

# Méthode sans grille pour la localisation de sources à partir de mesures asynchrones

Nathan ITARE   Gilles CHARDON   José PICHERAL

Université Paris-Saclay, CNRS, CentraleSupélec, Laboratoire des signaux et systèmes, 91190, Gif-sur-Yvette, France

**Résumé** – Les caractéristiques d’une ou plusieurs sources acoustiques comme leurs positions et leurs puissances sont des informations importantes et nécessaires pour de nombreux problèmes. Les mesures asynchrones, consistant à réaliser plusieurs acquisitions en déplaçant un réseau de microphones, ou à utiliser plusieurs réseaux non synchronisés, permettent de pallier certaines limitations par rapport à l’utilisation d’un seul réseau. Cette étude présente une méthode basée sur un problème d’ajustement de la matrice de covariance (covariance matrix fitting) formulé avec une parcimonie continue et en utilisant des mesures asynchrones. La parcimonie continue permet d’effectuer une localisation en trois dimensions sans avoir une discrétisation de l’espace qui demanderait beaucoup de ressources en terme de calculs et mémoire. L’algorithme *Sliding Frank-Wolfe* est utilisé pour résoudre le problème d’optimisation en dimension infinie correspondant. Dans cette étude, la formulation en dimension infinie appliquée au cas asynchrone est présentée puis les performances de la méthode sont évaluées et comparées avec les bornes de Cramér-Rao dans différentes configurations.

**Abstract** – The characteristics of one or many acoustic sources like their positions and their powers are important and necessary information for numerous problems. Asynchronous measurements, which consists in realizing several acquisitions at different locations with a same antenna or to use several antennas not synchronized, enable to overcome some limits in comparison to a single antenna. This study presents a method based on a covariance matrix fitting problem with a gridless formulation and using asynchronous measurements. The gridless formulation enables to perform a localization in three dimensions without fine space discretization which is demanding in terms of computational resources. The *Sliding Frank-Wolfe* is used to solve the corresponding infinite dimension optimization problem. In this study, the infinite dimension formulation applied to the asynchronous case is presented then the performances of the method are evaluated and compared with Cramer Rao bounds in different configurations.

## 1 Introduction

Les performances de localisation de sources avec un réseau de microphones peuvent être limitées par la taille du réseau ainsi que son nombre de microphones. L’utilisation de mesures asynchrones est une alternative permettant de pallier ces limitations. Les mesures asynchrones sont des mesures effectuées par plusieurs réseaux non synchronisés ou par un seul réseau déplacé à différents endroits donnant donc plusieurs acquisitions non synchronisées. Les méthodes de l’état de l’art peuvent se classer en deux catégories : la première consiste à compléter la matrice de covariance en estimant les termes d’inter-corrélation entre les réseaux utilisés pour les acquisitions asynchrones [4, 7], la deuxième consiste à effectuer un traitement sur chaque acquisition puis à combiner les estimations obtenues [6]. Ces travaux ont montré que les mesures asynchrones permettent d’élargir la plage de fréquence de fonctionnement du réseau tout en gardant un nombre raisonnable de microphones. Ainsi en déplaçant un réseau 2D pour réaliser plusieurs mesures asynchrones, il est possible d’améliorer la précision de la localisation et de réaliser une localisation 3D.

Les méthodes de localisation comme la formation de voies utilisent généralement une discrétisation de l’espace de recherche des sources, ce qui peut être rapidement coûteux en terme de calculs pour des grands espaces 3D ou pour des grilles avec un maillage fin. Une technique de localisation formulée avec parcimonie continue a été présentée dans [2] avec de la formation de voies et dans [1] avec des méthodes de covariance matrix fitting (CMF) (ajustement de la matrice

de covariance). Cette dernière étude explore une méthode de CMF appelée COMET2 (COvariance Matrix Estimation Technique) qui est une variante de la méthode présentée dans [5] et qui présente des performances meilleures par rapport à l’état de l’art. [6] utilise également un estimateur COMET appliqué à des mesures asynchrones mais en dimension finie. L’objectif de cette présente étude est donc d’étendre la variante COMET2 au cas de mesures asynchrones en utilisant une formulation en dimension infinie et d’évaluer les performances au travers de simulations. Dans un premier temps, une présentation de la méthode COMET2 formulée avec parcimonie continue est donnée. Contrairement aux méthodes de l’état de l’art, la méthode proposée traite les mesures asynchrones conjointement, sans passer par une étape intermédiaire de reconstruction de matrice de covariance complète. Dans un deuxième temps, les performances de la méthode sont évaluées en terme d’erreurs quadratiques moyennes et comparées avec les bornes de Cramér-Rao.

## 2 Formulation du modèle

On se place à une fréquence d’étude  $f$  (en pratique on utilise une transformée de Fourier à court terme des signaux acquis). On considère un réseau de  $M$  microphones. On réalise  $J$  acquisitions asynchrones, pour chacune de ces acquisitions le réseau est déplacé à une position différente. Enfin, chaque acquisition comporte  $S$  échantillons temporels. En présence de  $K$  sources, les signaux  $\mathbf{p}_{j,s}$  à l’échantillon  $s$  pour la  $j^{\text{ième}}$

acquisition asynchrone peuvent s'exprimer tel que :

$$\mathbf{p}_{j_s} = \sum_{k=1}^K a_{k_j_s} \mathbf{g}_j(\mathbf{x}_k) + \mathbf{n}_{j_s}, \quad (1)$$

avec  $a_{k_j_s}$  les amplitudes complexes de la source  $k$ ,  $\mathbf{g}_j(\mathbf{x}_k)$  le vecteur des fonctions de Green entre la source placée à la position  $\mathbf{x}_k$  et les microphones, et  $\mathbf{n}_{j_s}$  le bruit de mesure supposé blanc en espace et en temps. Par exemple, le vecteur des fonctions de Green en champ libre entre une position  $\mathbf{x}$  et les microphones placés en  $\mathbf{y}_j$  est donné par :

$$\mathbf{g}_j(\mathbf{x}) = \frac{\exp(-i\kappa\|\mathbf{x} - \mathbf{y}_j\|_2)}{\|\mathbf{x} - \mathbf{y}_j\|_2}, \quad (2)$$

avec  $\kappa$  le nombre d'onde. La matrice de covariance  $\Sigma_j$  de la  $j^{\text{ème}}$  mesure asynchrone  $\mathbf{p}_{j_s}$  peut être décrite dans le cas de sources décorréliées entre elles et avec le bruit de mesure :

$$\Sigma_j = \sum_{k=1}^K q_k \mathbf{g}_j(\mathbf{x}_k) \mathbf{g}_j(\mathbf{x}_k)^H + \sigma^2 \mathbf{I}, \quad (3)$$

avec  $q_k$  la puissance de la  $k^{\text{ème}}$  source,  $(\cdot)^H$  l'opérateur de conjugué Hermitienne, et  $\sigma^2$  la variance du bruit de mesure. Cette matrice de covariance peut être estimée à partir des mesures :

$$\hat{\Sigma}_j = \frac{1}{S} \sum_{s=1}^S \mathbf{p}_{j_s} \mathbf{p}_{j_s}^H. \quad (4)$$

Cet estimateur converge vers la vraie matrice de covariance lorsque que le nombre d'échantillons  $S$  augmente.

Dans le cadre de la parcimonie continue, la distribution de sources est décrite par une mesure  $\mu$  définie sur le domaine  $\Omega$  où les sources sont supposées être placées. Ainsi la distribution des  $K$  sources ponctuelles peut être écrite avec une mesure de Dirac  $\delta_{\mathbf{x}}$  par :

$$\mu = \sum_{k=1}^K q_k \delta_{\mathbf{x}_k}, \quad (5)$$

Le problème d'estimation des positions  $\mathbf{x}_k$  et des puissances  $q_k$  peut donc être remplacé par une estimation de la mesure  $\mu$ .

Avec cette description des sources, la matrice de covariance pour les paramètres  $\mu$  et  $\sigma^2$  peut s'écrire :

$$\mathbf{R}_j(\mu, \sigma^2) = \int_{\Omega} \mathbf{C}_j d\mu + \sigma^2 \mathbf{I}, \quad (6)$$

avec  $\mathbf{C}_j = \mathbf{g}_j(\mathbf{x}) \mathbf{g}_j(\mathbf{x})^H$ . En combinant l'estimateur COMET en dimension finie pour mesures asynchrones de [6] et le problème COMET en dimension infinie de [1], on peut estimer  $\mu$  de la manière suivante :

$$\hat{\mu} = \operatorname{argmin}_{\mu \in \mathcal{M}} \sum_{j=1}^J \operatorname{tr}(\mathbf{R}_j) + \operatorname{tr}(\hat{\Sigma}_j \mathbf{R}_j^{-1} \hat{\Sigma}_j). \quad (7)$$

$\mathcal{M}$  désigne l'ensemble des mesures de Radon sur  $\Omega$ . Nous utilisons ici le critère COMET modifié pour prendre en compte les cas où le nombre d'échantillons est réduit.

Le problème d'optimisation (7) est résolu avec l'aide de l'algorithme *Sliding Frank Wolfe* (SFW) [3]. L'algorithme est

détaillé dans [1]. Pour chaque itération  $k$ , une nouvelle source est introduite à l'emplacement du maximum global de

$$\eta_k(\mathbf{x}) = \sum_{j=1}^J \mathbf{g}_j(\mathbf{x}) (\mathbf{R}_{j,k-1}^{-1} \hat{\Sigma}_j \hat{\Sigma}_j \mathbf{R}_{j,k-1}^{-1}) \mathbf{g}_j(\mathbf{x}) - \|\mathbf{g}_j(\mathbf{x})\|^2, \quad (8)$$

avec  $\mathbf{R}_{j,k-1}$  la matrice de covariance obtenues avec les paramètres estimés à l'itération  $k-1$ . Les paramètres sont ensuite optimisés localement. L'algorithme est stoppé soit à convergence, soit quand un nombre prédéfini de sources est identifié. Dans ce cas, l'algorithme paraît mieux se comporter avec des vecteurs sources  $\mathbf{g}(\mathbf{x})$  normalisés.

### 3 Évaluation des performances

Deux configurations de réseaux sont considérées ici, un premier cas avec des réseaux parallèles à une distance de 8 m, et un deuxième cas avec des réseaux orthogonaux. Chaque réseau possède 128 microphones et les placements sont visibles sur la Figure 1. Les sources sont placées respectivement à (0.11 -0.03 2.3) et (0.52 -0.32 5.71) avec des puissances de 1 Pa<sup>2</sup> à un mètre de distance, et un niveau de bruit  $\sigma^2 = 0.94$  Pa<sup>2</sup>. Les puissances des sources et le niveau de bruit sont supposées être identiques pour les deux mesures asynchrones. En complément du cas d'étude asynchrone, on considère deux autres configurations, d'une part le cas d'un réseau seul qui correspond à une acquisition avec un seul des deux réseaux, et d'autre part le cas synchrone qui correspond à une acquisition avec deux réseaux placés à des endroits différents mais dont les mesures sont synchronisées ce qui est équivalent à ne considérer qu'un seul grand réseau rassemblant les microphones des deux réseaux. Afin d'avoir un temps d'acquisition total constant, 250 échantillons sont utilisés par réseau dans le cas asynchrone, et 500 échantillons sont utilisés dans le cas synchrone ou lorsqu'un seul réseau est considéré.

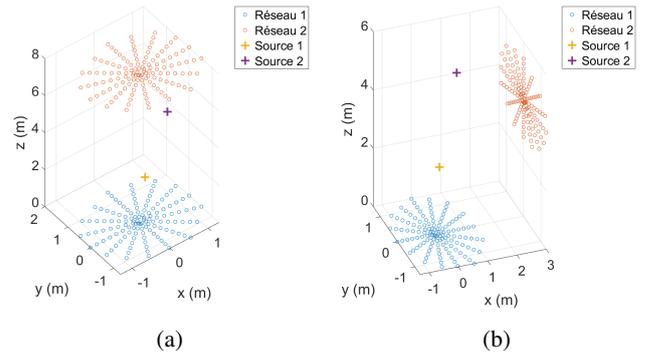


FIGURE 1 : Placements des réseaux et des sources pour les simulations : (a) réseaux opposés, (b) réseaux orthogonaux.

#### 3.1 Cas de réseaux parallèles

Les Bornes de Cramér-Rao (BCR) pour chaque cas, réseaux seuls, réseaux synchrones et asynchrones, en fonction de la fréquence et dans le cas de réseaux opposés sont présentées à la Figure 2. Les BCR sont obtenues en inversant la matrice d'information de Fisher, elle-même obtenue en sommant les

matrices de Fisher de chaque antenne [6]. Les BCR représentées dans les figures correspondent aux sommes des BCR de chaque source. La Figure 2 montre également le ratio des BCR entre le cas synchrone et le cas asynchrone, c'est à dire la perte de performance engendrée par l'absence de synchronisation entre les réseaux, et entre les réseaux seuls et le cas asynchrone, qui rend compte du gain de performance obtenu en déplaçant le réseau, pour un temps de mesure total constant.

Les bornes en position et en puissance montre une décroissance avec l'augmentation de la fréquence. Comme attendu, le cas synchrone présente une meilleure performance par rapport aux autres cas due à la quantité d'information disponible plus importante. Le placement du deuxième réseau permet d'avoir des meilleures estimations de la position par rapport au premier réseau. En effet, la première source est plus éloignée mais d'amplitude plus forte et la deuxième source est d'amplitude plus faible mais plus proche. Le cas asynchrone permet d'obtenir un gain de performance par rapport aux réseaux utilisés séparément, à la fois pour l'estimation des positions et des puissances. Pour les puissances, les bornes se rejoignent toutes en hautes fréquences.

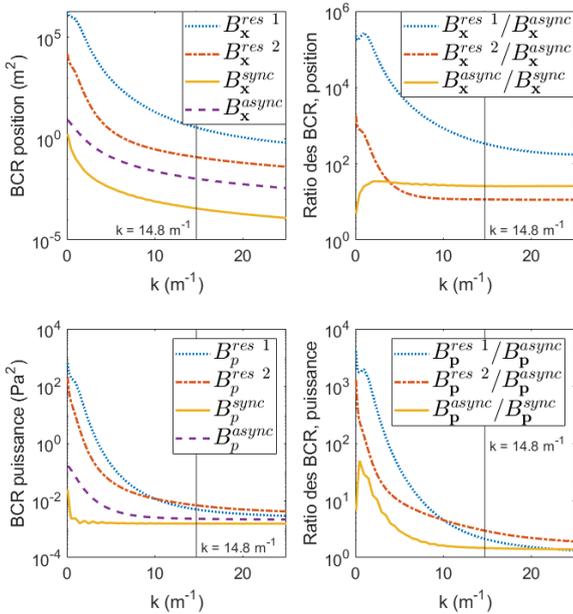


FIGURE 2 : Bornes de Cramér-Rao en position et en puissance dans le cas de réseaux parallèles avec 2 sources, pour les réseaux seuls, le cas synchrone et le cas asynchrone.

### 3.2 Cas de réseaux orthogonaux

La Figure 3 montre les BCR en position et en puissance en fonction de la fréquence dans le cas de réseaux orthogonaux et pour les mêmes configurations que précédemment. La Figure montre également les différents ratios présentés dans le cas des réseaux opposés. La décroissance en fréquence est également visible pour les positions et puissances. Le ratio entre les cas synchrones et asynchrones pour les BCR en positions est plus faible par rapport aux réseaux opposés.

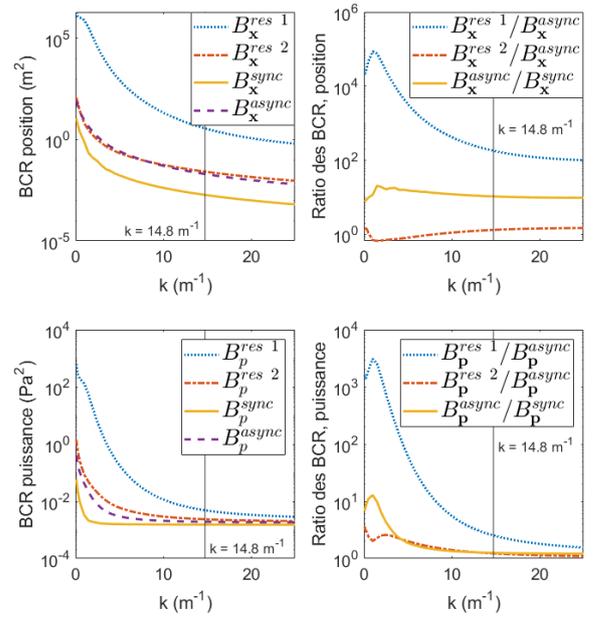


FIGURE 3 : Bornes de Cramér-Rao en position et en puissance dans le cas de réseaux orthogonaux avec 2 sources, pour les réseaux seuls, le cas synchrone et le cas asynchrone.

## 3.3 Simulations

### 3.3.1 Cas de réseaux parallèles

Afin de caractériser la performance de la méthode COMET2 formulée avec parcimonie continue en fonction du bruit, des simulations ont été réalisées. La méthode est comparée au cas synchrone ainsi qu'aux deux réseaux utilisés séparément. Les mêmes positions de sources que précédemment sont utilisées mais avec des puissances respectives de  $0.5 \text{ Pa}^2$  et  $0.1 \text{ Pa}^2$ . La grille d'initialisation de recherche des sources est définie telle que  $-1 \leq X \leq 1$ ,  $-1 \leq Y \leq 1$  et  $2 \leq Z \leq 7$ , avec un pas de discrétisation de  $0.05 \text{ m}$  dans chaque direction. Dans le cas de réseaux synchrones, la discrétisation dans les directions  $x$  et  $y$  permettent de tomber près du bon maximum local mais la discrétisation en  $z$  n'est pas suffisante. Afin de pallier à cette limitation, plusieurs optimisations de  $\eta(\mathbf{x})$  sont effectuées à partir des positions en  $x$  et  $y$  obtenues avec la première grille d'initialisation, et diverses positions en  $z$ . La fréquence utilisée est de  $800 \text{ Hz}$  ( $k \approx 14.8 \text{ m}^{-1}$ ). La Figure 4 présente l'évolution de l'erreur quadratique moyenne (EQM) en position et en puissance en fonction de la variance du bruit  $\sigma^2$  lorsque les réseaux sont opposés, ainsi que les BCR correspondantes. Chaque point sur les graphes correspond à la moyenne des EQM des deux sources et du bruit pour 500 réalisations. Les résultats montrent une décroissance de l'EQM avec le bruit à la fois pour la position et pour la puissance. La méthode proposée permet bien d'atteindre des performances proches des BCR pour chacune des configurations en position et en puissance.

### 3.3.2 Cas de réseaux orthogonaux

Les résultats avec les réseaux orthogonaux sont présentés à la Figure 5 avec les EQM et les BCR correspondants. Le même

## 4 Conclusion

Cette étude présente une méthode de localisation employant un estimateur COMET avec une formulation sans grille appliqué à des mesures asynchrones. Afin de tester la méthode avec des mesures asynchrones, un réseau de microphones est utilisé à deux positions différentes avec deux configurations, une première où les réseaux sont parallèles et une deuxième où les réseaux sont orthogonaux. Les bornes de Cramér-Rao permettent d'évaluer les performances de ces configurations et de les comparer avec l'utilisation des réseaux seuls et des réseaux synchrones. Les résultats montrent que dans les cas considérés ici, l'utilisation de mesures asynchrones permet une meilleure performance par rapport à un réseau seul et avec un temps de mesure constant. Plusieurs simulations de sources dans les différentes configurations montrent que les erreurs quadratiques moyennes des estimations en positions et en puissances se rapprochent des bornes de Cramér-Rao avec la méthode présentée.

## Références

- [1] Gilles CHARDON : Gridless covariance matrix fitting methods for three dimensional acoustical source localization. *Journal of Sound and Vibration*, 551:117608, mai 2023.
- [2] Gilles CHARDON et Ulysse BOUREAU : Gridless three-dimensional compressive beamforming with the Sliding Frank-Wolfe algorithm. *The Journal of the Acoustical Society of America*, 150(4):3139–3148, octobre 2021.
- [3] Quentin DENOYELLE, Vincent DUVAL, Gabriel PEYRÉ et Emmanuel SOUBIES : The Sliding Frank-Wolfe Algorithm and its Application to Super-Resolution Microscopy. *IOP Publishing*, 36(1):014001, décembre 2019.
- [4] Lourenço Tércio LIMA PEREIRA, Roberto MERINO-MARTÍNEZ, Daniele RAGNI, David GÓMEZ-ARIZA et Mirjam SNELLEN : Combining asynchronous microphone array measurements for enhanced acoustic imaging and volumetric source mapping. *Applied Acoustics*, 182:108247, novembre 2021.
- [5] B OTTERSTEN, P STOICA et R ROY : Covariance Matching Estimation Techniques for Array Signal Processing Applications. *Digital Signal Processing*, 8(3):185–210, juillet 1998.
- [6] Wassim SULEIMAN, Pouyan PARVAZI, Marius PESAVENTO et Abdelhak M. ZOUBIR : Non-Coherent Direction-of-Arrival Estimation Using Partly Calibrated Arrays. *IEEE Transactions on Signal Processing*, 66(21): 5776–5788, novembre 2018.
- [7] Liang YU, Jerome ANTONI et Quentin LECLERE : Spectral matrix completion by Cyclic Projection and application to sound source reconstruction from non-synchronous measurements. *Journal of Sound and Vibration*, 372:31–49, juin 2016.

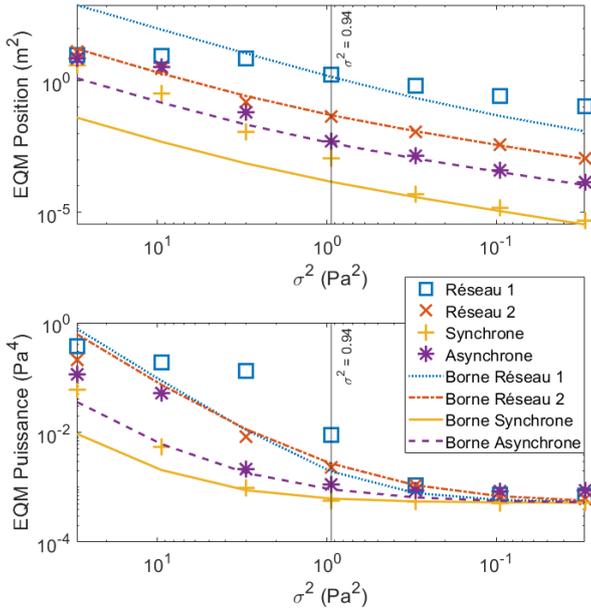


FIGURE 4 : Évolution des EQM en position et en puissance en fonction de la variance du bruit dans le cas de réseaux parallèles.

comportement de décroissance avec le bruit est visible. Les EQM atteignent également les BCR en position et en puissance. Les EQM des cas synchrones et asynchrones avec les réseaux orthogonaux présentent des erreurs plus importantes qu'avec les réseaux opposés.

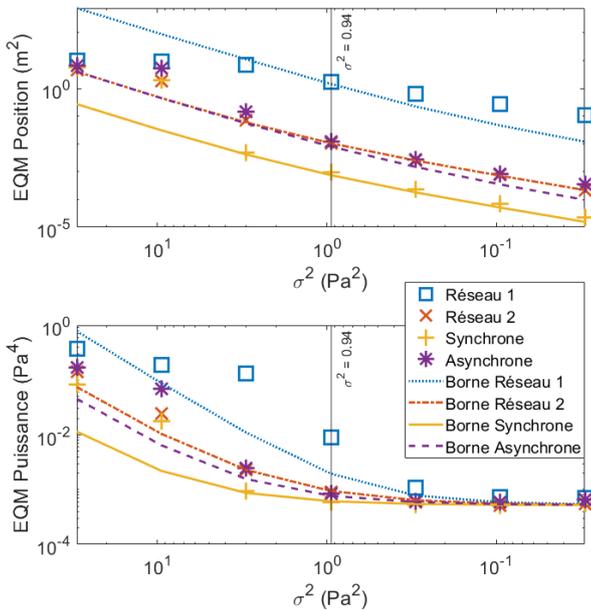


FIGURE 5 : Évolution des EQM en position et en puissance en fonction de la variance du bruit dans le cas de réseaux orthogonaux.