BeBeC-2022-D13

GRIDLESS BEAMFORMING: THEORETICAL ANALYSIS OF THE ONE SOURCE CASE, AND SPARSITY BASED METHODS FOR MULTIPLE SOURCES

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

Abstract

We present a theoretical analysis of the beamforming method in the frequency domain for the estimation of the position and power of an acoustical source. In particular, the bias and variance of the estimation of the parameters of the source are considered, allowing a definitive assessment of the steering vector formulations encountered in the literature. In the Gaussian source and noise case, an asymptotically unbiased and minimal variance estimation is obtained by combining two formulations. The case of multiple sources is tackled by a infinite dimensional version of the Covariance Matrix Fitting problem, the optimization problem implicitly considered when using the DAMAS algorithm. This problem is solved by the Sliding Frank-Wolfe algorithm, without constraints on the array shape or source model, demonstrated with numerical and experimental results in 3D settings.

1 INTRODUCTION

Beamforming is one of the most popular and simple source localization methods. Its output is an estimation of the acoustical power coming from points in a region of space where sources are searched. From a propagation model, steering vectors are computed, allowing to estimate the power coming from a given direction or point in space. Several formulations of the steering vectors have been proposed, each with their advantages and drawbacks [9]. However, a clear and definitive assessment of the different definitions of the steering vectors was not available. We present here recently published results [2]. In particular, the variances of the estimation of the power and position of an acoustical source are analyzed, allowing a complete comparison between the formulations. Additionally, we show that maximum likelihood estimation of the parameters of the sources, an unbiased and statistically efficient estimator, is obtained by combining two formulations.

In the second part of the paper, a gridless beamforming deconvolution method is proposed. Indeed, beamforming is limited in resolution, as closely spaced sources cannot be separated, especially a low frequencies. In such cases, deconvolution methods can be used to reveal the sources. Several methods have been proposed, based on iterative algorithms, or optimization problems. A famous case is the DAMAS algorithm [1], recently shown to be equivalent [4] to the CMF method [12].

We propose here an infinite dimensional version of the CMF optimization problem, solved using the Sliding Frank-Wolfe algorithm [5]. Simulation and experimental results show the superior performances of this method compared to CLEAN-SC, OMP [8] and the finite dimensional CMF method.

In both sections, emitted or received acoustical signals are modeled as random variables, and the data used to estimate the positions and powers of the sources is an estimation of the spatial covariance matrix of the measurements $\hat{\Sigma}$. Sources are assumed to be independent, and the measurement noise is supposed to be white, with known power σ^2. The acoustical signals are segmented in short windows, and a Fourier analysis is then performed, yielding a vector of complex amplitudes of the signals at each microphone $\mathbf{s}_j$, at a given frequency of interest and for the j-th segment. An estimation of the spatial covariance matrix is obtained by

$$\hat{\Sigma} = \frac{1}{J}\sum_{j=1}^{J} \mathbf{s}_j \mathbf{s}_k^H \tag{1}$$

where J is the number of segments and $\cdot^H$ denotes Hermitian conjugation. The sound propagation is modeled by the array manifold $\mathbf{a}(\mathbf{x})$, containing the complex amplitudes of the acoustical signals at the microphones of the array for a source of unit amplitude at position $\mathbf{x}$. Sources are assumed to be located in a region Ω of the space.

2 LOCALIZATION OF ONE SOURCE

In cases where only one source is present, localization and quantification of the power of the source can be achieved by beamforming. Given a steering vector $\mathbf{h}(\mathbf{x})$, the estimated position $\hat{\mathbf{x}}$ of the source is usually obtained by solving the optimization problem

$$\hat{\mathbf{x}} = \operatorname*{argmax}_{\mathbf{x}\in\Omega} \mathbf{h}(\mathbf{x})^H \hat{\Sigma} \mathbf{h}(\mathbf{x}) \tag{2}$$

and the estimated power of the source is given by

$$\hat{p} = \mathbf{h}(\hat{\mathbf{x}})^H \left(\hat{\Sigma} - \sigma^2 \mathbf{I}\right) \mathbf{h}(\hat{\mathbf{x}}) \tag{3}$$

assuming that the variance of the noise σ^2 is known.

Cases where multiple sources are present, with sufficient distance between the sources, can also be tackled by this method. We present here the main results of the analysis. Further results and details are available in [2].

2.1 Steering vector formulations

While the choice of the steering vector is evident in the case of far-field sources, several formulations of the steering vector have been proposed for three-dimensional localization of sources. We recall here the four formulations frequently used in the literature [9]

Formulation I This formulation is similar to the far-field case, using only the phase of the source vector:

$$h_n^{\mathrm{I}}(\mathbf{x}) = \frac{1}{N}\frac{a_n(\mathbf{x})}{|a_n(\mathbf{x})|}. \tag{4}$$

Formulation II Here, amplitudes are also compensated:

$$h_n^{\mathrm{II}}(\mathbf{x}) = \frac{1}{N}\frac{a_n(\mathbf{x})}{|a_n(\mathbf{x})|^2}. \tag{5}$$

Formulation III In this formulation, the source vector is divided by its squared ℓ_2 norm:

$$\mathbf{h}^{\mathrm{III}}(\mathbf{x}) = \frac{\mathbf{a}(\mathbf{x})}{\|\mathbf{a}(\mathbf{x})\|_2^2}. \tag{6}$$

Formulation IV Finally, in formulation IV, steering vectors are normalized:

$$\mathbf{h}^{\mathrm{IV}}(\mathbf{x}) = \frac{1}{\sqrt{N}}\frac{\mathbf{a}(\mathbf{x})}{\|\mathbf{a}(\mathbf{x})\|_2}. \tag{7}$$

2.2 Performances of the steering formulations

For a source with known position the following theorem [2] gives the bias and variance of the estimation of the power of the source, in function of the steering vector used for the estimation.

Theorem 1 *The estimated power $\hat{p}$ of a source with known position $\mathbf{x}_s$ has the following mean and variance:*

$$\mathrm{E}(\hat{p}) = |\langle \mathbf{a}(\mathbf{x}_s), \mathbf{h}(\mathbf{x}_s)\rangle|^2 p \tag{8}$$

$$\mathrm{Var}(\hat{p}) = \frac{1}{S}\left(|\langle \mathbf{a}(\mathbf{x}_s), \mathbf{h}(\mathbf{x}_s)\rangle|^2 p + \|\mathbf{h}(\mathbf{x}_s)\|_2^2\sigma^2\right)^2. \tag{9}$$

Values of the terms appearing in the expectation and variance of the estimation are given in Table 1, showing that, as already observed, formulations II and III yield unbiased estimations of the power of the source, assuming that its position is known. Additionally, the variance of formulation III can be shown to be lower than the variance of formulation II by an application of the Cauchy-Schwarz inequality. Consequently, formulation III yields the lowest mean squared error (MSE) for the estimation of the power of the source.

	$\langle \mathbf{a}(\mathbf{x}_s), \mathbf{h}(\mathbf{x}_s) \rangle$	$\|\mathbf{h}(\mathbf{x}_s)\|_2^2$
I	$\frac{1}{N}\sum_{n=1}^{N}\frac{r_0}{r_n}$	$\frac{1}{N}$
II	1	$\frac{1}{N^2}\sum_{n=1}^{N}\frac{r_n^2}{r_0^2}$
III	1	$\frac{1}{\sum_{n=1}^{N} r_0^2/r_n^2}$
IV	$\sqrt{\frac{1}{N}\sum_{n=1}^{N}\frac{r_0^2}{r_n^2}}$	$\frac{1}{N}$

Table 1: Values of $\langle \mathbf{a}(\mathbf{x}_s), \mathbf{h}(\mathbf{x}_s) \rangle$ and $\|\mathbf{h}(\mathbf{x}_s)\|_2^2$ for the four formulations.

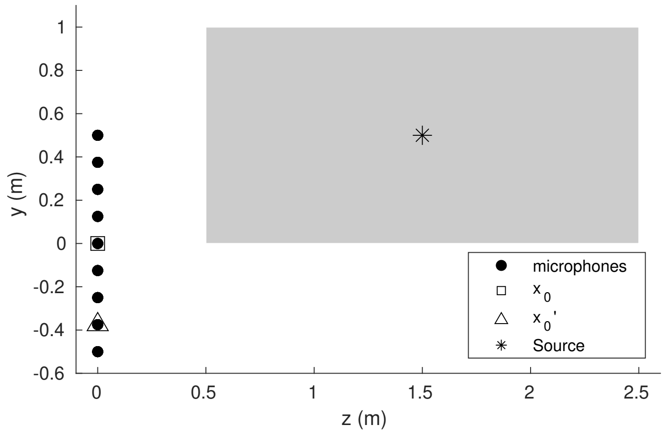

Figure 1: Geometry of the problem. Region of interest for Figure 2 indicated by the grey rectangle.

Performances of the estimation of the position of the source are now analyzed, using numerical simulations. The geometry of the problem is given on Figure 1. Microphones are placed on a regular square grid in the *xy* plane. Two reference points $\mathbf{x}_0$ and $\mathbf{x}_0'$ will be used.

Figure 2 shows the value of the beamforming criterion for the four formulations and the two reference points. As already known, formulations I and IV yield unbiased estimations, while formulations II and III are biased. In addition, it is shown here that the bias depends on the reference point (cf. II' and III'), and estimated positions can be further away from the array than the actual position of the source. Formulations I and IV are not affected by the choice of the reference point.

Additional simulations using a cubic array showed that formulation IV has a lower variance than formulation I.

Figure 3 gives the bias, variance, and MSE of the estimated position in function of the frequency. Formulations I and IV yield the lowest MSE. Low variances of formulations II and III

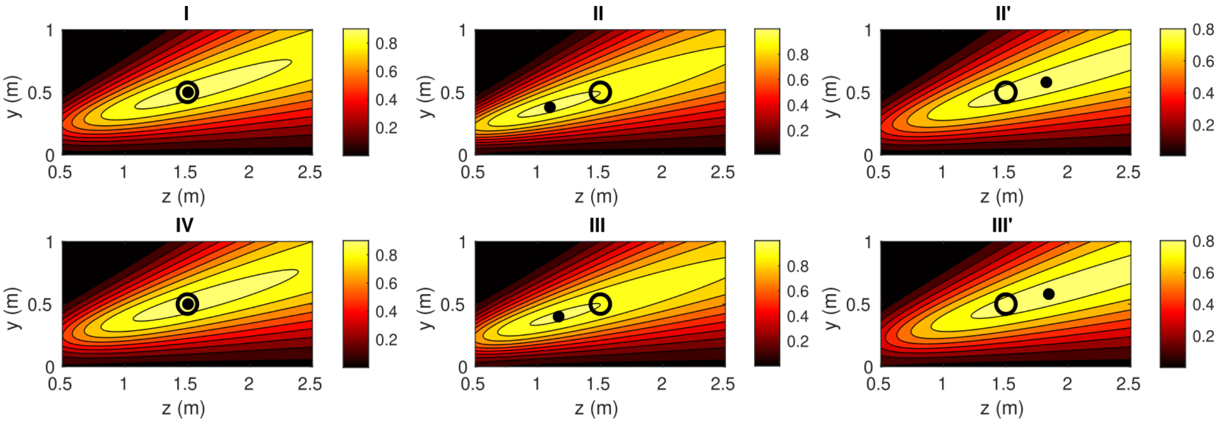

Figure 2: Value of the beamforming criterion for the four formulations (linear scale) in a planar region. For formulations II and III, results for reference points $\mathbf{x}_0$ (II and III) and $\mathbf{x}'_0$ (II' and III') are given. Position of the source indicated by a circle, maximum value of the criterion by a disk.

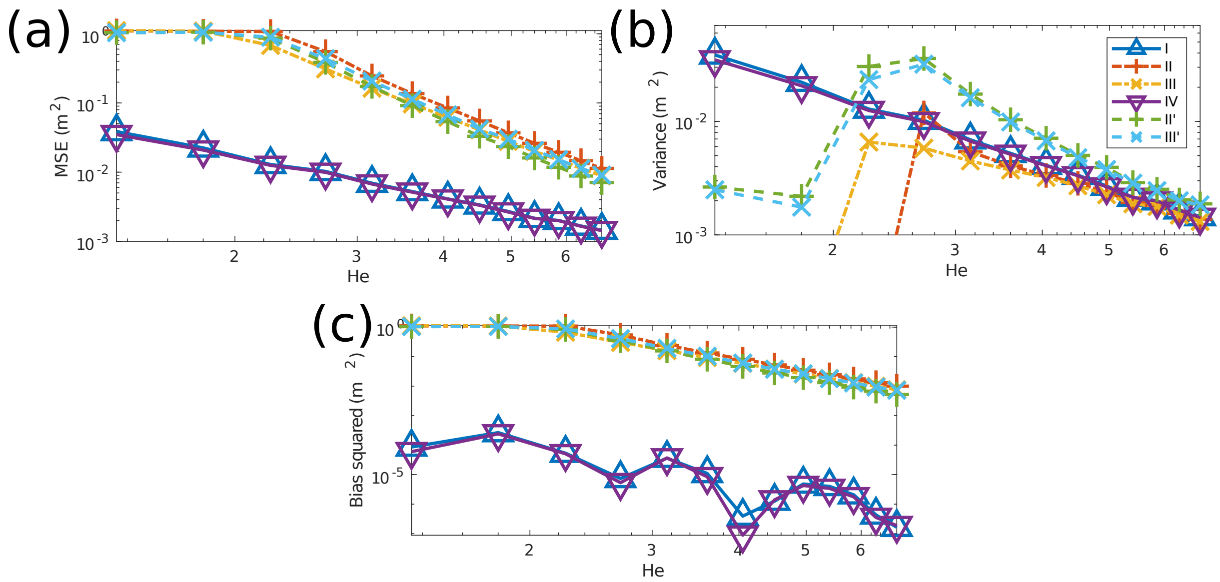

Figure 3: (a) MSE, (b) variance and (c) squared norm of the bias of the estimation of the position, in function of H_e.

at low frequencies are explained by the fact that the estimated position is constrained in the grey zone shown on Figure 1.

Finally, performances of the estimation of the power, with unknown position, are considered. The left panels of Figure 4 shows the bias, variance, and MSE of the estimation of the power received at the reference point $\mathbf{x}_0$. As expected, formulations I and IV are biased. Although formulations II and III where shown to be unbiased in cases where the source position is known, the power is here overestimated. This is actually expected, as the position is selected as the one maximizing the beamforming criterion, with a value necessarily higher than the value at the

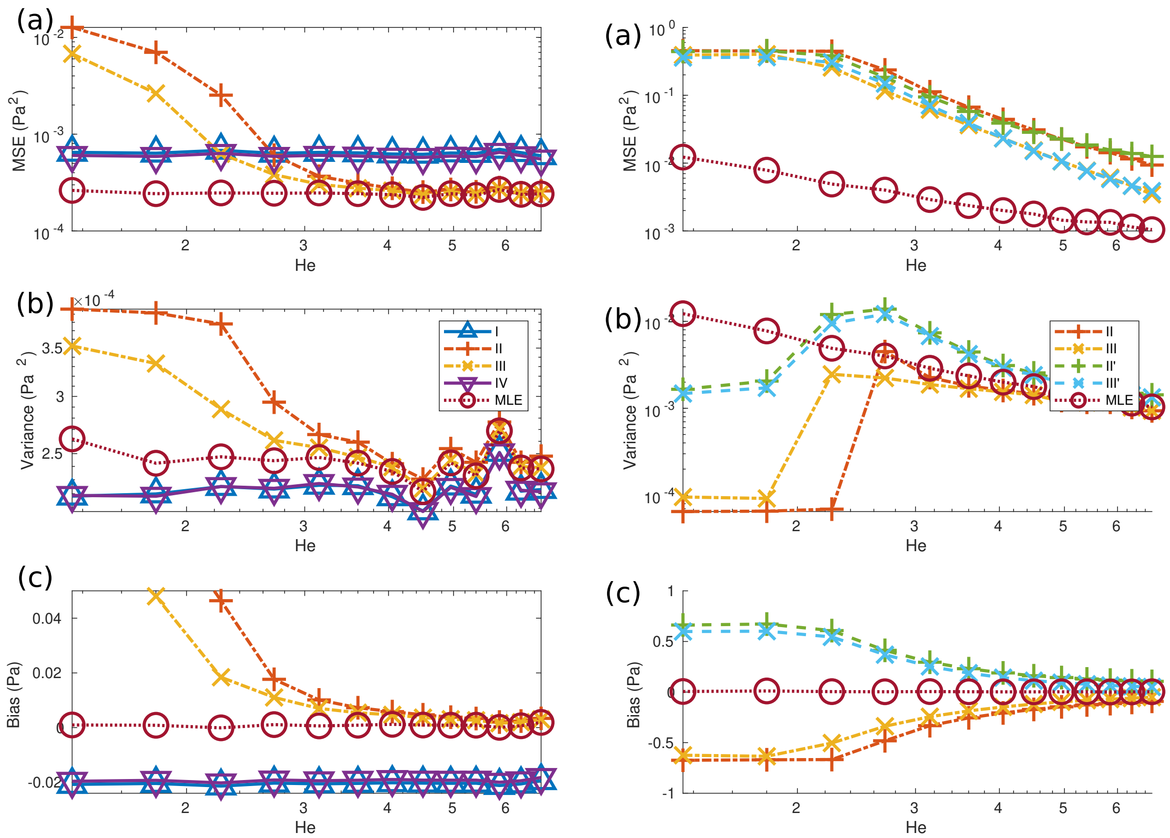

Figure 4: (a) MSE, (b) variance and (c) bias of the estimated amplitude at the reference point $\mathbf{x}_0$ *(left), and comparison with reference point* $\mathbf{x}'_0$ *(right)*

actual position, which is unbiased.

Performances of formulations II and III are also given for the estimated power at a fixed distance to the source (which is proportional to the power emitted by the source), for the two reference points on the right panels of Figure 4. Here, power is underestimated for the reference point $\mathbf{x}_0$. As the distance between the source and the reference point is underestimated, so is the power emitted by the source. Conversely, the power is overestimated for the reference point $\mathbf{x}'_0$ (cf. II' and III').

As a conclusion, while unbiased power estimation in the known position case, and unbiased estimation of the position are possible, unbiased estimation of the power when the position is unknown is not possible using a simple beamforming method as used here. The situation is even worse when the power emitted by the source is to be estimated, which is sensitive to position estimations errors.

2.3 Maximum likelihood estimation

In the far-field case, beamforming can be shown to be a maximum likelihood estimation (MLE) [7] of the power and direction of a source, assuming a Gaussian model for the source and the noise. Maximum likelihood estimation can also be applied to the three dimensional case, as shown by the following theorem [2].

Theorem 2 *With*

$$B(\mathbf{x}) = \frac{\mathbf{a}(\mathbf{x})^H \hat{\Sigma} \mathbf{a}(\mathbf{x})}{\|\mathbf{a}(\mathbf{x})\|_2^2} \tag{10}$$

$$= \mathbf{h}^{\mathrm{IV}}(\mathbf{x})^H \hat{\Sigma} \mathbf{h}^{\mathrm{IV}}(\mathbf{x}), \tag{11}$$

the MLE for the position of the source and its power are given by

$$\hat{\mathbf{x}} = \operatorname*{argmax}_{\mathbf{x}\in\Omega} \frac{B(\mathbf{x})}{\sigma^2} - \log B(\mathbf{x}) \tag{12}$$

$$\hat{p} = \frac{1}{\|\mathbf{a}(\hat{\mathbf{x}})\|_2^2} \left(B(\hat{\mathbf{x}}) - \hat{\sigma}^2\right) \tag{13}$$

$$= \mathbf{h}^{\mathrm{III}}(\hat{\mathbf{x}})^H \left(\hat{\Sigma} - \sigma^2 \mathbf{I}\right) \mathbf{h}^{\mathrm{III}}(\hat{\mathbf{x}}) \tag{14}$$

In practice, the optimization problem (12) has the same solution as maximizing $B(\mathbf{x})$.

Maximum likelihood estimation is known to be asymptotically unbiased and efficient (in the sense that its variance reaches the Cramér-Rao bound). Therefore, unbiased and efficient estimation of the position and power of a source is achieved by combining formulation IV to estimate the position, and formulation III to estimate the power. Figure 4 shows that the estimated power by MLE is unbiased and has the lowest MSE in all cases (received and emitted power, choice of the reference point).

We conclude this section by remarking that these results are obtained under a Gaussian assumption on the source and the noise. Difference source and noise models will imply different ML estimators, not necessarily similar to beamforming.

3 LOCALIZATION OF SEVERAL SOURCES

We now consider the localization of several sources. In this case, maximum likelihood for the estimation of the parameters of the sources is intractable, as it necessitates to solve a high-dimensional non-convex problem. While the output of beamforming can be used to localize multiple sources in cases where they are sufficiently separated, more advanced methods are usually necessary.

3.1 Deconvolution based methods

Deconvolution methods aim at solving an inverse problem in order to reveal the distribution of sources from the beamforming map. We note that except in particular cases, these methods do not perform deconvolution in a strict sense, as the direct problem cannot be modeled as a convolution.

Two classes of methods can be found in the literature :

- optimization based methods, such as CMF [12] or SEM, DAMAS-NNLS, SODIX, etc.
- iterative methods, such as DAMAS [1], CLEAN-SC [10], OMP [8], etc.

In particular, the DAMAS algorithm sparked the interest in such methods. While its performances are appreciated, it has the drawback of a large computational complexity. It was recently proved [4] that the DAMAS algorithm actually solves the CMF optimization problem

$$\underset{\mathbf{P}\in\mathscr{D}_+^N}{\operatorname{argmin}} \|\mathbf{APA}^H + \sigma^2\mathbf{I} - \hat{\Sigma}\|_F^2 \tag{15}$$

where $\mathbf{A}$ is the matrix collecting the values of $\mathbf{a}(\mathbf{x})$ on a discrete grid, $\mathbf{P}$ the diagonal matrix containing the estimated powers of the sources on the grid, and σ^2 is the noise level. $\mathscr{D}_+^N$ is the set of diagonal matrices with nonnegative coefficients. Algorithms more efficient than DAMAS exist to solve this problem [4]. It was also shown that the nonnegativity constraint of CMF is sufficient to ensure the sparsity of its solution [11], retrospectively explaining the performances of DAMAS for sparse distributions of sources.

3.2 Gridless deconvolution

The proposed gridless deconvolution method is based on a infinite-dimensional formulation of the CMF problem, where instead of a finite dimensional vector containing the powers of sources located on a finite grid, a measure is used to model the distribution of sources. We recall here that a measure is a function mapping a set to a real positive number, with additivity properties. Here, the measure modeling the distribution of sources maps a subset of the domain of interest to the total power of the sources contained in the set. A particular example of measure is the Dirac mass $\delta_{\mathbf{x}_0}$ modeling a punctual source at $\mathbf{x}_0$ of unit power.

The covariance matrix of the measurements can be written as the integral

$$\Sigma = \int_\Omega \mathbf{a}(\mathbf{x})\mathbf{a}(\mathbf{x})^H d\nu(\mathbf{x}) \tag{16}$$

which, for a finite number K of punctual sources at positions $\mathbf{x}_k$ and with powers p_k, modeled by

$$\nu = \sum_{k=1}^K p_k \delta_{\mathbf{x}_k}, \tag{17}$$

reduces to

$$\Sigma = \sum_{k=1}^K p_k \mathbf{a}(\mathbf{x}_k)\mathbf{a}(\mathbf{x}_k)^H \tag{18}$$

The infinite dimensional CMF problem is the formulated as

$$\mu^\star = \underset{\mu\in\mathscr{M}^+}{\operatorname{argmin}} \left\| \int_\Omega \mathbf{a}\mathbf{a}^H d\mu + \sigma^2\mathbf{I} - \hat{\Sigma} \right\|_{\text{Fro}}^2 . \tag{19}$$

where $\mu^\star$ is the measure modeling the distribution of sources (usually a discrete measure as in Eq. (17)), σ^2 the noise level, and $\mathscr{M}^+$ the set of measures defined on Ω.

To solve this problem, we used the Sliding Frank-Wolfe algorithm[5], which consists of the following iterations:

1. identify a new source by picking the maximum of the beamforming map,
2. optimize the powers of the identified sources,
3. optimize their powers and positions jointly,
4. remove their contributions from the covariance matrix,
5. repeat until convergence.

In most cases the algorithm stops in a finite number of iterations. One can also stop the algorithm when a fixed number of sources have been identified.

While the problem is formulated without discretization, the algorithm still uses a discretized grid for the initialization of the identification of a new source in the first step of the algorithm. Under the condition that the initialization grid is fine enough to identify the global maximum of the beamforming map, the grid has no influence on the result of the algorithm.

We note that no regularization parameters, or number of steps, etc., have to be set.

3.3 Results : experimental data

The method is tested on experimental data. Four sources at a distance of approximately 4.4m from a microphone array are emitting uncorrelated noises. The array consists of 128 microphones, with an aperture of 2.6m. The central frequency of the Fourier analysis is 1818Hz. More details on the experiment are available in [3].

The infinite dimensional CMF (CMF-SFW) method is compared with beamforming, CMF and OMP.

On the beamforming map computed in the source plane, two of the sources cannot be separated. Results of CMF, with a grid step of 0.05m, and 26901 nodes, show that the sources are spread over multiple points. For OMP (with four iterations, in a gridless formulation), estimation of the positions of the sources is inaccurate, in particular for the two central sources. Finally CMF-SFW is able to localize the sources with reasonable accuracy, and limited spread of the sources over multiple points. Running times are 0.27s for OMP, 17.8s for CMF, 29.3s for CMF-SFW.

3.4 Results : simulations

Simulations were also used to assess the performances of CMF-SFW in comparison to CLEAN-SC and OMP (both in gridless formulations). Four uncorrelated sources are simulated, with positions drawn randomly in a cube of side 2m, with center at a distance of 4m to the center of the array.

Performances were evaluated using measures similar to those proposed by Herold and Sarradj [6], with the modification that performances are measured in power, not in dB levels (this avoids infinite errors when sources are not correctly localized). Mean squared powers of sources

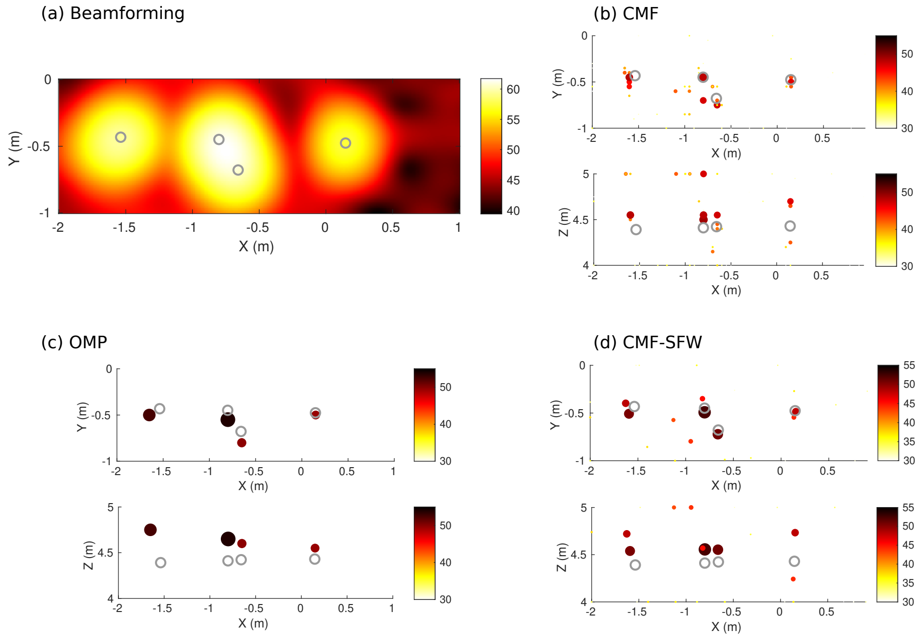

Figure 5: Experimental results of (a) beamforming, (b) CMF, (c) OMP, (d) CMF-SFW.

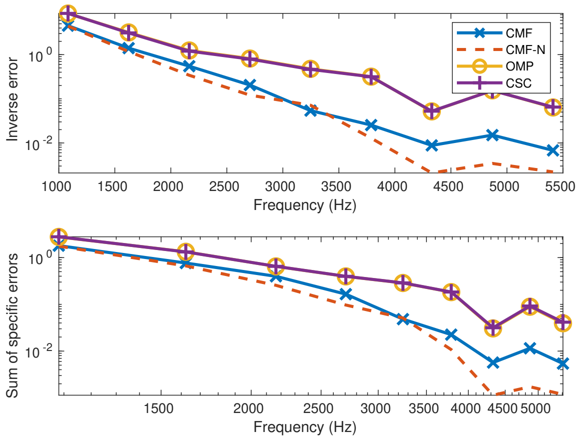

Figure 6: Performances of CMF, CMF-N, OMP and CLEAN-SC in function of the frequency. Top : inverse level error. Bottom: sum of the specific level errors.

localized outside of the union of spheres of radius $r = 0.05$m centered on the actual sources are plotted on figure 6-top, in function of the frequency. The power of a source is estimated by summing the powers of the estimated sources that are closest to the source in a sphere of radius r. Mean squared errors summed over the four sources are plotted on figure 6-bottom. In both cases, CMF-SFW with four iterations has the best performances. CMF-SFW run until convergence has better performances than CLEAN-SC and OMP, while these two methods are run knowing the number of sources.

4 CONCLUSION

A definitive assessment of steering vector formulations for beamforming is given, with the main result that asymptotically unbiased and efficient estimation of the position and power of a source is obtained by combining two beamforming steering vector formulations. These results were obtained under the assumption that the measurement noise is white and its power is known.

A gridfree method for multiple source localization is proposed, based on a infinite dimensional formulation of the CMF problem. First results show that this method has good performances. Further evaluations of the method with the state of the art will be performed.

References

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

[2] G. Chardon. "Theoretical analysis of beamforming steering vector formulations for acoustic source localization." *Journal of Sound and Vibration*, 517, 116544, 2022. doi: https://doi.org/10.1016/j.jsv.2021.116544.

[3] G. Chardon, F. Ollivier, and J. Picheral. "Localization of sparse and coherent sources by orthogonal least squares." *The Journal of the Acoustical Society of America*, 146(6), 4873–4882, 2019. doi:10.1121/1.5138931.

[4] G. Chardon, J. Picheral, and F. Ollivier. "Theoretical analysis of the DAMAS algorithm and efficient implementation of the covariance matrix fitting method for large-scale problems." *Journal of Sound and Vibration*, 508, 116208, 2021. doi:10.1016/j.jsv.2021.116208.

[5] Q. Denoyelle, V. Duval, G. Peyré, and E. Soubies. "The sliding Frank–Wolfe algorithm and its application to super-resolution microscopy." *Inverse Problems*, 36(1), 014001, 2019. doi:10.1088/1361-6420/ab2a29. Publisher: IOP Publishing.

[6] G. Herold and E. Sarradj. "Performance analysis of microphone array methods." *Journal of Sound and Vibration*, 401, 152–168, 2017. doi:10.1016/j.jsv.2017.04.030.

[7] E. Lehmann and G. Casella. *Theory of Point Estimation*. Springer, second edition, 1998. Chapter 6.

[8] T. Padois and A. Berry. "Orthogonal matching pursuit applied to the deconvolution approach for the mapping of acoustic sources inverse problem." *The Journal of the Acoustical Society of America*, 138(6), 3678–3685, 2015. doi:10.1121/1.4937609.

[9] E. Sarradj. "Three-Dimensional Acoustic Source Mapping with Different Beamforming Steering Vector Formulations." *Advances in Acoustics and Vibration*, 2012, e292695, 2012. doi:10.1155/2012/292695. Publisher: Hindawi.

[10] P. Sijtsma. "CLEAN Based on Spatial Source Coherence." *International Journal of Aeroacoustics*, 6(4), 357–374, 2007. doi:10.1260/147547207783359459. Publisher: SAGE Publications.

[11] M. Slawski and M. Hein. "Non-negative least squares for high-dimensional linear models: Consistency and sparse recovery without regularization." *Electronic Journal of Statistics*, 7, 3004–3056, 2013. doi:10.1214/13-EJS868.

[12] T. Yardibi, J. Li, P. Stoica, and L. N. Cattafesta. "Sparsity constrained deconvolution approaches for acoustic source mapping." *The Journal of the Acoustical Society of America*, 123(5), 2631–2642, 2008. doi:10.1121/1.2896754.