfin, la mise à profit de la parcimonie du signal sur la totalité de sa bande passante est envisageable.

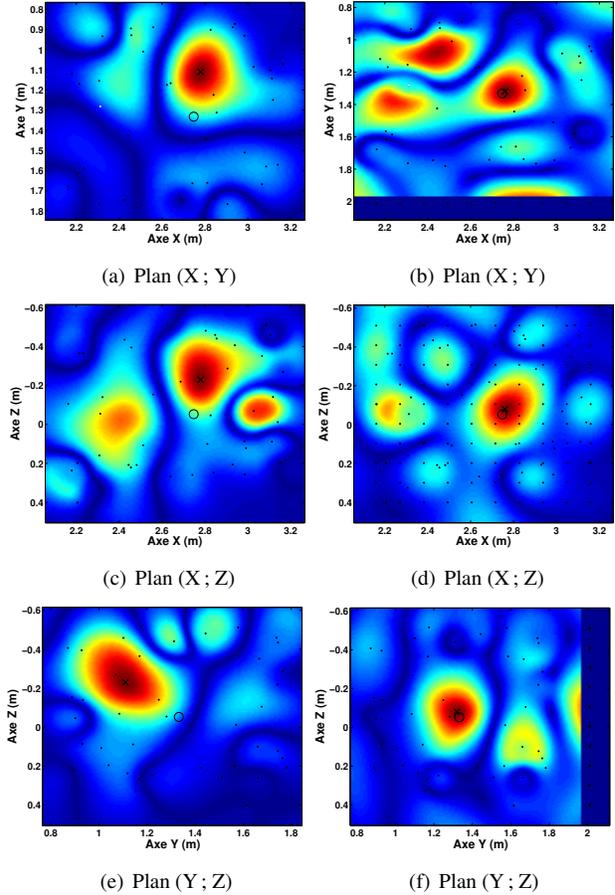


FIGURE 5 – Coupes selon les trois plans de l’espace au point de corrélation maximale entre mesures et dictionnaire \mathbf{S} après retrait de la composante diffuse. (a), (c), (e) Sans mesures virtuelles, erreur de localisation : $\varepsilon = 0,28$ m. (b), (d), (f) Avec mesures virtuelles, erreur de localisation : $\varepsilon = 0,03$ m.

Références

- [1] G. Chardon, L Daudet. *Source Localization in an unknown Reverberant Environment Using Compressive Sampling in the Frequency Domain*. Proc. ICASSP, 2012.
- [2] G. Chardon, A Cohen, L Daudet. *Reconstruction of Solutions to the Helmholtz Equation from Punctual Measurements*. Proceedings of Sampta, 2013.
- [3] M. Vetterli, I. Dokmanic. *Room Helps : Acoustic Localization with Finite Elements*. Proc. ICASSP, 2012.
- [4] S. Shaobing Chen, D.L. Donoho, M.A. Saunders. *Atomic Decomposition by Basis Pursuit*. SIAM Review, Vol. 43, No 1, pp. 129-159, 2001.
- [5] J.A. Tropp, A.C. Gilbert. *Signal Recovery from Random Measurements via Orthogonal Matching Pursuit*. IEEE Trans. On Information Theory, Vol. 53, No 12, pp 4655-4666, 2007.
- [6] M. Crocco, A. Del Blue, M. Bustreo, V. Murino. *A Closed Form Solution to the Microphone Position Self-Calibration Problem*. Proc. ICASSP, 2012.