Gridless covariance matrix fitting methods for three dimensional acoustical source localization

Gilles Chardon

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

Abstract

In this article, we consider gridless source localization based on the spatial covariance matrix of acoustical data collected by an array of microphones. Covariance matrix fitting problems are formulated in infinite-dimensional settings, and solved by the Sliding Frank-Wolfe algorithm. The proposed method does not impose any constraint on the geometry of the array, the propagation model or the domain of interest, and does not necessitate a training phase. It is tested on simulated and experimental measurements for the localization of sources in a three-dimensional domain. Performances are compared to the state of the art, showing in particular a better resolution than MUSIC (MUltiple SIgnal Classification) at low SNR.

Keywords: source localization, gridless methods, Frank-Wolfe

1. Introduction

The localization of noise sources and the quantification of their power from acoustical data from a microphone array, in particular, from the spatial covariance matrix (SCM), also known as the cross-spectral matrix, of the measurements has attracted a sizable corpus of research[1]. Beamforming is simple to implement and efficient in the case of a unique source[2], but its limited resolution prevents accurate characterization of the acoustical sources, in particular at low frequencies. Several beamforming deconvolution methods have been proposed to alleviate this limitation, based on the deconvolution of the beamforming map (DAMAS (Deconvolution Approach for the Mapping of Acoustic Sources)[3], DAMAS-NNLS (DAMAS-NonNegative Least Squares)[4], etc.), or iterative algorithms (CLEAN and CLEAN-SC (CLEAN-Source Coherence)[5], HR-CLEAN-SC (High Resolution-CLEAN-SC), [6], Orthogonal Matching Pursuit (OMP)[7], etc.). Localization of the sources can also be performed using the SCM directly. MUSIC[8] (MUltiple SIgnal Classification) and the Covariance Fitting Method (CMF)[9] methods are examples of such methods. In particular, CMF is based on a least-squares fit of the estimated SCM to a theoretical SCM. It has recently been proved to be equivalent to the DAMAS method, as the DAMAS algorithm solves the CMF optimization problem[10].

The methods cited above are all subject to the basis mismatch problem[11]. As they necessitate the discretization of the domain of interest, they cannot accurately represent acoustical sources that are not exactly on the grid. This implies a spreading of the acoustical sources on neighboring grid nodes, and a bias in the estimation of their powers. It should be noted that refining the grid cannot be expected to alleviate this problem[12].

Email address: gilles.chardon@centralesupelec.fr (Gilles Chardon)

Gridless methods for source localization are an attractive alternative to grid-based methods, as they are not subject to these limitations. Several methods have been recently proposed for the model considered here (the so-called *unconditional model*[13]), based on deep learning[14], global optimization by differential evolution[15, 16], or in particular settings, semi-definite programming[17]. However, as will be shown later, these methods have limitations that prevent their use in general cases.

In order to perform gridless source localization, we propose a method based on the formulation of an infinite dimensional optimization problem, where the distribution of sources is no longer described by a finite dimensional vector, but by a measure. The problem we consider is a infinite dimensional version of the CMF problem. This optimization problem is solved using the Sliding Frank-Wolfe algorithm (SFW)[18], initially proposed for the Beurling LASSO (Least Absolute Shrinkage and Selection Operator)[19]. A variant of the CMF problem, known as COMET (COvariance Matching Estimation Technique), [20, 21], will also be considered. A similar method was used recently by the author in the case of deterministic sources [22]. This model (*conditional model*[13]), and the associated state of the art, are out of scope of the present study, where sources are assumed to be random.

The proposed method does not impose limitations on the propagation model, the configuration of the array or of the sources, and can be used without a priori knowledge on the number of sources. Performances in term of mean squared errors (MSE) of the proposed method are compared to several other methods: CLEAN-SC, HR-CLEAN-SC, OMP, orthogonal beamforming (OBF) and MUSIC. These experiments demonstrate the superiority the proposed method in comparison to CLEAN-SC, HR-CLEAN-SC and OMP, and better performances than MUSIC at low Signal to Noise Ratio (SNR). The

superiority of the proposed method in this case is explained by a better resolution than MUSIC.

The paper is organized as follows. Section 2 introduces the source localization model, and discusses the state of the art. The optimization problems used for the estimation of the parameters of the sources are formulated in section 3. In section 4, the SFW algorithm is recalled, and its adaptation to the CMF and COMET problems is introduced. Experimental and numerical results are given in sections 5. Section 6 concludes the paper.

A preliminary version of this work was presented at ICSV27[23]. Code and data necessary to reproduce the presented results are available online[24].

2. Model and state of the art

We consider an array of M microphones, located at positions $\mathbf{y}_m$, measuring acoustical data. Complex amplitudes of the measurements at a given frequency f are obtained at S times t_s (in practice, time domain measurements are analyzed by a Short Time Fourier Transform). Assuming the presence of K sources at positions $\mathbf{x}_k$, with complex amplitudes a_{ks}, the measured data $\mathbf{p}_s$ at time t_s can be decomposed as

$$\mathbf{p}_s = \sum_{k=1}^{K} a_{ks}\mathbf{g}(\mathbf{x}_k) + \mathbf{n}_s \tag{1}$$

where $\mathbf{g}(\mathbf{x}_k)$ is the vector collecting the values of the Green function from the source at $\mathbf{x}_k$ to the sensors, and $\mathbf{n}_s$ is a measurement noise, assumed to be white in space and time. In free field conditions, the vector $\mathbf{g}(\mathbf{x})$ is given by its coefficients

$$g_m(\mathbf{x}) = \frac{\exp\left(-\mathrm{i}\kappa\|\mathbf{x}-\mathbf{y}_m\|_2\right)}{\|\mathbf{x}-\mathbf{y}_m\|_2}, \tag{2}$$

where κ is the wavenumber. This will be the case in the simulations and experiments. However, the theory of the method does not assume any particular form of $\mathbf{g}$.

The amplitudes a_{ks} are assumed to be realizations of random variables, and the covariance matrix $\mathbf{\Sigma}$ of the measurements $\mathbf{p}_s$, themselves random variables, is considered. Assuming that the sources amplitudes are mutually uncorrelated, and uncorrelated with the measurement noise, the covariance matrix $\mathbf{\Sigma}$ is given by

$$\mathbf{\Sigma} = \sum_{k=1}^{K} q_k \mathbf{g}(\mathbf{x}_k)\mathbf{g}(\mathbf{x}_k)^\star + \sigma^2\mathbf{I} \tag{3}$$

where $\cdot^\star$ denotes Hermitian conjugation, q_k is the power of the k-th source, and σ^2 is the variance of the measurement noise. The data used to estimate the positions and powers of the source are the coefficients of the SCM

$$\hat{\mathbf{\Sigma}} = \frac{1}{S}\sum_{s=1}^{S} \mathbf{p}_s\mathbf{p}_s^\star \tag{4}$$

which approaches $\mathbf{\Sigma}$ when the number of snapshots S increases.

2.1. State of the art

Several methods have been proposed to perform gridless localization of acoustical sources from the data $\hat{\mathbf{\Sigma}}$. Moreover, some methods of the literature can be easily extended to gridless localization.

Beamforming. When only one source is present, beamforming can be used to locate the source and estimate its power. Beamforming localizes a source by maximizing

$$B(\mathbf{x}) = \frac{\mathbf{g}(\mathbf{x})^\star\hat{\mathbf{\Sigma}}\mathbf{g}(\mathbf{x})}{\|\mathbf{g}(\mathbf{x})\|_2^2}. \tag{5}$$

In the case of Gaussian source and noise, it can be interpreted as a maximum likelihood estimator[2], which is asymptotically (when the number of snapshots S increases) unbiased and efficient (its variance reaches the Cramér-Rao lower bounds). However, extending maximum likelihood estimation to several sources necessitates to solve a non-convex problem in high dimension (in 3D, $4K$ where K is the number of sources), which is intractable.

Deep learning. Application of deep learning to sound source localization from $\hat{\mathbf{\Sigma}}$ was recently investigated in a 2D setting[14]. A major limitation of this method is the necessity of learning a neural net for every combination of microphone array, frequency of interest, domain of interest, and number of sources. Additionally, the parametrization adopted in this article, with the sources ranked in order of decreasing power, implies that the function to be learned has discontinuities (e.g. when the power of the n-th source is increased above the power of the $n-1$-th source). However, Lipschitz regularity is a desirable property for neural nets, ensuring robustness with respect to errors in the inputs[25]. This regularity is, by definition, incompatible with the discontinuity of the estimator to be learned. Because of these limitations, deep learning will not be considered here.

Global optimization. Alternatively, global optimization can be used for the localization of several sources[15, 16], using a differential evolution algorithm. This method also requires the number of sources, and is computationally intensive. This method will be tested with experimental measurements. However, its large computational time will prevent its inclusion in the Monte-Carlo simulations.

Iterative algorithms. Additionally, methods initially proposed for grid-based source localization can be easily adapted to gridless localization. CLEAN-SC, as well as the similar OMP method, are iterative methods that, at each iteration, identify a new source by maximizing a criterion on a grid. in HR-CLEAN-SC, a further post-processing is applied to the results of CLEAN-SC.

A gridless version of these algorithms can be obtained by optimizing this criterion in the domain where source are searched, e.g. by a Newton method.

Subspace based methods. MUSIC, where sources are found as local maxima of a pseudo-spectrum, can also be extended to gridless localization. Orthogonal beamforming can also be extended by performing the beamforming step gridlessly[26].

Far field sources. In the case of linear microphone arrays and distant sources, methods leveraging properties of complex exponentials can be used to perform gridless direction of arrival estimation. In particular, atomic norm[27] and the COMET optimization problem[17] were recently used in this context. These methods cannot deal with the 3D settings considered in this article.

3. Infinite dimensional covariance matrix fitting problems

In this section, we formulate the infinite dimensional problems solved to estimate the positions and powers of the sources. We first briefly recall the CMF problem in finite dimension. Given a grid of L potential sources locations $\mathbf{z}_l$, and $\mathbf{q} \in \mathbf{R}^L$ the vector of the powers of the source (most being zero if the distribution of source is sparse), Eq. (3) can be written

$$\mathbf{\Sigma} = \mathbf{G}\mathrm{diag}(\mathbf{q})\mathbf{G}^\star + \sigma^2\mathbf{I} \tag{6}$$

with $\mathbf{G}$ with columns $\mathbf{g}(\mathbf{z}_l)$.

In the Covariance Fitting Matrix method, the finite dimensional vector $\hat{\mathbf{q}}$ modeling the estimated distribution of sources in space is obtained by solving the optimization problem

$$\hat{\mathbf{q}} = \underset{\mathbf{q}\in\mathbf{R}_+^L, \sigma^2\in\mathbf{R}_+}{\mathrm{argmin}} \left\| \hat{\mathbf{\Sigma}} - \left(\mathbf{G}\mathrm{diag}(\mathbf{q})\mathbf{G}^\star + \sigma^2\mathbf{I}\right)\right\|_{\mathrm{Fro}}^2 \tag{7}$$

which is a least-squares fit of the covariance matrix of the model $\mathbf{\Sigma}$ to the spatial covariance matrix of the measurements $\hat{\mathbf{\Sigma}}$. The Frobenius norm of a matrix $\|\cdot\|_{\mathrm{Fro}}$ is the ℓ_2 norm of its coefficients.

This is a nonnegative least-squares problem, which can be solved efficiently using the Lawson-Hanson algorithm and optimizations specific to this problem[10]. However, as pointed out above, gridded problems are subject to the basis mismatch problem, which cannot be solved by refining the grid.

In order to formulate a gridless source localization, we represent a distribution of sources by a measure μ, that is, a function mapping a subset E of Ω to the positive real numbers, measuring the total power of the sources located in E. A particular example of measure is the Dirac mass $\delta_{\mathbf{x}}$ modeling a punctual source of unit power, with $\mu(E) = 1$ is E contains the source, else $\mu(E) = 0$. A distribution of K punctual sources of powers q_k and positions $\mathbf{x}_k$ can be represented by a discrete measure

$$\mu = \sum_{k=1}^{K} q_k \delta_{\mathbf{x}_k}. \tag{8}$$

With $\mathbf{C}(\mathbf{x}) = \mathbf{g}(\mathbf{x})\mathbf{g}(\mathbf{x})^\star$, Eq. (3) can be rewritten as

$$\mathbf{\Sigma} = \int_\Omega \mathbf{C} d\mu + \sigma^2\mathbf{I}, \tag{9}$$

which serves as an infinite dimensional version of Eq. (6).

The infinite dimensional version of CMF is the following optimization problem:

$$\hat{\mu} = \underset{\mu\in\mathcal{M}}{\mathrm{argmin}}\, f_{\mathrm{CMF}}(\mu, \sigma^2). \tag{10}$$

where $\mathcal{M}$ is the set of Radon measures on Ω, and

$$f_{\mathrm{CMF}}(\mu, \sigma^2) = \left\|\mathbf{R}(\mu, \sigma^2) - \hat{\mathbf{\Sigma}}\right\|_{\mathrm{Fro}}^2 \tag{11}$$

$$\mathbf{R}(\mu, \sigma^2) = \int_\Omega \mathbf{C} d\mu + \sigma^2\mathbf{I} \tag{12}$$

Likewise, the two COMET1 and COMET2 covariance matrix fitting problems [20, 21] are obtained by setting

$$f_{\mathrm{COMET1}}(\mu, \sigma^2) = \left\|\mathbf{R}^{-1/2}\left(\mathbf{R} - \hat{\mathbf{\Sigma}}\right)\hat{\mathbf{\Sigma}}^{-1/2}\right\|_{\mathrm{Fro}}^2 \tag{13}$$

$$= \mathrm{tr}\left(\hat{\mathbf{\Sigma}}^{-1}\mathbf{R}\right) + \mathrm{tr}\left(\hat{\mathbf{\Sigma}}\mathbf{R}^{-1}\right) + C_1 \tag{14}$$

$$f_{\mathrm{COMET2}}(\mu, \sigma^2) = \left\|\mathbf{R}^{-1/2}\left(\mathbf{R} - \hat{\mathbf{\Sigma}}\right)\right\|_{\mathrm{Fro}}^2 \tag{15}$$

$$= \mathrm{tr}\left(\mathbf{R}\right) + \mathrm{tr}\left(\hat{\mathbf{\Sigma}}\mathbf{R}^{-1}\hat{\mathbf{\Sigma}}\right) + C_2 \tag{16}$$

where C_1 and C_2 are constant with respect to the parameters to be estimated. COMET1 is obtained as an approximation of the maximum likelihood estimator. COMET2 is a variant, which cannot be considered as such an approximation, but does not necessitate the inversion of the covariance matrix $\hat{\mathbf{\Sigma}}$ and therefore can be used in cases where the number of snapshots is smaller than, or comparable to, the number of sensors.

Dependency of $\mathbf{R}$ on the parameters μ and σ^2 is here left implicit for the sake of readability. In the remainder of the text, f will refer to one of these three objective function.

4. The Sliding Frank-Wolfe algorithm

The Sliding Frank-Wolfe was originally proposed to solve the Beurling LASSO problem[18] for infinite dimensional sparse recovery problems. We will use here this algorithm to solve the infinite dimensional CMF and COMET problems in Eqs. (10), (14) and (16).

Details of the algorithm are given in Table 1. The notation $\mathbf{X}$ denotes a tuple of k positions $(\mathbf{x}_1, \dots \mathbf{x}_k)$ and $\mathbf{q}$ is the k-dimensional vector containing their powers, $(\mathbf{X}, \mathbf{x}_\star)$ denotes the insertion of $\mathbf{x}_\star$ into the tuple $\mathbf{X}$. F is defined as

$$F(\mathbf{X}, \mathbf{q}, \sigma^2) = f\left(\sum_{i=1}^{k} q_i \delta_{\mathbf{x}_i}, \sigma^2\right). \tag{17}$$

The iterations of the SFW algorithm can be summarized as follows:

- a source is first added, by solving the global optimization problem in Eq. (23). In practice, a local search is performed, initialized by a search on a finite grid.
- powers of the sources are updated the problem in Eq. (24).

- Then, powers and positions are jointly optimized in the problem in Eq. (25). Here, local optimality is sufficient. A quasi-Newton method is used, initialized at the positions and powers obtained at the previous step.

- The algorithm is stopped when adding a source does not improve the objective, or when a given number of sources are identified.

We used the Matlab 2021a function `fmincon` to solve the problems in Eqs. (23), (24) and (25), with the sequential quadratic programming algorithm.

The identification of a new source in Eq. (23) and the stopping criterion are derived as follows.

Following the principles of the Sliding Frank-Wolfe algorithm, at each step a source (modeled by a Dirac measure) is added. The location of this new source is chosen by maximizing the decrease of the objective function for a source with infinitesimal power, that is

$$\mathbf{x}_\star = \operatorname*{argmax}_{\mathbf{x}\in\Omega} \eta(\mathbf{x}) \tag{18}$$

where

$$\eta(\mathbf{x}) = -\frac{\mathrm{d}}{\mathrm{d}\alpha} f(\mu + \alpha\delta_{\mathbf{x}}, \sigma^2). \tag{19}$$

Routine computations yield the following expressions of the criterion η for the three problems:

$$\eta_{\mathrm{CMF}}(\mathbf{x}) = \mathbf{g}(\mathbf{x})^\star(\hat{\mathbf{\Sigma}} - \mathbf{R})\mathbf{g}(\mathbf{x}) \tag{20}$$

$$\eta_{\mathrm{COMET1}}(\mathbf{x}) = \mathbf{g}(\mathbf{x})^\star(\mathbf{R}^{-1}\hat{\mathbf{\Sigma}}\mathbf{R}^{-1} - \hat{\mathbf{\Sigma}}^{-1})\mathbf{g}(\mathbf{x}) \tag{21}$$

$$\eta_{\mathrm{COMET2}}(\mathbf{x}) = \mathbf{g}(\mathbf{x})^\star(\mathbf{R}^{-1}\hat{\mathbf{\Sigma}}\hat{\mathbf{\Sigma}}\mathbf{R}^{-1})\mathbf{g}(\mathbf{x}) - \|\mathbf{g}(\mathbf{x})\|^2. \tag{22}$$

Evaluation of η for the COMET problems necessitates to invert the matrix $\mathbf{R}$. In cases where the number of sources is small compared to the number of sensors, the matrix $\mathbf{R}$ can be efficiently inverted with the Woodbury matrix identity. In addition the matrix $\mathbf{\Sigma}$ has to be inverted, once, for COMET1.

After the local update of the positions and powers of the identified sources, the objective function can be decreased only by adding a new source. When $\eta(\mathbf{x}) \leq 0$ for all positions $\mathbf{x}$, the objective function cannot be decreased further, and the algorithm is stopped.

5. Results

The method is first compared to the state of the art using simulations. Performances of the methods are assessed by the MSE of the estimation of the position and power of the sources. Then, the methods are compared on experimental data.

The tested methods are MUSIC, CLEAN-SC, HR-CLEAN-SC, OMP and OBF. Global optimization by differential evolution is not considered in the simulations because of its computational demands (computational times larger by two orders of magnitude compared to the other methods).

In the case of MUSIC, powers of the sources are estimated by the diagonal terms of the least-squares fit of the covariance matrix

$$\hat{\mathbf{Q}} = \operatorname*{argmin}_{\mathbf{Q}\in\mathbf{R}^{K\times K}} \left\| \hat{\mathbf{\Sigma}} - \left(\hat{\mathbf{G}}\mathbf{Q}\hat{\mathbf{G}}^\star + \hat{\sigma}^2\mathbf{I}\right)\right\|_{\mathrm{Fro}}^2 \tag{26}$$

$$= \hat{\mathbf{G}}^\dagger\left(\hat{\mathbf{\Sigma}} - \hat{\sigma}^2\mathbf{I}\right)\hat{\mathbf{G}}^{\dagger\star} \tag{27}$$

where $\hat{\sigma}^2$ is the average of the $N - K$ smallest eigenvalues of $\hat{\mathbf{\Sigma}}$, and $\hat{\mathbf{G}}$ is the $N \times K$ matrix containing the vectors $\mathbf{g}(\hat{\mathbf{x}}_k)$ for the estimated positions $\hat{\mathbf{x}}_k$ and $\hat{\mathbf{G}}^\dagger$ its pseudo-inverse. This estimator of the powers of the sources will also be used as a post-processing to the results of the COMET problems.

In the experimental and simulation results, an array of 128 microphones is used, with positions shown on figure 1, in the plane $Z = 0$.

Algorithm 1 Sliding Frank-Wolfe algorithm

$\mu^{[0]} \leftarrow 0$, $\mathbf{X}^{[0]} \leftarrow ()$
for $k = 1, \ldots, K$ **do**
  Identify a new source by solving

$$\mathbf{x}_\star = \operatorname*{argmax}_{x\in\Omega} \eta^{[k]}(\mathbf{x}) \tag{23}$$

  if $\eta(\mathbf{x}_\star) \leq 0$ **then**
    Stop
  else
    $\mathbf{X}^{[k-1/2]} = (\mathbf{X}^{[k]}, \mathbf{x}_\star)$
    Optimize the amplitudes:

$$\mathbf{q}^{[k-1/2],\sigma^2} = \operatorname*{argmin}_{\mathbf{q}\in\mathbf{R}_+^k,\sigma^2} F(\mathbf{X}^{[k-1/2]}, \mathbf{q}, \sigma^2) \tag{24}$$

    Optimize the amplitudes and positions:

$$(\mathbf{X}^{[k]}, \mathbf{q}^{[k]}, \sigma^2) = \operatorname*{argmin}_{\mathbf{X}\in\Omega^k,\mathbf{q}\in\mathbf{R}_+^k,\sigma^2} F(\mathbf{X}, \mathbf{q}, \sigma^2) \tag{25}$$

    $\mu^{[k]} \leftarrow \sum_{n=1}^k q_n^{[k]}\delta_{\mathbf{x}_n^{[k]}}$
  end if
end for

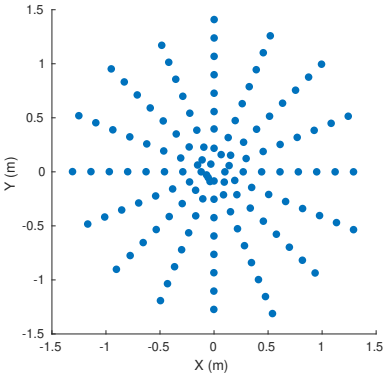

Figure 1: Layout of the microphone array

Source	X (m)	Y (m)	Z (m)	Power (Pa^2)
1	0.01	-0.13	4.12	1
2	0.11	-0.03	3.92	1
3	0.52	-0.32	4.71	0.5
4	-0.43	0.32	3.43	0.1

Table 1: Coordinates and powers of the simulated sources

5.1. Simulations

The number of sources $K = 4$ is here assumed to be known. OMP, CLEAN-SC (abbreviated as CSC), HR-CLEAN-SC (abbreviated as HRCSC), and the proposed methods are run with as many iterations as sources. OBF and MUSIC are used with the dimension of the signal subspace equal to the number of sources.

In the case of MUSIC, which is not guaranteed to return as many sources as requested, results are given only when MUSIC returns four sources for at least 90% of the samples, which are then used to compute the MSE.

The domain Ω is defined by $-1 \leq X \leq 1$, $-1 \leq Y \leq 1$, $3 \leq Z \leq 5$ (when not specified, coordinates are given in meters). Except when specified otherwise, $S = 500$ snapshots are used and the SNR is -10 dB.

Four sources are simulated at fixed positions and powers given in table 1. Performances are estimated by Monte-Carlo simulations, by averaging over 100 realizations of the source signals and noise.

Signal to noise ratio. The performances in function of the SNR are plotted on figure 2, at frequency $f = 2700$ Hz. Performances of OBF, OMP and CLEAN-SC do not improve when SNR increases. For OMP and CLEAN-SC, this is caused by inaccurate identification of the sources at earlier iterations, which cannot be corrected at later iterations. For OBF, as the sources have similar powers, the eigenvectors are linear combinations of the source vectors, which prevents an accurate estimation of the position and power of the sources. For the sake of clarity and given their poor localization performances, results of power estimation for OBF, OMP and CLEAN-SC will not be plotted.

At low SNR, CMF and COMET2 have the best performances, both in position and power estimation. While performances of COMET1 in position estimation are on par with the other methods, performances in amplitude estimation are limited.

At low SNR, MUSIC is unable to find all sources in most of the realizations. For SNR higher than 0 dB, all sources are identified, and MUSIC slightly outperforms the proposed methods at high SNRs. However, using the same power estimation method with, e.g., the positions estimated by COMET2 yields better performances (COMET2LS). HR-CLEAN-SC has similar performances as CMF for position estimation, but is not as accurate for power estimation. Several values of the parameters μ of HR-CLEAN-SC were tested, with slightly different results, and no effect on the comparison between the methods. Results are given here and in the two following simulations for $\mu = 0.25$.

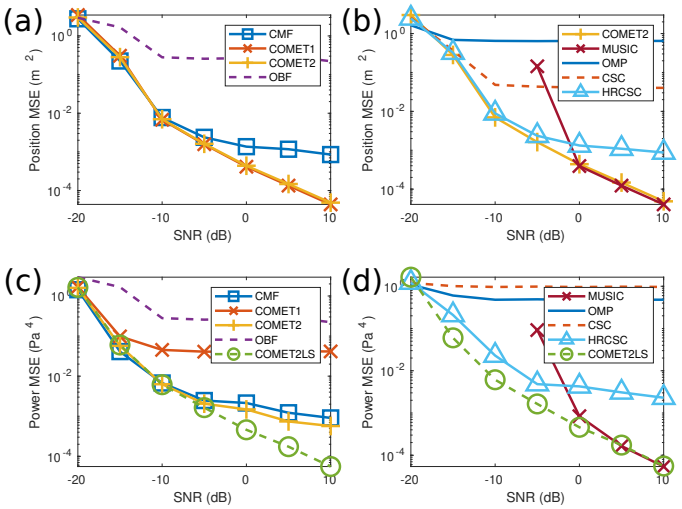

Figure 2: Simulations. MSE of position, (a)-(b), and power estimation, (c)-(d), in function of the SNR

Further experiments will be conducted at SNR = -10dB, which is not uncommon in wind tunnels[28].

Frequency. Here, performances are given in function of the frequency, with $f \in [1080, 5400]$

As above, best performances are obtained by COMET2. Performances of MUSIC are poor at this noise level. Results of OMP and CLEAN-SC improve as the frequency increases, but do not reach the performances of COMET2. Performances of HR-CLEAN-SC compared to the other methods are similar to COMET2 in position, but not as good in power.

Number of snapshots. The performances of the methods are here compared for varying number of snapshots, i.e. varying measurement duration, from $S = 1$ to $S = 5000$, with $f = 4330$ Hz.

MUSIC necessitates a large number of snapshots compared to the other methods to yield accurate estimations. COMET1 cannot be used with less snapshots than microphones. At high number of snapshots, performances of COMET2, CMF and MUSIC are comparable, while HR-CLEAN-SC is less accurate, in particular for power estimation.

Resolution. Resolution, the ability to identify two closely spaced sources, is tested by applying the methods on a scene consisting of two sources of equal power, separated by a varying distance δ.

The resolution in the X axis, parallel to the array, is first considered, at $f = 2700$Hz, on figure 5. The sources have coordinates $[\delta/2, 0, 4]$ and $[-\delta/2, 0, 4]$, and the X coordinates of the estimated sources are plotted in function of δ. The actual positions of the sources are indicated by dashed lines. The size of the markers is proportional to the estimated power.

Results of CMF and COMET1 are similar to COMET2 and not pictured, likewise for OMP, with similar results to CLEAN-SC. COMET2 and HR-CLEAN-SC have here the best resolu-

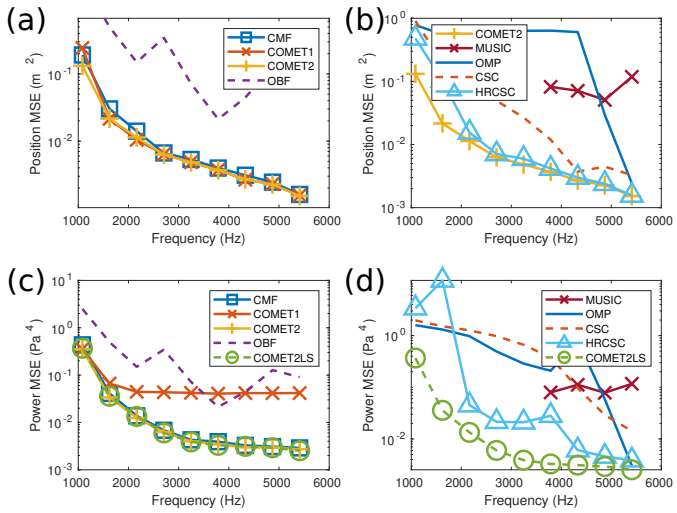

Figure 3: Simulations. MSE of position, (a)-(b), and power estimation, (c)-(d) in function of the frequency

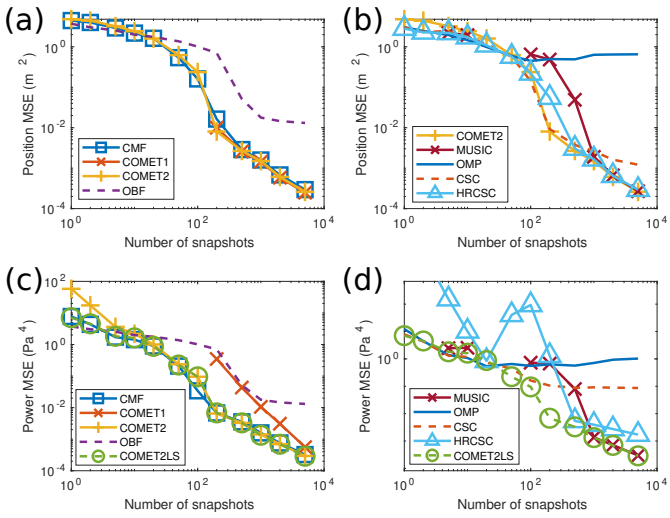

Figure 4: Simulations. MSE of position, (a)-(b), and power estimation, (c)-(d) in function of the number of snapshots

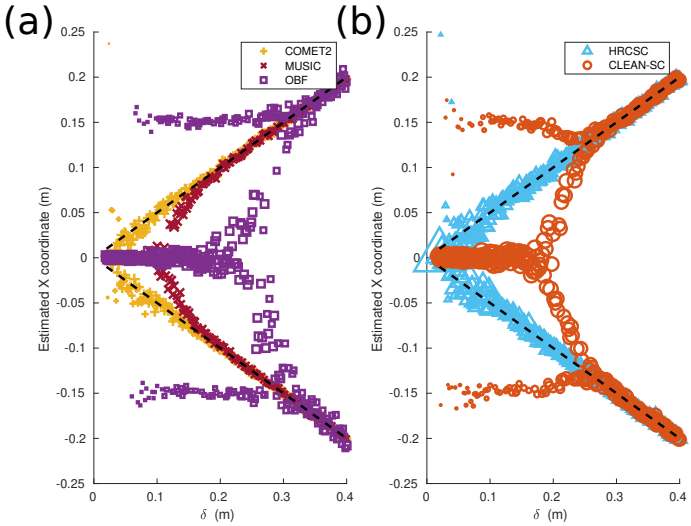

Figure 5: Simulations. Resolution in a plane parallel to the array. (a) COMET2, MUSIC, OBF, (b) CLEAN-SC, HR-CLEAN-SC. The estimated X coordinates are plotted in function of the space δ between the sources.

tion, as the estimated positions correspond to the actual coordinates, even for small gaps δ. The parameter μ of HR-CLEAN-SC was here set at 0, which yielded the best results. MUSIC, OMP, and OBF are unable to identify the two sources and return a source between the actual sources for δ smaller than 0.11, 0.18, and 0.25 respectively.

The resolution in range is considered on figure 6, at f = 27000Hz (a higher frequency is used as the resolution in range is less precise). Similar results are obtained. We note that resolution performances degrade around d = 0.2. However, while COMET2 resolution performances are not as good as in the X coordinate, its performances remain better than the other methods. HR-CLEAN-SC, while having better resolution as CLEAN-SC, MUSIC and OBF, does not reach the performances of COMET2.

For a more precise evaluation of the resolution performances of COMET2 and HR-CLEAN-SC, Figures 7 and 8 shows the average estimation error of the position of the two sources for varying δ along the X and Z axes respectively, as well as the first and third quartiles, indicated by the error bars, meaning that 50% of the errors fall between the errors bars. COMET2 appears to have both a smaller bias and errors more concentrated around 0 that HR-CLEAN-SC.

As a conclusion of the numerical experiments, COMET2 is shown to have better performances than MUSIC at low SNR, which is mainly explained by its better resolution at low SNR. COMET2 dominates CLEAN-SC, HR-CLEAN-SC, OBF and OMP at all regimes.

5.2. *Experimental results*

In the experiment, four sources (Visaton-BF32 omnidirectional loudspeakers) are used, pictured on Fig. 9, emitting white noise. The acoustical field is measured using an array of 128 MEMS microphones (INVENSENSE–INMP441) located

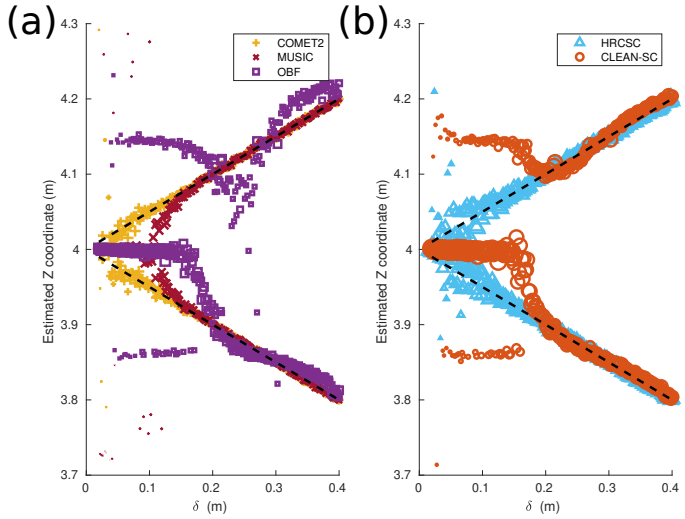

Figure 6: Simulations. Resolution in range. (a) COMET2, MUSIC, OBF, (b) CLEAN-SC, HR-CLEAN-SC. The estimated Z coordinates are plotted in function of the space δ between the sources.

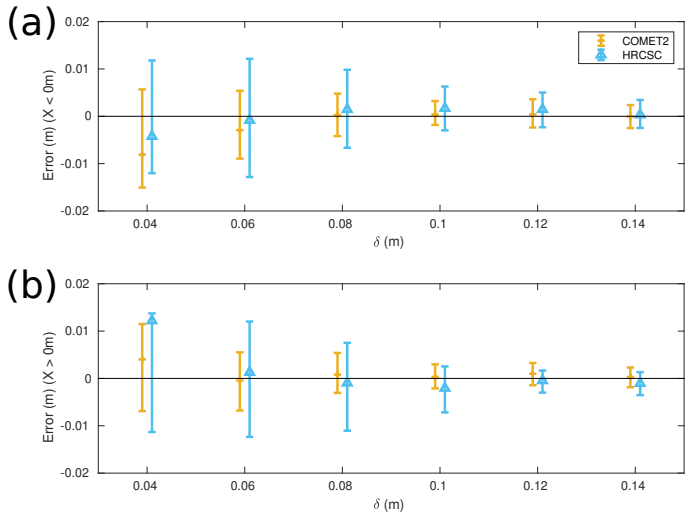

Figure 7: Simulations. Resolution in a plane parallel to the array. Distribution of the errors of COMET2 and HR-CLEAN-SC for varying δ. (a) Source at negative X, (b) source at positive X.

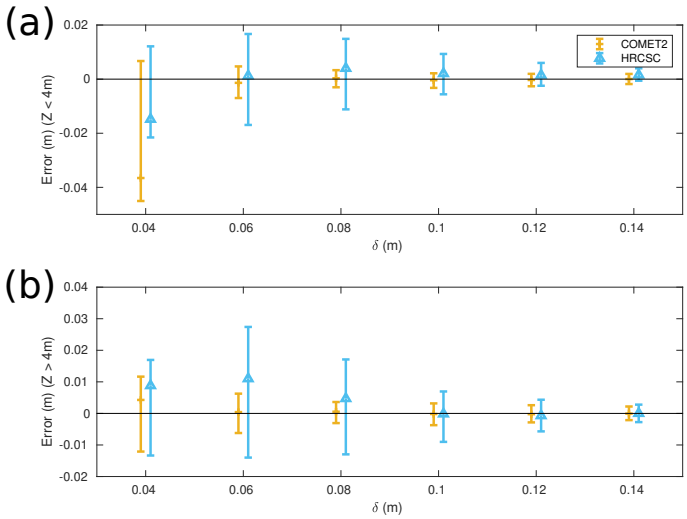

Figure 8: Simulations. Resolution in range. Distribution of the errors of COMET2 and HR-CLEAN-SC for varying δ. (a) Source closer to the array, (b) source further from the array.

in the plane $Z = 0$, with positions shown on Fig. 1. The sampling frequency is $F_s = 50$ kHz, and signals are analyzed by a Short-Time Fourier Transform, with Hann window of duration 82ms (4096 samples) and 50% overlap. More details on the experimental setup are found in Ref. [29]. The anechoic room at Institut Jean le Rond d'Alembert, Sorbonne Université, Paris, France, has a background noise level of 16 db(A), with a cut-off frequency of 80 Hz. The SNR of the measurement will thus be assumed to be high.

The domain of interest is the box defined by $-2 \leq X \leq 1$, $-1 \leq Y \leq 0$, $4 \leq Z \leq 5$, containing the four sources. Results are given for the frequency $f = 1818$ Hz. The signals have a length of 8s, which corresponds to $S = 196$ snapshots.

Figure 10 shows the result of conventional beamforming (using formulation IV, for position estimation [2]) in the plane $Z = 4.6$, where the sources lie approximately. Resolution of beamforming is not sufficient to identify the two central sources.

Sources located by COMET2, with an overestimated number of sources of $K = 8$ are plotted on figure 11, and compared with the results of MUSIC, CLEAN-SC, HR-CLEAN-SC, and differential evolution (DE) with four sources. The size of the markers is proportional to the estimated power of the sources. Computation times are 11s for COMET2, 0.4s for MUSIC, 1s for CLEAN-SC, 42s for HR-CLEAN-SC and 284s for differential evolution. COMET2 and MUSIC are able to locate the four sources. CLEAN-SC is unable to estimate the positions of the two central sources accurately. Estimation of the positions of the sources with differential evolution is not as accurate as COMET2 or MUSIC. HR-CLEAN-SC splits each of the central sources as two less powerful sources.

On figure 12, the same data is used, after adding a uncorrelated Gaussian noise of SNR = -10dB. While COMET2 is still able to locate the four sources, MUSIC cannot separate the two

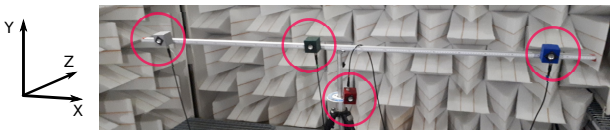

Figure 9: Experiment. Acoustical sources and coordinate frame

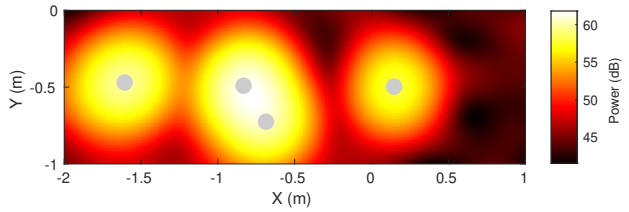

Figure 10: Experiment. Output of beamforming in the source plane. Actual positions are indicated by grey disks.

central sources, confirming the results obtained above at low SNR. CLEAN-SC yields results similar as in the noiseless case. HR-CLEAN-SC is here unable to locate the sources accurately.

6. Conclusion

Gridless acoustical source localization was performed with infinite dimensional versions of the CMF and COMET optimization problems, solved by the Sliding Frank-Wolfe algorithm.

Estimation performances were shown to be better compared to the state of the art, except MUSIC at high SNR. The better performances of the proposed methods as low SNR compared to MUSIC are explained by a better resolving power. In terms of computational time, the SFW algorithm lies between the faster MUSIC or CLEAN-SC methods, and the slower HR-CLEAN-SC and differential evolution algorithms.

References

[1] R. Merino-Martínez, P. Sijtsma, M. Snellen, T. Ahlefeldt, J. Antoni, C. J. Bahr, D. Blacodon, D. Ernst, A. Finez, S. Funke, T. F. Geyer, S. Haxter, G. Herold, X. Huang, W. M. Humphreys, Q. Leclère, A. Malgoezar, U. Michel, T. Padois, A. Pereira, C. Picard, E. Sarradj, H. Siller, D. G. Simons, C. Spehr, A review of acoustic imaging methods using phased microphone arrays: Part of the "Aircraft Noise Generation and Assessment" Special Issue, CEAS Aeronautical J. 10 (1) (2019) 197–230. doi:10.1007/s13272-019-00383-4.

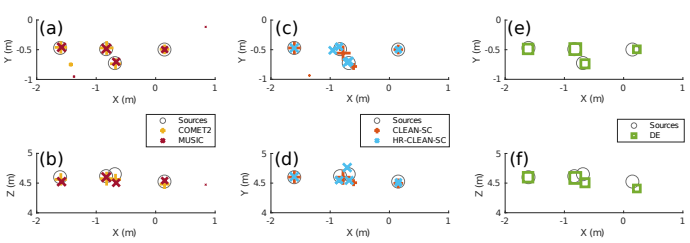

Figure 11: Experiment. Positions of the sources obtained by (a)-(b) COMET2, MUSIC, (c)-(d) CLEAN-SC and HR-CLEAN-SC, (e)-(f) differential evolution, noiseless measurements. Front (top) and top (bottom) views. The size of the markers is proportional to the estimated power of the sources.

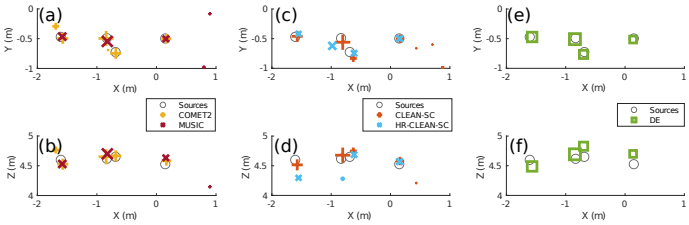

Figure 12: Experiment. Positions of the sources obtained by (a)-(b) COMET2, MUSIC, (c)-(d) CLEAN-SC and HR-CLEAN-SC, (e)-(f) differential evolution. Front (top) and top (bottom) views. The size of the markers is proportional to the estimated power of the sources.

[2] G. Chardon, Theoretical analysis of beamforming steering vector formulations for acoustic source localization, J. Sound Vib. 517 (2022) 116544. doi:10.1016/j.jsv.2021.116544.

[3] T. F. Brooks, W. M. Humphreys, A deconvolution approach for the mapping of acoustic sources (DAMAS) determined from phased microphone arrays, J. Sound Vib. 294 (4) (2006) 856–879. doi:10.1016/j.jsv.2005.12.046.

[4] K. Ehrenfried, L. Koop, Comparison of Iterative Deconvolution Algorithms for the Mapping of Acoustic Sources, AIAA J. 45 (7) (2007) 1584–1595. doi:10.2514/1.26320.

[5] P. Sijtsma, CLEAN based on spatial source coherence, Int. J. Aeroacoustics 6 (2007) 357–374. doi:10.1260/147547207783359459.

[6] P. Sijtsma, R. Merino-Martinez, A. M. Malgoezar, M. Snellen, High-resolution CLEAN-SC: Theory and experimental validation, Int. J. Aeroacoustics 16 (4-5) (2017) 274–298, publisher: SAGE Publications. doi:10.1177/1475472X17713034.

[7] T. Padois, A. Berry, Orthogonal matching pursuit applied to the deconvolution approach for the mapping of acoustic sources inverse problem, The J. Acoust. Soc. of Am. 138 (6) (2015) 3678–3685. doi:10.1121/1.4937609.

[8] R. Schmidt, Multiple emitter location and signal parameter estimation, IEEE Trans. on Antennas Propag. 34 (3) (1986) 276–280. doi:10.1109/TAP.1986.1143830.

[9] T. Yardibi, J. Li, P. Stoica, N. S. Zawodny, L. N. Cattafesta, A covariance fitting approach for correlated acoustic source mapping, The J. Acoust. Soc. of Am. 127 (5) (2010) 2920–2931. doi:10.1121/1.3365260.

[10] G. Chardon, J. Picheral, F. Ollivier, Theoretical analysis of the DAMAS algorithm and efficient implementation of the Covariance Matrix Fitting method for large-scale problems, J. Sound Vib. (2021) 116208doi:https://doi.org/10.1016/j.jsv.2021.116208.

[11] Y. Chi, L. L. Scharf, A. Pezeshki, A. R. Calderbank, Sensitivity to Basis Mismatch in Compressed Sensing, IEEE Trans. on Signal Process. 59 (5) (2011) 2182–2195. doi:10.1109/TSP.2011.2112650.

[12] V. Duval, G. Peyré, Sparse spikes super-resolution on thin grids II: the continuous basis pursuit, Inverse Probl. 33 (9) (2017) 095008. doi:10.1088/1361-6420/aa7fce.

[13] P. Stoica, A. Nehorai, Performance study of conditional and unconditional direction-of-arrival estimation, IEEE Trans. on Acoust., Speech, Signal Process. 38 (10) (1990) 1783–1795. doi:10.1109/29.60109.

[14] P. Castellini, N. Giulietti, N. Falcionelli, A. F. Dragoni, P. Chiariotti, A neural network based microphone array approach to gridless noise source localization, Appl. Acoust. 177 (2021) 107947. doi:10.1016/j.apacoust.2021.107947.

[15] A. M. N. Malgoezar, M. Snellen, R. Merino-Martinez, D. G. Simons, P. Sijtsma, On the use of global optimization methods for acoustic source mapping, The J. Acoust. Soc. of Am. 141 (1) (2017) 453–465, publisher: Acoustical Society of America. doi:10.1121/1.4973915.

[16] B. von den Hoff, R. Merino-Martínez, D. G. Simons, M. Snellen, Using global optimization methods for three-dimensional localization and quantification of incoherent acoustic sources, JASA Express Lett. 2 (5) (2022) 054802. doi:10.1121/10.0010456.

[17] Y. Park, P. Gerstoft, Gridless sparse covariance-based beamforming via alternating projections including co-prime arrays, The J. Acoust. Soc. of

Am. 151 (6) (2022) 3828–3837. doi:10.1121/10.0011617.
[18] Q. Denoyelle, V. Duval, G. Peyré, E. Soubies, The sliding Frank–Wolfe algorithm and its application to super-resolution microscopy, Inverse Probl. 36 (1) (2019) 014001. doi:10.1088/1361-6420/ab2a29.
[19] Y. de Castro, F. Gamboa, Exact reconstruction using Beurling minimal extrapolation, J. Math. Analysis Appl. 395 (1) (2012) 336–354. doi:10.1016/j.jmaa.2012.05.011.
[20] B. Ottersten, P. Stoica, R. Roy, Covariance Matching Estimation Techniques for Array Signal Processing Applications, Digit. Signal Process. 8 (3) (1998) 185–210. doi:10.1006/dspr.1998.0316.
[21] P. Stoica, P. Babu, J. Li, SPICE: A Sparse Covariance-Based Estimation Method for Array Processing, IEEE Trans. on Signal Process. 59 (2) (2011) 629–638. doi:10.1109/TSP.2010.2090525.
[22] G. Chardon, U. Boureau, Gridless three-dimensional compressive beamforming with the Sliding Frank-Wolfe algorithm, The J. Acoust. Soc. of Am. 150 (4) (2021) 3139–3148. doi:10.1121/10.0006790.
[23] G. Chardon, Gridfree source localization using the sliding frank-wolfe algorithm, in: 27th International Congress on Sound and Vibration, ICSV 2021, Online, 2021.
[24] G. Chardon, gilleschardon/SFWCMF, doi:10.5281/zenodo.7573626 https://zenodo.org/record/7573626 last accessed 26/01/2023,.
[25] A. Virmaux, K. Scaman, Lipschitz regularity of deep neural networks: analysis and efficient estimation, in: S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, R. Garnett (Eds.), Advances in Neural Information Processing Systems, Vol. 31, Curran Associates, Inc., 2018.
[26] E. Sarradj, Three-dimensional gridless source mapping using a signal subspace approach, in: BEBEC 2022, Berlin, Germany, 2022.
[27] Y. Zhang, Y. Yang, L. Yang, Y. Wang, Direction-of-arrival estimation for coherent signals through covariance-based grid free compressive sensing, JASA Express Lett. 1 (9) (2021) 094801. doi:10.1121/10.0006389.
[28] C. VanDercreek, R. Merino-Martínez, P. Sijtsma, M. Snellen, Evaluation of the effect of microphone cavity geometries on acoustic imaging in wind tunnels, Appl. Acoust. 181 (2021) 108154. doi:10.1016/j.apacoust.2021.108154.
[29] G. Chardon, F. Ollivier, J. Picheral, Localization of sparse and coherent sources by orthogonal least squares, The J. Acoust. Soc. of Am. 146 (6) (2019) 4873–4882. doi:10.1121/1.5138931.